2013

NFL Betting Market: Using Adjusted Statistics to Test Market Efficiency and Build a Betting Model

James P. Donnelly
Claremont McKenna College

Claremont McKenna College

NFL Betting Market: Using Adjusted Statistics to Test Market Efficiency and Build a Betting Model

SUBMITTED TO

PROFESSOR HEATHER ANTECOL

BY

J. PATRICK DONNELLY

FOR

SENIOR THESIS

FALL 2013

DECEMBER 2, 2013

Acknowledgements

I would first like to thank my reader, Professor Heather Antecol for making this paper possible. You agreed to read my paper and, that alone, is worth appreciating. I honestly would not be in this position if it was not for your help. I would also like thank my parents for being so supportive and believing in me no matter what. Making you happy and remembering your advice was my driving force behind this project. Finally, I would like to thank my brothers. Connor, your dedication to helping me was valuable. Ross, thanks for finding ways to take my mind off this paper.

Abstract

The use of statistical analysis has been prevalent in the sports gambling industry for years. More recently, we have seen the emergence of "adjusted statistics", a more sophisticated way to examine each play and each result (further explanation below). And while adjusted statistics have become commonplace for professional and recreational bettors alike, little research has been done to justify their use. In this paper the effectiveness of this data is tested on the most heavily wagered sport in the world – the National Football League (NFL). The results are studied with two central questions in mind: Does the market account for the information provided by adjusted statistics? And, can this data be interpreted to create a profitable betting strategy? First, the Efficient Market Hypothesis is introduced and tested using these new variables. Then, a betting model is built and tested.

Table of Contents

1. Introduction

The amount of money wagered in the sports gambling market is enormous, with the NFL at the top as the most popular sport. The National Gambling Impact Study Commission estimated $3.45 billion was wagered legally in Nevada and an additional $380 billion illegally in the overall sports market for 2012 (Association 2012). To put this figure in perspective, it is about equal to the annual GDP of South Africa (GDP (current US$) 2013). Of the legal amount wagered, around 40% is on NFL games alone. The sheer amount wagered has caught the attention of academic researchers and the general public alike, as they test the market and try to develop models to exploit potential inefficiencies.

The most popular type of wagers is point-spread bets. A point-spread, or line, is a specific point handicap given to the "underdog" against the "favorite." Technically, the spread is to make the contest between two teams an even matchup. As an example, consider the New York Giants as an underdog against the Seattle Seahawks and receiving a +9 point handicap.[1] For betting purposes, this adds nine points to the Giants final score. Betting on the Giants wagers they will lose by less than nine points, tie, or win the game. Placing a bet on the Seahawks simply wagers they will win by more than nine points. A tie, which in this case means the Seahawks win by exactly 9 points, results in a "push" and the amount wagered is returned. The sportsbooks (linemakers or bookmakers) are a casino or, more recently, offshore online site create the point spread for games.[2] The

[1] The spread can also be expressed as the Seattle Seahawks receiving a -9 handicap.
[2] Online sports gambling are illegal in most states. The reason being, the NFL, NCAA and other organizations do not want money to be involved with games by influencing outcomes. Further, they do not want to deal with any speculation of outside influence on games.

spreads for each individual linemaker are almost always identical to the others, reducing the risk for arbitrage.

The Balanced Book Hypothesis states that linemakers set lines to equally split the amount of money wagered on each team in a given matchup. This eliminates the risk for bookmakers by guaranteeing they receive the "vigorish." A bettor must put up $11 in order to receive $10 for a successful prediction; the $1 difference is the vigorish. Essentially, this is the commission earned by the sportsbook for handling a transaction. However, no money is earned by the sportsbook if the bet wins. Sportsbooks are ideally trying to create a no-risk situation in which the money wagered is split equally between the underdog and the favorite. This way they collect the "vigorish" and do not have liability should an either outcome occur. When there is more than 50% of the money on one side, the sportsbook is exposed should that team cover the spread. To combat this, they will adjust the spread to invite wagers on the other side, and will continue doing this until the betting is split 50/50. The 11-10 creates a situation where a betting strategy must win more than 52.38% of bets to be profitable (Paul and Weinbach 2007). It also must be kept in mind linemakers do not want any strategy to be consistently found to make a profit, as vigorish from losing bets are their primary source for income.

Now that the basics of the NFL betting market are covered, the purpose of this paper is to us adjusted statistics to look for a betting strategy to consistently earn a profit, exposing market inefficiency. Even with the substantial amount of money put on the line, there is uncertainty about whether profitable strategies exist and, if so, can be implemented. Not surprisingly, there are individuals who have dedicated their lives to studying the betting markets and looking for ways to make a profit. There are subscriptions to “expert picks”, programs to help create and test betting strategies, and an overwhelming amount of information on the internet explaining betting theories.

As more sophisticated statistics have been developed, they have become a focal point for bettors in their search of profitable strategies. While a further explanation exists in a later section, adjusted statistics is a method for evaluating teams in which each play is compared to a league-average baseline based on the conditions that exist during that play and adjusted for the quality of the opponent. Using these statistics for the NFL has become a common topic within the betting world. However, little has been done in academic literature to test its effectiveness. This paper looks to fill this gap and test whether adjusted statistics are reflected in the spread and the actual outcome of the game, whether they have an influence on the team’s success ATS (Against The Spread), and if they can help create a betting model from in-sample data to test on external seasons.

2. Literature Review

The NFL betting market has been a hot topic going back to 1964, with the point-spread market being the main subject (Pankoff 1968). There are academic papers supporting and opposing the theory that there are opportunities for profit in the NFL betting market. Authors take different paths to investigate this theory and draw different conclusions. The difficulty in finding a concrete answer could be the driving force for past and future studies. In the following sections three topics are explored. First, the importance of understanding the current market and why advanced studies are necessary. Second, an introduction of the Efficient Market Hypothesis and the theories used to test it. And last, I will develop a new perspective of advanced statistics which has yet to be investigated.

2.1 Current NFL Betting Market

From an outsider's perspective, creating a line that equally splits the betting money is a daunting task. With the abundance of money involved and the prospect for large profits by bookmakers, it makes sense the most advanced techniques are employed to set lines—in addition to the comprehensive knowledge these individuals have about the sport. Simple analysis of the effectiveness of betting lines is important to show the make-up of the betting market and call for more advanced analysis. According to SportsInsights.com, from the 2004 to the current NFL seasons, simply betting on the home team has won 48.9% of the time against-the-spread and simply betting on the favorite has won 49.6% of the time ATS. Both of these percentages are very close to 50% and show how well bookmakers recently predict, or split, outcomes. More importantly, it demonstrates betting on the home/visitor or favorite/underdog alone will not result in a profitable strategy over 52.4%.

Song et al. (2007) article further developed the success of the bookmakers' lines and the difficulties of profitably predicting the outcome. They recorded the predictions from 52 expert forecasters and a diverse group of statistical systems which predict point-spreads for the seasons 2000/2001 and 2001/2002.[3] Expert picks and the statistical systems success rates were compared to a naïve model, or randomly picking winners.

[3] These experts' picks were published in the local newspaper for cities with an NFL team, the USA today newspaper, weekly national sports magazines, and websites of two national sports networks. The models were found across the internet.

Experts were slightly worse predictors ATS and the statistical systems were slightly better predictors ATS. Even more telling, only one expert and none of the models had a statistically significant success rate greater than the profit margin. The difficulty of betting ATS is demonstrated in the lack of success by both groups of predictors, highlighted by the lack of success amongst individuals who are paid “experts.”[4]

2.2 Efficient Market Hypothesis

While linemakers set very effective lines, this does not mean hope is lost for sports bettors. The most common question tested in prior literature is: are betting markets efficient? The Efficient Market Hypothesis (EMH) assumes all publicly available information, including prior or current statistics and news, is reflected in the price or, in this case, the spread (Wever and Aadland 2010). It was originally constructed for financial markets, but proved hard to test because the true value of an investment is hard to determine and there is no exact endpoint for an investment.[5] Researchers have found betting markets have well-suited conditions to test this hypothesis as these problems are easy to determine (Gray and Gray 1997). There have been two classifications of an efficient market in past literature. The first definition of an efficient market is that no independent variable should be significantly different than zero, besides the spread. If this was the case, it would suggest: (1) there is a discrepancy between the actual outcome and line and (2) other information can be used to predict the actual margin (Gray and Gray, 1997). The equation representing this definition is given as:

$$A_i = \beta_0 + \beta_1 S_i + \beta_2 X_i$$

[4] The experts are held to high standards because of the large amount of readers they reach. They also receive salaries for the information published. Although, it is found their predictions are not of much help.
[5] Stock prices are constantly changing.

A_i is the actual outcome of game *i*, β_0 is the y-intercept, $\beta_1 S_i$ is the spread for game *i* and its coefficient, and $\beta_2 X_i$ is any other publicly known information for game *i*. For this particular designation of EMH, the null hypothesis asserts: the intercept should equal zero, the coefficient for the spread should equal 1, and the coefficient for all other information should equal zero. This definition is very restrictive and it is not surprising it does not hold—linemakers would need to set almost perfect lines. However, most authors will not accept this as evidence of an inefficient market and take the EMH a step further.

According to the second definition of the EMH, no NFL betting strategy should be able to create abnormal returns compared to the market using any of the information mentioned above.[6] Most authors assume an amount significantly different than the profit margin, 52.3%, are “abnormal returns.” Therefore, divergences must first be found between the actual outcome and point-spread AND a profitable betting strategy must take advantage of this divergence. Although significantly beating 50% may be a more academic measure of (in)efficiency, a profitable betting strategy is the ultimate goal.

The considerable amount of participants involved and money wagered in the NFL betting market leads one to think there must be a profitable strategy. However, many studies have found this may not be the case. Most papers look at specific strategies thought to take advantage of different biases inherent in bettors’ and/or linemakers’ decisions.

2.3 Strategies

---

[6] A condition for the EMH is that the risks of the investments must be the same. This does not apply to the NFL point-spread market. All bets are considered equally risky—money is lost or doubled.

Warren Buffett is an example of someone who contradicted the EMH in financial markets by finding investments that are undervalued. This tactic can be translated to the betting markets as bettors look for tendencies for spreads to undervalue or overvalue a multitude of categories and, as a result, be mispriced.[7] The most popular inefficiency tested is the home-underdog bias and it holds an important position in the academic literature on this subject. The theory is that home teams whom are also underdogs are underrated. Gray and Gray (1997) used a probit model to test for some simple biases that may exist. Their tests show home-underdogs are undervalued and there is a market overreaction to recent performances. This manifests in teams who perform poorly (well) recently, but played well (poorly) overall. Testing for the home-underdog bias goes back to an important assumption about the sports betting market: lines are not created to be accurate predictions of games, but take bettors tendencies into consideration. If it is to balance the books or maximize profits, authors want to find the biases that are being accounted for.

The home-underdog bias is the most prominent example of what makes up most of the literature on the NFL betting market. Inefficiencies are found in a specific period, usually a season or two, and assumptions are made for its existence. The theories on undervaluation and inefficiencies are applied to many other game characteristics—the playing surface, weather, the day of the week the game is played, bye weeks, and others (Paul and Weinbach 2011) (Borghesi 2004). In relation to this paper, previous findings of inefficiencies support further tests.

[7] Some examples include: home-field advantage, parity in the NFL, crowd factors, weather, etc.

Studies that argue the NFL betting market is efficient show inefficiencies found in other articles would dissipate within the same season or in the next. This change in dynamic could be due to bookmakers correcting profitable inefficiencies or the high variability of NFL teams' performances.[8] Sung and Tainsky (2012) found simple strategies of betting on home/away or favorite/underdog to have small periods of profitability.[9] However, there is no discernible pattern when these periods take place. A low amount of variation in actual scores and point-spreads exists and it is mostly unexplained (Paton and Williams, 2005).

The testing of different biases is an important bridge between what bettors use in the field and academic literature. While these papers are very interesting and important, Sauer (2005) observes the study of finding inefficiencies is stale as new studies seek to find small time frames with evidence showing there is an exploitable bias. Studies have become an avenue for authors to support behavioral analysis and their effects on lines and less about testing market efficiency. This concern may be valid as many strategies look at very small samples to support findings. By focusing on biases, there is not a more general theory that can be applied to more than just a small amount of observations. This problem with studies has led to countless contradictions within the academic literature. There is a new route that has been slightly explored, although not completely.

2.4 Gap

There is a gap in the literature that is interesting. Many of the strategies tested use independent variables tied with the current season or state of the NFL. One such study

[8] Upsets are very common in the NFL. There are many human factors and "luck" that can drastically change the outcome of a game, which is not reflected in statistical data.
[9] In 2005, betting on favorites was profitable, but was successful 42.97% for the following season.

created a straightforward strategy of using past score differentials for each team within that season to predict game margins (Osborne 2001). The conclusions drawn from this paper are the reasoning behind the make-up of this paper. Osborne states it is possible to use past information on team strengths to earn positive profits. Linemakers incorporate some of the information in the spread, but not completely. However, NFL performance has a high variability from season to season—which is reflected in the models large variance of success from one year to the next. It is important to look at many seasons when developing and testing a model.

Statistics are important to show the current condition of the NFL, but are ineffective at predicting future success ATS. To highlight this, current teams score more points and gain more total yards than in the past. A team who scores forty points now may be just as good, against-the-spread, as a team who scored twenty points ten years ago. Many teams play in “weaker” divisions and face “weaker” opponents, making raw statistics bad measures of a team’s ability. So, these statistics predictive ability for one year could be much different in the next or even within the same year. Testing team statistics has its issues, but it fixes some of the problems Sauer brought up and moves in a different direction than the past.

A short peruse of online betting sites will reveal the importance of a more unique form of statistics. Dr. Bob, who has been featured on ESPN, is a professional sports bettor who owns his own subscription website. He is a leader in the sports betting industry and Las Vegas casinos are known to move point-spreads after he distributes his predictions to subscribers. His website highlights the use of "compensated statistics."[10] Given this, it is surprising studies have not tried to incorporate adjusted statistics in their models. Many of these adjusted statistics are available on the internet and come with in-depth explanations of how they are calculated. There are many advantages to these statistics. It would make statistics relevant for the overall quality in the league at that time and easily comparable to other years. Further, it moves away from testing different biases bettors or linemakers may have and tests a model that is both consistent and easily determinable. This is extremely important because it makes basically every game a valid observation. Testing these strategies would be the first step to closing this gap and opening the doors to similar tests in the future.

Boulier et. al (2006) paper was the only one found to use adjusted statistics, in this case overall team power rankings, to predict the outcome ATS. The rankings were for the seasons from 1994-2000 and created by the New York Times. Teams are ranked 1-32 and a difference in the rankings is the "predicted spread" for a matchup. They found the power rankings did not add to the information contained in the spread, which is evidence for an efficient market. Although on the right track, there are drawbacks to this method. First off, it is very basic. Although it would be very satisfying for a simple strategy like this to work, the bookmakers would not be in business if this was the case. Secondly, by attaching consecutive whole numbers to each team does not take in account any differences between the rankings. The difference between the team ranked one and two may be small, while, the difference between 15 and 16 may be large. These power

[10] I refer to them as adjusted statistics.

rankings do not incorporate these differences. These are problems new statistics may be able to account for.

3. Data

There are two categories of information that are essential for this paper: (1) point-spreads and actual outcomes for each game and (2) adjusted statistics for each game. Even with the popularity of testing the NFL betting market, point-spread archives are very difficult to find. While game statistics archives are much easier to come by, extracting the information from every game is not recommended. These circumstances most likely contributed to the small sample sizes most papers dealt with, as extracting lines and statistics when they come out for every game would take a long time. Further, any other variables an author wished to include would face the same time consuming issue.[11] Unfortunately, this probably limited the scope of many papers. The gathering of data for this paper faced some of the same issues, which are explained below.

The data was gathered from three different sources. The scores, lines, and raw team statistics for each game were taken from Warren Repole's website. Every game is included from the 2003/2004 to 2011/2012 seasons. Spread data is collected from scoresandodds.com and game statistics from cbssports.com. The adjusted statistics were manually removed from www.footballoutsiders.com. A premium subscription was required that included a small fee. The stats from *Football Outsiders*' website were combined with the spread and game data. This was done by hand, for each week of each

[11] For example, gathering weather data for past game dates to test its effect on team performance ATS would be a huge hassle.

season. There were many checks in place to locate and fix any problems that occurred with the input.

3.1 Point-Spreads and Actual Outcomes

For any test on the NFL betting market, the most important information is the point-spreads and actual outcome for each game. These alone are the key for prior academic work. For the most part, no other variables were needed to test for certain biases outside of these two values and basic information—home, away, favorite, underdog.[12] Finding these figures was the first and most important step.

This data is pulled from Warren Repole's website (www.repole.com). The datasets were already in Microsoft excel format, making extraction and any necessary changes very easy. In addition to the spreads and outcomes, he includes stats for each game going back to the season starting in 2000 and ending with the most current season. For this paper, the season starting in 2003 to the season starting in 2012 are included—a total of 10 seasons. The first 7 seasons are tested to create the models, leaving the last three for out-of-sample testing of the betting strategies.

3.2 Adjusted Statistics

NFL seasons are broken up into 17 weeks of games. Every team plays one game each week, besides each team's "bye week." For the adjusted statistics, all information from the previous games is included in that current week. For example, at week *t*, the statistics for weeks *t*-1 are included in the statistics. This is an essential feature because this is the information available when the decision to bet, or not, is made in real world situations.

[12] Favorites and underdogs are derived from the spread. So, no other information is needed.

When creating the models, it makes sense the sample replicates the same circumstances a bettor would face. The importance of current season information for the models led to the first two weeks of a season to be removed from the sample. For the first week of games, there is not information on any team. The second week was removed because one game is not enough game play for the adjusted statistics to become useful. Bye weeks begin no earlier than the 3rd week and end by the 12th. Therefore removing the 1st and 2nd weeks of the season maintains information for two games is collected for every team.[13] Also the 17th week was removed because many teams have already established a specific playoff position and the games do not matter. The starting players for the teams may not play, making past statistics irrelevant.

As stated earlier in the review of literature, including adjusted statistics in betting strategies is common practice in the betting market. Although some forms of raw and adjusted statistics have been tested in academic literature—power rankings, scoring differentials, game statistics—more advanced forms have not. This lack of testing was the inspiration for this paper and a place where it can add something new to the current research and literature. The choice to use statistics from *Football Outsider's* was based on three factors: their formula for creating the statistics, the ability to go back to past years week by week, and the general popularity of the statistics.

The statistics that were recorded for each week are total, offense, defense, and special teams Defense-adjusted Value Over Average (DVOA). To fully explain how the

---

[13] The first week essentially has no statistical information on the game, besides pre-season games. However, it is common knowledge that these games are more of a tryout for back-ups looking to make the team.

DVOA statistics are calculated would take many paragraphs.[14] However, a basic understanding of how they differentiate from other statistics will show their complexity and usefulness. Excerpts from *Football Outsiders* website best describes the process they use:

> Conventional NFL statistics value plays based solely on their net yardage. The NFL determines the best players by adding up all their yards no matter what situations they came in or how many plays it took to get them…. [For DVOA] Every single play run in the NFL gets a “success value” based on this system, and then that number gets compared to the average success values of plays in similar situations for all players, adjusted for a number of variables. These include down and distance, field location, time remaining in game, and the team’s lead or deficit in the game score…The biggest variable in football is the fact that each team plays a different schedule against teams of disparate quality. By adjusting each play based on the opposing defense’s average success in stopping that type of play over the course of a season, we get DVOA, or Defense-adjusted Value Over Average. Rushing and passing plays are adjusted based on down and location on the field; passing plays are also adjusted based on how the defense performs against passes to running backs, tight ends, or wide receivers. Defenses are adjusted based on the average success of the offenses they are facing… The final step in calculating DVOA involves normalizing each year's ratings. As you may know,

[14] A full reference can be found on: http://www.footballoutsiders.com/info/methods

> offensive levels in the NFL have gone up and down over the years. Right now, the overall level of offense in the league is probably at its highest level of all time. Therefore, we need to ensure that DVOA in a given season isn't skewed by league environment... DVOA is a percentage, so a team with a DVOA of 10.0% is 10 percent better than the average team, and a quarterback with a DVOA of -20.0% is 20 percent worse than the average quarterback. Because DVOA measures scoring, defenses are better when they are negative.

The underlying feature of all the steps described above is everything is adjusted to become more informative. The actual values of DVOA statistics is described later in the variable definitions section.

The simple feature of obtaining past years alone made these specific statistics valuable. The availability of past years by week was crucial for effective simulation, explained in the first paragraph of this subsection. The extraction was very meticulous and time consuming as each week had to be copied and combined with the lines and statistics. However, this dataset is very valuable for many reasons: the large amount of observations over many years, the quality of the information—the setup of the data and the formula used—and the overall popularity of this information.

It was important to have data that is accepted as credible. More recently, the use of adjusted statistics has seen a leap in importance for baseball. The movie “Moneyball” depicts the first use and success of such statistics in Major League Baseball. It is now essential that all programs use such statistics when making up a team. The same advanced statistics renaissance is beginning in the NFL and *Football Outsiders* is at the head of all public information (fantasylistings.com). While the formula alone suggests this is very advanced and beneficial information, the large amount of positive reviews found on the internet is reassuring (twominutewarning.com, footballperspective.com). The founder of Football Insiders also serves as a consultant to various NFL teams, further validating the work they have done (http://www.footballperspective.com/interview-with-aaron-schatz/). In one interview, the founder of *Football Outsiders* talked about the consulting he does with NFL teams.

There are drawbacks to adjusted statistics. First is the accumulation of games and statistics it takes for DVOA information to be beneficial. This is not a problem faced by the biases tests because there was no further information needed outside of the spread and actual outcome. Second, changing data through adjustments can be detrimental. The formula and process used is very important and must be checked for accuracy. Incorrectly weighting certain values or other changes can reduce the significance of a statistic, making it less useful.

3.3 Variable Definitions

After data was extracted from the two different sources and combined, the most informative variables needed to be chosen and/or created. No changes were made to the actual outcome or spread. The line is given as *Line*. It is recorded in reference to the home team. A “-5” spread indicates the home team is favored by 5 points and this will be deducted from their final score. A positive value would be added to the home team’s final score. The actual outcome is given as *Actual Outcome*. It is recorded in the same way as the point-spread. The equation is:

$$ActualOutcome = Away\ Points - Home\ Points$$

If the home team was to score more points in a game, the outcome would be negative. The line would show the home team is favored if it, also, is negative. This set up makes the comparison between the two variables values easy.

The raw game statistics were not necessary for this paper—they are not ex ante information and not good indicators of an outcome. Further, the adjusted statistics include most of this information making the independent variables correlated and increasing multicollinearity. This affects the p-values of variables and can cause them to be statistically insignificant. It would be interesting to look at the differences between raw and adjusted statistics for this type of tests, but the required information was not obtained and the comparison of the two is not the goal of this paper.

As described in section 2.3, the values of the DVOA statistics are set-up as percent values. There are weighted and unweighted values for the categories: total, offense, defense, and special teams. A value of zero percent for a statistic is average, a positive percentage is better than average, and a negative percentage is worse than average. The opposite rules are true for the defense DVOA stat—a negative value means the defense is better than average. The percentages were changed to whole numbers so the coefficients are easier to understand. This makes all matchups between two teams comparable at many different levels. Even more important, it makes comparison between separate seasons possible because everything is relative to the average for its exclusive season. The structure makes teams easily comparable, but it was taken to another level.

To make any given matchup more comparable and make the tests more logical, the differences between the DVOA statistics were used instead of the each teams own values for each category. An example will highlight the need for the differences to be used. Suppose there are two teams with high above average offense DVOA values. Individually, this would suggest the actual outcome would be in both teams favor for most games. However, the matchup is relevant as the difference can still show which team is superior. Therefore, the difference is more important than the individual values.

To reduce problems that result from multicollinearity, total DVOA has been left out and the values for offense, defense and special teams are kept. The total DVOA is a derivative of these three values. Also, this makes it so the impact of different facets of a game can be compared to the others—by looking at their coefficients and significance levels. A paper by Dare and Dennis (2001) found all biases are not equal and opposite in all situations.[15] Instead, lines have very “specific biases.” They may overvalue the home team, but not the away team. The idea of breaking down the different areas where possible inefficiencies, biases, exist is important and incorporated into this paper.

The formula for these difference variables is simple:

$$WeightedOFFDifference = WeightedOFFHome - WeightedOFFAway$$
$$WeightedDEFDifference = WeightedDEFHome - WeightedDEFAway$$
$$WeightedSTDifference = WeightedSTHome - WeightedSTAway$$

[15] A bias on team-one does not result in an equal and opposite bias for team-two.

The variable on the left for each equation is the differences between the home and away team's individual values— the variables on the right side of the equations. Still, there are some differences to point out. Because an increase in the offense and special team statistics reflects the team is better in that category, I want the difference to refer to the home team. Therefore, the away team's value is subtracted from the home team's value. A positive (negative) outcome shows the home team is better (worse) in that category. Explained earlier, a decrease in the defense statistic means the team is better in that category. Once again, I want the difference to refer to home team. A negative (positive) outcome shows the home team is better (worse) in that category.

3.4 Summary Statistics

Table 1 in the appendix shows the averages of all pertinent variables for each individual year and all years combined. This is done for both the sample data and out-of-sample data. The reasons for showing the different variables are different for the two categories of information.

Comparing the values of actual outcomes for games and the line is important for testing market efficiency and the starting point for most papers on this subject. As shown in Table 1, the mean actual outcome is slightly more negative than the average line at -2.70 and -2.59 points, respectively. This difference is very small, which is unfavorable for bettors as this could make inefficiencies hard to find. It does seem the differences between the two variables for each specific year are larger than the average for all years combined. However, this does not tell the whole story. The standard deviation for the actual outcome is consistently much larger than that for the line. The actual outcomes deviate from the mean much more than the line does. While this may just be a factor of the unpredictable nature of the NFL, it is worth noting. A successful strategy would be able to consistently account for these deviations. However, it is most likely impossible to account for all of the variability because professional sports are greatly influenced by human error. Not much else can be taken from this data except the point spread seems to underestimate the home team's ability on average. However, there are years when the bookmakers overestimate the home team's ability on average. This goes back to the problems other authors have seen with inefficiency tests because they do not account for the changes that occur season to season.

A somewhat discouraging result of the summary is the difference in the means of the actual outcome for the in and out-of-sample data. The out-of-sample data has an actual outcome mean of only -1.59 compared to the -2.70 for sample years. Besides 2006, there is no other year the mean is larger than -2. For most years, the average is actually closer to -3. This suggests the models derived from the sample years may not be effective on the out-of-sample data because there is a major change in the dependent variable. An unfortunate side of statistical modeling is the past must be able to predict the future for the model to be effective. Or, at least some of the trends should continue. The effect of the changes in the NFL cannot be determined until the data is tested.

In the description of DVOA statistics it was stated the average for every category in the league is given as zero. Table 1 is a summary of the averages for the values for all years combined and each year individually. The values do not end up equaling exactly zero, but are very close. Most of the averages deviate less than one point from zero. The high standard deviations can be expected for these figures because the average is meant to be zero while also containing outliers on each side of the spectrum.

4. Empirical Strategy & Results

This section restates the assumptions of the Efficient Market Hypothesis and analyzes if it holds for the sample seasons. Models used for the data are defined and the resulting variables are looked at. The models are then tested on the out-of-sample and the success is recorded.

4.1 Efficient Market Hypothesis and Model Creation

There are two steps for testing the NFL betting market and are directly related to the two definitions of the Efficient Market Hypothesis. First, the sample seasons are tested and a model is built. Second, this model is tested on the out-of-sample data and its success rates are reviewed. The equation that will be created looks to predict the actual outcome of games. It is the same equation that was seen earlier, when going over the Efficient Market Hypothesis. The variables are changed to match what is used for this paper. The equation is given as:

$$
\begin{aligned}
ActualOutcome_i \\
&= \beta_0 + \beta_1 Line_i + \beta_2 WeightedOFFDifference_i \\
&+ \beta_3 WeightedDEFDifference_i + \beta_4 WeightedSTDifference_i + \varepsilon_i
\end{aligned}
$$

According the first definition of EMH explained previously, the added variables should not add to the equation. This means the coefficients for the weighted difference variables should be equal to zero and the line is predictive of the actual outcome. In short, no added variables based on public information should help in predicting the actual outcome of a game. When the actual outcomes are regressed on the line, and weighted difference variables for all years, the equation comes out to:

$$\begin{aligned} &ActualOutcome_i \\ &= (-0.41) + 0.91Line - 0.05WeightedOFFDifference \\ &+ 0.01WeightedDEFDifference - .10WeightedSTDifference \end{aligned}$$

Seeing the equation makes the coefficients easy to interpret. All of the coefficients have the expected effects on the actual outcome. Holding all variables constant, the model predicts the home team will win by -0.41, although the value is not significant at $p<0.05$. This can be viewed as the "home-field advantage." In practice, some stadiums are thought to have more of an advantage than others. However, this equation does not take this account and normalizes the value of playing at home. For every one point increase in the line, the actual outcome increases by 0.91. This coefficient is significant. This alone shows there may be some inefficiency in the NFL betting market because the null hypothesis has been betrayed.

The question behind betting market efficiency is: does the line capture all available information? Because of this question, the line was regressed on the adjusted variables that were to be used for the model. By doing this, it can be analyzed if the statistics can be used to predict the line and, more importantly, if bookmakers use adjusted statistics in their line creation models. The results are in Table A. The combined sample is only shown—there is no year-by-year information. The results suggest that adjusted statistics are definitely used when creating the line. All of the values are significant at $P<0.001$ and the R-squared value is very high at 0.6795. Although the adjusted statistics are used to create lines, they may not be fully accounted for. If this was the case, they could be used to better predict the actual outcome when combined with the information in the line.

Table 2 shows all of the values and standard deviations. For the adjusted statistics, the actual increase or decrease by a one point increase in the dependent variable are not as important as the direction of the sign. However, the effects will still be reviewed. A one point increase in the offense difference leads to a 0.05 decrease in the actual outcome. This makes sense. If the difference is increasing, it reflects the home team is increasingly better in that category. Remember, a negative outcome means the home wins—it is reflected the same way as the spread. An increase in the defense statistic works in the opposite direction. A more negative number is means the home team is better. Therefore, an increase in the difference should lead to an increase in the actual outcome, which it does. Special teams work the same way as the offense and has a surprising large effect on the actual outcome—it is also statistically significant. An increase in the difference leads to a decrease in the actual outcome.

Comparing the effects of the different facets of a game on the actual outcome, special teams unexpectedly has the largest influence on the actual outcome from a one point increase in the difference—in terms of the coefficient value. In terms of game play, the special team seems the most insignificant of the three categories. The only somewhat credible speculation is the parity in the NFL is so high and any small edge a team has is very beneficial.

4.2 Analysis of Variable Variation between Seasons

The problem faced by past papers due to inconsistency between years was the inspiration to list the variable values by year and combined. To take it a step further and compare the sample data to the out-of-sample data, the same coefficient values were found for the last three seasons. Table 2 lists the variable values by year and shows if they are statistically significant. The issue of NFL variation seems to be present in the data used for this paper.

The biggest problem that shows up with the constant is the difference between the in and out-of-sample data. The model that was created from the sample has a negative value for the constant, -0.410. However, the constant value for the out-of-sample data is positive, 0.247. This value is not large but may result in some bad predictions. Unfortunately, for NFL bets, there is only two options—win or lose. So, small discrepancies can lead to a bad bet. Also to note, the constant is never statistically significant.

The coefficients for line change much more than expected from one year to the next. To make a hypothesis about why the variation exists, Table 1 needs to be reexamined. The average line value does not change much from year to year. The variation is probably not derived from this. It is most likely due to the volatility in the actual outcome—seen by the high standard deviations in Table 1. Luckily, the constant for in and out-of-sample data are similar. The values for *line* are statistically significant for every year besides one.

There are small differences between adjusted statistics from year-to-year, but nothing seems to be much of a problem. Additionally, the in and out-of-sample values are similar. The only issue is the lack of statistically significant values. This could be due to multicollinearity discussed earlier. The adjusted statistics were found to be significantly correlated with the line. Including them all as independent variables can cause the standard deviation to increase.

The last piece of surprising information is the difference in the R-Squared from year-to-year. The difference between the minimum and maximum value is large—0.1053 in year 2006 and 0.2968 in 2005. However, the combined in and out-of-sample data values are very similar. Overall, the lack of consistency in any of the variables makes creating an effective model difficult to do.

4.3 Out-of-Sample Testing

To test the second definition of the Efficient Market Hypothesis, the model is tested on the 2010-2012 seasons—a total of three seasons. Due to the 11-10 rule, a betting strategy must win 52.38% of the bets. While this does not seem like a significant hurdle, it has been a tough rate to pass. The model that was created in section 4.1 is used to create an estimate for the actual outcome. This value is compared to the point-spread and a bet is made on whichever team the predicted outcome favors. For example, if the predicted outcome is -10 and the line is -5 for a specific game, then the home team would be bet on. A bet is placed on every game.

Table 3 shows the success of the model by year and overall for the out-of-sample data. Over the three years, the model correctly bet on 317 games and incorrectly bet on 290 games. This comes out to a total of 52.22% percent—below the 52.38% profit level. The difference in success from year-to-year was an issue. Only 44.55% were correctly predicted in 2010, 54.95% were correctly predicted in 2011, and 57.14% were correctly predicted in 2012. Therefore, besides the season during 2010, the betting strategy was profitable for the two other years. However, a strategy which consistently earns profit must be found. From this sample, this strategy does not.

The reasons for the models lack of success is hard to determine. From the tables it seems 2010 has some similar characteristics to the model, but ended up being the worst year. There are similarities in the values of the weighted statistic coefficients for the model and 2010. However, the constant is much different— 0.410 for the model and 0.866 for the 2010 season. Another difference is seen in the average outcome. For 2010, it was only -1.13 points. This is very different than the -2.70 points average for the sample. All of these subtle differences could lead to drastic changes in the success of a betting model. Further analysis could be done to try and find the exact reasons for bad, and good, seasons.

Unfortunately, more seasons were not tested. A better picture of the variation from one year to the next would have been helpful to see and determine if this model rejects the second definition of the EMH. If only 2011 and 2012 were to be tested, the model may have shown the market was inefficient. If only 2010 were to be tested, the test would conclude the market is efficient.

An important feature for building statistical models is the past is similar to the future. For NFL football, this may not be the case. The common problem with betting strategies is they lack consistency. While it could be due to a problem with the model or strategy, it could be an outcome of the unpredictability of the NFL. There are so many factors of the game that can easily sway the outcome. Yet, statistical or theoretical strategies cannot account for these “chance” occurrences.

To see what would happen if bets were actually made based on this model, an amount earned section is included in Table 3. It is assumed $10 is bet on every game and a losing bet pays $11—one dollar for the vigorish. Overall, the bettor would have lost a total of $20. A whopping $332 would have been lost in 2010 alone. In 2011 and 2012 the bettor would have realized a profit total of $109 and $203, respectively.

5. Conclusion

Adjusted statistics are an accepted part of many NFL betting strategies used in the betting market. However, there is a noticeable gap between the academic literature and what is used in real life situations. Most papers looked to support behavioral theories about the bettors and linemakers in the NFL betting market. While basic statistical analysis has received some attention, there has yet to be through research done exploring adjusted statistics. This paper looks to add to the current academic literature and use adjusted statistics to test the different forms of the Efficient Market Hypothesis (EMH) and create a successful betting model.

The first, and weaker, form of the EMH was found to not hold. The point-spread market is not a prediction market as discrepancies are found between game outcomes and the spread. On one hand, adjusted statistics were found to be used by linemakers to create spreads. On the other hand, adjusted statistics were found to have predictive ability of the actual outcome the spread does not incorporate.

A betting strategy was developed using the coefficients created by the 7 sample seasons. The model incorporates the line and different adjusted statistics to create predictive game outcomes for each specific game. Next, this model is tested on 3 seasons outside of the data used to build the model. The predictive outcome is compared to the actual outcome and a hypothetical bet is wagered and the result is computed.

The betting strategy was found to predict slightly below the profit margin needed, but had very high variability from one year to the next. After analyzing the summary statistics and coefficients from each year, large disparities are observed from one season to the next. This adds to the current hypothesis that a consistently profitable strategy is hard to find because of the unpredictable nature of the NFL.

Although the strategy was "unsuccessful", it is a stepping stone for similar tests in the future. There are many more paths to be explored by using adjusted statistics in the NFL betting market. In this study alone, there were alternate routes that could have been taken—a change in what statistics were tested, what type of model was used, or what was analyzed would have completely changed the make-up of this paper.

The biggest problem this, and most other previous, papers face is the lack of large in and out-of-sample data. The hope is the datasets created by writing this paper can be expanded upon. Both by using the data currently in them and by supplementing the data, more tests on this subject can be explored.

Another adjusted statistic was recorded for this paper, but it is not included. It would be interesting to compare and contrast the ability of different adjusted statistics. Because they are not just raw statistics and a theory for certain adjustments is necessary, different theories and formulas can be compared.

Further comparison can be made at model level. A probit models success can be compared to the linear regression model used in this paper. A feature of the model that sets up nicely with the betting market is the binary outcome. The betting markets also have a binary outcome—win or lose the bet. Instead of creating a predictive score outcome, it creates a predictive probability of success. Other papers have explored using the probit model, but not with adjusted statistics included as dependent variables. I see this as a logical next step for adjusted statistics testing in the NFL betting market.

The elusive 53% will be a goal that continues to be tested and retested in the future. The strategy in this paper unsuccessfully hurdled this mark, but hoped to shed light on the use of adjusted statistics. With the advancement of new statistics which hold much more information than anything seen before, there is no doubt that more studies will look to incorporate them.

Tables

Table 1: Summary of Variable Means by Year — **Out-of-Sample Data**

	Combined	2003	2004	2005	2006	2007	2008	2009	**Combined**	**2010**	**2011**	**2012**
Actual Outcome	-2.70 (15.38)	-3.17 (14.79)	-2.97 (15.49)	-3.46 (14.59)	-1.55 (14.58)	-3.22 (16.46)	-2.08 (15.22)	-2.61 (16.70)	-1.59 (15.35)	-1.13 (15.48)	-1.76 (15.56)	-1.41 (15.47)
Line	-2.59 (6.21)	-2.56 (5.24)	-2.66 (5.45)	-2.62 (5.80)	-2.70 (5.98)	-2.43 (6.99)	-2.70 (6.21)	-2.36 (7.55)	-2.26 (5.99)	-2.36 (5.70)	-2.45 (6.62)	-1.90 (5.70)
Weighted Offense DVOA Home	-0.66 (18.18)	-0.71 (14.29)	-0.47 (19.69)	-1.23 (20.57)	-0.66 (15.35)	-1.07 (19.96)	-0.37 (16.91)	-0.62 (19.75)	-0.99 (17.16)	-0.93 (17.94)	-0.92 (16.67)	-1.29 (16.90)
Weighted Defense DVOA Home	-0.21 (13.91)	-0.42 (13.18)	0.17 (13.79)	0.18 (14.22)	-0.33 (14.26)	-1.09 (13.18)	-0.16 (15.44)	0.33 (13.23)	-0.08 (12.51)	0.17 (12.27)	0.36 (11.14)	-0.70 (13.92)
Weighted Special Teams DVOA Home	-0.44 (5.45)	-0.35 (5.55)	-0.54 (5.10)	-0.38 (5.22)	-0.63 (5.11)	-0.69 (6.45)	-0.38 (5.59)	-0.13 (5.00)	-0.27 (5.93)	-0.13 (6.12)	-0.69 (5.93)	0.07 (5.66)
Weighted Offense DVOA Away	-0.04 (18.26)	0.50 (14.81)	-.055 (19.32)	-0.39 (20.68)	0.56 (16.08)	0.15 (19.08)	-0.67 (17.95)	(0.26 (19.17)	0.03 (16.82)	0.86 (17.87)	-0.58 (17.45)	0.15 (15.30)
Weighted Defense DVOA Away	-0.41 (13.86)	0.35 (13.02)	-0.24 (14.09)	-0.27 (13.36)	-1.15 (14.37)	-0.25 (13.20)	-0.37 (15.97)	-0.91 (12.77)	-0.34 (13.27)	-0.05 (12.68)	-0.37 (11.84)	-0.49 (14.97)
Weighted Special Teams DVOA Away	-0.12 (5.45)	0.2965 (5.12)	0.23 (5.20)	-0.53 (5.50)	-0.78 (5.13)	0.14 (6.37)	0.01 (5.29)	-0.19 (5.36)	-0.04 (5.84)	0.15 (5.83)	-0.59 (5.90)	0.30 (5.69)
Observations:	1408	200	200	200	202	202	202	202	607	202	202	203

Table 1 shows the means of the actual outcome, line and DVOA statistics for the years combined and each year individually. The same information is displayed for the out-of-sample data.

Table 2: Sample Variables (Coefficient and Standard Error) **Out-of-Sample Data**

	All Years Combined	2003	2004	2005	2006	2007	2008	2009	**All Years Combined**	**2010**	**2011**	**2012**
Constant	-0.410	0.967	-1.412	0.418	0.46	-1.894	0.527	-1.347	0.247	0.866	0.611	-0.720
	(.46)	(1.27)	(1.31)	(1.13)	(1.36)	(1.17)	(1.24)	(1.26)	(0.71)	(1.27)	(1.24)	(1.24)
Line	0.909**	1.615**	0.66*	1.460**	0.74*	0.716**	0.975**	0.622*	0.852**	0.867**	0.970**	0.529
	(.11)	(0.31)	(.31)	(0.25)	(0.30)	(0.24)	(0.27)	(0.28)	(0.17)	(0.28)	(0.28)	(0.35)
Weighted Offense Difference	-0.047*	0.136*	-0.117	-0.021	0.004	-0.158**	0.042	-0.127	-0.067	-0.046	-0.025	-0.188*
	(0.02)	(0.06)	(0.06)	(0.04)	(0.07)	(0.05)	(0.06)	(0.06)	(0.04)	(0.06)	(0.07)	(0.07)
Weighted Defense Difference	0.013	-0.103	0.126	-0.119*	0.015	-0.052	0.0643	0.0695	0.005	0.006	-0.002	0.038
	(0.02)	(0.07)	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.07)	(0.03)	(0.06)	(0.07)	(0.06
Weighted Special Teams Difference	-0.100*	-0.121	-0.206	-0.119	-0.107	-0.0209	0.007	-0.010	-0.050	0.095	-0.016	-0.206
	(0.05)	(0.13)	(0.14)	(0.12)	(0.13)	(0.11)	(0.13)	(0.13)	(0.06)	(0.12)	(0.12)	(0.11)
R-Squared	0.1937	0.1691	0.2057	0.2968	0.1053	0.2906	0.1619	0.2672	0.1739	0.1507	0.1987	0.2144
Observations	1408	200	200	200	202	202	202	202	607	202	202	203

*p<0.05, **p<0.01

Table 2 shows the coefficients and accompanied standard deviation for the relationship between the DVOA variables and the actual outcome of a game.

Table 3: Model Success on Out-of-Sample Data

	All Years Combined	**2010**	**2011**	**2012**
Correctly Predicted	317	90	111	116
Incorrectly Predicted	290	112	91	87
Total	607	202	202	203
Percentage Correct	52.22%	44.55%	54.95%	57.14%
Amount Earned	$-20	$-332	$109	$203

Table 3 shows the success of the betting model created as a betting strategy.

The amount earned is based on a $10 bet for each game.

Table A: Adjusted Statistics and Line (Coefficient and Standard Error)

	All Years Combined
Constant	-2.731***
	(0.09)
Weighted Offense Difference	-0.160***
	(0.01)
Weighted Defense Difference	0.119***
	(.01)
Weighted Special Teams Difference	-0.056***
	(0.01)
R-Squared	0.6795
Observations	1408

***p<0.001

Table A shows the relationship between the DVOA adjusted statistics and the point spread value

References

Association, American Gaming. *Sports Wagering.* 6 30, 2012. http://www.americangaming.org/industry-resources/research/fact-sheets/sports-wagering (accessed 25 13, 10).

Betting Strategy. www.sportsinsights.com (accessed 10 15, 2013).

Borghesi, Richard. "Weather Biases in the NFL Betting Market: Explaining the Home Underdog Effect." 2004: 1-27.

Boulier, Bryan L., H.O. Stekler, and Sarah Amundson. "Testing the efficiency of the National Football League Betting Market." *Applied Economics*, 2006: 279-284.

Dare, William H., and Steven A. Dennis. "A Test for Bias of Inherent Characteristics in Betting Markets." *Journal of Sports Economics* 12, no. 6 (2011): 660-665.

Footballoutsiders.com Review. 2013. http://fantasylistings.com/reviews/FootballOutsiders_com_Review-155 (accessed 11 15, 2013).

GDP (current US$). 12 01, 2013. http://data.worldbank.org/indicator/NY.GDP.MKTP.CD (accessed 10 20, 2013).

Gray, Philip K., and Stephen F. Gray. "Testing Market Efficiency: Evidence From The NFL Sports Betting Market." *The Journal of Finance*, 1997: 1725-1737.

Osborne, Evan. "Efficient Markets? Don't Bet on It." *Journal of Sports Economics* 2, no. 1 (2001): 50-61.

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

Paton, David, and Leighton Vaughan Williams. "Forecasting Outcomes in Spread Betting Markets: Can Bettors Use 'Quarbs' to Beat the Book?" *Joutnal of Forecasting* 24, no. 2 (2005): 139-154.

Paul, Rodney J., and Andrew P. Weinbach. "An Analysis of the Last hout of Betting in the NFL." *International Journal of Sport Finance*, 2011: 307-316.

Paul, Rodney J., and P. Andrew Weinbach. "Does Sportsbook.com Set Pointspreads to Maximize Profits? Tests of the Levitt Model of Sportsbook Behavior." *Journal of Prediction Markets*, 2007: 209-218.

Paul, Rodney, Robert Simmons, and Andrew Weinbach. *Chapter 13: Gambling and the NFL.* Vol. 2, chap. 13 in *The Economics of the National Football League*, 225-242. 2012.

Pro Football Prospectus by Footbaloutsiders.com. 2013. http://www.twominutewarning.com/pro_football_prospectus.htm (accessed 11 15, 2003).

Sauer, Raymond D., Vic Brajer, Stephen P. Ferris, and Wayne M. Marr. "Hold Your Bets: Another Look at the Efficiency of the Gambling Market for National Football League Games." *Journal of Political Economy*, 1988: 206-213.

Song, ChiUng, Bryan L. Boulier, and Herman O. Stekler. "The comparative accruacy of judgmental and model forecasts of American Football Games." *International Journal of Forecasting*, 2007: 405-413.

Sung, Yoon Tae, and Scott Tainsky. "The National Football League Wagering Market: Simple Strategies and By week_related Inefficiencies." *Joutnal of Sports Economics*, 2012: i-20.

Wever, Sean, and David Aadland. *Herd Behavior and Underdogs in the NFL.* February 2010.