2019

NFL Betting Market Efficiency: Finding a Profitable Betting Strategy

Spencer Anderson
Skidmore College, sanderso@skidmore.edu

NFL Betting Market Efficiency: Finding a Profitable Betting Strategy

Spencer Anderson

May 2, 2019

This thesis is submitted in partial fulfillment of the requirements for the course Senior Seminar (EC375), during the Spring Semester of 2018

While writing this thesis, I have not witnessed any wrongdoing, nor have I personally violated any conditions of the Skidmore College Honor Code.

Name: Spencer Anderson

Signature: ___________________

Abstract

This paper uses pregame spread data to attempt to analyze the efficiency of the NFL betting market. I look at hot hand betting, performance during prime-time games, home field advantage, and favorites to test for inefficiencies in the market, thereby uncovering a profitable betting strategy. Using OLS regression analysis, I find no evidence of a profitable betting strategy betting on teams on streaks or home underdogs. However, my results suggest that it may be profitable to bet on the home team in prime-time games, regardless of their favorite or underdog status.

I. Introduction

In sports economics literature, it is almost universally agreed upon that the betting markets are efficient in the long run. This is because informed bettors offset any bias created by uninformed bettors (Paul & Weinbach, 2005). As a result, the point spread (the expected point differential right before the start of the game) is an efficient, unbiased indicator of the general expected outcome. Thus, sports betting markets provide an easy way to study market efficiency due to the known payouts and well-defined “end” period.

However, there may exist certain short-run inefficiencies that bettors can exploit for profits. The purpose of this paper is to attempt to find these inefficiencies that can result in a consistent profitable betting strategy in the NFL betting market. For instance, some papers determine that bettors overvalue home teams because of the “home-field advantage” phenomenon (Dare & Dennis, 2011). Conversely, other papers find that bettors overvalue the favorites, and that betting on the home underdogs is the way to go. A common misconception of bettors is that “hot” teams will continue to do well, thus operating under the hot hand fallacy. Other papers look at the strategy of betting against teams who are favored by a large margin (Aadland & Wever, 2012). In my study, I look at the accuracy of the spread line as a predictor of game outcome, hot hand betting behavior, home field advantage, favorites and underdogs, and prime-time performance to test for various strategies.

In the last decade, sports betting has exploded in the United States. The most popular sport to bet on is by far the National Football League (NFL). According to Statista.com, the market cap is nearing $1 trillion. The rise of daily fantasy sports (DFS) companies like DraftKings and FanDuel has paved the way for consumers to view sports in a different way; one that places emphasis on the individual performance of players, and offers payouts to bettors who select the best lineup of players in a given game. Many states are starting to legalize sports betting, which has expanded the market even further. The rise of social media has given much more exposure to the gambling implications of games (such as a team covering the spread in the last minute), as opposed to just the final score of the game. More importantly, these technologies allow information to spread much faster and thus make bettors more knowledgeable about their wagers. As a result of all these factors, many papers have attempted to explore the market in hopes of discovering profitable betting strategies. In economic terms, this can be seen as attempting to find inefficiencies in the market.

Many economists tend to look at the stock market to study market efficiency. However, the NFL betting market is much simpler. Thus, economists look to it as a more accessible way to study market efficiency. There is a plethora of literature on the efficiency of the stock market, and it is generally accepted and assumed that it is efficient in the long run. It is also assumed that stock prices reflect all relevant information available to the public (Fama 1970). However, the NFL market is a bit different, in that there is a known outcome of every game which generally reflects the effort and ability of each team during said game. Thus, the NFL betting market is a much more intuitive market to study market efficiency and given its vast increase in popularity and exposure, is very much worth exploring in the fields of sports economics and finance.

While there are many relevant wagers one can put on a single NFL game, there are three core bets that make up the majority of the betting market: the spread, the money line, and the over/under. The spread is a predicted score differential (away-home) between the two teams. For example, if the pregame spread is -7, then the home team is favored to win by a margin 7 or more points. Therefore, if one places a wager on the home team against the spread, he or she wins the bet if the team wins by 7 or more points and loses if the team either loses outright or wins by fewer than 7 points. Payouts for spread betting are one-to-one – that is, a successful bet will double the initial wager. Additionally, if the actual score differential is the same as the spread, then the bet is a "push" and the bettor gets his or her money back. As a result, bookmakers typically use 0.5 spread lines to combat this. Bookmakers also tend to set the spread such that there is about an equal number of bets on either side to minimize risk of taking a loss and maximize revenue. If the public bets heavily on one side, and that outcome occurs, they lose a lot of money. Thus, bookmakers constantly change spreads based on betting volume, new information, and bettor sentiment throughout the week

The money line is a bet on a team to win the game outright. It is categorized by either a positive or negative number, similar to the spread. If the number is positive, then the team is an underdog, and thus a riskier bet. The money line number itself represents the payout on a $100 bet. For instance, if a team is +200, then the bettor wins $200 (plus the initial $100 wager) if the team wins. Conversely, if a team is favored to win, their money line number will be negative, meaning it is a safer bet. If a team is -200, the bettor wins $50 on a $100 bet (oddsshark.com).

The over/under line is a bet not on who wins the game, but the total amount of points scored by both teams. Similar to the spread, the over/under line is typically set in 0.5 terms (i.e. 47.5 total points). Payouts are one-to-one here as well.

For the purposes of my study, I will focus just on the money line and the spread, as they focus more on the outcome of the game, rather than just the total number of points scored. The spread and money line "opening" and "closing" lines. The opening line is typically released the day after the previous week's game is played, comparable to a stock's price at the market opening time. As teams finish their games and look toward their next opponent, so do bettors and bookmakers. Throughout the week, these lines move around based on betting volume and increased information about the teams (injuries, roster changes, etc.), similar to stock prices. The closing line is determined and locked in at the start of kickoff.

As mentioned earlier, there may exist certain inefficiencies in the NFL betting market that bettors can exploit for profits. In this paper, I will explore a few of these scenarios. First, I will conduct an analysis to examine how well the spread predicts the outcome of the game. Second, I will examine the role of hot hand betting. When teams are on winning or losing streaks, bettors might tend to follow the hot hand. This phenomenon is common in other gambling settings (for example, Croson et al., 2005), so I want to see if it has a significant impact on the efficiency of the NFL betting market. Third, I will analyze home and away underdogs to see if there is a profitable betting strategy in this regard. Some economists claim that betting on home underdogs is more profitable than away underdogs (Dare & Dennis, 2011). Fourth, I am interested in the impact of a team's performance during prime-time games. Prime time games (Thursday Night Football, Sunday Night Football, and Monday Night Football) tend to have the highest viewership of all other games in a given week, and thus might garner more attention from bettors. A team's performance in these games may impact their opening spread for their next game. Furthermore, the increased exposure may incentivize players and coaches to put forth more effort than they otherwise might.

The sequence of this paper will occur as follows: Section II is a comprehensive literature review of all relevant economic theories related to my research, as well as similar market efficiency and NFL betting market works. Section III discusses my data set and analytical framework of my study, followed by empirical results in Section IV. Section V is a discussion of these results, including betting strategy recommendations based on my results. Section VI discusses the limitations of this study and provides insights for future research in this field.

II. Literature Review

Efficient Market Hypothesis (EMH)

The Efficient Market Hypothesis was developed by Fama (1970). It states that stock prices fully and accurately represent all information to the public. It assumes that current stock prices are impacted only by current information, and historical performance of stocks is not a factor in its present price. It also assumes perfect (and free) information. EMH literature often notes the random walk of prices, meaning that no single investor can "beat the market" without taking on extra risk because he has no information that is not also available to the public.

The EMH notes three types of efficiency: weak form, semi-strong form, and strong form. Weak form is defined as using historical returns and stock prices. Semi-strong form refers to new information, such as earnings reports, media coverage, etc.. Strong form is shown through something called "monopolistic information," i.e. information that is not available to the public. Generally, Fama (1970) notes that the first two forms generally give us the strongest insights with which to test market efficiency.

There are many drawbacks to the EMH. Nath (2015) critiques this theory for a few reasons. Firstly, the 2008 Financial Crisis resulted in many of the traditional finance theories being challenged due to "lack of perspective on the markets (Nath, 2015). Nath claims that if the assumptions about efficient markets were held, the Financial Crisis would not have occurred. Aside from the applications of EMH, the theory itself has also been challenged. Nath (2015) claims that there is significantly more information available (social media, investment news on the internet, etc.) that no single investor is able to consume every piece of information about a certain market. Thus, decisions are influenced more by sentiment and emotion than rationality. As a result, the relatively new field of behavioral finance have attempted to explain the anomalies that cannot be explained by traditional finance theory.

Another paper that challenges the EMH is Howden (2009). He notes that a common assumption of market efficiency literature is that all individuals interpret information the same way. In reality, people value stocks in different ways, and new information might not affect one investor as heavily as it might another. He notes that the transaction costs (fees) associated with stock trading hurts the efficiency of the stock market. As a result, a profitable strategy has to account for the commission cost on top of the amount invested. Third, the time it takes between locking in a trade and the processing time of the transaction often results in a difference between the buying price. The main takeaway of this paper is that the EMH is not realistic enough to apply to financial markets. However, it nonetheless serves as a benchmark to analyze markets, and applies quite well to the NFL betting market despite these limitations.

The EMH and the NFL Betting Market

In the NFL betting market, under the assumption of the EMH, there would be no consistently profitable betting strategies, under the same assumption that no one can exploit insider information. Additionally, the general commission for sportsbooks is about five percent per wager (oddsshark.com). Thus, over the long run, there should be a net loss of around five percent. However, over the short run, profitable strategies may prove to be successful as I have hypothesized.

One of the more prominent papers that looks at the overall efficiency of the NFL betting market is Gray & Gray (1997). Again, the NFL betting market is a simpler, more accessible means to study market efficiency because payoffs are known with certainty in advance, which clears up much of the ambiguity of other markets like the stock market. The authors include a variable in their probit regression analysis that accounts for past performance of the teams. One of the main reasons for this is that people tend to follow the hot hand or “herd” in gambling (Aadland & Wever, 2012), meaning more people bet on the teams that are hot, or on a winning streak. Gray & Gray (1997) find a few profitable betting strategies from their model, such as betting on teams that have had strong overall seasons but have had played poorly in recent weeks (also known as contrarian strategy). In other words, betting against the public is a potential way to exploit this inefficiency in the market.

This paper builds on ideas from Golec & Tamarkin (1991), who were among the first to apply market efficiency tests to the National Football League. They look at data for the 1973-1987 seasons. Their main contribution is that they add two dummy variables – the favorite status and home-away status of the team. Under the assumption of the Efficient Market Hypothesis, these variables would not add bias to the population regression function. However, their results show the only profitable strategy over the course of their sample was to bet the underdog consistently. According to the results, bets on the underdog won roughly 56% of the time. Thus, they believe that the NFL betting market is not completely efficient. However, this paper is fairly old and the increase in technology (thus more information available to the public) since then might affect the results found here.

Richard Zuber is one of the leading researchers in the field of sports gambling, especially its market efficiency. His paper with Gandar and Bowers (1985) is one of the first to use NFL point spread data to test the efficiency of the betting market. Their model is an extension of that of Pankoff's (1968). They incorporate several performance statistics of both teams, based on the assumption that the outcome of the game is a function of the collective efforts of both teams. They do a piecewise regression function, one for each week of the season. All variables were significant with the expected sign (i.e. positive or negative). They conclude that their model can be used to predict NFL point spreads due to the "speculative inefficiencies" in the market. However, these do not imply market inefficiency as a whole.

A few years later, Sauer et al. (1988) contradicted the results of Zuber et al. (1985), stating that i) the 16-part model created biased and inaccurate results and ii) was too small a sample size, as it was only for one season. Sauer et al. (1988) perform the same betting strategies proposed in Zuber et al. (1985), and find opposite results (they saw large losses as opposed to the profits seen in the Zuber et al. paper).

Zuber et al. (1988) responded to the criticism by building on their previous research. Their sample includes data on the 1980-1985 NFL seasons, using the same variables as their previous work, finding no significant inefficiencies in the closing betting lines. Instead of performing standard regression tests for efficiency, they use tests for rationality to see if people make rational decisions when betting. They create seven tests in total, three of which focus on the rationality of betting behavior. Of these tests, none produced profitable results. Therefore, they conclude that the market is generally dominated by uninformed bettors, contradicting the claim that the informed bettors offset the uninformed bettor inefficiency. This differs from Paul & Weinbach (2005), who state that the informed bettors offset the inefficient shift in spreads created by uninformed bettors.

Amundson et al. (2006) extends the work done by Zuber et al. (1985, 1988). However, the main variable of interest in their study is the surface on which the game is played. Similarly to before, they perform an OLS regression function of the realized point difference on the point spread, including a dummy variable for if the field is grass or turf. Because bookmakers consider overall record (both outright and against the spread), variables for these are included as well. Teams whose home field is turf may have trouble playing on grass. However, the results were statistically insignificant, meaning that NFL players do not struggle to transition between field surfaces.

Continuing with field conditions, Borghesi (2007) attempts to find a profitable betting strategy when considering the weather during football games. The model uses the same methodology as Amundson et al. (2006), but includes a variable that captures the difference in average temperatures in the away team's home town and average temperature of the city in which the game is played. This is relevant because many bettors believe teams that typically play in warmer weather will struggle when on the road against cold-weather opponents. The results show support for this belief, and there seems to be an inefficiency in the market here. Betting on home teams the northern U.S. when facing an away team from the south seems to be profitable according to this study. As a result, weather is something I will include in my analysis.

Home Field Advantage and Favorites

One of the more common factors that bettors and bookmakers alike consider is the home field advantage phenomenon. Historically, the home team tends to perform better than their season statistics may reflect. A common strategy in NFL betting is considering the home-underdog bias. Essentially, this theory assumes that road favorites struggle to cover the spread. Thus, bettors may overvalue the success of road favorites. Dare & Dennis (2011) test this theory with the goal of uncovering a profitable betting strategy betting on home underdogs (i.e. against the road favorites). Using data from the 2005-2011 seasons, they analyze the difference between the spread and the actual outcome of the game (mean forecast error). After doing this, they run a t-test against 0 to find where there might be bias towards home teams and favorites. Their results are statistically significant, indicating a potential profitable betting strategy betting on home teams.

Humphreys et al. (2013) critique the work of Dare & Dennis (2011) to test if the models hold true over a larger sample size. Their results are similar, but Humphreys et al. (2013) find that the results in both studies are part of a bigger phenomenon going on here. They were able to reject the null hypothesis of a 0 mean forecast error for away favorites, but failed to reject the same null hypothesis for home underdogs. Thus, this study suggests bettors favor teams that have been successful in covering the spread (who tend to also be the favorites). They assume (based on behavioral economic theory) that bettors tend to bet on better teams (here defined as teams that score a lot) because there is higher utility involved with watching higher scoring games. Thus, Humphreys et al. (2013) construct a new model using betting percentages on favorites. They try to find a profitable betting strategy by analyzing the percentage of winning bets, but find no such strategy betting against favorites with this new model. To conclude, bettors prefer betting on favorites, but favorites do not cover enough to consistently make profits.

Fodor et al. (2014) examines holdover bias in the NFL betting market to explore overbetting of the favorites early on in the season. During week one of the season, most information on spreads relies on the team's performance last season. Using data from the 2004-2012 seasons, the study looks at playoff teams from the prior season playing against non-playoff teams early in the season. In the study, the playoff teams only covered the spead 35.7% of the time. These results suggest that it may be profitable betting the underdogs early on in the season based on the lack of information from early-season play.

These findings are confirmed by Davis et al. (2015), who claim that week 2 of the season is the most likely week for bettors to "beat the market." Similarly to Fodor et al. (2014), they assume that there is less information available early in the season. If a 1-0 team is favored against an 0-1 team, then the point spread shifts up to a mean of 5.83 points, which is quite high. In this study, they find that underdogs exceed expectations and are thus undervalued. However, the EMH is a long-run theory and we cannot draw any conclusions about inefficiencies in the market from 1-2 weeks of the season. Nonetheless, it is one strategy that can be analyzed further in the future.

The final betting strategy I want to explore is a team's performance during prime-time games. This element of betting is seldom discussed in NFL betting literature due to the relatively small sample size, but one paper that does examine this is Shank (2018). The sample used in this study is more up-to-date than those of the previous literature I have discussed to this point. The purpose of this more recent sample is to investigate if some of the previous inefficiencies have been corrected in the market over time. Shank (2018) finds that it is profitable to bet on the home teams in prime time games (Sunday Night, Monday Night, and playoff games). These games, per Shank (2018), yielded a 60% winning percentage, with a 67% rate if the home team is the underdog. Thus, Shank (2018) provides insights that the home-underdog strategy noted in Dare & Dennis (2011) is magnified even further when the games are played in prime-time slots. However, this paper does not propose any form of regression model, it is simply a historic analysis. To my knowledge, no other prominent paper with a similar methodology to mine accounts for performance during prime-time. As a result, I hope to contribute to the economic discussion of NFL betting market efficiency in this light.

Hot Hand and Representative Heuristics

Representative heuristics are the mental shortcuts we take to judge probabilities (Tversky & Kahneman 1974). Simply put, this can be anything that speeds up the time it takes for us to make a decision. However, this increased speed often hinders the accuracy of the decision. These shortcuts often lead to unfavorable decisions because they cause the individual to overestimate the likelihood of a certain outcome based on previous events. Both the gambler's fallacy and the hot hand fallacy are examples of this.

Broadly defined, the hot hand fallacy is the false or exaggerated belief that an individual's performance varies systematically over the short run. In other words, individuals that act under this phenomenon believe that the next outcome of a certain event is likely to be similar to previous outcomes (i.e. that outcome is "hot"). For example, in basketball, many fans, players, commentators, etc., believe that there is systematic variation in shooting percentage. If the player is on a streak of making shots, he is perceived as "hot" and thus his teammates will continue to pass him the ball on offense. Applying this to sports betting, bettors may favor teams that have won a few games in a row, and are thus "hot."

The literature for the hot hand in both sports and gambling is shown in a variety of ways. Most of it stems back to a paper written by Tversky & Kahneman (1974) titled "Judgment Under Uncertainty." This work was the first documented discovery of these representative heuristics. A plethora of papers followed in the wake of this paper. One of the most groundbreaking was Gilovich et al. (1985), who were among the first to discover the hot hand fallacy. Using shot data on the Philadelphia 76ers players, they found no evidence of a positive correlation between successive shots, thus determining the "fallacy."

One of the first field studies on this topic was Croson et al. (2005), who used video footage of a roulette table at a casino in Reno, NV as their empirical setting. This paper found evidence of bettors exhibiting hot hand behavior, in that they were more likely to place subsequent bets when they previously won compared to having previously lost. However, this research design does not control for the income (i.e. "house money") effect – the idea that people who have more chips than when they started are more likely to bet more money. Additionally, the authors were not able to control for the amount of chips that people bet, as the data came from security tapes and thus could not capture the amount of chips being wagered.

Xu & Harvey (2014) examine the hot hand fallacy in online sports gambling. They explored three things – whether or not the hot hand fallacy is actually present, the bettor's beliefs in this behavior, and if there is a causal relationship between the gambler's fallacy (the belief that if something happens more frequently than normal it will happen less frequently in the future) and the hot hand fallacy. The authors found more evidence of the hot hand fallacy than the gambler's fallacy, consistent with the findings of Croson et al. (2005) as mentioned above. When people won, they continued betting on events with "safer" odds. When people lost, they bet on events with riskier outcomes. However, the interesting thing here is that people actually believed in the gambler's fallacy. Thus, by operating under the gambler's fallacy, they actually developed the "hot hand." By this logic, the authors argue that the gambler's fallacy created the hot hand effect in gambling.

Applying the hot hand to sports betting markets, Paul & Weinbach (2005) assess the effect of hot hand betting in the NBA. In their study, they use data from the 1995-2002 seasons. They claim that misconceptions of the public can be exploited for profits (i.e. against the hot hand). This was first hypothesized by Camerer (1989), who found that it could be profitable to bet against teams on winning streaks. Under his hypothesis, point spreads are influenced by the hot hand, but the outcomes are not. In Paul & Weinbach (2005), the results indicate that the public tends to overbet teams on winning streaks. Furthermore, they find that betting on NBA teams with winning streaks of four or more reject a fair bet, and betting *against* teams on a win streak of two or more games was found to be profitable on average. Furthermore, the betting public for this sample seems to not pay as much attention to losing streaks as they do winning streaks. This might be because there is less utility involved with rooting for a losing team. The main takeaways from this paper are that big underdogs, especially home underdogs, are the most profitable teams to bet on.

The main insight that Paul & Weinbach (2005) contribute is that bettors tend to overvalue teams on winning streaks, since they are “hot” and will likely continue to succeed in their eyes. One of the first papers to apply this to the NFL was Woodland & Woodland (2000). However, they found no profitable betting strategies betting against streaks. Thus, there is no inefficiency betting against the hot hand. However, a limitation of this claim is the assumption that bookmakers construct spreads such that they attract an equal amount of bets on each side of the line. As a result, spreads for teams on winning streaks are unusually high.

There is a large number of variables that bookmakers consider when generating their point spreads, over/unders, and money lines. Such variables include field surface (i.e. grass or turf), weather, temperature, etc. Amundson et al. (2006) look at field surface, using an OLS regression model with observed point difference as the dependent variable and predicted spread as the main independent variable, using a dummy variable for if the field’s surface is grass or not. Other variables in their model include overall season record and total accumulated point spread for the season. Both these variables have been shown to play a large role in bookmakers decisions, so their model can analyze the effect of the field on the teams playing in the game. For instance, a team whose home field is on turf might have a hard time playing an away game on grass. However, the results were insignificant and no profitable betting strategy was found from this analysis.

III. **Data and Analytical Framework**

My data set contains every NFL game from the 2006-2014 seasons, including preseason, regular season, and postseason contests. Certain variables of interest include the result of the game, pregame spread, and opening and closing money lines. It was obtained from Professor Mike Lopez in the Skidmore College Mathematics Department. I chose to omit preseason games for two reasons. Firstly, people rarely bet on them, so in that regard it is not worth analyzing. Second, teams play them differently than regular season games, as they treat them as practice games where they can rest their starters to avoid injury and play their less-impactful players to give them a shot at making the roster. This method of purposeful omission is also implemented by several NFL betting papers such as Gray & Gray (1997), Zuber et al (1988), among many others. I also chose to omit Super Bowls, since they are played at a neutral site, eliminating the home field advantage phenomenon, which I want to analyze in my study. Finally, not all games were designated an opening and/or closing money line in the data set. This is potentially because they are so-called "pick-em" games, where the teams are evenly matched so bookmakers don't set a line on them. For this reason, I omitted all those games, leaving me with a total of 2,355 observations over eight seasons.

Table (1) contains descriptive statistics for my data set. The average home spread was -2.94, meaning that on average, the home team was favored to win by a margin of roughly three points. In other words, if one were to bet on the visiting team to cover the spread, they would either have to i) win outright, or ii) lose the game by fewer than 3 points. The average Over/Under line was 43.3 points. However, average total points (home team points + away team points) was 44.27. Thus, the O/U line slightly underestimates the total points scored in a game in my sample. On average, visiting teams scored roughly 21 points (about three touchdowns per game), where the home teams averaged around 23 points. Home teams were generally favored on the money line, consistent with the "home-field advantage" phenomenon shared by bettors and bookmakers. There is a shift from the opening to closing money line for both the home and away teams. This can most likely be attributed to the release of new information throughout the week, such as injuries, changes in field conditions, or anything that might impact the outcome of the game in the eyes of bettors. Finally, average temperature was 58 degrees in the outdoor stadiums. Temperature was not recorded for games played in domes or closed roofs.

My first step in my methodology is to run a simple t-test against 0 to see if the difference between the spread and point differential is statistically significant:

H_0: *HomeSpread – PointDifference = 0*

H_A: *HomeSpread-PointDifference* $\neq 0$

I find that the p-value is greater than 0.05, so I fail to reject the null hypothesis. This means that on average, the spread is a decent indicator that the spread predicts the score rrelatively efficiently.

As we can see in Table (1), the standard deviations are quite high for the money lines. To mitigate this effect on my regression models, I will convert them into a probability function that the home team will win, putting it on a 0 to 1 scale. This model was developed by Cortis (2015), and is used in various sport economics papers, (Ge 2018, Mani, 2018, among others). The equation for this function is shown in equation (1):

$$Money\ Line = \begin{cases} +100\left(\frac{1-p_i}{p_i}\right), p_i \leq 0.5 \\ -100\left(\frac{p_i}{1-p_i}\right), p_i > 0.5 \end{cases}$$

In this equation, p_i represents the probability of team X winning game i. This probability is calculated by the bookmakers given all available information. Using these probabilities, I regress the observed point difference (here defined as $PointDiff_{it}$) in each game (away score-home score) on the win probability. This will allow me to predict the spreads for each game. According to Card & Dahl (2011), the best fit for this model is a third-order regression function. From this regression function, I estimate the spread from the money lines. I use this technique to create predicted spreads as a proxy to the spreads themselves, as my sample does not distinguish between opening and closing spreads. However, it does include opening and closing money lines, so I use this regression to get a predicted opening and closing spread. Equation (2) is shown below.

$$PointDiff_{it} = \beta_0 + \beta_1 ProbWin_{it} + \beta_2 ProbWin_{it}^2 + \beta_3 ProbWin_{it}^3 + \varepsilon_{it}$$

My first regression function is a similar model to those of Amundson et al. (2006) and Borghesi (2007). It is shown in equation (3). The main dependent variable is the observed point differential, while my independent variable is the home point spread. The dependent variable *PointDiff* is assumed to be a function of the collective effort of both teams, as well as the exogenous measures impacting that effort (Pankoff, 1965). These exogenous measures serve as control variables in my function. They are the stadium type, field surface, and the temperature at the start of the game. Stadium type, labeled as *Dome* in my regression function, is a dummy variable that takes a value of 1 if the field is a dome or closed roof, and 0 if otherwise. Field surface, labeled as *Grass*, takes a value of 1 if the field on which game *i* is played is a grass surface, and 0 if it is any type of artificial turf.

For the *Dome* variable, I hypothesize that teams whose home fields are in domes or closed roofs will tend to score more. Thus, I expect the sign of the coefficient to be positive. However, I do not expect the results to be statistically significant, as the controlled environment can be advantageous to the visiting team as well. Additionally, I expect that these teams will tend to struggle on the road, as they are not as used to playing outdoors as the home team. Similarly, I expect that teams whose home fields are on grass as opposed to turf will perform better based on the findings of Amundson et al. (2006). However, I do not expect the opposite effect (grass teams playing on the road on turf) to be as impactful. This model is shown below in equation (3):

$$PointDiff_{it} = \beta_0 + \beta_1 Spread_{it} + \beta_2 Dome_{it} + \beta_3 Grass_{it} + \beta_4 Temp_{it} + \varepsilon_{it}$$

For game *i* in season *t*, if $\beta_1 > 1$, the favorite team tends to cover the spread. This can be interpreted as a one-point change (increase or decrease in the spread) is linked to a more than one-point change in the observed score difference. If β_1 is less than one, then the favorite team tends not to cover the spread. If this is the case, a one-point change in the spread is linked to less than a one-point change in the actual score difference. I hypothesize that β_1 will be positive, meaning that the spread slightly over-predicts the observed score.

Next I will explore another model derived by Borghesi (2007) that more closely tests the impact of temperature on the outcome of a game. In this model, he simply regresses point differential on the pregame spread with the same control variables while adding in a temperature squared variable. The model is shown below in equation (4):

$$PointDiff_{it} = \beta_0 + \beta_1 Spread_{it} + \beta_2 Dome_{it} + \beta_3 Grass_{it} + \beta_4 Temp_{it} + \beta_5 Temp_{it}^2 + \varepsilon_{it}$$

However, when checking for multicollinearity for this model, I find that the variance inflation factors (VIFs) for *Temp* and $Temp^2$ are above 5, so there is multicollinearity here. Thus, I do not include $Temp^2$ in the rest of my methodology.

The next part of my analysis is looking at the home-field advantage phenomenon and how it plays into certain betting strategies. It is generally accepted that playing at home is more advantageous to playing on the road. Historically, home teams tend to perform better. Thus, the spreads generally reflect this sentiment. In my sample, the home team is usually favored by around 3 points. Furthermore, in the sample, home teams are the favorite approximately 66% of the time (1,571 of 2,355). To go about assessing the home field advantage effect, I include a dummy variable for when the home team is the favorite to equation (3). This variable takes a value of 1 if team A is the home team, and is also the favorite to win the game, and 0 if otherwise. The home-field advantage regression function is shown in equation (5):

$$PointDiff_{it} = \beta_0 + \beta_1 Spread_{it} + \beta_2 HomeFavorite_{it} + \beta_3 Dome_{it} + \beta_4 Grass_{it} + \beta_5 Temp_{it} + \varepsilon_{it}$$

My hypotheses for this function come from the results of Dare & McDonald (1996) and Aadland & Weaver (2010) that the spread coefficient (b_1) in this regression function will be positive and greater than 1. This would mean that the home favorites tend to cover the spread more often than not. I also expect b_1 to be greater than the b_1 in equation (3), meaning that home favorites tend to cover the spread more often than the general favorites do.

For my prime-time analysis, I incorporate a dummy variable to control for if the game takes place during a prime-time slot. I also include the *HomeFavorite* variable from Equation (5) because according to Shank (2018), it is profitable to bet on the home team in prime-time games, especially if the home team is an underdog. These games take a value of 1 in my sample if the game is nationally broadcasted on Thursday Night Football, Sunday Night Football, or Monday Night Football (and 0 if otherwise). I expect similar spread coefficients to my previous models, and a significant coefficient for the prime-time variable. The model is below in Equation (6):

$$PointDiff_{it} = \beta_0 + \beta_1 Spread_{it} + \beta_2 HomeFavorite_{it} + \beta_3 Dome_{it} + \beta_4 Grass_{it} + \beta_5 Temp_{it} + \beta_6 PrimeTime_{it} + \varepsilon_{it}$$

Finally, I will look at hot hand betting behavior by once again regressing the point differential on the spread with my controls. However, I will also include winning and losing streak variables for both home and away teams as well. If the home team is on a 1-game winning streak, then the variable takes a value of 1. When the home team is on a winning streak of 2, the variable takes a value of 2. When the home team is on a streak of 3 or more wins, then the variable takes a value of 3. Conversely, when the home team is on a 1-game losing streak, then the variable takes a value of -1. The same methodology is applied to the rest of home losing streaks, as well as the away team winning and losing streaks. Given the possibility of two teams on the same streak playing each other, I expect the signs to vary and not have any pattern. However, I cannot refer to literature for this, as pretty much all sports economics papers that focus on hot hand betting tend to have access to betting volume data instead of just analyzing the outcome of the game. The hot hand fallacy is a "fallacy" for a reason, and thus I expect no profitable betting strategy from betting against teams on streaks. Furthermore, I expect no significant results. The hot hand model is shown below in equation (7):

$$PointDiff_{it} = \beta_0 + \beta_1 Spread_{it} + \beta_2 HomeFavorite_{it} + \beta_3 Dome_{it} + \beta_4 Grass_{it} + \beta_5 Temp_{it} + \beta_6 HomeLoseStreak_{it} + \beta_7 HomeWinStreak_{it} + \beta_8 AwayLoseStreak_{it} + \beta_9 AwayWinStreak_{it} + \varepsilon_{it}$$

IV. Empirical Results

Table 2 shows the initial results of my population regression function shown in Equation (3). The spread coefficient was positive at 1.036, meaning a one point change in the spread is associated with a 1.036 point increase in observed score differential. The p-value is statistically significant at the 1% level. The r-squared value is .177, meaning that 17.7% of the variability in my sample is explained by my regression function. This is fairly strong. Due to the uncertain nature of sports, I doubt that I will be able to come up with a model that consistently predicts the outcome of a given game since so much can happen to change the odds of it. The other significant variable was *Grass*, meaning that teams whose home field is played on grass as opposed to turf tend to perform better against the spread. The coefficient for *Grass* is 2.523, which is quite large. The temperature and dome variables were not significant.

Table (3) shows the results of Equation (4), where I extend my initial regression function to include an additional control for temperature. Again, the spread coefficient is significant at the 1% level. Here, a one point change in the visiting spread is associated with a 1.036 point decrease in observed score differential. The R-squared is the same as in table 3, and the temperature-squared term is insignificant (p-value = 0.553). Thus, we cannot say for certain that temperature has an impact on the outcome of the game. Furthermore, this model has multicollinearity per my robustness check, so I do not include $Temp^2$ in the rest of my models.

Table (4) shows the regression results for equation (5). The home spread had a positive significant coefficient of 1.120, meaning a one-point change in the spread is associated with a 1-point increase in point differential. The home favorite coefficient was -1.298. To interpret this, we would say that if a team is the home team *and* the favorite, they tend to score 1.298 fewer points than expected. However, the p-value is 0.273, so the coefficient is not statistically significant. Consistent with my previous models, the *Grass* coefficient was statistically significant with a coefficient of 2.531. My *Temperature* and *Dome* coefficients were not significant. The R^2 was 0.177, similar to my previous models.

In Table (5), the results for the prime-time analysis are shown. *Home Spread* was positive and significant at the 1% level with a 1.040 coefficient. *Grass* was also significant at the 1%

level with a 2.503 coefficient. *Dome* and *Temperature* were insignificant with coefficients of 0.878 and 0.012, respectively. *PrimeTime* had a negative coefficient of -1.510 and was significant at the 10% level. This indicates that for the most part, the score difference is closer in prime-time games than otherwise. Since the score difference is categorized as away score minus home score, this result shows that the home team tends to perform better in prime-time games. This result is consistent with the findings of Shank (2018), confirming my hypothesis. Perhaps the p-value will become more significant with a larger sample size. After, there are significantly fewer prime-time games played per season.

Table (6) shows the results for the model that incorporates the home favorite term from equation (5), paired with the Prime-Time analysis from table (5). Similarly, *Home Spread* and *Grass* were statistically significant at the 1% level. *PrimeTime* was again statistically significant at the 10% level. *HomeFavorite* had a positive coefficient of 0.661 but was not statistically significant. This suggests that the home teams tend to do better in prime-time games, regardless of if they are the favorite or not. Shank (2018) finds a significant winning percentage of home underdogs in prime time games, so my results differ from him in this light.

Table (7) shows results for the hot hand model. As expected, *Home Spread* and *Grass* were statistically significant and positive as with all other models. The coefficients for all home and away winning and losing streaks had inconsistent signs and were not statistically significant. I expected teams on winning streaks to perform better and teams on losing streaks to perform worse. This would mean the coefficients for home winning streak and away losing streak would be negative, while away winning streak and home losing streak would be positive. However, my results show that home teams on two or three (or more) winning streaks to favor the away team. Additionally, the results favor away teams on one or three game winning streaks. The home team is favored when they are a two game losing streak or when the away team is on a one or two game winning streak. However, the coefficients are not statistically significant, so we cannot conclude that hot hand betting is a profitable strategy.

For the above models, I perform a few robustness checks. First, I check for heteroskedasticity. I perform a Breusch-Pagan test for heteroskedasticity for my models. For all models, the p-value for the test was greater than 0.05, so I fail to reject the null hypothesis of constant variance, confirming that there is no heteroskedasticity. Since I am using panel data, I test for multicollinearity by obtaining the VIF's for my regression function. The VIF's are shown below in table (8). They are all fewer than 5, meaning that there is no imperfect multicollinearity in my model. I also perform a Hausman test to decide whether I should use a fixed effects or random effects model. I fail to reject the null hypothesis, so I use a random effects model.

V. Discussion of Results

Equation (3) is a standard OLS regression that examines the relationship between observed point differential and pregame spread, stadium type, field conditions, and temperature. Results are displayed in Table (2). The significant coefficients are b_1*(spread)* and b_3(*grass)*. Both are significant at the 1% level. Per the EMH, if the NFL market is efficient, then the coefficient for b_1 should be 1. In other words, the relationship between the spread (usually made by Nevada bookmakers) and the actual point differential should be a 1:1 ratio. Thus, there do in fact exist inefficiencies in this market.

The *dome* coefficient (b_2) is positive, which implies that playing in a dome actually favors the away team (spread is calculated as away points – home points), holding all other variables constant. It is statistically insignificant, however. The grass (b_3) coefficient is statistically significant with a positive coefficient of 2.52. This is one of the more valuable insights of this paper. Road teams tend to perform significantly better when they play on grass as opposed to turf.

Temperature (b_4) was also positive, which means that increases in temperature favor the away team while decreases in temperature give the home team the advantage. However, we cannot draw any reasonable conclusions from this, as the p-value is quite high at .573. This result differs from the results of Borghesi (2007), who find statistically significant results for changes in temperature on the outcome of the game.

In Equation (4) I implement a temperature-squared term to further assess the impact of temperature on the outcome of the game. The coefficient is quite small and is insignificant, so no insightful conclusion can be drawn from this model, other than that temperature is not a great indicator of the score of the game.

Equation (5) looks at the home field advantage phenomenon as it relates to the spread. Similar to my previous models, *Dome* and *Temperature* variables are statistically insignificant. $HomeFavorite_{it}$ variable is -1.298, indicating that when the home team is the favorite, the score should decrease and when the road team is the favorite, the score should increase. The p-value is statistically insignificant at 0.273, but it is an interesting phenomenon nonetheless.

My hot hand model analyzes how teams on streaks perform relative to teams not on streaks. The signs for all coefficients for home and away winning/losing streaks were inconsistent across the board and insignificant, suggesting that there is not a profitable betting strategy betting on or against teams on streaks. This paper does not consider betting volume, but other papers that analyze hot hand betting focus more on the behavior rather than the actual results. Thus, I cannot say for certain how bettors perceive teams on streaks; I can only suggest that there is not a conclusive strategy when considering the streakiness of a team.

My prime-time model yielded a statistically significant *PrimeTime* coefficient, suggesting that the home team tends to benefit from playing their games in the national spotlight. However, I could not draw any conclusions about the favorite or underdog status of the team due to the high p-value of *HomeFavorite*. Nonetheless, there could be a profitable betting strategy when considering prime time games. However, bookmakers might be aware of this and set their odds such that uninformed bettors lean more toward the away team.

While much of the literature looks at the opening and closing spreads to look for inefficiencies, I did not have that data available to me. As a proxy, I want to replicate this methodology using the opening and closing money lines in the future. This is important because these changes over the course of the week can provide valuable insights into bettor sentiment with the release of new information. For instance, in finance theory, investors tend to overreact to new information (Howden, 2009). Thus, when the spread (or money line) shifts, it is due to new information and can cause bettors to overreact, thus shifting the lines more than they should.

VI. Limitations and Extensions

One of the primary limitations of my study is that I did not have access to betting volume data. This limited me from being able to observe betting behavior with spread changes, prime time games, hot hand betting, etc. Access to betting volume would bring light to whether bookmakers use imbalanced books as a tool for profit maximization. Another limitation that my work does not account for is the opening and closing spreads. To account for this, I used closing money lines as a proxy to predict the spreads in Equation (1). As a result, these are not the true spreads, which could result in some bias. Additionally, I did not look at the shift in spreads throughout the week. This is because an assumption of my study is that the closing spreads are the best indicator of the outcome of the game. Furthermore, I did omit all Super Bowls, which are among the most heavily bet sporting events in the world. However, this is not as big a limitation as the previous two.

When including my streak variable, I was unable to control for the "bye week," which is the one week in the season between weeks 4 and 12 (each team has a different one) where the team does not play. I was unable to create a variable to control for this, as my data set was constructed in a way that did not display when the bye week was.

I used OLS for my analysis which has its benefits. However, many bettors are only interested in if their bets win or lose, and not the margin of victory or defeat. Thus, there is a strong case for a probit model being a more appropriate means of analyzing efficiency. This is because probit models do not account for the excess points by which a team covers the spread. Some papers do use probit models for their analytical frameworks, but I chose to go with OLS for reasons stated in section III.

Finally, I found multicollinearity in one of my models (results shown in table (4)) with the addition of the $temperature^2$ term, so I was unable to include this in the rest of my analysis. According to Borghesi (2007), the paper on which this model is based, the weather favors teams playing in the north. However, I was unable to continue with my analysis of this due to the multicollinearity here.

Despite the limitations of my study, I believe it can open some doors for potential future research. First, one of the main findings in this study is that field surface has a significant impact on the outcome of the game. Only one paper lightly touches on this (Amundson et al., 2006). Thus, more work can be done on this to see if the visiting opponent's home surface is significant as well. Second, my prime-time coefficient was positive but insignificant. However, Shank (2018) notes a 60% covering percentage on home teams during prime time games. A potential avenue for future research is relating the hot hand betting to prime-time performance. I would suggest analyzing the opening spread for a team the week following a prime-time game. They might be overvalued if they do well and undervalued if they get beat.

The prime-time and hot hand biases could be linked together in future research as well. In most literature that looks at the shift in spread throughout the week, hot hand biases are more often observed early in the week. We could see some this effect magnified with teams who played on prime-time the week before. Further research using early-week spreads might shed some more light on bettor perception of teams on winning or losing streaks.

There is a large volume of literature relating home field advantage and favorites. Some proves that the home team's ability to score is underestimated by bettors, while some find the opposite. Some papers suggest that large underdogs receive more hype than not (meaning a lower spread than expected). The point is that the literature is inconclusive on the home field advantage phenomenon and how to bet favorites. Finally, applying a different methodology (probit and mean difference error are the two most common across sports betting literature) could net different results.

VII. **Conclusion**

This paper uses spread and money line betting data to expose short-run inefficiencies in the NFL betting market. I convert the closing money line into a predicted spread line. I test multiple conditions that bettors tend to pay attention to when considering placing a bet on a game. These are prime-time games, betting on teams on streaks (hot hand), and home field advantage/favorites. Most of the literature and economic theory behind this paper suggest that the closing lines are the best indicator of the outcome of the game since the most information is available by that time. My results indicate that the spread overpredicts the observed point differential, but not enough to expose a long-run inefficiency. However, there is a profitable betting strategy betting on home teams in prime-time games. Additionally, home teams whose fields are played on grass tend to perform better than those who play on turf. This finding, while statistically significant in my analysis, requires further research to investigate.

The NFL betting market provides a more accessible (and interesting) alternative method to test market efficiency. It serves as a good proxy to something like the stock market because it is extremely difficult to test the validity of the EMH due to the vast amounts of information, the true value of a stock, or the uncertainty of decision-making. In the NFL betting market, the outcomes and payouts are known. Furthermore, the probability distributions are known (via the money line), so bettors have less ambiguity when making a bet than investors do when considering a stock.

In the stock market, investors tend to overreact to price shocks, but the market tends to correct itself over time. However, the NFL betting market does not have time to correct itself, as the bets are locked at the start of the game. Overreaction in the NFL is usually attributed to new information coming out about the teams. If this information can be distributed more efficiently, then maybe the market can run more efficiently. The inefficiencies found in the NFL betting market could also be a driving force behind why it remains illegal in many states. The uncertainty and unpredictability of sports, while providing common spectators with an entertaining leisure activity, is the main reason why gambling on is wildly addictive and can often lead to huge losses.

Through testing the efficiency of the NFL betting market, I conclude that it is overall efficient in the long run, but there are certain strategies that give a better chance at yielding profits. One caveat of my research is that I pooled all seasons together, instead of conducting my analysis on a season-by-season basis. Certain strategies might be profitable in one season but not another (like those found in Zuber et al. 1988). This might be the case, as football has so much uncertainty to it. Any team can beat any team on a given day. One explanation of the overall efficiency is the recent increase in technology and information between the teams and the public. Social media and the internet provide instantaneous information for the fans and bettors, so things like injury updates and field conditions come at the click of a button, which was not the case in previous literature. This means that with more information available, bettors are better informed and thus make the market more efficient.

Appendix: Tables of Results

Table 1: Descriptive Statistics

Variable	Obs	Mean	St. Dev	Min	Max
Spread	2355	2.94	5.89	-14.5	26.5
O/U	2355	43.3	4.56	31.5	59
Visiting Points	2355	20.96	10.14	0	59
Home Points	2355	23.3	10.5	0	62
Opening VML	2355	66.86	209.73	-625	1250
Opening HML	2355	-95.68	222.07	-2500	565
Closing VML	2355	63.64	216.2	-850	1500
Closing HML	2355	-93.75	227.66	-2400	671
Total Points	2355	44.27	14.05	9	99
Temperature	1936	58.85	16.24	-1	99

Table (2)

Variable	Coefficient
HomeSpread	1.036*** (0.053)
Dome	0.890 (1.261)
Grass	2.523*** (0.708)
Temperature	0.014 (0.020)
Constant	-2.043 (1.268)
R-Squared	0.177
Observations	2355

Standard errors in parentheses
***p<0.01, ** p<0.05, *p<0.1

Table (3)

Variable	Coefficient (St.Err)
Home Spread	1.036*** (0.053)
Dome	0.934 (0.74)
Grass	2.537*** (0.709)
Temperature	0.014 (0.020)
$Temperature^2$	0.000 (0.000)
Constant	-1.616 (0.848)
R-Squared	0.177
Observations	2,355

Standard errors in parentheses
***p<0.01, ** p<0.05, *p<0.1

Table (4)

Home Spread	1.120*** (0.094)
Home Favorite	-1.298 (1.183)
Dome	0.859 (1.227)
Grass	2.531*** (0.718)
Temperature	0.014 (0.020)
Constant	-1.409 (1.362)
R-Squared	0.177
Observations	2,355

Standard errors in parentheses
***p<0.01, ** p<0.05, *p<0.1

Table (5)

Variable	Coefficient
Home Spread	1.040*** (0.054)
Dome	0.878 (1.226)
Grass	2.503*** (1.095)
Temp	0.012 (0.020)
PrimeTime	-1.510* (0.847)
Constant	-1.691 (1.248)
R-Squared	0.178
Observations	2355

Standard errors in parentheses
***p<0.01, ** p<0.05, *p<0.1

Table (6)

Variable	Coefficient
Home Spread	1.078*** (0.079)
Dome	0.846 (1.227)
Grass	2.510*** (0.718)
Temp	0.012 (0.020)
PrimeTime	-1.499* (0.847)
HomeFavorite	0.661 (0.993)
Constant	-1.340 (1.355)
R-Squared	0.178
Observations	2355

Standard errors in parentheses
***p<0.01, ** p<0.05, *p<0.1

Table (7)

Variable	Coef
Home Spread	1.036*** (1.261)
Dome	0.890 (1.261)
Grass	2.523*** (0.708)
Temp	0.014 (0.020)
1HomeLoseStreak	0.026 (0.858)
2HomeLoseStreak	-0.729 (1.122)
3HomeLoseStreak	0.634 (1.085)
1HomeWinStreak	-0.102 (0.858)
2HomeWinStreak	0.593 (1.155)
3HomeWinStreak	0.389 (1.152)
1AwayLoseStreak	0.871 (0.813)
2AwayLoseStreak	-0.510 (1.149)
3AwayLoseStreak	0.723 (1.254)
1AwayWinStreak	-0.689 (0.885)
2AwayWinStreak	-0.536 (1.117)
3AwayWinStreak	0.166 (1.120)
Constant	0.227 (1.493)
R-Squared	0.177
Observations	2355

Standard errors in parentheses
***p<0.01, ** p<0.05, *p<0.1

Table (8): VIF Results (Equations 3-7)

Equation (3)

	VIF	1/VIF
Home Spread	1.020	0.977
Dome	1.240	0.804
Grass	1.250	0.801
Temp	1.080	0.977
Mean VIF	1.150	

Equation (4)

	VIF	1/VIF
Home Spread	1.023	0.977
Dome	1.245	0.804
Grass	1.256	0.796
Temp	28.969	0.035
Temp^2	28.957	0.035
Mean VIF	12.29	

Equation (5)

	VIF	1/VIF
Home Spread	2.21	0.452
Dome	1.245	0.803
Grass	1.249	0.801
Temp	1.08	0.926
Home Favorite	2.191	0.456
Mean VIF	1.595	

Equation (6)

	VIF	1/VIF
Home Spread	1.025	0.976
Dome	1.243	0.804
Grass	1.249	0.80
Temperature	1.079	0.927
Prime Time	1.003	0.997
Mean VIF	1.12	

Equation (6) w/ added Home Favorite term

	VIF	1/VIF
Home Spread	2.211	0.452
Dome	1.245	0.803
Grass	1.249	0.80
Temperature	1.081	0.925
Prime Time	1.004	0.996
Home Favorite	2.192	0.456
Mean VIF	1.497	

Equation (7)

	VIF	1/VIF
Home Spread	1.022	0.978
Dome	1.246	0.802
Grass	1.25	0.800
Temperature	1.07	0.934
1HomeLoseStreak	1.002	0.998
2HomeLoseStreak	1.112	0.899
3HomeLoseStreak	1.003	0.997
1HomeWinStreak	1.052	0.950
2HomeWinStreak	1.026	0.974
3HomeWinStreak	2.012	0.497
1AwayLoseStreak	1.024	0.976
2AwayLoseStreak	1.046	0.956
3AwayLoseStreak	1.092	0.915
1AwayWinStreak	1.01	0.990
2AwayWinStreak	1.72	0.581
3AwayWinStreak	1.1	0.909
Mean VIF	1.1741875	

References

Aadland, D., & Wever, S. (2012). Herd behaviour and underdogs in the NFL. Applied Economics Letters, 19(1-3), 93-97.

Amundson, S., Boulier, B. L., & Stekler, H. O. (2006). Testing the efficiency of the national football league betting market. Applied Economics, 38(3), 279-284.

Ayton, Peter, and Ilan Fischer. "The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness?." Memory & cognition 32, no. 8 (2004): 1369-1378.

Borghesi, R. (2007). The home team weather advantage and biases in the NFL betting market. Journal of Economics and Business, 59(4), 340-354.

Camerer, C. (1989). Does the Basketball Market Believe in the Hot Hand? American Economic Review, 79(5), 1257-1261.

Card, D., Dahl, G.B., 2011. Family violence and football: the effect of unexpected emotional cues on violent behavior. Q. J. Econ. 126 (1), 103–143.

Cortis, D., 2015. Expected values and variances in bookmaker payouts: a theoretical approach towards setting limits on odds. J. Predict. Mark. 9 (1), 1–14

Croson, Rachel, and James Sundali. "The gambler's fallacy and the hot hand: Empirical data from casinos." Journal of risk and uncertainty 30, no. 3 (2005): 195-209.

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.

Davis, J. L., L. McElfresh, K. Krieger, and A. Fodor. 2015. "Exploiting Week 2 Bias in the NFL Betting Markets." The Journal of Prediction Markets 9 (1): 53–67.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383-417.

Fodor, A., Difilippo, M., Krieger, K., & Davis, J. (2013). Inefficient pricing from holdover bias in NFL point spread markets. *Applied Financial Economics, 23*(17), 1407-1418.

Ge, Q., 2018. Sports sentiment and tipping behavior. Journal of Economic Behavior & Organization, 13(1), 1-15.

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.

Howden, D. (2009). Fama's efficient market hypothesis and mises's evenly rotating economy: Comparative constructs. Quarterly Journal of Austrian Economics, 12(2), 3-12.

Humphreys, B. R., Paul, R. J., & Weinbach, A. P. (2013). Bettor biases and the home-underdog bias in the NFL. International Journal of Sport Finance, 8(4), 294-311.

Mani, Shankar, "NFL Betting Market Efficiency: A Closer Look at the Final Day of Betting" (2018). *Economics Student Theses and Capstone Projects*. 94.

Nath, T. (2015, October 15). Investing Basics: What Is The Efficient Market Hypothesis, and What Are Its Shortcomings? Retrieved from https://www.nasdaq.com/article/investing-basics-what-is-the-efficient-market-hypothesis-and-what-are-its-shortcomings-cm530860

Pankoff, L. D. (1968). Market efficiency and football betting. Journal Of Business, 41203-214

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.

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.

Shank, C. (2017). Is the NFL Betting Market Still Inefficient? *Journal of Economics and Finance*, *42*(4), 818-827. doi:10.2139/ssrn.3022567

Tversky, Amos, and Daniel Kahneman. “Judgment under uncertainty: Heuristics and biases.” science 185, no. 4157 (1974): 1124-1131.

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.