

NFL Betting Market: Home Teams in Abnormal Weather

Nate Ackert

Econometrics

Prof. Rick Eichhorn

Due: 12/03/2018

Introduction

In 1997, Philip K. Gray and Stephen F Gray published *Testing Market Efficiency: Evidence from the NFL Sports Betting Market*. The article examines market efficiency in NFL betting markets from the perspective of fixed trading strategies. Essentially, the authors take a variety of betting strategies, and using their economic modelling, they examine which of strategy produces the best returns. In addition, this article produces its own probit model based on the traditional OLS (ordinary least squares) model. Efficiency is tested by searching for betting strategies that yield significantly positive returns, on average. If the market is efficient, there exists no such strategy, as the NFL game spread captures all relevant information.

Gray et al. found that the most efficient fixed betting strategy was to bet on home teams that were underdogs. In games from 1976 to 1994, this strategy had a winning percentage of 52.51 percent, which is only slightly above 52.4 percent, the minimum winning percentage requirement to break even.

This research concludes that a strategy that suggests betting on home-team underdogs will not yield consistent profit. Additionally, the authors acknowledge limitations of their model in their conclusion, stating that more exogenous variables, namely weather, could be used to improve the model. Intuitively, abnormal weather conditions in an NFL game should be an advantage to the any home team (including underdogs), as the conditions provide a less-familiar game environment for both teams. However, the geographical environment is more familiar to the home team, which theoretically adds an advantage of familiarity. Therefore, I hypothesize that betting on NFL home-teams in games of abnormal weather conditions, regardless of underdog or favorite status, will lead to positive returns. This basic hypothesis is illustrated mathematically in Figure 1

below, where HWP represents Home Team Winning Percentage, and AWC represents abnormal weather conditions.

Figure 1: Relationship Hypothesis

$$HWP = f(AWC) \quad (1)$$

Modeling

To begin, in order to examine the relationship between HWP and AWC , we must populate AWC with different aspects of abnormal weather. Figure 2 below reveals the population regression function for abnormal weather conditions, where $Wind$ represents abnormal wind speeds, $Humidity$ represents abnormal humidity, and $Temp$ represents abnormal temperatures. u_i will represent the error term.

Figure 2: Abnormal Weather PRF

$$AWC = \beta_0 + \beta_1 Wind + B_2 Humidity + B_3 Temp + U_i \quad (2)$$

In order to determine which weather conditions are "abnormal," we must define our terms. For purposes of this paper, abnormal wind will be defined as wind speeds above 15 miles per hour, abnormal humidity will be defined as over 80 percent humidity, and abnormal temperatures will be defined as under 30 degrees Fahrenheit or over 75 degrees Fahrenheit, as elements of cold and heat undoubtedly affect NFL players. Game score data as well as game-day weather data will be collected from 1,632 NFL games from 2007 to 2013.

Data

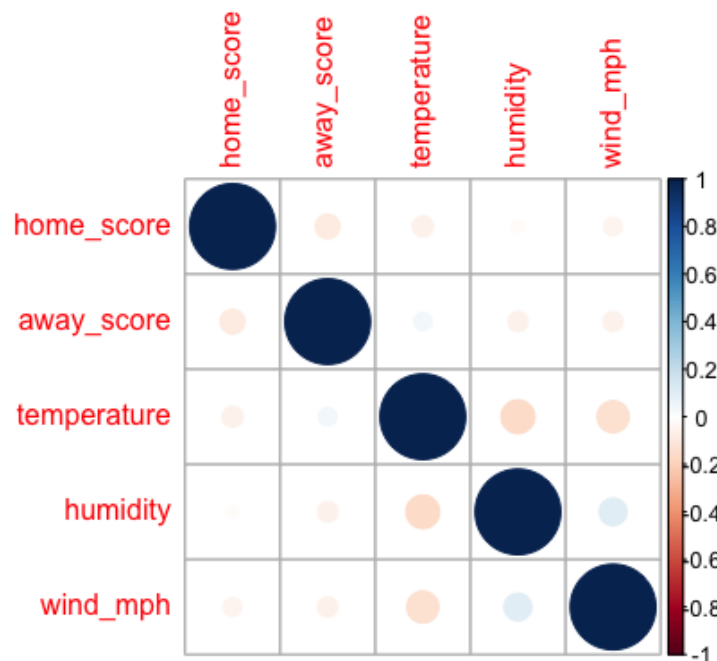
To test the hypothesis illustrated in the population regression function in Figure 2, this study will use NFL game data as well as weather data to evaluate the relationship between different abnormal weather conditions and the performance of home teams in NFL football games.

Of the 1,632 games covered by the data, 581 of exhibited at least 1 of the pre-established conditions of an "abnormal weather" game.

The data¹ for this paper comes from NFLsavant.com, a prominent NFL database. The data contains measures of home and away team scoring, humidity, wind, temperature and other weather notes from every NFL game. The weather data will be used to correspond with the game score data in order to identify any relationship, particularly with regard to performance of home NFL teams.

Diagnostics

From an initial look at my data, there appears to be no deterministic relationship between my variables, eliminating the possibility of serial correlation. The correlation matrix below illustrates the weak correlations of the system variables for this study.



Additionally, this as the figures in the Hypothesis Testing section of this paper will show, there was little evidence of heteroskedasticity, or a pattern of widening variance as the X variables increase.

Finally, one concern I have about my regression is the possibility of multicollinearity

¹NFL Savant: Advanced NFL Database

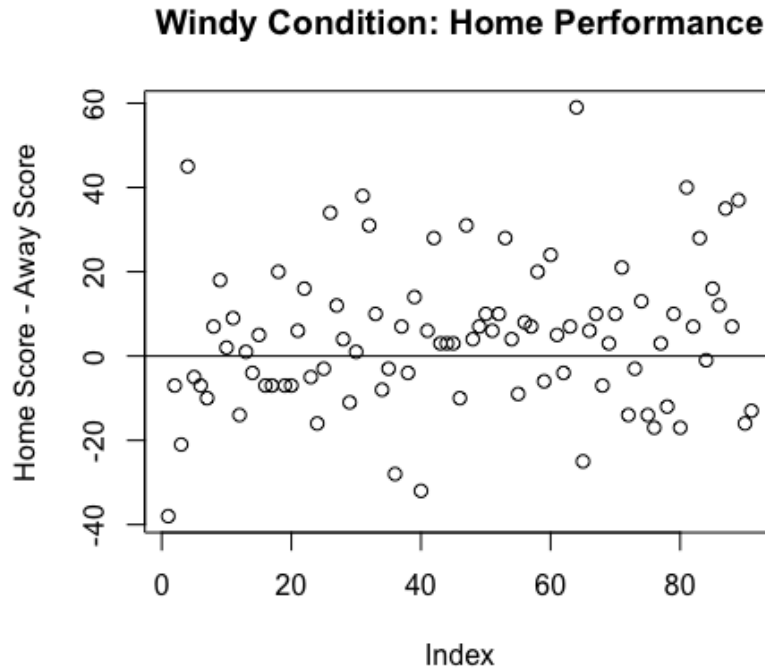
with regard to humidity and temperature. Multicollinearity occurs when two variables in a regression are correlated. Independent variables must be completely independent in order to determine their full effect in a regression, otherwise the correlated variables send the same signal to OLS. I worry that in some game conditions, particularly games played in humid climates such as Florida or North Carolina, it is possible that humidity and hot temperatures may have an existing relationship. However, this measure is not necessarily a deterministic relationship, and addressing the scientific relationship between coastal heat and humidity is beyond the scope of this paper. Therefore, these system variables will be left alone for the purpose of this paper.

Hypothesis Testing

In order to test the validity of my hypothesis, this paper will analyze each of the variables listed in the Abnormal Weather Conditions (*AWC*) population regression function in order to see if home team winning percentage could indeed a function of abnormal weather, and weather not a fixed betting strategy based on these variables could produce consistent returns. In order to measure consistent returns, this section will utilize the benchmark of a 52.4 percent breakeven point, the same metric that was used in the Gray et al. study that this paper cites.

To begin, we will examine the variable of wind. This study defines "abnormal" wind conditions as winds exceeding 15 miles per hour. Through 1,632 games of data, a total of 91 games experienced "abnormal" wind conditions. Figure 3 below illustrates the performance of home and away teams in games of abnormal wind conditions.

Figure 3: Windy Conditions



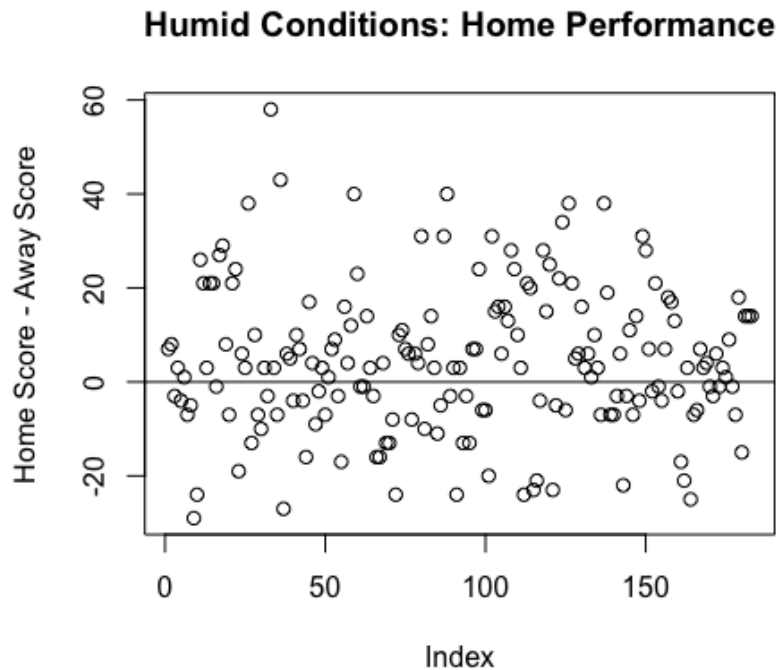
In the figure above, measurements above the horizontal axis (0) indicate wins, and measurements below the horizontal axis indicate losses. From Figure 3, We see that subtracting away team scores for each individual game leaves the data skewed slightly above the horizontal axis, indicating that home teams have slightly overperformed in abnormally windy games. Of the 91 games recorded with abnormal wind conditions, 60.4 percent of the games were won by the home team. Therefore, if a better were to bet on the home team every time a game was played in abnormally windy conditions, they would predict the outcome 60.4 percent of the time. The figure significantly outperforms Grey et al.'s 52.4 benchmark for breakeven, signifying that this may be a viable betting strategy for consistent returns.

Intuitively, it follows logically that home teams perform better in the wind than away teams. This is because home players are more familiar with their surroundings in their home stadium, including visual cues (stadium decor, fan seating arrangements, etc.) that can be used for throwing, catching, and performing other actions that become more difficult in abnormal wind conditions.

The next variable of our population regression function that we will test is the variable

of humidity. For this study, "abnormal" humidity is defined as a humidity percentage above 80 percent. Of the 1,632 games contained in the data, 183 games experienced abnormal humidity. Figure 4 below illustrates how home teams fared in games with abnormal humidity.

Figure 4: Humid Conditions



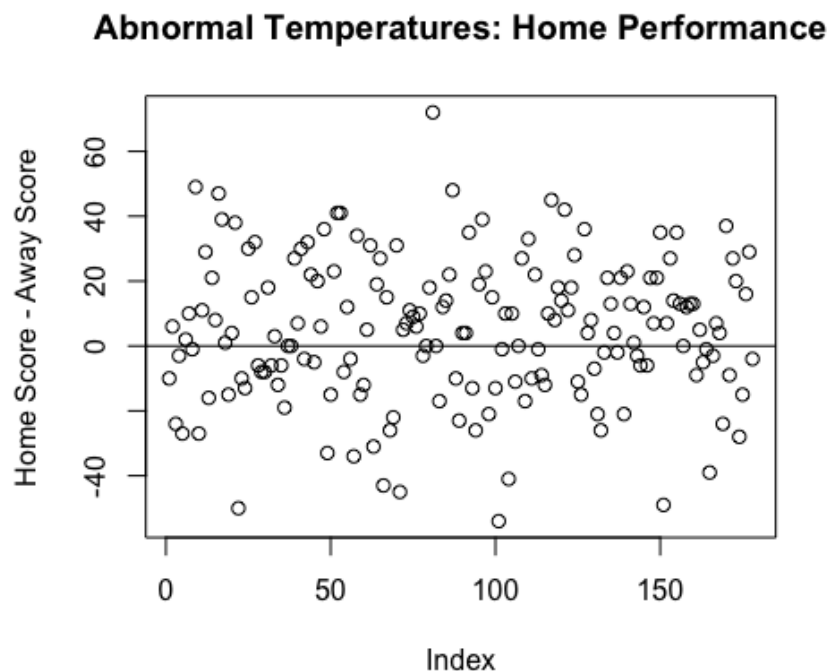
From looking at Figure 4 above, we see similar results to Figure 3, with home results slightly skewed above the horizontal axis. Of the 183 games recorded with abnormal humidity, the home team won 58.4 percent of the time. Using Gray et al.'s metric of 52.4 percent for breakeven, this figure is slightly above the 52.4 percent mark, but likely not enough to warrant a consistent fixed betting strategy.

It follows logically that a home team may perform slightly better than their opponents in games of abnormality. This may be because athletes are more familiar with hydration procedures at their home stadium - where the hydration stations are, familiarity with on-site trainers, and other procedure-related familiarities.

Next, we will examine the third variable of our Abnormal Weather Condition PRF: temperature. For the sake of this study, abnormal temperatures are defined as tem-

peratures above 75 degrees Fahrenheit or under 30 degrees Fahrenheit. Figure 5 below illustrates home team performance in abnormal temperatures.

Figure 5: Abnormal Temperatures

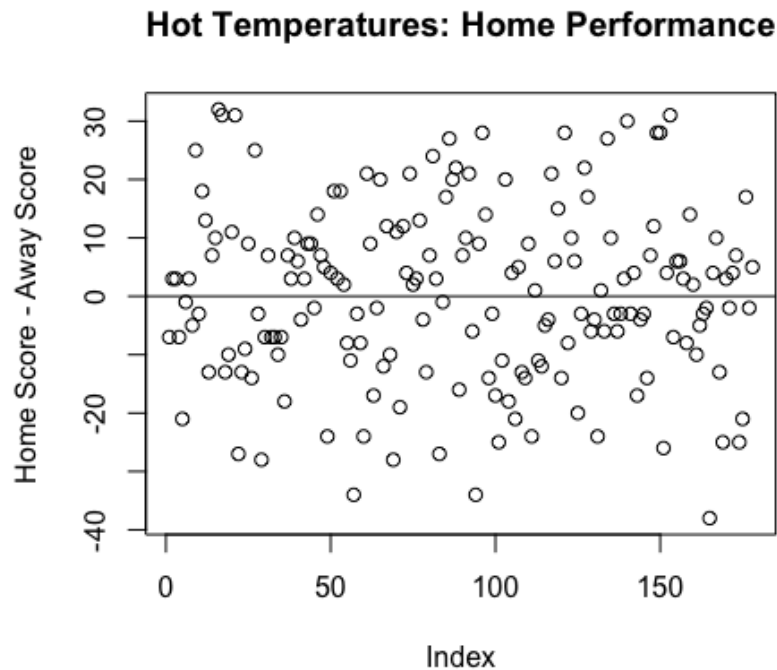


As we see from the figure above, the horizontal axis essentially splits the data, indicating no real overperformance by either home or away teams in abnormal temperatures. Out of 307 games played in abnormal temperatures, the home team won 53.42 percent of the games. While this is slightly above the 52.4 percent breakeven point, this hardly warrants a fixed betting strategy based on this variable.

While betting on home teams in abnormal temperatures may not be an optimal fixed betting strategies, there are two components to abnormal weather, and it is possible that one component of abnormal weather may provide more of an advantage to home teams than the other.

One component of abnormal temperatures is "hot" temperatures, which consists of games played in heat greater than 75 degrees Fahrenheit. Of the 1,632 games recorded, 178 of them were played in abnormally hot temperatures. Figure 6 below illustrates home team performance in abnormally hot games.

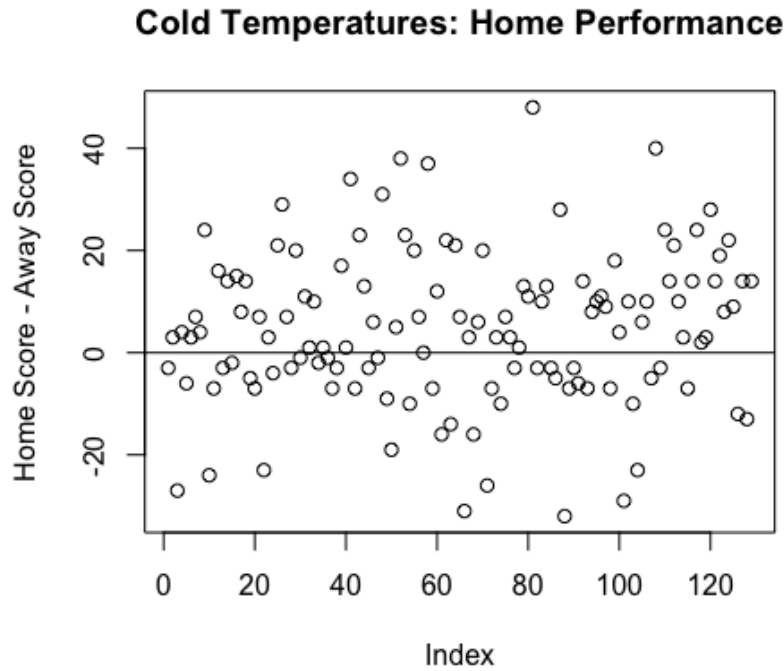
Figure 6: Hot Conditions



Looking at the plot above, we see that the horizontal axis once again essentially splits the data in half, revealing no real advantage for home or away teams in abnormally hot games. Of the 178 abnormally hot games, the home team won 47.78 percent of the time, which does not reach the 52.4 percent benchmark of a breakeven betting strategy. In fact, this measure of abnormal weather slightly favors the away team. This measure does not warrant a fixed betting strategy based on hot temperatures.

The second component of the abnormal temperature measure is "cold" temperatures, which consists of games played in sub-30 degrees fahrenheit temperatures. Figure 7 below illustrates home team performance in abnormally cold games.

Figure 7: Cold Temperatures



Of the data collected from 6 years of NFL games, 129 games were played in abnormally cold temperatures. As we see in Figure 7, the data is skewed above the horizontal axis, indicating an advantage for home teams in cold weather. Of the 129 games in cold weather, the home team won 79 of them, which amounts to a 61.29 winning percentage. This greatly exceeds the breakeven mark of 52.4 percent, indicating that a fixed betting strategy involving cold weather game may result in consistent returns.

Final Results

From our hypothesis testing of different abnormal weather conditions with regard to home team performance in NFL games, our final results are mixed. Figure 8 below displays the winning percentage of home teams under each of the abnormal weather conditions listed in the population regression function, as well as an average total for the home model. Because this paper's hypothesis is based on a fixed betting strategy of betting on strictly home teams, the home teams' winning percentage is effectively the percentage of games the model correctly predicts.

Figure 8: Results Table

Breakeven	Weather Total	Heat	Cold	Humidity	Wind	Total
52.4%	53.42%	47.78%	61.29%	58.40%	60.40%	56.11%

As the table above shows, of the five total variables included in the population regression function, four of them exceeded the breakeven point, indicating that these variables could contribute to positive returns if a betting strategy was put in place. That being said, among the 4 variables that exceeded the breakeven point, only two of them had significant-enough differences from the breakeven point to warrant a long-term fixed betting strategy: abnormally cold temperatures and abnormally windy conditions. As a whole, when combining all variables with relation to home team winning percentage, the model predicts the correct outcome 56.11 percent of the time.

Conclusions and Limitations

Based on NFL game and weather data, this model creates a betting strategy that correctly predicts the outcome of NFL games 56.11 percent of the time in the given sample. While this number is slightly above the 52.4 percent breakeven point, it is unlikely that, given the small sample size relative to the history of the NFL, that this model is an effective predictor of NFL games, and therefore does not provide a fixed betting strategy that will certainly produce positive returns.

The data shows that while all measures of weather do not necessarily provide advantages to home teams, there is reason to suggest that certain weather conditions are more advantageous to home teams than others. For example, cold weather had a much higher rate of success for home teams than warm weather. Cold weather and windy weather provided the strongest evidence of inefficiencies in the NFL betting market, inefficiencies which could be capitalized on for profit in the future.

While this model covers a large amount of weather and game data, there are limitations to this study. One limitation is the amount of games measured. While six seasons and 1,632 games creates a significant amount of data, this model could be improved with more data, as the NFL dates back to 1920.

Another limitation to this study is the absence of precipitation. While humidity covers the amount of water in the atmosphere, it fails to capture the affect of rain or snow on the outcome of NFL games. The absence of precipitation data is likely due to the difficulty of measurement when it comes to precipitations. For instance, it would be difficult to compare rainfall to snowfall in a single metric, as it is likely that snow and rain affect football teams in different ways.

Overall, due to the inability of the all of the variables in the PRF to significantly exceed the 52.4 percent benchmark of correct prediction, a fixed strategy of betting on home teams in abnormal weather conditions would not likely yield consistent returns. Therefore, we reject our null hypothesis that $HWP = f(AWC)$. Moving forward, this study could be improved with more data, data regarding precipitation, and by including more aspects of an outdoor environment, such as altitude, local vegetation, oxygen levels, coastal-relative location, and more.

References

- [1] Philip K. Gray, Stephen F. Gray *Testing Market Efficiency: Evidence From The NFL Sports Betting Market*. The Journal of Finance, September 1997.
- [2] NFLsavant.com: Advanced NFL Statistics, All Weather Data, 2013