

Skidmore College

Creative Matter

Economics Student Theses and Capstone
Projects

Economics

2017

NFL Betting: Is the Market Efficient?

Mathew Marino
Skidmore College

Follow this and additional works at: https://creativematter.skidmore.edu/econ_studt_schol



Part of the [Other Economics Commons](#)

Recommended Citation

Marino, Mathew, "NFL Betting: Is the Market Efficient?" (2017). *Economics Student Theses and Capstone Projects*. 27.

https://creativematter.skidmore.edu/econ_studt_schol/27

This Thesis is brought to you for free and open access by the Economics at Creative Matter. It has been accepted for inclusion in Economics Student Theses and Capstone Projects by an authorized administrator of Creative Matter. For more information, please contact dseiler@skidmore.edu.

NFL Betting: Is the Market Efficient?

By

Mathew Marino

A Thesis Submitted to

Department of Economics

Skidmore College

In Partial Fulfillment of the Requirement for the B.A Degree

Thesis Advisor: Qi Ge

May 2, 2017

Abstract

The Purpose of this research paper is to analyze the NFL point spread and Over/Under betting market and determine if it follows the efficient market hypothesis. This research paper uses data from armchairanalysis.com for the 2000 NFL season through the 2015 NFL season including playoffs. I use OLS and probit regressions in order to determine NFL betting market efficiency for the point spread betting market and the Over/Under betting market. The results indicate, that as a whole, the NFL betting market appears to be efficient. However, there may be behavioral tendencies that can be taken advantage of in order to make consistent betting profits.

I. Introduction

According to Statista.com (2015), the sports gambling market has been estimated to be worth between \$700 billion and \$1 trillion, but the exact size is difficult to measure because it is not legal everywhere. The most popular sport to bet on is professional American football, which is known as the National Football League (NFL). The increase in the amount of money being wagered and popularity has led financial economists to become more interested in the sports industry and study the efficiency of the market (Borghesi, 2006). The Efficient Market Hypothesis (EMH), as formalized by Fama (1970), is a theory that states financial markets are efficient if the prices of stocks contain all publicly available information. That is, stocks are always traded at their fair value and it is impossible to buy undervalued stocks for profit or sell overpriced stocks for profit in the long run. In more basic terms, it should be impossible for an individual to “beat the market” because the price of a good should include all publicly available information.

In recent years, the efficient market hypothesis has been applied to sports betting markets because testing for efficiency in the financial markets is very difficult, if not impossible. The main difficulty with testing the efficient market hypothesis in the financial markets is that there is no direct test for market efficiency because the true value of the investment is never revealed (Fama, 1970, Gray & Gray, 1997). A long period of time exists between the purchase of a stock and the sale of that stock, making it very difficult to measure the efficiency. Due to this inability to accurately test the efficiency of the financial markets, many economists have turned to sports betting to test for market efficiency. The sports betting market provides a perfect testing ground for market efficiency because there are many similarities between the two markets. There are many participants in both markets, large sums of money are exchanged, and the betting lines are formed based on publicly available information, similar to stock prices. Bookmakers act in a

similar fashion as stock brokers; both charge fees for matching buyers and sellers (Borghesi, 2006). Other analogous variables exist in both the stock market and the NFL betting market. Gray and Gray (1997) cite that a team's recent record is similar to the recent performance of the stock market performance. Both stocks and teams that have been performing well as of late, attract more attention and more people would like to bet or invest on that team or stock. In addition, according to the efficient market hypothesis, past performance of stocks should not help investors predict the future performance. In terms of football, if Team A won their game last week, that should have no effect on whether they win their game this week. The essence of the efficient market hypothesis is to determine if an investor could earn abnormal returns compared to the market returns by looking at a stock's recent performance or other related variables. In football betting, could a bettor earn profits based on the recent performance of the team or other related variables that could affect the game outcome such as weather or injuries.

By studying the 16 NFL seasons from 2000 through 2015, this study contributes to the discussion on NFL betting and market efficiency. This paper examines the point spread betting market that was first analyzed by Pankoff (1968) by using performance measures (Gray & Gray, 1997, Zuber et al. 1985, 1988) as well as exogenous factors such as stadium characteristics (Boulier et al. 2007) and weather (Borghesi, 2006). My paper will also investigate the Over/Under NFL betting market using similar variables that were used to test the point spread market. To my knowledge, through reading related literature, no paper has touched upon this area, and this is certainly a feature of NFL betting that is interesting to analyze.

My findings for the point spread betting market are consistent with those of previous literature. By testing the NFL point spread market for 16 seasons, I was unable to find variables that would allow a bettor to earn abnormal returns and profit in the long run, leading me to the

conclusion that the NFL betting market is indeed efficient. My paper does find a statistical inefficiency in the Over/Under betting market that may allow a bettor to earn abnormal returns. The $Dome_{it}$ variable appears to increase the likelihood of the score of the game going over the Over/Under betting line. However, this single variable may not be enough to base a betting strategy on.

The rest of this paper is organized as follows. Section 2 discusses the NFL betting market and explains the nuances that exist. Section 3 examines existing literature that is relevant to this topic. Section 4 describes the data that will be used in this study. Section 5 introduces four econometric models, two standard OLS regressions and two probit regressions. This section will also explain the variables used in the models. Section 6 discusses the results from the two econometric models. Section 7 offers my concluding remarks.

II. The NFL Betting Market

The betting market contains jargon that must be known to fully understand the rest of this paper and interpret the results. In NFL betting, there are two types of bets, the money line and the point spread. A bet on the money line is when you bet that Team A will beat Team B straight up. You expect Team A to beat Team B, or vice versa, there is no point spread involved. A bet on the point spread is when you “give” a team points in an effort to level the playing field, to make bets for either team equally attractive. A point spread is produced by bookmakers when one team is considered better than another team, in order to equalize the betting volume on each team. The point spread is negative for the favorite and positive for the underdog. The favorite has to win by at least the spread margin and the underdog has to win outright or lose by less than the margin. For example, Team A is the favorite with a spread of -3 and Team B is the underdog with a spread of

+3. Team A has to win the game by more than 3 points in order to win against the spread, also called covering the spread. Inversely, Team B has to lose the game by 3 points or less or win the game outright. If the difference in the score of the game is exactly 3 points, the result is a push and all money is returned to the bettors. In an effort to avoid ‘pushes’, the spread is usually set as a half number, for example the spread would be -2.5 for Team A or +2.5 for Team B.

The point spread and the Over/Under line are set at the beginning of the week, known as the opening line. The line may adjust throughout the week based on new information, such as injury reports or weather reports, until right before the game starts, known as the closing line. The same holds true for the Over/Under betting line. One important aspect of football betting is that once you place a bet, your bet is locked in at that spread. Unlike horse betting, where no matter when you placed your bet, your bet is based on the closing odds. The closing line is the important betting line, in the sense that it should be the most unbiased predictor of the game outcome, and is the betting line that I use in the paper.

Bets under the spread follow the “Bet \$11 to win \$10” rule. You must risk \$11 to win \$10 in the point spread betting market (Humphreys, 2011). This 10% commission is called *vigorish* and is the commission the bookmaker collects on all bets. In order for a bettor to break-even in the betting market, they must win 52.4% of their bets. This is due to the standard break-even condition used in sports betting. The break-even condition is as follows: $p(10) = (1 - p)11$, when you solve for p , you get 0.524 or 52.4% (Lacey, 2001).

III. Literature Review

Eugene Fama (1970) is one of the earliest authors to write about market efficiency. Fama said a market is considered efficient when the price fully reflects all publicly available

information. Fama (1970) is the first to consider the joint hypothesis problem. The joint hypothesis problem stems from the fact that it is unknown exactly what an efficient market looks like, thus it is hard to prove whether or not the markets are efficient. There is no norm to test the market against and determine if it is efficient. Hamburger and Platt (1975) investigate the idea of the efficient market hypothesis (EMH) or the expectations hypothesis as they called it. The expectations hypothesis states that “the forward rates of interest implicit in the yield curve provide unbiased estimates of the market’s expectations of future spot rates” (Hamburger and Platt, 1975). In more basic terms, does the market efficiently incorporate all publicly available information, such as estimates of income and liquidity (Hamburger and Platt, 1975), when forming expectations of the future rates? Hamburger and Platt (1975) find that the market presents weak form efficiency, meaning the market is relatively efficient and past rates have no effect on future rates. The Efficient Market Hypothesis also has applications in the stock market, in which the prices of securities should accurately reflect all publicly available information, and investors cannot systematically outperform the market (Lee, 1998). Investors should not be able to find stocks that are overpriced or underpriced relative to their true value.

a. EMH in Sports Betting Market

Most current literature focuses on the efficiency of the NFL betting market. Football is by far the most popular sport in the US and attracts the most betting activity among the other major US sports. The largest segment of the sports betting industry’s output is the NFL, which is valued at around \$17 billion (Borghesi 2006). The American Gaming Association estimated in 2015 that \$95 billion was wagered on NFL and college football games. The sports betting market is continuously growing, and could become even larger if it becomes legal throughout the country. If the NFL point spread betting market were to be considered efficient, the spread would

be an unbiased predictor of the actual game outcome. Bettors should not be able to find consistent profitable betting strategies such as always betting on the home team or always betting on the underdog. The results are mixed with authors finding the market has weak form efficiency while some have claimed to have found profitable betting strategies. Authors have found that, in a basic sense, the betting market is efficient, but there are scenarios in which bettors could take advantage of human tendencies to earn profits from betting (Zuber et al. 1988).

Pankoff (1968) analyzes the football point spread betting market during all 856 NFL games, from 1956 to 1965. Pankoff employs a very basic model by regressing the actual point spread outcomes on to the betting lines set by the bookmakers. Pankoff (1968) estimates that the intercept and slope parameters for the regression would be zero and one respectively. Pankoff's (1968) model, while elementary, found the betting lines are an unbiased predictor of the game outcomes and set the ground work for other authors to build upon. More recent literature has recognized how elementary Pankoff's (1968) model is, and believe a more sophisticated model that incorporates more variables is needed.

Golec and Tamarkin (1991) analyze the NFL betting market for efficiency from all regular season and playoff games from the 1973 – 1987 seasons. In addition to previous literature that simply tested for market efficiency, Golec and Tamarkin (1991) introduce two dummy variables for home or away and favorite or underdog. These variables took the value of one if the team was home or if the team was the favorite and a value of zero if the team was away or if the team was the underdog for that game. If the efficient market hypothesis is to hold true, these added variables would have no value and would not create a bias in determining the spread of the game because they would have already been factored into the spread originally. Their results show that during the 15 NFL seasons from 1973-1987, the only statistically

significant winning strategy was to bet on the home underdog. Betting on the home underdog was a profitable strategy, winning 55.6% of the time, which is enough to cover the standard break even condition discussed earlier. During a shorter period from 1973-1979 the winning percentage of betting on the home underdog was even higher at 58.1%.

Around the same time, Dare and MacDonald (1996) also test for market efficiency in the NFL. They disagree with the models that Golec and Tamarkin (1991) use, stating that their model makes one of the dummy variables in the system a unit vector. Dare and MacDonald (1996) claim this makes the model fall short of a complete solution and can lead to biased coefficients. Dare and MacDonald (1996) argue the variables home or away, and favorite or underdog are interdependent and can result in restrictions in the model. More often than not, the home team is also the favorite to win the game, thus there is interdependence between the two variables. (Dare and Holland, 2004). Including variables for both home/away and favorite/underdog, can produce skewed results. In their study, Dare and MacDonald (1996) ignore games that were played at a neutral site as well as games that were considered a pick'em. A pick'em is when the two teams playing are evenly matched and a point spread is not published for the game. To solve the problem with Golec and Tamarkin's (1991) model, Dare and MacDonald (1996) impose restrictions on the estimates to account for interdependence between the home dummy variable and the favorite variable. After creating these restrictions, Dare and MacDonald (1996) find that the NFL point spread market is efficient and disagree with the findings of Golec and Tamarkin (1991).

b. Performance measures

Zuber, Gandar, and Bowers (1985) use data from the 1983 NFL regular season and run a similar regression to Pankoff (1968) and find efficiency in the market. However, Zuber et al.

(1985) believe Pankoff's (1965) model was too simple, so they created a different model that incorporated more variables because they believed the actual result of any football game can be viewed as the net outcome of the simultaneous efforts of the two teams involved (Zuber, Gandar, and Bowers, 1985, p.3). They divide the test into 16 parts, one part for each week of the season. The independent variable is a vector of coefficients that included variables such as rushing yards, passing yards, wins prior to the game, fumbles, interceptions, number of penalties and the number of rookies. All the variables were statistically significant and have their expected sign. Zuber, Gandar, and Bowers (1985) conclude that their model could be used to predict the point spreads of NFL games because speculative inefficiencies exist in the market. They note however, even though their results suggest speculative inefficiencies exist in the market, that does not imply market inefficiency. Sauer (1988) disagrees with Zuber et al. (1985) and argues that the results they found that suggest the NFL betting market is inefficient are inaccurate. Sauer (1988) claims that because Zuber et al. (1985) divided the test into 16 parts, they increased the sample variance of the estimators making it harder to reject the efficiency hypothesis. Sauer (1988) claims that splitting a large sample into many small samples creates a weaker test. Sauer (1988) argues if Zuber et al. (1985) kept the sample size at 224 rather than the 16 that they used, they would not have found any speculative inefficiencies. Sauer (1988) also tests the betting strategies, that Zuber et al. (1985) claim were profitable, during the 1984 NFL season and found that the same betting strategies did not return profits, and instead they incurred substantial losses. I agree with Sauer (1988) that splitting the test into 16 parts produced inaccurate results. Furthermore, the results found Zuber et al. (1985) may be inaccurate because they only tested one NFL season. The same results may not hold true over a larger sample size.

Zuber, Gandar, O'Brien, and Russo (1988), build upon Zuber, Gandar, and Bowers (1985) just a few years later by testing the market efficiency for the 1980-1985 NFL regular seasons. They use the same variables as Zuber et al. (1985) and find no significant bias in the closing lines. Interestingly, Zuber et al. (1988) note that the order of teams in the differencing in score has an impact on the regression results, meaning the spread in terms of the home team, the away team, or the favorited team etc. Zuber et al. (1988) suggest that regression tests are too weak to reject market efficiency, but economic tests could reject market efficiency. The economic tests they created were to test for market rationality. Essentially, they wanted to test if people were making rational decisions when betting. They create seven tests for market efficiency, four based on technical rules developed by Vergin and Scriabin (1978), and three rules based on rationality of betting behavior. Vergin and Scriabin (1978) found the four mechanical rules to produce winning percentages above the break-even point of 52.4%, but Zuber et al. (1988) did not find that to be the case during their sample. However, all three behavioral rules did produce profitable results. The three behavioral rules that were constructed earned abnormal profits relative to the break-even point. These results lead Zuber et al. (1988) to believe that the point spread betting market tends to be dominated by unsophisticated bettors rather than knowledgeable bettors. Paul and Weinbach (2011) also find similar profitable betting strategies. They find that many uninformed bettors tend to dominate the last hour of betting in the NFL. Betting on the underdog, when the percentage bet on the favorite increases in the last hour, wins over 60% of the time (Paul & Weinbach, 2011). This shows that while market efficiency cannot be rejected, there appears to be behavioral tendencies that can be exploited to earn abnormal betting profits. One limitation of this paper is that they only analyze six NFL seasons. This limits the amount of games they are able to examine and may not create a

completely accurate result. Also, they wait until week nine of every season in order to create the performance variables that they use. This also decreases the number of observations they have in their sample and will not produce the most accurate result. Interestingly, by including more seasons they find different results than their previous paper and cannot not conclude that the market is inefficient.

Gray and Gray (1997) test the efficiency of the NFL point spread market over a sample of 4,219 NFL games from 1976 to 1994. They use a probit regression model similar to Golec and Tamarkin (1991) employing dummy variables for home or away and favorite or underdog as well. To capture the effect of perceived runs in a series of games, Gray and Gray (1997) include the past performance of the teams to analyze if it is efficiently included in the point spread. They offer two reasons for including this variable. First, momentum or contrarian investment literature suggests that past performance of a stock can be a signal of strength and increase demand which will increase price in the future. Second is the psychological idea of people believing in streaks. People tend to follow the 'herd' or the hot hand, as discussed in Wever and Aadland (2012), and more people bet on the teams that are playing well at the moment. This leads to bettors having a bias towards the hot teams and not recognize the parity between NFL teams. Gray and Gray (1997) estimate a probit model with five variables; the overall winning percentage for the home and team the away team in the current season, the number of times the home team and away team has beaten the spread in the past four games, and a dummy variable that takes the value of one if the home team is the favorite, and zero otherwise. For the probit regression that includes variables for recent success, Gray and Gray (1997) eliminate the first few games of each season to create the recent form variables. The model finds that over the course of their study, the home underdog betting strategy only generated returns in excess of the break-even points 52.4%, two

times in a seven-season span and three times in an eleven-season span. These results show support for the findings of Golec and Tamarkin (1991). Using the probit model, Gray and Gray (1997) find that betting on teams that have had strong overall seasons relative to recent success also produce a profitable betting strategy. As mentioned earlier, people tend to overreact to recent success and want to bet on the hot teams. Betting on teams that have had a strong overall season but that have performed poorly in recent weeks generates a profitable betting strategy. This is consistent with the contrarian strategy, in which an investor buys poorly performing stocks and sells them once they begin to perform well. In one of their models, Gray and Gray (1997) use a dummy variable for the team that is the favorite and a dummy variable that is for the home team. Dare and MacDonald (1996) have noted the similarity and interdependence of these two variables. In the NFL, the home team is more likely than not also the favorite team for that matchup. The relationship with these two variables can produce inaccurate results.

The idea of following the herd and betting on the hot teams has been well known and Wever and Aadland (2010) add to the established literature. Wever and Aadland (2010) analyze 5,976 NFL games during 24 seasons from 1985 to 2008, and test for market efficiency using a regression model that includes variables for the home favorites and the home underdogs. Similar to Gray and Gray (1997), Wever and Aadland (2010) find that bettors tend to overvalue recent success and will bet on the teams that are playing well at the moment. After testing for market efficiency, Wever and Aadland (2010) find that profitable behavioral betting strategies occur. From 2000 to 2008, betting on underdogs with a large closing line produces winning percentages of almost 60%, significantly greater than the required 52.4% break even condition. Krieger, Pace, Clarke and Girdner (2015) find similar results to Wever and Aadland (2010) and Gray and Gray (1997), that bettors tend to favor teams that have been playing well as of late, or have been

'hot'. Krieger et al. (2015) analyze the wagering market of NFL and NBA games when one of the teams has clinched a playoff position. Krieger et al. (2015) find that the opening betting lines are systematically set too aggressively in favor of the teams that have already clinched a playoff spot. This is to take advantage of the unsophisticated bettors that Zuber et al. (1988) have noted, dominate the market. Krieger et al. (2015) acknowledges this systematic bias towards the teams that have clinched a playoff spot presents a potential profitable betting strategy. They conclude that the market does correct some as the week progresses and the closing line is set, but statistical evidence shows the behavioral tendencies of bettors to bet on the playoff team still presents profitable betting strategies, even after trading costs. This idea builds upon the previous idea of betting on teams with strong success in recent weeks. However, the difference is that these teams have already clinched a playoff spot, thus they do not have much to play for because they are already in the postseason. This is partly due to uncertainty that bettors do not know if the team will rest their star players or not, which will greatly affect the outcome of the game. This is relevant to my paper because it further shows that bettors favorite teams that are playing well even when the team has nothing to play for in terms of clinching a postseason position. This could possibly create an opportunity for profitable betting strategies.

c. Exogenous measures

One area the literature has not touched upon to this point is weather conditions and the effect that may have on the game outcome. This factor may play an important role in the spread betting market as football is played in the fall and winter, when weather is normally inclement. According to Borghesi (2006), much research has been done regarding the impact temperature conditions have on human performance. It has been noted that gross and fine motor skills are significantly impacted by exposure to low temperatures (Borghesi 2006). He also notes that

warm weather affects human performance as well. In order to differentiate the affect each type of weather has on performance, Borghesi (2006) measures each variable separately. To measure the baseline temperature for each area, Borghesi (2006) took the average temperature for the five days leading up to the game. He then took the difference between the mean temperature and the realized game temperature and hypothesized that the team with the smaller temperature acclimation should have the relative performance advantage during the game. Borghesi's (2006) data set is from the 1981 NFL season through the 2004 season and includes 5,748 games. His model includes the closing line and two dummy variables for the home/away, and favorite/underdog (Golec & Tamarkin, 1991; Gray & Gray, 1997). Borghesi (2006) creates a control for the weather acclimatization for the home team because they are used to playing in the weather at their home stadium. Borghesi (2006) also takes into consideration the fact that not all changes in weather are equal. For instance, a drop from 50°F to 45°F does not have the same effect as a drop from 30°F to 25°F. To control for these effects, Borghesi (2006) squares the acclimatization variables. Borghesi (2006) finds that colder weather has a greater impact on visitor's performance than heat does, which is consistent with prior research. Other conclusions Borghesi (2006) finds are that bettors tend to overestimate home team performance in hot weather games and underestimate home team performance in cold weather games. Therefore, because home teams outperform the expectations in cold weather games, bets on those teams are underpriced and present a profitable betting opportunity (Borghesi, 2006). One caveat with this paper is that although the team may play in warm weather normally, the players that make up the team are from different places. For example, the Miami Dolphins play in the warm Florida weather but they have players on their team from the North East which gets much colder than it

does in the south. The affect the colder weather has on the team may not be as big as it appears to be because players are used to playing in that weather being that they are from the North East.

Building upon Borghesi (2006), Boulier, Stekler, and Amundson (2007) test to see if the characteristics of the stadium and the relative strengths of the teams have been incorporated into the spread betting market. Boulier et al. (2007) uses data from the 1994-2000 NFL seasons. The relative strength of each team is a computer-generated score measuring the abilities of each team, published by *The New York Times*. The power score includes the win loss record of both teams and the aggregate difference in points scored for each team. The characteristics of the stadium includes whether the playing surface is grass or turf and whether the stadium is enclosed, also called a dome. Both of these variables are dummy variables. The addition of the dome dummy variable is important because an enclosed stadium will echo sounds and thus produce a louder environment to play in, which will affect the visiting team when they are communicating. Boulier et al. (2007) find that while the dummy variable for playing surfaces was significant at the 6% level, overall it does not appear to be a profitable betting strategy. Similarly, Boulier et al. (2007) conclude that the power score and stadium characteristics added no additional information to help explain the outcome of games. The playing surface variable may have been an important aspect to look at back in 2007 when this paper was written, but in the NFL today, most playing surfaces are turf. There are only a handful of fields that are actual grass and most are transitioning to all turf in the next few years.

The literature thus far has touched on many important factors that could contribute to the outcome of the game. None of the literature I have read has incorporated all of the factors into one model. I think combining the recent performance of teams (Gray & Gray, 1997; Boulier et al. 2007), the characteristics of the stadium (Boulier et al., 2007), the game statistics (Zuber et al,

1985; Wever & Aaland, 2012; Krieger et al. 2015) and the weather (Borghesi, 2007) would paint a more complete case proving if the NFL betting market truly is efficient. I also believe these factors will be important factors in analyzing the over under betting market, something that has not been researched as in depth as the point spread market. When betting on football, you have many different betting options. You can bet on a game using the spread, you can bet on the game straight up with the money line, and you can bet on the Over/Under total. I believe the Over/Under betting line is just as important to analyze for market efficiency. One implication of playing in a dome that I think is important, that Boulier et al. (2007) didn't address, is the weather conditions. A dome allows you to control weather conditions, creating a perfect environment to play in. This creates a premium playing condition in terms of temperature and wind and should be advantageous to the offense, which in turn should result in a higher Over/Under prediction.

IV. Data

The data is a cross-section for sixteen seasons from the 2000 NFL season through the 2015 NFL season, including the post season, and includes 4,255 games. The data is a large compilation rather than divided up on a season by season basis. The data has been obtained from Armchairanalys.com. I have created two data sets, one for the point spread betting market and one for the Over/Under betting market. For the spread betting market data set, I removed 178 games from the sample, for a total of 4,077 observations, for the following reasons. I omitted Super Bowls, because the game is played at a neutral site, therefore no home or visiting team designation is possible. I omitted pushes, where neither team covered the point betting spread and I omitted pick'em games, which are games that are played between two evenly matched

teams so no point spread is published for the game. For the Over/Under betting market data set, I removed 87 games for a total of 4,168 observations. I omitted Super Bowls, again because no home or away designation is possible and I removed games that ended in a push, which is when the total amount of points scored by both teams was equal to the Over/Under betting line. I created variables for the Away and Home scores in the previous week, a dummy variable for whether the Home and Away team beat the spread in their game the previous week, if the game was played in a dome (enclosed stadium) or outside, whether the game was played on turf or grass, and whether the final score of the game was larger than the Over/Under betting line.

Taking a look at the summary statistics in Table 1, paints a picture that has been consistent among most papers. On average the home team scores more points than the away team, which is consistent with the belief of home field advantage. The average actual difference in score is almost identical with the average point spread line published by the bookmakers. This shows, on average, that the point spread betting market is efficient. If you take a closer look, you see that the standard deviations for these variables are 15.10 and 5.90, respectively. The large standard deviations of these variables show that, while on average the spread is an accurate predictor of the actual difference in score, there is a large amount of deviation from the mean. This may show evidence of inefficiencies in the market, and present an opportunity to exploit those inefficiencies. The max and min for the two variables are vastly different as well. For example, the max for the point spread is 26.5 points while the largest actual difference in score is 46. Thus, it appears there could be some inefficiency in the points spread market. The same holds true for the Over/Under market. The average Over/Under and the average total amount of points scored are very close to identical. However, the standard deviations for these variables are quite large as well, 4.87 and 14.16 respectively. Once more, the max and min for the Over/Under are vastly

different than the max and min for the total points scored. The max for the Over/Under is 62 whereas the max for the total points scored is 106, again showing there may be the presence of market inefficiency.

V. Methodology

Current literature attempts to discover any inefficiencies in the NFL point spread betting market and finds that weak form efficiency does exist, but there are behavioral betting strategies that would produce profitable betting strategies (Zuber et al. 1988). This is due to sportsbooks taking advantage of human behavioral tendencies and the large number of uninformed bettors participating in the market (Humphreys, 2011). If the sportsbooks attempt to create spreads to take advantage of the uninformed bettors, the informed bettors should be able to counter and bet the opposite and thus find profitable betting strategies. The same scenario holds true for the Over/Under betting line.

I agree with the conclusions of most of the recent work regarding market efficiency in the NFL betting market. I hypothesize that overall the NFL betting market is efficient, meaning the point spread that is published for each game accurately incorporates all publicly available information. If this is the case, the market is efficient and there will not be a way to systematically earn abnormal returns from betting. Similar to other authors, I believe that my results will hold true only to my sample and may not be the same during another sample. This may be because during some time periods, betting trends may be more prevalent than in other periods and that could present slight inefficiencies in the market. Zuber et al. (1985) find profitable betting strategies in their data sample, while Sauer (1988) find that those same

strategies did not hold true in his data sample. Furthermore, I also believe the Over/Under betting market will be efficient as well for the same reasons stated above.

a. Probit vs OLS

There is much debate in the literature about which type of regression provides the imprecise results when testing for market efficiency. Some authors, such as Golec and Tamarkin (1991), and Zuber et al. (1988) use a standard OLS regression. They regressed the actual difference in score against the published spread along with other variables they deemed important. Others such as Gray and Gray (1997) and Borghesi (2007) use a probit model because they claim a standard OLS regression has the potential to overweight outliers and produce unreliable results. The dependent variable is a binary choice dummy variable for whether the spread was covered in that game. This is due to the fact that bettors are more concerned about the binary outcome of the bet, that is win or lose, not about the margin in which they win the bet. For example, if the spread is -3, a bettor is just as happy whether the team wins by 4 points or 5 points. In either scenario, the bettor wins the bet. There is no extra payment for beating the spread by more points. Due to this binary outcome, a probit model seems like a logical choice. I will also use a probit model for the point spread betting market as well as the Over/Under betting market. In this paper, I will construct both a standard OLS regression as well as a probit model regression. The reason for this is twofold; I want to see if there are any differences in the results of the two tests in terms of market efficiency. While the probit model and the OLS model use similar variables, the two models are distinct and different from each other. Furthermore, if there are any differences I want to see how large the differences are and why those differences may have occurred. Another reason it makes sense to use a probit regression rather than an OLS regression is due to the $[0, 1]$ binary choice. If I were to use an OLS regression with a binary dependent

variable, it would be possible to receive a predicted probability that is outside of the [0,1] range. This would produce unreasonable results that would not make sense in my model.

b. Econometric Models

The OLS models I will construct are based on those implemented by Pankoff (1968) and Golec & Tamarkin (1991). These models regress the actual difference in score onto the predicted point spread and a host of variables that I believe may have an effect on the point spread. The same idea is used for the Over/Under OLS model. In addition, I will construct a probit model in an attempt to negate the affects that outliers may have on the results. Probit regressions are still affected by outliers, but the magnitude is smaller than that of an OLS model because of the normal distribution assumption of the error term. The probit models I have constructed are based upon the authors before that have used a probit model, such as Gray and Gray (1997), Dare and Holland (2004), and Borghesi (2007). These probit models have a binary output as the dependent variable. The models are as follows:

$$Diff_{it} = \alpha + \beta_1 PS_{it} + \beta_2 Dome_{it} + \beta_3 Field_{it} + \beta_4 HomeBSP_{it} + \beta_5 AwayBSP_{it} + \beta_6 AwayPoints_{it} + \beta_7 HomePoints_{it} + \beta_8 Temp_{it} + e_{it} \quad (1)$$

$$Pr(WPS_{it}) = \Phi(\beta_1 HomePoints_{it} + \beta_2 AwayPoints_{it} + \beta_3 HomeBSP_{it} + \beta_4 AwayBSP_{it} + \beta_5 Temp_{it} + \beta_6 Field_{it} + e_{it}) \quad (2)$$

Equation 1 and 2 will analyze how accurately set the point spread betting line is. Equation 1 is a standard OLS regression where $HomeBSP_{it}$ and $AwayBSP_{it}$ are dummy variables for whether or not the home or away team, respectively, beat the spread in their

previous game. Equation 2 is a probit model where WPS_{it} is a dummy variable that takes the value of 1 if the away team beat the spread in their previous game and 0 otherwise.

$$Total_{it} = \alpha + \beta_1 OU_{it} + \beta_2 Dome_{it} + \beta_3 AwayPointsPrev_{it} + \beta_4 HomePointsPrev_{it} + \beta_5 Temp_{it} + \beta_6 Field_{it} + e_{it} \quad (3)$$

$$Pr(WOU_{it}) = \Phi(\beta_1 Dome_{it} + \beta_2 AwayPointsPrev_{it} + \beta_3 HomePointsPrev_{it} + \beta_4 Temp_{it} + e_{it}) \quad (4)$$

Equation 3 and 4 will investigate how accurately the Over/Under betting line is set. Equation 3 is a standard OLS regression where $AwayPointsPrev_{it}$ and $HomePointsPrev_{it}$ are the amount of points scored by each team respectively in their game the previous week. Equation 4 is a probit model where WOU_{it} is a dummy variable that takes the value of 1 if the total amount of points scored in the game was over the Over/Under total, and 0 otherwise.

After testing the point spread market and the Over/Under market, I then created tests to analyze if there is any difference amongst the best teams or the worst teams. I created two sub groups, the top five teams over the past decade from 2000-2010 and the bottom five teams over the past decade, in terms of winning percentage, playoff appearances, number of Super Bowls and a few other performance variables. I did this because the top five teams over the past decade should be consistently more favored in their matchups. Inversely, the bottom five teams should be consistently less favored in their matchups. Testing this will show whether there is a systematic bias against the top five or bottom five teams. These teams were chosen by NBC Sports based on a number of factors. They included the teams win-loss percentage over the

decade, the amount of playoff appearances they had, how well those teams played in the playoffs, how many losing seasons they had and the number of Super Bowls they won, where applicable. A losing season is defined as having more losses than wins in a season. The top five teams over the past decade are; New England Patriots, New York Giants, Indianapolis Colts, Philadelphia Eagles, and the Pittsburgh Steelers. The bottom five teams are; Cincinnati Bengals, Cleveland Browns, Houston Texans, Detroit Lions, and the Buffalo Bills. After choosing these teams, there were 1,273 observations for the point spread data set and 1,307 observations for the Over/Under data set. One caveat about the Houston Texans is that they joined the NFL in 2002 as an expansion team, meaning they had two less seasons than all other teams on the list. I created a dummy variable *Ranking* that takes the value of 1 if the team is one of the top five teams listed above, and 0 if the team is one of the bottom five teams listed above. Another caveat is that I only used the games in which the specified teams above were the home team in the game. I ran the same four regressions as stated earlier in the paper.

c. Variables

In equation 1, $Diff_{it}$ is the actual difference in score, represented as the away teams score minus the home teams score. In equation 2, $Total_{it}$ is the total amount of points scored in the game by both the away and home teams. In equation 3, WPS_{it} is a dummy variable that equals 1 if the away team covered the spread in the previous week, and 0 if they did not cover. Much of the literature has discussed the issue with which spread to use. Golec and Tamarkin (1991) recognize that there is no correct way to choose the team of record. The team of record is defined as the team whose perspective the spread and outcome is taken from. The spread is positive for the underdog and negative for the favorite. There are three ways to determine the team of perspective; choose the home team, choose the away team, or randomly choose a team. In this

paper, the team of record is the away team and the spread is from their perspective. In equation 4, WOU_{it} is a dummy variable that takes the value of 1 if the total amount of points scored in the game in question was over the Over/Under line, and 0 if it was under. These variables are distinct to their respective models. The following variables are the same across all four models.

PS_{it} is the closing point spread line defined with respect for the away team. Again, the away team is the team of record in this paper. OU_{it} is the closing Over/Under line. $Dome_{it}$ is a dummy variable that equals 1 if the game was played in a dome or a stadium with a closed roof and 0 otherwise. $Field_{it}$ is a dummy variable equal to 1 if the game was played on turf and 0 otherwise. $AwayBSPrev_{it}$ takes the value of 1 if the away team covered the spread in their previous game, 0 if they did not cover and 3 if the game ended in a push. $HomeBSPrev_{it}$ takes the value of 1 if the home team covered the spread in their previous game, 0 if they did not cover and 3 if the game ended in a push. $AwayPointsPrev_{it}$ is the amount of points the away team scored in their previous game. If they did not play a game the previous week, there will be no value. $HomePointsPrev_{it}$ is the amount of points the home team scored in the previous week, again if they did not play the previous week, there is no value. There is no recent performance variable for the first week of every season because there are no stats to look at. This means there are 16 weeks that there is no recent performance variable. $Temp_{it}$ is a dummy variable to measure the temperature at the start of the game. $Temp_{it}$ takes the value of 0 if the temperature is 32° Fahrenheit or below, 1 if the temperature is between 32° Fahrenheit and 50° Fahrenheit, and 2 if the temperature is 50° Fahrenheit or above. These temperature cutoffs are completely arbitrary. Lastly, e_{it} is the error term that will include any variables not taken into account in the model. All variables have an it subscript, i is for the i th game in season t .

VI. Results

a. Probit results

Results are presented in Table 2. The coefficient of the probit model is the marginal effect the variable has on the dependent binary choice variable. In column 4, the coefficient is the marginal effect each variable has on the game going over the Over/Under betting line. When I first ran the regression for equation 4, I included the $Field_{it}$ variable. That equation did not produce any statistically significant results. The $Field_{it}$ variable is approaching significance at the 5% level. Altering the equation slightly and removing the $Field_{it}$ variable does, however, produce statistically significant results for the $Dome_{it}$ variable at the 10% level. This may be because a majority, if not all stadiums that have domes also have a turf playing surface. This relationship between the two variables may affect the significance of the $Dome_{it}$ variable. Some of the significance of the $Dome_{it}$ may have been absorbed by the $Field_{it}$ variable because of the relationship the two have. Removing this variable allows the coefficient of $Field_{it}$ to be more accurate. Also, most NFL stadiums now have turf, so the $Field_{it}$ variable may not be as important now as it was in previous studies. According to my results, playing the game inside a dome increases the likelihood of the total score of the game going over the Over/Under betting line by 5.5 percentage points. This is in line with the results I expected to receive because playing inside a dome creates optimal playing conditions which would positively affect the offensive output in the game. This may have implications for the NFL Over/Under betting market, but the betting on NFL games simply because they are played in a dome may not be enough to create a profitable betting strategy in the long run. All other variables in this model are not statistically significant, which is in line with the efficient market hypothesis.

In column 2, the coefficients are the marginal effects the variable has on whether the away team covers the spread or not. Equation 2 produced three statistically significant results,

two at the 10% level and one at the 5% level. $AwayPointsPrev_{it}$ is significant at the 5% level but the coefficient is -0.002 which will have little to no effect on the outcome of the game. This result shows that for every additional point the away team scored in their previous game, decreases the likelihood that the away team covers the spread by 0.2 percentage points. Interestingly the coefficient is negative. I would have expected the amount of points the away team scored in the previous week to be positively related to whether they cover the spread in the current week because usually the more points a team scores the better they are playing. There are however other factors in the game that determine if the team will cover the spread, such as defense. $Field_{it}$ is statistically significant at the 10% level. The marginal effects of playing on a turf field decreases the likelihood that the away team covers the spread by 4.14 percentage points. I do not believe actually playing the game on turf affects the away team but the $Field_{it}$ variable is unique to the home team. The inverse of this result means the home team increases the likelihood that they cover the spread by 4.14 percentage points. This appears to be indicative of a slight home field advantage rather than the playing field affecting the spread. Also considering most stadiums have turf, meaning the away team should not be at a disadvantage by playing on turf, further signals this may be a home field advantage indicator. Similar to the magnitude of the $Dome_{it}$ variable above, I do not think this is indicative of a profitable betting strategy in the long run. $AwayBSPrev_{it}$ is statistically significant at the 10% level with a coefficient of 0.0282 or 2.82%. This means that if the away team beat the spread in their previous game, they have a 2.82% higher likelihood of beating the spread in the current week. This may show evidence to support the idea of betting on the hot team as discussed by Gray and Gray (1997) and Aaland and Wever (2010). Again, I do not believe these results show an

inefficiency which a bettor could take advantage of to earn abnormal returns over a long period of time.

b. OLS results

Column 1 shows the results of the econometric model that tests the actual difference in score against several variables. This regression produced four statistically significant variables; $AwayPointsPrev_{it}$, $Field_{it}$, PS_{it} , and $AwayBSPrev_{it}$. $AwayBSPrev_{it}$ is significant at the 5% level with a coefficient of 0.961. If the away team beat the spread in their previous game, the spread for the current week should increase by almost a point. $AwayPointsPrev_{it}$ and $Field_{it}$ are significant at the 10% level. The coefficient for $Field_{it}$ is -1.177, meaning if the game is played on turf, the spread should decrease by 1.177 points. The fact that the game is played on turf should not affect the performance of either team because almost all stadiums have turf. Again, I believe this is more of a home field advantage indicator rather than a measure of the effect the playing surface has on the outcome of the games. The coefficient of $AwayPointsPrev_{it}$ is -0.055, which is negligible and will not affect the outcome of the game. Lastly, PS_{it} is significant at the 1% level and has a coefficient of -1.073. The interpretation of this is as the point spread is decreased by one point so too is the actual difference in the game. This variable is not an indicator of anything, rather, it is just a measure of accuracy for the spread. The expected value of this variable is one. The fact that four variables are statistically significant is interesting because if the NFL betting market was efficient, none of the variables would be expected to be statistically significant. The theory is that if the market is truly efficient, all of these variables would have been taken into consideration when the point spread was published and should not have an effect on the outcome of the game. The $Dome_{it}$ was not significant, which is not what I expected. An away team playing in an opposing stadium, especially one that

is a dome, is more susceptible to noise effects. I expected the $Dome_{it}$ variable to be positive and statistically significant because I believed playing in a dome would add points to the spread for the home team, meaning they would win by more due to the noise effects Boulier et al. (2007) discussed in his paper.

Column 3 shows the results for the efficiency of the NFL Over/Under betting market. Three variables are statistically significant, two at the 1% level and one at the 5% level. The variables are; OU_{it} , $Field_{it}$, and α . Most notably α is statistically different from zero at the 1% level, with a large coefficient of 6.471, which goes against what I expected and what recent literature has expected. $Field_{it}$ is statistically significant at the 5% level with a coefficient of 1.268. This means that if the game is played on turf, the Over/Under betting line will increase by 1.268 points. This is not in agreement with my expectations. As I stated previously, almost every stadium has turf, so the field characteristics should not affect the Over/Under betting line. Surprisingly, $Dome_{it}$ was not significant in my model, even when I altered the model and removed the $Field_{it}$ variable like I did in equation 2. The OU_{it} variable is statistically significant at the 1% level with a coefficient of 0.870. This means as the Over/Under betting line increases by one point, the total amount of points scored in the game increases by 0.870 points. This makes sense because as the betting line increases, the total amount of points should increase too if the market is to be efficient. Interestingly, both $AwayPointsPrev_{it}$ and $HomePointsPrev_{it}$ are not statistically significant. This is interesting because I hypothesized that the amount of points scored in previous weeks would have an impact on the Over/Under line set for the current week. Once again, this leads me to believe that the market is efficient as it has incorporated this data into the betting lines.

Overall, while some variables are statistically significant, I do not believe there is enough evidence to determine that the NFL betting market is efficient. Therefore, I conclude that the NFL betting market accurately incorporates all available information into the creation of the betting lines, thus creating an efficient betting market.

c. Top 5 Teams vs. Bottom 5 Teams

I used a probit model to test whether there was a systematic bias amongst the top five teams or the bottom five teams in both the point spread and Over/Under betting market, over the past decade. These results can be seen in Table 3 through Table 6. None of the variables were statistically significant for either the point spread betting line or Over/Under betting line for both the top five teams or the bottom five teams. There appears to be no evidence of a systematic bias against the top five or the bottom five teams. These results are in line with the previous results of the entire NFL betting market.

Although none of the variables were statistically significant when testing the top five teams and the bottom five teams, there are differences in the coefficients which are interesting. The results are shown in Table 3. In the regressions testing the Over/Under, the coefficient of the $Dome_{it}$ variable for the bottom five teams was .0202 while it was 0.131 for the top five teams. This means that if a bottom five team plays in a dome, the likelihood that the game goes over the Over/Under increases by 2 percentage points, whereas if a top five team plays in a dome, the likelihood increases by 13 percentage points. Although these results are not statistically significant, it is interesting there is such a large discrepancy. Another interesting difference is with the $AwayBSP_{it}$ variable. When testing the bottom five teams, the coefficient is 0.0124, while for the top five teams it is 0.000956. This means that if the away team beat the spread in their previous game, they have a 1.24% better likelihood of covering the spread against a bottom

five team, compared to a 0.09% chance to beat the spread against a top five team. This result, although not statistically significant, is interesting because it shows that top five teams are better than bottom five teams at covering the spread, a result that was expected.

d. Comparing Probit vs OLS

As stated earlier, I used a probit model and an OLS model because of the different results I was expecting to receive. One surprising difference is the $Dome_{it}$ in the models that test the Over/Under betting market. The dome variable is not statistically significant in the OLS model but is statistically significant at the 10% level in the probit model. Another interesting finding is that in both the OLS model and the probit model that tests the point spread betting market, $AwayBSPrev_{it}$, $AwayPointsPrev_{it}$, and $Field_{it}$ are all statistically significant. I do not know why some results are statistically significant for the OLS and not the probit, and vice versa. Analyzing why the results are different would be an interesting extension to this paper.

VII. Conclusion

The purpose of this paper is to investigate the point spread betting market, as well as the Over/Under betting market in the National Football League, and determine if any inefficiencies exist that would allow a bettor to make consistent profits. The NFL betting market provides an interesting alternative to the financial markets when testing for efficient markets. It is extremely difficult to test for the validity of the efficient market hypothesis in the financial markets due to the complexity of the markets and the uncertainty of the true value of a stock/security. Thus, the NFL betting market presents a great foundation for testing the efficient market hypothesis.

Through testing the NFL point spread and Over/Under betting markets for efficiency, I conclude that, overall, the NFL betting market is efficient. While some variables, such as

$Dome_{it}$ and $Field_{it}$, are statistically significant, the coefficients are not large enough to warrant me to conclude that the market is inefficient and determine profitable betting strategies exist.

One caveat of the research is that all of the NFL seasons are pooled together rather than investigated on a season by season basis. There may be some strategies that are profitable during one season and not another season, as found in Zuber et al. (1988) and Sauer (1989). This could mean there are some efficiencies but none that are consistent enough to persist through multiple seasons.

I believe the reason for efficiency of the market is the increase in technology and the increase in transparency between teams and the public. There is more transparency between teams and the public, so more people are aware of potential injuries, the starting lineup and other team related news that may not have been publicly available during previous years in other literature. An increase in technology also plays a role in the efficiency of the betting market. There is more historical data available and more information that is publicly available that, in theory, should be incorporated into the point spread or Over/Under and create a more accurate predictor of the game outcome. For instance, OddShark.com is an online betting database that has historical betting information from the 1980's up to the most recent games. You can look at every game played between Team A and Team B in the last 30 years, what the point spread lines were for those games, who won the game, whether they covered the points spread, and so much more. This allows bettors to see the history between teams as well as the recent trends that occur, allowing them to become more knowledgeable bettors. Inversely, the increase in technology has lowered the barriers to entry for sportsbooks (Spinosa, 2014). This increased competition could make bookmakers alter their point spread lines or Over/Under lines in order to increase the amount of bets they get. Sportsbooks will create more favorable lines for bettors in the hopes of

generating more bets and revenue. Altering the lines away from what the actual line should be based on all the publicly available information to generate more bets, could create inefficiencies in the market. Humphreys (2011) noted this in his paper when he discovered bookmakers alter lines in order to take advantage of bettor tendencies. For example, if the New York Giants are playing well as of late, bookmakers might increase the spread from -3.5 to -5.5, to take advantage of the fact that bettors like to bet on the hot teams. Altering the lines in this manner could create inefficiencies that knowledgeable bettors can take advantage of if they are aware of these tendencies.

VIII. Limitations

Some limitations of this paper are that the recent performance variable I created only looks at the past game rather than multiple weeks. If the variable was able to capture the recent performance for multiple weeks, it would create a more accurate representation of the team's recent performance compared to just the most recent week. Furthermore, if the team did not play a game in the previous week, there is no value for the variable. Ideally, the variable would be for the most recent game rather than the most recent week, but I was unable to create a variable that was able to do that. Another limitation of this paper is that the data is one large dataset that spans 16 seasons. Ideally, I would have preferred to test the betting markets for all 16 seasons individually as well as a whole. This might reveal market inefficiencies on a season by season basis. For example, during one season there may have been a betting strategy that was profitable that did not exist in another season. Again, I was unable to detect that because of the dataset that I used.

IX. Future Work

Many of the papers I have read have created economic tests in addition to the statistical tests I have created, in order to find profitable betting strategies. These tests are beyond my ability and time limit. Future researchers should examine the Over/Under betting market and create these economical tests that are discussed in the literature review, such as betting on the home underdog, or the team that becomes less favored as the week progresses. The literature on the NFL point spread betting market is concrete and much work has already been done on it. The Over/Under betting market, however, is widely untouched and presents an interesting new area to investigate. In addition, I think it would be interesting if future researchers could look at the Over/Under betting market on a team by team basis and see if there are any inefficiencies on a team level.

Table 1: Summary Statistics

| Variables | Observations | Mean | Std. Dev | Min | Max |
|---------------------|---------------------|-------------|-----------------|------------|------------|
| Away Points | 4077 | 20.445 | 10.192 | 0 | 59 |
| Home Points | 4077 | 23.043 | 10.504 | 0 | 62 |
| Spread | 4077 | 2.571 | 5.896 | -20 | 26.5 |
| Difference | 4077 | -2.598 | 15.100 | -59 | 46 |
| Over Under | 4077 | 42.834 | 4.868 | 30 | 62 |
| Total Points | 4077 | 43.488 | 14.158 | 3 | 106 |

Table 2: Probit and OLS regression results

| VARIABLES | (1) Diff (OLS) | (2) AwayBeatSpread (Probit) | (3) TotalPoints (OLS) | (4) OU (Probit) |
|----------------|-----------------------|-----------------------------------|-----------------------------|-------------------------|
| Spread | -1.073*** (0.0480) | | | |
| Dome | 1.147 (1.017) | 0.0258 (0.0358) | 0.228 (0.928) | 0.0550* (0.0314) |
| AwayBSPrev | 0.961** (0.424) | 0.0282* (0.0158) | | |
| HomeBSPrev | -0.473 (0.448) | -0.0130 (0.0157) | | |
| AwayPointsPrev | -0.0553* (0.0284) | -0.00208** (0.000995) | 0.0103 (0.0268) | -0.000419 (0.000909) |
| HomePointsPrev | 0.0176 (0.0298) | 0.00111 (0.001000) | -0.0195 (0.0264) | -0.00108 (0.000906) |
| Temp | 0.0180 (0.0169) | 0.000941 (0.000602) | -0.00643 (0.0163) | -0.000233 (0.000584) |
| Field | -1.177* (0.616) | -0.0414* (0.0221) | 1.268** (0.612) | |
| OverUnder | | | 0.870*** (0.0595) | |
| Constant | -0.113 (1.338) | | 6.471*** (2.454) | |
| Observations | 2,783 | 2,783 | 2,847 | 2,847 |
| R-squared | 0.174 | | 0.090 | |

Robust Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Top 5 and Bottom 5 teams

| VARIABLES | (1) Top 5 OU | (2) Top 5 Spread | (3) Bottom 5 OU | (4) Bottom 5 Spread |
|----------------|-----------------------|-----------------------|-----------------------|---------------------------|
| Dome | 0.131 (0.167) | -0.101 (0.169) | 0.0202 (0.0691) | 0.0145 (0.0744) |
| AwayBSPrev | | 0.000956 (0.0397) | | 0.0124 (0.0460) |
| HomeBSPrev | | -0.0721 (0.0451) | | 0.0488 (0.0406) |
| AwayPointsPrev | 0.000605 (0.00223) | -0.00167 (0.00240) | 0.000338 (0.00238) | 0.000335 (0.00274) |
| HomePointsPrev | -0.00302 (0.00217) | 0.00204 (0.00247) | -0.00404 (0.00251) | -6.89e-05 (0.00274) |
| Temp | 0.000759 (0.00159) | 0.00301* (0.00163) | 0.00275 (0.00183) | -0.00144 (0.00186) |
| Field | | -0.0426 (0.0493) | | (0.0744) 0.0291 |
| Observations | 456 | 449 | 402 | 389 |

Robust Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

References

- Boulier, B. L., Stekler, H. O., & Amundson, S. (2006). Testing the Efficiency of the National Football League Betting Market. *Applied Economics*, 38(3), 279-284.
- Borghesi, R. (2007). The Home Team Weather Advantage and Biases in the NFL Betting Market. *Journal Of Economics And Business*, 59(4), 340-354.
- Dare, W. H., & MacDonald, S. S. (1996). A Generalized Model for Testing the Home and Favorite Team Advantage in Point Spread Markets. *Journal Of Financial Economics*, 40(2), 295-318.
- Golec, J., & Tamarkin, M. (1991). The Degree of Inefficiency in the Football Betting Market: Statistical Tests. *Journal Of Financial Economics*, 30(2), 311-323.
- Gray, P. K., & Gray, S. F. (1997). Testing Market Efficiency: Evidence from the NFL Sports Betting Market. *Journal Of Finance*, 52(4), 1725-1737.
- Hamburger, M. J., & Platt, E. N. (1975). The Expectations Hypothesis and the Efficiency of the Treasury Bill Market. *Review Of Economics And Statistics*, 57(2), 190-199.
- Krieger, K., Pace, R. D., Clarke, N., & Girdner, C. (2015). Anchoring, Affect, and Efficiency of Sports Gaming Markets around Playoff Positioning. *Financial Services Review*, 24(4), 313-329.
- Lee, J. (1998). The Market Efficiency Hypothesis on Stock Prices: International Evidence in the 1920s. *Applied Financial Economics*, 8(1), 61-65.
- Pankoff, L. D. (1968). Market efficiency and football betting. *Journal Of Business*, 41, 203-214.
- Paul, R. J., & Weinbach, A. P. (2011). An Analysis of the Last Hour of Betting in the NFL. *International Journal Of Sport Finance*, 6(4), 307-316.
- Russo, B., Gandar, J. M., & Zuber, R. A. (1989). Market Rationality Tests Based on Cross-Equation Restrictions. *Journal Of Monetary Economics*, 24(3), 455-470.
- Sauer, R. D. (1988). Hold Your Bets: Another Look at the Efficiency of the Gambling Market for National Football League Games: Comment. *Journal Of Political Economy*, 96(1), 206-213.
- Zuber, R. A., Gandar, J. M., & Bowers, B. D. (1985). Beating the Spread: Testing the Efficiency of the Gambling Market for National Football League Games. *Journal Of Political Economy*, 93(4), 800-806.