Rationality and Efficiency in NFL Gambling Markets

Research Thesis

Presented in partial fulfillment of the requirements for graduation
with research distinction in Economics in the undergraduate
colleges of The Ohio State University

by

Alexander Williams

The Ohio State University
August 2021
Project Advisor: Assistant Professor John Rehbeck, Department of Economics

Abstract

Football pointspread betting permits unusually direct examination of market behavior compared to traditional asset markets. We present analysis of the pointspread market as a proxy for testing efficiency and rational expectations in other financial markets. This paper proposes novel statistical and economic tests to empirically appraise the rationality and efficiency of the National Football League pointspread gambling market. Econometric analysis of a 2012-2019 sample reveals several biased trends in pointspreads and deviations from rational expectations but fails to give evidence against efficiency. These results indicate that rational expectations might not be necessary for markets to retain long-run efficiency and suggest potential compatibility between decision models incorporating bounded rationality and those which assume neoclassical market efficiency.

Keywords: Efficient Markets Hypothesis, rational expectations, pointspread, American football, irrationality
JEL Classifications: G12, G14, L83

1 Introduction

It is often useful to empirically test concepts in economic theory. Developments in behavioral economics, combined with the 2007-2008 financial crisis, have cast considerable doubt on the theories of rational expectations and efficient markets. However, empirically testing the rationality and efficiency of traditional financial markets is often difficult or impossible.

The NFL pointspread gambling market's finite-horizon structure permits direct comparisons of observable ex post asset values with ex ante prices, enabling per se empirical asset pricing tests. The pointspread market thus avoids the "joint hypothesis problem" present in other financial markets. This advantageous feature, combined with the repeatedly noted (Pankoff 1968, Gandar et al. 1988, Levitt 2004, Lacey 1990, Gray & Gray 1997) similarities in structure between the pointspread and securities markets, offers an ideal environment for examining the theories of rational expectations and efficient markets.

To date, no firm conclusion exists concerning the rational expectations or efficiency of the pointspread market. Some authors claim to detect profit opportunities, while others offer evidence to the contrary. This paper seeks to ascertain what, if any, biased trends exist within the behavior of prices in the NFL pointspread betting market. We present, to date, one of the largest published datasets of NFL pointspread betting action, and test more betting strategies than any other authors. Furthermore, we explore the shortcomings of Pankoff's (1968) "partial rationality" model and offer improvements on his testing methods.

The data indicate that the NFL pointspread betting market's behavior is consistent with efficiency, despite the presence of nonrational expectations. Both pure statistical and direct economic tests fail to give decisive evidence of market inefficiency, while simultaneously indicating nonrational expectations and biased gambling behavior. Ordinarily, rational expectations of agents are a prerequisite for economic efficiency. However, given the persistent failure of portfolio managers to outperform market indexes despite well-documented behavioral biases present in financial markets, along with the work of authors such as Summers (1986) and Gulko (2005), these results support the growing body of literature that suggests irrationality might not be mutually exclusive with an efficient markets hypothesis (EMH) economic framework.

The sections of this paper proceed in the following manner. Section 2 describes the operation of the pointspread market. Section 3 offers statistical and economic tests of rationality and efficiency in the pointspread gambling market. Section 4 concludes the paper.

2 The NFL Gambling Market

Sports gambling is a large and high-profile industry in the United States. The American Gaming Association estimates annual gross betting on American football games to exceed \$95 billion (Purdum 2015), which National Football League pointspread wagering primarily constitutes.

The pointspread system operates in the following manner. At the start of the betting period, bookmakers establish uniform odds on an upcoming game, consisting of a betting line, PS, called the "pointspread", which estimates the game's result[1], and payout odds for each team, ω, which determine the ratio between money wagered and net payouts for a bet on that team[2]. Gamblers then bet on teams to "cover the spread", that is, to outperform the established line. During the betting period, bookmakers may nonuniformly adjust both the offered pointspreads and payout odds until the beginning of the respective games, after which no wagers are taken. The net profit to a gambler who wagers δ dollars on a team with a margin of victory MV will be:

$$\pi = \begin{Bmatrix} (\omega - 1)\delta & MV > PS \\ 0 & MV = PS \\ -\delta & MV < PS \end{Bmatrix} \quad \{1\}$$

As a direct consequence of Eq. {1}, the wins-to-bets ratio p required (ignoring pushes, i.e., where $MV = PS$) for a gambler to break even, assuming wager amounts are a constant δ over n wagers, will be:

$$\begin{aligned} E(\pi) = np(\overline{\omega} - 1)\delta + n(1-p)(-\delta) \\ = (p\overline{\omega} - 1)n\delta = 0 \Rightarrow \\ p\overline{\omega} - 1 = 0 \Rightarrow p = \frac{1}{\overline{\omega}} \end{aligned} \quad \{2\}$$

Notice that p is just the implied probability for a given outcome when ω is fair. So, a gambler that generates a long-run $\hat{p} > \frac{1}{\overline{\omega}}$ has discovered pointspread mispricing, as $\hat{\omega} = \frac{1}{\hat{p}} < \overline{\omega}$ is that the odds offered for the wagered-on outcomes exceed those implied by the ex post observed outcome probabilities.

To date, previous authors have apparently ignored real-world bookmaking arrangements, where ω is allowed to fluctuate, instead preferring to fix $\omega = \frac{21}{11}$. While this is the *modal* value for ω in our sample, nearly 70 percent (2836 / 4096) of possible wagers had a different value of ω. This heretofore unconsidered source of variation in pointspread pricing is a crucial factor in wager profitability, as we show above. Therefore, we incorporate ω into our empirical tests found in Section 3 as to develop a more robust evaluation of rationality and efficiency in the pointspread market. Note that throughout our analysis, we exclude "pushed" games from the data, as is common in the pointspread literature.

3 Empirical Methodology and Results

We present pooled cross-sectional data consisting of pointspreads, payout odds, wager proportions, final scores, and additional fundamental factors (which we discuss in Section 3.2) from each game in the 2012-2019 NFL regular seasons, a total sample size of 2048 games. We use pointspreads and payout odds quoted from Bovada, one of the largest online sportsbooks. We use wager proportions reported by *Wunderdog.com*, which aggregates real-money betting action data across the highest-trafficked online sportsbooks, including Bovada. Game dates, times, and final scores were taken from *NFL.com*'s official schedule webpage. We also utilize *FiveThirtyEight.com*'s NFL Elo ratings, which are an index of team strength.

3.1 Direct Economic Tests

The development of systematic profit opportunities in the pointspread market violates the efficient markets hypothesis. The most direct means of evaluating market efficiency is to conduct profitability tests of hypothetical gambling strategies.

Previous work in the pointspread literature typically proposes several novel strategies, and evaluates their profitability strictly in-sample. This standard approach immediately encounters several issues. Firstly, it is impossible to examine the entire set of possible pointspread gambling strategies. Failure to detect profitable betting rules may not indicate market efficiency, but rather that the authors selected the wrong strate-

[1] Typically, game results are positive (i.e., the winner has a positive margin of victory), and betting lines are negative (i.e., the favorite has a negative pointspread). To avoid confusion, this paper refers to betting lines as positive, in the same manner as game results.

[2] Payout odds can be expressed in several different units; we report odds ratios using the decimal system. For an excellent summary and comparison of the decimal and other competing odds units, see Cortis (2015).

gies to test. However, the question of which strategies to test is not altogether straightforward. Testing a wide enough array of strategies will inevitably result in a few being significantly profitable due to pure statistical variance. Some authors (e.g. Tryfos et al. 1984) attempt to address this by analyzing the rejection rate of the entire set of strategies they test as compared to various α levels, but this does not entirely resolve the issue either; such rejection rates are likewise greatly affected by selection bias. For instance, suppose the pointspread market currently underprices home teams, but an author only tests 1 strategy related to home-field advantage and 100 others relating to factors which are correctly priced- the author will have accuractely detected mispricing, but may incorrectly conclude, due to the low rate of discovery of significantly profitable strategies, that the pointspread market is efficient. Additionally, some strategies may overlap (e.g., betting exclusively on underdogs, and exclusively on home underdogs), and others may be able to be combined from or split into multiple distinct strategies (e.g. betting on teams with $PS \in [5, 10]$ might be split into the separate strategies betting on teams with $PS \in [5, 7]$ and with with $PS \in (7, 10]$, respectively). For many such strategies, there is no a priori reason to treat these together or separately; testing many narrowly-applicable strategies is likely to result in small sample size and low statistical power, while testing a few broad strategies may miss profitable rules which are only detectable at finer resolutions. Worse still, basing our conclusions for or against efficiency upon rejection rates forces them to principally depend upon detection of profitable strategies *and* that we do not select too many unprofitable strategies ex ante.

A more rigorous method of evaluating strategy profitability which avoids the aforementioned issues would be to pre-screen each betting rule's profitability with a sample separate from the main analysis. Those strategies which rejected the null hypothesis in the first sample would then be subject to testing in the main sample. If the rejection rate of the primary analysis exceeds α levels, then we may be confident to have detected a violation of the EMH. Alternatively, if the observed rejection rate is less than α, then we conclude that any previous detection of "profitable" strategies was only due to statistical chance. Performing tests of efficiency in this manner will guarantee that ex ante strategy specification does not bias rejection rates.

A wide survey of the pointspread-EMH literature reveals 141 distinct strategies proposed by previous authors[3], 39 of which yield in-sample wins-to-bets ratios significantly different from pure randomness (i.e., $p = 0.5$) under a two-sided binomial test. There is considerable variety in both the technical mechanisms these strategies use and their theoretical justifications (see further discussion below). These past results will serve as a sufficient screening process for the methodology described above. For an exhaustive list of the qualifying strategies, and the respective results of previous authors' tests, see Table 1.

The qualifying strategies generally attempt to exploit one or more of four main speculated inefficiencies: A) overreaction to historical performances, B) biased price movements that develop during the betting period, C) underpricing of home teams, and D) underpricing of underdogs. Variations on contrarian betting rules comprise a large class of strategies tested in the literature: Vergin & Scriabin (1978), Vergin (2001) and Woodland & Woodland (2000) note several profitable strategies of this variety, while Fodor et al. (2013) and Davis et al. (2015) identify sizable inefficiencies due to early-season overreactions by the pointspread market. Price-movement strategies are effectively contrarian as well, since they bet against trends in pricing and betting action, although they are more concerned with present market sentiment than the former class of betting rules, which target teams that underperform historical market pricing. Paul & Weinbach (2007, 2011) and Gandar et al. (1988) all discovered multiple price-movement strategies which were profitable over their respective samples. Underdog and home team strategies have a long, but more mixed track record, with some authors (Golec & Tamarkin 1991, Gray & Gray 1997, Levitt 2004) reporting small but significant underpricing, while others (Vergin & Scriabin 1978, Gandar et al. 1988, Tryfos et al. 1984) find more conflicting evidence.

It is well-known that the number of successful wagers s which a particular strategy yields over n total

[3] We review strategies proposed by Badarinathi & Kochman (1996), Gandar et al. (1988), Gray & Gray (1997), Golec & Tamarkin (1991), Lacey (1990), Levitt (2004), Tryfos et al. (1984), Vergin & Scriabin (1978), Paul & Weinbach (2007, 2011), Vergin (2001), Fodor et al. (2013), Davis et al. (2015), Shank (2018) and Woodland & Woodland (2000), excepting those which utilize pointspread advantage, a technique that place bets across several sportsbooks simultaneously, and has to date only been explored theoretically in the pointspread literature.

Table 1- Previous authors' results

Rule	Authors	n	$\hat{p}$	H_2 p-value
Bet on teams that are:				
underdogs	Golec & Tamarkin (1991)	3154	0.524	0.01^{***}
	Gray & Gray (1997)	4042	0.526	0.00^{***}
	Levitt (2004)	4793	0.518	0.01^{**}
	Tryfos et al. (1984)	1391	0.534	0.01^{**}
favorites by ≤ 3 points	Levitt (2004)	1559	0.522	0.08^{*}
favorites by ≤ 5 points	Gandar et al. (1988)	689	0.543	0.03^{**}
	Tryfos et al. (1984)	656	0.534	0.09^{*}
underdogs by > 5 but ≤ 10 points	Vergin & Scriabin (1978)	407	0.543	0.09^{*}
underdogs by > 5 points	Tryfos et al. (1984)	735	0.535	0.07^{*}
underdogs by > 15 points	Vergin & Scriabin (1978)	50	0.640	0.06^{*}
underdogs, and have $\leq 45\%$ of betting action	Paul & Weinbach (2007)	195	0.574	0.04^{**}
underdogs, and have $\leq 35\%$ of betting action	Paul & Weinbach (2007)	136	0.610	0.01^{**}
	Paul & Weinbach (2011)	290	0.566	0.03^{**}
underdogs, and have $\leq 30\%$ of betting action	Paul & Weinbach (2007)	103	0.641	0.01^{***}
	Paul & Weinbach (2011)	234	0.585	0.01^{**}
underdogs, and have $\leq 25\%$ of betting action	Paul & Weinbach (2011)	162	0.580	0.05^{**}
underdogs, in the first half of the season	Levitt (2004)	2209	0.523	0.03^{**}
underdogs, and are playing at home	Gray & Gray (1997)	1307	0.546	0.00^{***}
	Golec & Tamarkin (1991)	1115	0.556	0.00^{***}
	Levitt (2004)	1483	0.533	0.01^{**}
playing at home	Golec & Tamarkin (1991)	3154	0.515	0.09^{*}
Bet on teams whose pointspread, during the betting period:				
decreases	Gandar et al. (1988)	874	0.549	0.00^{***}
decreases, when in the previous week a majority of pointspreads were more accurate at the end of the betting period than at the beginning	Gandar et al. (1988)	365	0.570	0.01^{***}
decreases by ≥ 1 point, and are not playing at home	Shank (2018)	599	0.563	0.00^{***}
Bet against teams which in their previous game:				
won by ≥ 10 points	Vergin (2001)	1507	0.478	0.10^{*}
won by ≥ 20 points	Vergin (2001)	591	0.464	0.08^{*}
	Lacey (1990)	122	0.590	0.06^{*}
lost, whose opponents lost their previous game, and are favorites, in Week 2	Davis et al. (2015)	58	0.707	0.00^{***}
covered by ≥ 10 points as a favorite, and are favorites	Gandar et al. (1988)	167	0.581	0.04^{**}
Bet on teams which in their previous game:				
failed to cover by ≥ 15 points	Vergin (2001)	761	0.531	0.10^{*}
lost, and are underdogs	Woodland & Woodland (2000)	1540	0.526	0.04^{**}
failed to cover, and are underdogs	Woodland & Woodland (2000)	1410	0.525	0.07^{*}
lost, failed to cover, and are underdogs	Woodland & Woodland (2000)	1204	0.527	0.06^{*}
Bet against teams which in their previous 2 games:				
won twice	Woodland & Woodland (2000)	1303	0.526	0.07^{*}
won twice, and are favorites	Woodland & Woodland (2000)	861	0.532	0.07^{*}
covered twice	Lacey (1990)	320	0.578	0.01^{***}
Bet on teams which in their previous 2 games:				
lost twice	Lacey (1990)	320	0.425	0.01^{***}
lost twice, and are not playing at home	Shank (2018)	502	0.540	0.08^{*}
Bet against teams averaging the highest margin of victory over:				
their previous 2 games	Gandar et al. (1988)	79	0.608	0.07^{*}
their previous 4 games	Vergin & Scriabin (1978)	57	0.333	0.02^{**}
either their previous 1, 2, 3, 4 or 5 games	Vergin (2001)	954	0.466	0.04^{**}
Bet against teams averaging the highest cover margin over:				
their previous 2 games	Vergin (2001)	189	0.418	0.03^{**}
their previous 3 games	Gandar et al. (1988)	71	0.620	0.06^{*}
their previous 4 games	Vergin & Scriabin (1978)	59	0.373	0.07^{*}
either their previous 1, 2, 3, 4 or 5 games	Vergin (2001)	940	0.451	0.00^{***}
Bet against teams which in the previous year:				
covered 7 more games than they failed to cover, or whose opponents covered 7 less games than they failed to cover	Vergin (2001)	259	0.564	0.05^{**}
made the playoffs and whose opponents failed to make the playoffs	Lacey (1990)	327	0.450	0.08^{*}
made the playoffs and whose opponents failed to make the playoffs, in Week 1	Fodor et al. (2013)	59	0.644	0.04^{**}

$^{*}\ p < 0.1,^{**}\ p < 0.05,^{***}\ p < 0.01$

wagers follows a binomial distribution. What is less clear, however, is what null wins-to-bets ratio p_0 should be used to generate said distribution. The standard choices utilized by previous authors are $p_0 = 0.5$ and $p_0 = 11/21$. However, these fail to consider the variation in ω as discussed in Section 2. Since the p_0 set under a null hypothesis of gambler unprofitably is directly dependent upon the offered odds ratios, we posit that correct specification of p_0 should instead be made based upon the distribution of ω. This claim is supported by the data, as a chi-squared goodness-of-fit test was unable to reject that the observed cover probabilities were equivalent to those implied by ω (p = 0.15).

The most appropriate mechanism for modeling odds ratios such as ω is to fit a beta prime distribution to the data such that $\omega - 1 \sim \beta'(\hat{\alpha}, \hat{\beta})$, which would imply (by Eq. {2}) that $p \sim \beta(\hat{\alpha}, \hat{\beta})$[4]. Conveniently, the resulting beta-binomial tests are also tests of profitability, since if the observed $\hat{p}$ is significantly greater than the p_0 implied by the offered payout odds, then bookmakers must be offering ω greater than optimal and allowing bettors profit opportunities.

Under the assumed beta-binomial distribution, when each wager uses a constant $\delta = 1$, the probability of observing a given betting strategy which wins s out of n wagers under a null hypothesis of unprofitability is:

$$P(s \mid H_0 : \pi \leq 0) = \sum_{k=s}^{n} \binom{n}{k} \frac{B(\hat{\alpha} + k, \hat{\beta} + n - k)}{B(\hat{\alpha}, \hat{\beta})} \quad \{3\}$$

In addition to a beta-binomial test of H_0, we present for comparison binomial tests against the hypotheses $H_1 : \hat{p} \leq \frac{11}{21}$ and $H_2 : \hat{p} = \frac{1}{2}$. H_1 is equivalent to H_0 under the hypothetical scenario where all $\omega = \frac{21}{11}$, and H_2 is that the given strategy does not explain any variation in $MV - PS$. Binomial tests are more precise than the Z-tests utilized in much of the literature. While binomially distributed random variables do converge to normality as $n \to \infty$, modern computers and statistical software enable fast and accurate computation of binomial p-values, an advantage authors of the more dated parts of the pointspread literature did not enjoy. Therefore, we prefer to utilize binomial probabilities for significance testing of pointspread betting strategy profitability. The results of our tests are offered below in Tables 2 and 3, which report betting strategy profitability against lines at the start and end of the betting period, respectively.

Of the 71 total strategies tested (32 against opening lines, 39 against closing lines), none were profitable at the $\alpha = 0.05$ level, and only one was profitable at the $\alpha = 0.1$ level. Under the null hypothesis that *no* strategy is profitable in the long-run, we expect $(\alpha)(n)$ rejections of H_0 due to statistical noise. At 90 and 95 percent confidence, the expected number of rejections are 7.1 and 3.6, respectively. These results are clearly consistent with market efficiency and cast significant doubt on previous authors' claims of discovering profitable betting strategies. Interestingly, the single strategy which was significantly profitable at the 90 percent confidence level, betting closing lines against teams with the highest $\overline{MV}$ over their previous 4 games, was insignificantly so when using the standard binomial test against H_1. Indeed, tests of H_0 generally yield lower p-values than tests of H_1, indicating that our beta-binomial test may be more sensitive to successful betting rules than the regular binomial tests. However, this difference, while significant ($p < 0.001$), is relatively small in most cases, including for the aforementioned profitable strategy, and does not affect the overall harmony of the data with a conclusion of market efficiency.

The outlook of the data is considerably murkier when considering the (non)randomness of wagered-upon outcomes under a given betting strategy. Binomial tests rejected the null of H_2 at the $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.1$ levels 0, 3, and 10 times, respectively. While at the 95 percent (3.6 expected rejections) and 99 percent confidence (0.7 expected rejections) levels, we reject H_2 fewer times than expected, at 90 percent confidence (7.1 expected rejections), we observe over 40% more rejections than expected. Further complicating matters, every strategy which yielded a $\hat{p}$ significantly different from randomness was only so when betting against closing lines. While it would be statistically inappropriate to re-specify our selection of betting strategies *post hoc*, if we had limited consideration of strategies to betting against closing lines, the

[4]Upon investigation of the data, it is immediately clear that bookmaker odds-setting behavior is such that neither ω nor the implied cover probability $\frac{1}{\omega}$ follow any well-known probability distributions. We select the beta prime distribution for hypothesis testing due to its (relatively) closer approximation of ω as compared to alternative distributions, and for its well-established statistical properties for modeling probabilities and odds ratios.

Table 2- Betting Strategies vs. the Opening Line

				P-values		
Rule	n	$\hat{p}$	π/n	H_0	H_1	H_2
Bet on teams that are:						
underdogs	1911	0.509	-0.029	0.86	0.91	0.46
underdogs by ≤ 3 points	695	0.495	-0.055	0.91	0.94	0.82
underdogs by ≤ 5 points	1092	0.506	-0.033	0.84	0.88	0.69
underdogs by > 5 but ≤ 10 points	658	0.521	-0.004	0.54	0.57	0.29
underdogs by > 5 points	819	0.512	-0.022	0.73	0.77	0.53
underdogs by > 15 points	17	0.588	0.131	0.22	0.39	0.63
underdogs, in the first half of the season	900	0.517	-0.013	0.64	0.68	0.33
underdogs, and are playing at home	629	0.493	-0.061	0.92	0.94	0.75
playing at home	1985	0.487	-0.072	1.00	1.00	0.24
Bet against teams which in their previous game:						
won by ≥ 10 points	874	0.491	-0.061	0.96	0.98	0.61
won by ≥ 20 points	360	0.483	-0.075	0.92	0.94	0.56
lost, whose opponents lost their previous game, and are favorites, in Week 2	30	0.567	0.078	0.26	0.39	0.58
covered by ≥ 10 points as a favorite, and are favorites	253	0.482	-0.080	0.89	0.92	0.62
Bet on teams which in their previous game:						
failed to cover by ≥ 15 points	480	0.488	-0.069	0.93	0.95	0.62
lost, and are underdogs	1059	0.497	-0.052	0.94	0.96	0.85
failed to cover, and are underdogs	985	0.506	-0.034	0.84	0.88	0.75
lost, failed to cover, and are underdogs	883	0.502	-0.041	0.88	0.91	0.95
Bet against teams which in their previous 2 games:						
won twice	936	0.517	-0.012	0.63	0.67	0.31
won twice, and are favorites	615	0.528	0.009	0.40	0.42	0.17
covered twice	808	0.507	-0.031	0.79	0.83	0.70
Bet on teams which in their previous 2 games:						
lost twice	961	0.492	-0.060	0.96	0.98	0.65
lost twice, and are not playing at home	451	0.510	-0.025	0.70	0.74	0.71
Bet against teams averaging the highest margin of victory over:						
their previous 2 games	119	0.487	-0.066	0.76	0.81	0.85
their previous 4 games	106	0.566	0.082	0.17	0.22	0.21
either their previous 1, 2, 3, 4 or 5 games	308	0.503	-0.036	0.74	0.78	0.95
Bet against teams averaging the highest cover margin over:						
their previous 2 games	115	0.496	-0.050	0.69	0.76	1.00
their previous 3 games	109	0.514	-0.016	0.55	0.62	0.85
their previous 4 games	107	0.523	0.001	0.47	0.54	0.70
either their previous 1, 2, 3, 4 or 5 games	322	0.509	-0.025	0.67	0.72	0.78
Bet against teams which in the previous year:						
covered 7 more games than they failed to cover, or whose opponents covered 7 less games than they failed to cover	214	0.509	-0.025	0.64	0.69	0.84
made the playoffs and whose opponents failed to make the playoffs	942	0.485	-0.073	0.98	0.99	0.38
made the playoffs and whose opponents failed to make the playoffs, in Week 1	56	0.500	-0.046	0.59	0.69	1.00

The beta distribution used in tests of H_0 was $\beta(\hat{\alpha} = 1856.936, \hat{\beta} = 1687.538)$

$^{*}\ p < 0.1,^{**}\ p < 0.05,^{***}\ p < 0.01$

Table 3- Betting Strategies vs. the Closing Line

Rule	n	$\hat{p}$	π/n	P-values H_0	H_1	H_2
Bet on teams that are:						
underdogs	1952	0.520	-0.004	0.56	0.64	0.08*
favorites by ≤ 3 points	655	0.513	-0.018	0.64	0.72	0.53
favorites by ≤ 5 points	1067	0.522	0.000	0.52	0.56	0.16
underdogs by > 5 but ≤ 10 points	699	0.526	0.010	0.45	0.46	0.17
underdogs by > 5 points	885	0.518	-0.008	0.59	0.66	0.31
underdogs by > 15 points	26	0.500	-0.040	0.52	0.67	1.00
underdogs, and have $\leq 45\%$ of betting action	1377	0.524	0.001	0.50	0.52	0.08*
underdogs, and have $\leq 35\%$ of betting action	772	0.543	0.034	0.24	0.15	0.02**
underdogs, and have $\leq 30\%$ of betting action	491	0.536	0.019	0.34	0.32	0.12
underdogs, and have $\leq 25\%$ of betting action	234	0.547	0.034	0.26	0.26	0.17
underdogs, in the first half of the season	921	0.543	0.042	0.23	0.13	0.01**
underdogs, and are playing at home	727	0.516	-0.008	0.60	0.68	0.41
playing at home	1972	0.491	-0.057	0.92	1.00	0.46
Bet on teams whose pointspread, during the betting period:						
decreases	1577	0.521	0.000	0.55	0.61	0.11
decreases, when in the previous week a majority of pointspreads were more accurate at the end of the betting period than at the beginning	895	0.528	0.014	0.42	0.40	0.09*
decreases by ≥ 1 point, and are not playing at home	478	0.550	0.056	0.19	0.13	0.03**
Bet against teams which in their previous game:						
won by ≥ 10 points	870	0.503	-0.037	0.77	0.89	0.87
won by ≥ 20 points	364	0.505	-0.037	0.70	0.77	0.88
lost, whose opponents lost their previous game, and are favorites, in Week 2	31	0.548	0.036	0.33	0.46	0.72
covered by ≥ 10 points as a favorite, and are favorites	279	0.495	-0.055	0.78	0.85	0.90
Bet on teams which in their previous game:						
failed to cover by ≥ 15 points	465	0.501	-0.043	0.76	0.85	1.00
lost, and are underdogs	1095	0.516	-0.013	0.61	0.71	0.30
failed to cover, and are underdogs	1035	0.527	0.008	0.45	0.44	0.09*
lost, failed to cover, and are underdogs	904	0.524	0.004	0.48	0.50	0.15
Bet against teams which in their previous 2 games:						
won twice	937	0.530	0.016	0.39	0.36	0.07*
won twice, and are favorites	641	0.532	0.019	0.38	0.35	0.11
covered twice	777	0.530	0.017	0.40	0.37	0.10*
Bet on teams which in their previous 2 games:						
lost twice	952	0.504	-0.037	0.77	0.89	0.82
lost twice, and are not playing at home	449	0.521	-0.013	0.52	0.56	0.40
Bet against teams averaging the highest margin of victory over:						
their previous 2 games	122	0.541	0.041	0.34	0.39	0.42
their previous 4 games	105	0.590	0.134	0.09*	0.10	0.08*
either their previous 1, 2, 3, 4 or 5 games	313	0.534	0.020	0.37	0.39	0.26
Bet against teams averaging the highest cover margin over:						
their previous 2 games	117	0.564	0.087	0.19	0.22	0.20
their previous 3 games	111	0.532	0.018	0.41	0.47	0.57
their previous 4 games	106	0.528	0.013	0.43	0.50	0.63
either their previous 1, 2, 3, 4 or 5 games	329	0.541	0.033	0.29	0.28	0.15
Bet against teams which in the previous year:						
covered 7 more games than they failed to cover, or whose opponents covered 7 less games than they failed to cover	151	0.530	0.019	0.42	0.47	0.52
made the playoffs and whose opponents failed to make the playoffs	938	0.496	-0.051	0.85	0.96	0.82
made the playoffs and whose opponents failed to make the playoffs, in Week 1	58	0.517	-0.008	0.49	0.59	0.90

The beta distribution used in tests of H_0 was $\beta(\hat{\alpha} = 309.484, \hat{\beta} = 281.332)$

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

data would clearly support a conclusion that PS admits wagers which are significantly more likely to win than not. This does not indicate, however, that the probability of such wagers is high enough to ensure positive profits, as once we account for ω (such as in tests of H_0), the data instead supports the null of unprofitability. As it stands, whether bets against PS are equally likely win as to lose is wholly unclear, but we cannot ignore the possibility that closing lines are biased.

These findings are quite surprising given many previous authors' results. For example, the strong biases in early-season pointspreads as documented by Fodor et al. (2013) and Davis et al. (2015) are not replicated in our sample. Similarly, strategies that speculate on biased trends in betting action of the variety proposed by Gandar et al. (1988) and Paul & Weinbach (2007, 2011) are far less profitable (and more importantly, insignificantly so) in our sample as compared to the authors' original tests. This is not to say that such biases are not present in PS itself, as evidenced by the non-negligible number of rejections of H_2; however, simply considering the actual payouts of ω resolves these apparent "inefficiencies" discovered by previous authors. In this respect, our results indicate rigorous out-of-sample testing of betting strategies using real-market data is crucial when assessing the efficiency of the NFL pointspread market, lest we too hastily conclude the existence of profitable betting strategies.

3.2 A Regression Test of Efficiency

That pointspread gamblers behave in a manner consistent with rational expectations is that their conditional expectations are equivalent to true expectations. We explore the rationality (or lack thereof) of bettors in this sense via a regression test, which we present below.

The earliest and most common approach to recovering expectations from the pointspread data is a "partial rationality" test of the variety proposed by Pankoff (1968). The partial rationality method interprets PS as gamblers' conditional expectations, and MV as true expectations. Often, these partial rationality tests use a simple OLS regression test with MV and PS as the dependent and independent variables, respectively (e.g. Pankoff 1968, Gandar et al. 1988, Sauer 1998), although some authors also insert fundamental factors as independent variables into the partial rationality equation (e.g. Golec & Tamarkin 1991, Shank 2018, Sauer et al. 1988).

However, there are both theoretical and practical considerations which suggest these partial rationality tests are misspecified. Firstly, the pointspread gambler, unlike in other securities markets, faces a binary outcome: a wager, once placed, either wins, or loses, and offers a constant net payout in either case. No matter how large $MV - PS$ is, the gambler is only concerned that its value is positive, and profits equally for all such outcomes. Mathematically, the profit-maximizing gambler is interested in the value of the cumulative distribution function of MV evaluated at PS, not the expected values of MV and PS. Therefore, expectations in the pointspread market are better characterized by probabilities, not point-values. Additionally, Levitt (2004), supported by more recent authors (Simmons et al. 2011, Paul & Weinbach 2007, 2011), presents convincing evidence that bookmakers intentionally avoid setting market-clearing PS. Thus, PS may not accurately aggregate market sentiment, and instead only reflect bookmakers' expectations, not those of gamblers. Finally, these partial rationality tests have consistently failed to detect deviations from rational expectations, leading several authors to conclude that these tests are statistically weak (e.g. Gandar et al. 1988, Sauer et al. 1988, Sauer 1998).

Pointspread markets' function permits a more direct and better-suited measurement of rational expectations. We interpret the proportion of betting action a team receives BA as the pointspread market's conditional expectation of the likelihood that team covers their game, and the binary outcome CO of the game (i.e., 1 if the team covers, and 0 if the team fails to cover) as true expectations. So, on average, we expect CO and BA to be equivalent. Moreover, if gamblers have rational expectations, then no subset of publicly available information should be able to explain CO against BA. Stated differently, rational expectations dictates that no variable(s) should be able to predict $CO - BA$. Therefore, we derive the following regression equation:

$$CO - BA = b_0 + B \cdot \Theta + \epsilon \qquad \{4\}$$

Where b_0 is a constant coefficient, B is a vector of coefficients for the vector of explanatory variables Θ, and ϵ is a residual. Under our null hypothesis of rational expectations, b_0 and B should be just zero and the zero vector, respectively. The natural choice of significance test is therefore the F-test of each regression coefficient's joint equality with zero. If any independent variables separately or jointly explain variation in $CO - BA$, then pointspread gamblers have failed to fully and/or correctly incorporate available information into their bets, a violation of rational expectations.

We select several fundamental and technical factors for Θ. The technical variables we include are: PS, the net change in PS during the betting period, $\frac{1}{\omega}$ for both the home and away teams, and BA itself. The fundamental variable we include are: the implied pointspread generated by each team's FiveThirtyEight Elo rating (an index of team strength), the host's implied cover probability generated by the Elo pointspread[5], binary dummy variables for each team's participation in the previous season's playoffs, the number of days of rest of each team since their most recent game, the time zone of each team's home location (defined as 0 for teams in the Eastern Time Zone, 1 for Central Time, etc.), the weekday of the game, and a categorical variable for the EST start time of the game (i.e., morning, early afternoon, late afternoon, and evening). We present the results of our regression test in Table 4. Note that we regress for home teams only, to avoid letting the data enter our sample twice.

The results of our test strongly indicate nonrational expectations. PS, BA, and the Elo-implied cover probability are all significant and robust at the 95% confidence level, as is the model as a whole. The most striking result is the extremely negative coefficient on BA, which implies that with other factors considered, a 4% increase in betting action is associated with an approximate 1% decrease in a team's cover probability. While this is the clearest indication of nonrational gambler expectations, the other significant variables support a conclusion of nonrationality as well. Gamblers appear to not fully account for the likelihood of relative strengths of each team, as evidenced by the strongly positive Elo-implied cover probability coefficient, although curiously, the coefficient on the Elo-implied PS is actually negative, indicating both that the former coefficient may be biased upwards and that the Elo rating index interacts nonlinearly with $CO - BA$. Interestingly, the coefficient on PS is positive, indicating that pointspread gamblers may actually not be betting on favorites enough, all else considered. This is quite surprising given crowds' strong and well-documented favorite-bias in the pointspread market

Table 4- Rational expectations regression

Sample	2012-2019	2012-2015
Dependent variable	$CO - BA$	CO
PS	0.047**	0.029
	(0.024)	(0.026)
ΔPS	-0.004	-0.004
	(0.007)	(0.012)
Host $\frac{1}{\omega}$	-2.856	1.780
	(5.124)	(6.381)
Visitor $\frac{1}{\omega}$	-3.917	0.214
	(5.228)	(6.560)
BA	-1.204***	-1.262***
	(0.117)	(0.168)
Elo-implied PS	-0.045*	-0.029
	(0.024)	(0.026)
Elo-implied $P(MV > PS)$	1.424**	0.964
	(0.725)	(0.810)
Host made playoffs	0.022	0.064*
	(0.026)	(0.038)
Visitor made playoffs	-0.007	-0.041
	(0.027)	(0.039)
Host days of rest	-0.009	-0.007
	(0.006)	(0.009)
Visitor days of rest	-0.007	0.000
	(0.006)	(0.008)
Host Time Zone	-0.005	-0.012
	(0.013)	(0.019)
Visitor Time Zone	-0.017	-0.037**
	(0.011)	(0.016)
Wednesday	-0.603	-0.653
	(0.502)	(0.504)
Thursday	-0.053	-0.138
	(0.063)	(0.090)
Saturday	-0.132	-0.266
	(0.099)	(0.230)
Monday	-0.066	-0.096
	(0.061)	(0.086)
Morning	-0.168	0.235
	(0.135)	(0.251)
Late Afternoon	-0.023	0.030
	(0.034)	(0.048)
Evening	0.051	0.090
	(0.045)	(0.064)
b_0	3.523	-0.843
	(5.416)	(6.783)
N	1972	987
R^2	0.129	0.161
$F(b_0 = 0, B = \vec{0})$	14.48***	9.26***

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

[5] See Silver, Boice & Paine (2014) for a discussion of FiveThirtyEight's Elo methodology.

(e.g. Simmons et al. 2011, Gray & Gray 1997, Golec & Tamarkin 1991, Levitt 2004, Paul & Weinbach 2007, 2011); in our sample, favorites attracted an average of 57.6% of all betting action, despite that favorites only covered in 47.8% of their games. It may be instead that while bettors prefer betting on favorites, they simply choose the wrong favorites when gambling.

Identification of the factors which lead pointspread gamblers to overbet on suboptimal wagers is less straighforward. We have detected a few sources of variation in $CO - BA$, however, the model's relatively low R^2 value suggests that factors not observed in our model are responsible for crowds' failure to properly incorporate available information into their choice of wagers. The magnitude of the coefficient on BA reinforces this conclusion; a negative coefficient on BA indicates that crowds would still be betting incorrectly, even after adjusting for the other observed variables (assuming BA is non-causal on CO, which seems likely). Pointspread outcomes appear to be fairly noisy, so while additional exploration of the pointspread "factor zoo" could prove fruitful, discovery of additional explanainers of CO may be equally difficult. As a practical consideration, until such factors are identified, our results suggest that betting strategies of the variety proposed by Paul & Weinbach (2007, 2011) are the most likely to be profitable as compared to the currently available alternatives.

Table 5- Regression strategy profitability

Betting strategy	Bets	$\hat{p}$	π	H_0 p-value
In-sample				
$\hat{CO} \geq 0.50$	987	0.565	66.88	0.05*
$\hat{CO} \geq 0.52$	774	0.576	66.07	0.03**
$\hat{CO} \geq 0.54$	593	0.595	70.78	0.01***
$\hat{CO} \geq 0.56$	430	0.616	66.87	0.05***
$\hat{CO} \geq 0.58$	285	0.639	56.11	0.00***
$\hat{CO} \geq 0.60$	196	0.658	44.91	0.00***
Out-of-sample				
$\hat{CO} \geq 0.50$	985	0.517	-19.16	0.60
$\hat{CO} \geq 0.52$	777	0.510	-27.80	0.69
$\hat{CO} \geq 0.54$	601	0.516	-15.64	0.60
$\hat{CO} \geq 0.56$	424	0.512	-14.62	0.63
$\hat{CO} \geq 0.58$	308	0.516	-8.79	0.57
$\hat{CO} \geq 0.60$	207	0.512	-7.92	0.59

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

These above deviations from the behavior predicted by rational expectations, while strong, do not appear to create profit opportunities. We use the above linear probability model, re-estimated over an abbreviated 2012-2015 sample, to predict CO out-of-sample in the 2016-2019 seasons, and report our results in Table 5. While the model is successful in-sample, it fails to maintain its profitability out-of-sample. Each strategy is insignificantly greater than a 50% wins-to-bets ratio out-of sample, indicating that strategies based on our regression model are unlikely to perform better than random without the inclusion of additional significant factors, if they exist. This result runs counter to the findings of several previous authors (Gray & Gray 1997, Wever & Aadland 2012, Osborne 2001), who likewise discovered evidence of nonrational expectations, but simultaneously found that out-of-sample profit opportunities resulted. It may be that the pointspread market bookmakers have since adjusted their pricing strategies as to avoid profit opportunities, despite gamblers' continued biases, or that such findings were spurious.

Strong evidence of nonrational expectations are present in the data despite apparently effcient prices, a curious result when considering the typical consensus of rational expectations as a prerequisite for market effciency. The data clearly indicates nonrational expectations, but simultaneously offers no evidence of ineffciency; either the market is effcient despite crowd irrationality, and rational expectations of agents are not a prerequisite for market effciency, or the tests used thus far are too weak to detect ineffciency.

4 Conclusion

This paper presents several empirical tests of efficiency and rationality in the NFL pointspread betting market. We observe nonrational expectations of gamblers, but market behavior is consistent with a semi-strong EMH. Neither statistical nor economic tests reject market efficiency despite the discernable biases that we detect. We also fail to find evidence of supposedly profitable betting strategies proposed by other authors.

Our tests yield results quite similar to those of Summers (1986), who finds that even when simulating securities price movements with built-in nonrational expectations of agents, efficiency could not be rejected with standard statistical tests. Likewise, Gulko (2005) provides evidence that a semi-strong EMH holds in the bond market despite that market participants ignore relevant information when issuing interest rate forecasts. While Summers (1986) concludes that investors' failure to detect serially correlated mispricing is due to the weakness of the tests he employs, those presented here yield the same results despite comparably greater power against the null of efficiency. It seems rather more likely that irrational agents coexist with an efficient pointspread betting market. Considering the numerous similarities between securities trading and the pointspread market, traditional financial markets might likewise experience the efficiency-without-rationality phenomenon.

The observed disconnect between irrationality and inefficiency is surprising, although not inconsistent with patterns shown in other markets. Repeated studies have found significant biases and heuristics present in all manner of economic actors, yet long-run market efficiency tends to hold. Gambler behavior significantly differs from rational expectations, but long-run profitability is still close to or less than 0. Several explanations could exist: an elementary market structure, entry barriers, transaction costs imposed by market-makers, bettors' derivation of utility from nonmonetary sources, and discounting of price risk could all be factors. Nevertheless, the results of this study suggest that even in lieu of rational, profit-maximizing agents, certain markets might still be efficient.

Conventional economic thought assumes rational expectations of agents to be a necessary condition for the efficient markets hypothesis, however, the evidence we present suggests otherwise. It may be that the traditional neoclassical efficient markets framework is more compatible with irrationality than once thought. In any case, the evidence presented here from the NFL pointspread market certainly warrants a reexamination of the relationship between rational expectations and efficiency in financial markets.

5 Acknowledgements

The author thanks John Rehbeck, Matthew Weinberg, Ryan Ruddy, John Draper, and audiences at The Ohio State University for their invaluable and insightful feedback. The author is also grateful for the support this project received from the Ohio State University Sports and Society Initiative.

References

Badarinathi, & R., Kochman, L. (1996). Football betting and the efficient market hypothesis. *The American Economist*, **40**(2), 52–55.

Cortis, D. (2015). Expected Values and Variances in Bookmaker Payouts: A Theoretical Approach Towards Setting Limits on Odds. *The Journal of Prediction Markets*, **9**(1), 1–14.

Davis, J., McElfresh, L., Krieger, K., & Fodor, A. (2015). Exploiting Week 2 bias in the NFL betting markets. *The Journal of Prediction Markets*, **9**(1), 53–67.

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.

Gandar, J., Zuber, R., O'Brien, T., & Russo, B. (1988). Testing rationality in the point spread betting market. *The Journal of Finance*, **43**(4), 995–1008.

Golec, J., & Tamarkin, M. (1991). The degree of inefficiency in the football betting market. *Journal of Financial Economics*, **30**(2), 311–323.

Gray, P., & Gray, S. (1997). Testing market efficiency: evidence from the NFL sports betting market. *The Journal of Finance*, **52**(4), 1725–1737.

Gulko, L. (2005). Efficient irrational markets. *The Journal of Portfolio Management*, **31**(2), 64–72.

Lacey, N. (1990). An estimation of market efficiency in the NFL point spread betting market. *Applied Economics* **22**(1), 117–129.

Levitt, S. (2004). Why are gambling markets organised so differently from financial markets? *The Economic Journal*, **114**(495), 223–246.

Osborne, E. (2001). Efficient markets? Don't bet on it. *Journal of Sports Economics*, **2**(1), 50–61.

Pankoff, L. (1968). Market efficiency and football betting. *The Journal of Business*, **41**(2), 203–214.

Paul, R., & Weinbach, A. (2007). Does sportsbook.com set pointspreads to maximize profits? Tests of the Levitt model of sportsbook behavior. *The Journal of Prediction Markets*, **1**(3), 209–218.

Paul, R., & Weinbach, A. (2011). NFL bettor biases and price setting: further tests of the Levitt hypothesis of sportsbook behaviour. *Applied Economics Letters*, **18**(2), 193–197.

Purdum, D. (2015). Research shows U.S. could dominate global legalized sports betting market. Retrieved February 13, 2016, https://www.espn.com/chalk/story/_/id/13614240.

Sauer, R. (1998). The economics of wagering markets. *Journal of Economic Literature*, **36**(4), 2021–2064.

Sauer, R., Brajer, V., Ferris, S., & Marr, M. (1988). Hold your bets: another look at the efficiency of the gambling market for National Football League games. *Journal of Political Economy*, **96**(1), 206–213.

Shank, C. (2018). Is the NFL betting market still inefficient? *Journal of Economics and Finance*, **42**(4), 818–827.

Silver, N., Boice, J., & Paine, N. (2014). How Our NFL Predictions Work. Retrieved August 15, 2020, https://fivethirtyeight.com/methodology/how-our-nfl-predictions-work.

Simmons, J., Nelson, L., Galak, J., & Frederick, S. (2011). Intuitive biases in choice versus estimation: implications for the wisdom of crowds. *Journal of Consumer Research*, **38**(1), 1–15.

Summers, L. (1986). Does the Stock Market Rationally Reflect Fundamental Values? *The Journal of Finance*, **41**(3), 591–601.

Tryfos, P., Casey, S., Cook, S., Leger, G., & Pylypiak, B. (1984). The profitability of wagering on NFL games. *Management Science*, **30**(1), 123–132.

Vergin, R. (2001). Overreaction in the NFL point spread market. *Applied Financial Economics*, **11**(5), 497–509.

Vergin, R., & Scriabin, M. (1978). Winning strategies for wagering on National Football League games. *Management Science*, **24**(8), 809–818.

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

Woodland, B., & Woodland, L. (2000). Testing contrarian strategies in the National Football League. *Journal of Sports Economics*, **1**(2), 187–193.