Bettor biases and market efficiency in the NFL totals market

Kyle A. Kelly
West Chester University

Yanan Chen
West Chester University

ABSTRACT

Previous research found that the NFL totals betting market is inefficient for games with the highest posted totals and that bettors prefer betting the over in these games. We extend this and find that the contrarian strategy of betting the under on games with a posted total of 47.5 or higher wins 59.7 percent of the time during the 2001-2009 NFL seasons. This inefficiency disappears when covering the 2010-2018 seasons. We provide evidence that bettors still prefer betting the over and argue that the elimination of this simple betting rule can be explained by a change in sportsbook pricing behavior.

Keywords: Sports Betting Markets, Market Efficiency, NFL Totals
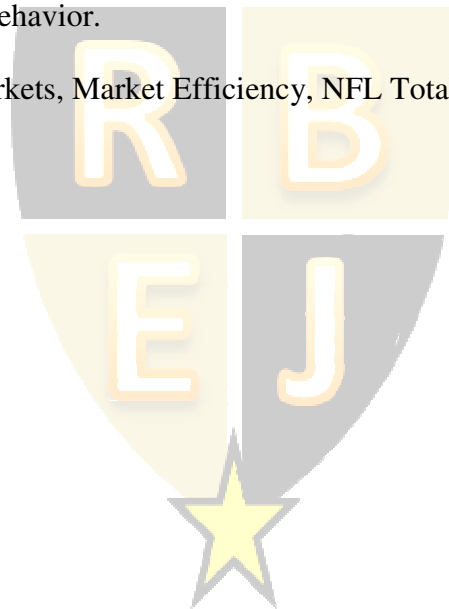
SECTION I: INTRODUCTION

Paul and Weinbach (2002) test for market efficiency in the NFL totals betting market for the 1979 to 2000 NFL seasons. While the authors conclude that the overall market is efficient, certain subsets of the market are shown to contain profitable betting opportunities. They argue that bettors prefer betting the over in games predicted to be high scoring and that sportsbooks respond to this bettor bias by setting the total higher than would be implied by market efficiency. Specifically, they find that totals set by the sportsbook of 47.5 or above are set higher than the efficient point and that a contrarian strategy of betting under these totals yields a statistically significant winning percentage of 58.7 percent.

Whether inefficiencies in betting markets are the exception or the rule is not yet settled in the literature. Studies of the NFL betting market by Tryfos et al. (1984), Sauer et al. (1988), Gray and Gray (1997), Boulier et al. (2006), Kochman and Goodwin (2006), and Sapra (2008) find that gambling markets are in general efficient and that profitable betting strategies are difficult to find. Sauer (1998) concludes that any inefficiencies that may exist are only temporary and disappear through a learning process by gamblers. Other studies of the NFL betting market yield different conclusions. Borghesi (2008) finds that certain weather events (hot, windy, and rainy weather) are not fully incorporated in the totals line and that profitable bets on the under can be made. Wever and Aaland (2012) find that large underdogs are mispriced and found significant out-of-sample profits. Fodor et al. (2014) find the mispricing of lines in Week 1 of NFL seasons. Teams that miss the playoffs in the prior season who meet playoff teams in Week 1 cover the spread more often than is implied by efficiency. Davis et al. (2015) extend this analysis to find that Week 2 underdogs when both teams are 0-1 are undervalued.

Implicit in these papers is that sportsbooks operate to "balance the book" meaning that lines are set in reaction to bettor behavior so that an equal amount of dollars is bet on each side of the wager. Under this assumption, sportsbooks earn risk-free profits on commissions charged. Recent research has challenged this assumption. Levitt (2004) argues that sportsbook set lines not to minimize risk but to maximize profits. He argues that sportsbooks are more skilled at setting lines than the public and can set prices that exploit bettor biases (for instance, preference for betting road favorites), thus generating higher profits. Using data from an online sportsbook handicapping tournament, he finds that in about half the games two-thirds of the wagers fall on one side of the wager. He estimates that by exploiting such biases the sportsbook increases its profits 20-30 percent higher than if it sets prices to balance the books. Studies by Paul and Weinbach (2007, 2011) use actual volume of betting dollars from Sportsbook.com to analyze bettor biases for the 2006 and 2007 NFL seasons. Their findings confirm Levitt's hypothesis in that sportsbooks do not set prices to attract equal action. They find that the percentage of dollars bet are highest for large favorites, road favorites, and high totals. Using percentage of bets (not dollars) on different wagers, Humphreys (2011) and Humphreys et al. (2013) find a larger percentage of bets occur on home teams, favorites, overs, and teams with higher winning percentages.

Given this background the authors revisit the findings of Paul and Weinbach (2002) that contrarian strategies of betting under the highest totals earn positive profits by extending their dataset to include the 2001-2018 NFL seasons. The authors look to answer several questions. First, to what extent to bettor biases last? Is their profitable betting rule an artifact of their sample or does it persist? Second, do bettors learn from previous inefficiencies and change their betting behavior over time? Third, has the price-setting behavior of sportsbooks of setting prices

to exploit bettor biases changed over time? On this latter point, Humphreys et al. (2013) show that over the 2005-2011 NFL seasons the road-favorite bettor bias persisted, but contrarian strategies of betting home-underdogs which once showed to be profitable no longer exist. They argue that greater competition among sportsbooks, the drastic reduction in costs for bettors obtaining information, and the presence of well-informed bettors has caused sportsbooks to change their price-setting behavior from one that exploits bettor biases to one that is an unbiased forecast of the outcome.

The authors find that the simple contrarian strategy of betting the under for games with a posted total of 47.5 or higher remained remarkably profitable over the 2001-2009 NFL seasons, covering at a statistically significant rate of 59.7 percent. This profitable strategy disappears across the 2010-2018 NFL seasons. Accounting for the increase in scoring and posted totals that occurred in the NFL over this later period did not yield profitable returns. Using data on the betting volumes in the same fashion as Humphreys (2011) and Humphreys et al. (2013) the authors find that bettor preferences for betting the over changed very little across the two periods. This casts doubt on the assumption that the bettors learn from previous inefficiencies and supports the argument that sportsbooks price-setting behavior has moved from exploiting known bettor biases and towards and unbiased forecast of the total score.

This paper is organized as follows. Section II looks at the profitability of betting under totals of 47.5 or more points during the 2001-2018 NFL seasons. Section III uses data on the percentage of bets in the totals market to assess if bettor biases have changed over time or if sportsbook pricing behavior has changed. Section IV concludes.

SECTION II: NFL TOTALS MARKET 2001-2018

A totals betting market begins when a bookmaker sets a "total" on the total points scored in a contest. For example, the opening game of the 2016 NFL season was a rematch of Super Bowl 50 which saw the Carolina Panthers visit the Denver Broncos. The posted total was 40.5 points. Since bookmakers typically follow the "11 for 10" rule, gamblers wishing to place a wager must risk $11 for every $10 won. Gamblers who believe that more than 40.5 points will be scored can bet "over" the posted total. Likewise, gamblers believing fewer than 40.5 points will be scored can bet "under" the posted total. The final score was Denver winning 21-20 for a total score of 41 points, making the game go over the posted total. Each $11 ticket bet on the over was now worth $21, while each $11 ticket on the under was now worth $0. For a gambler to break even when making such wagers, he must be successful in 11 out of 21 bets for a 52.4% winning percentage.

The authors check to see if the contrarian strategy of betting the under for all games with a posted total of 47.5 or more as outlined in Paul and Weinbach (2002) remains successful out-of-sample. The sample includes all NFL regular season games played between the 2001 and 2018, giving us a total of 4,600 observations. The data was collected from www.covers.com, a website devoted to sports gambling topics. Table 1 shows the over-under record for all games with a posted total of 47.5 or more broken down by season. The table includes the winning percentage on betting the under and the profitability for a gambler risking $110 on each bet. This simple strategy remained remarkably successful in the nine seasons following the original Paul and Weinbach (2002) study. From 2001 to 2009 eight of the nine seasons showed a positive profit with winning percentages greater than 52.4 percent. The second to last line in Table 1 shows the cumulative over-under record of 95-141 for a winning percentage of 59.7

percent and a profit of $3,650. The P-value from a one-sided binomial test of no profitability (null hypothesis: p = 0.524) is 0.014. This winning percentage is almost identical to that found in Paul and Weinbach's (2002) original sample (58.7 percent). The findings here suggest that bettor biases persist and that gamblers are not so quick to learn about market inefficiencies and adjust their betting strategies, as indicated in Table 1 (Appendix).

The success of this simple strategy vanishes when analyzing the 2010 to 2018 seasons. Table 1 shows that only half the seasons – 2012, 2015, 2017, and 2018 – had winning percentage above 50 percent; of that only the 2015, 2017, and 2018 seasons were profitable. The cumulative over-under record was 245-218 for a losing record on the under of 47.1 percent. This came with a cumulative loss of $4,350.

One problem with using 47.5 points as the threshold for evaluating the success of unders from 2010-2018 is that scoring in the NFL increased over time and with it bookmakers increased the posted total. Table 2 shows the summary statistics for total scoring and the posted totals for the sub-periods 2001-2009 and 2010-2018. Mean scoring increased by 2.5 points from 42.7 in 2001-2009 to 45.2 to 2010-2018. Median scoring increased a full two points across these time periods (from 42 to 44). Bookmakers responded to the higher scoring by adjusting upward the posted total. The mean posted total increased by 3.2 points from 41.6 during 2001-2009 to 44.8 during 2010-2018. The median posted total increased by three points (from 41.5 to 44.5). During the period studied by Paul and Weinbach (2002) the mean posted total was 40.3. Their original rationale of analyzing under bets for games with a posted total of 47.5 points or more is that this total was a full seven points above the mean. This is a reasonable starting point since seven points is a touchdown and extra point. In the sample period 2010-2018, 47.5 points is the third quartile, meaning 25 percent of all games played over this time had a posted total at or above this value. It could be that the disappearance of this simple profitable betting rule is that the rule is covering too many games. While the authors and others have shown that bettor biases do exists and that profitable rules can sometimes be found, it is hard to believe that such a simple rule would be profitable for one-fourth of all games in a market, as indicated in Table 2 (Appendix).

Given this narrative, the authors check to see if modifying the simple strategy can yield profitable returns in the 2010-2018 sample period. The authors use the same rule-of-thumb of checking the profitability of betting the under for all games where the posted total is seven points or higher than the posted total. This gives us a cutoff point of analyzing games with posted totals of 51.5 or more. The authors readily admit that if found profitable, this type of ex-ante search for a profitable betting strategy is hardly proof of inefficiencies in the market. The results in Table 3 show that these concerns are not warranted since no profits can be earned with this strategy. The table breaks down the over-under record, winning percentage, and profitability for each season. In only the 2017 season did the winning percentage exceed 50 percent. The cumulative over-under record for the eight seasons was 87-82 with a losing record on the under of 48.5 percent and a cumulative loss of $1,370, as indicated in Table 3 (Appendix).

SECTION III: BETTOR BIASES AND SPORTSBOOK BEHAVIOR

The results above show that the simple contrarian strategy of betting under on games with the highest posted total remained profitable during the 2001-2009 seasons, only to disappear and turn negative during the 2010-2018 seasons. What might have caused the change? Did bettors learn of this market inefficiency over time and change their preferences of betting over the

highest totals, resulting in sportsbooks adjust downward the posted totals on these games? To check this, the authors follow Paul and Weinbach (2011), Humphreys (2011), and Humphreys et al. (2013) and use betting volume data provided by Sports Insights. This company provides (for a fee) data on the percentage of bets placed on each side of a wager (over/under) going back to the 2004 NFL season. This gives us a sample of 3840 games. If gamblers observe profitable opportunities on betting under the posted total, then the percentage of bets on the over should decline over time.

Table 4 shows the summary statistics of the percentage of over wagers for the sub-periods 2004-2009 and 2010-2018. The values are very similar in the two time periods, with only a slight decline in the mean and median percentage of over wagers during the 2010-2018 season. The first quartile over percentages are 57 and 56, respectively, for the 2004-2009 and 2010-2018 sub-periods. This implies that in 75 percent of games played, the percentage of over wagers was 57 (56) percent or higher. Gamblers still show a strong preference for wagering on overs, as indicated in Table 4 (Appendix).

The summary statistics cannot tell us whether preferences for over wagers for games with the highest posted total has changed over time. To address this, the authors use OLS and regress the percentage of bets placed on the over (% Bet on Over) on the sportsbook posted total (Posted Total), a dummy variable (Time Period) that takes on the value of 0 from 2004-2009 and 1 from 2010-2018, and a dummy-interaction term with the posted total. The dummy variable and dummy-interaction term allows for the comparison of the effects of the posted total on the percentage of bets on the over across the two different time periods. The regression takes the form:

$$\% \ Bet \ on \ Over_i = \ \beta_0 + \beta_1 Posted \ Total_i + \alpha_0 Time \ Period_i + \ \alpha_1 Time \ Period_i * Posted \ Total_i + \ \varepsilon. \qquad (1)$$

Paul and Weinbach (2011) and Humphreys et al. (2013) estimate similar regressions. Each study finds a positive and significant effect on the posted total, indicating that the higher the total, the more bets placed on the over. The latter study analyzed data from the 2005-2011 NFL seasons. In the model above, a positive and significant sign on β_1 would indicate bettors prefer betting the over at higher totals across the 2004-2009 seasons. If bettor preferences have changed over time so that gamblers no longer prefer betting the over, the regression line should be flatter with a larger vertical intercept closer to 50 percent in the 2010-2018 time period. The coefficient on the dummy variable, α_0, should be positive, significant, and larger than β_0, while the dummy-interaction term, α_1, should have a negative and significant sign.

One weakness of the Sports Insight data is that it contains the percentage of bets placed on each side, not the percentage of dollars wagered. The percentage of dollars bet on either side is what determines the amount of risk the sportsbook is enduring on a specific wager. Numerous small bets on the over could be nullified by one large bet on the under. The data would fail to capture this. However, as Paul and Weinbach (2011) show, the percentage of bets placed and the percentage of dollars wagered are highly correlated, and so the results will give us useful insights into the reaction of sportsbooks to bettor behavior.

Table 5 shows the results of the estimates of the regression above. The estimated coefficient for the posted total is positive and significant, indicating that during the 2004-2009 seasons, a one-point increase in the posted total results in a 0.78 percent increase in the percentage of bets on the over. Given that the minimum posted total during this time was 30

points, the model predicts more than 50 percent bets on the over across all observations, as indicated in Table 5 (Appendix).

The estimate for the dummy-interaction term is also positive and significant. Combining this with the coefficient for the posted total gives us an estimated slope of 1.01 during the 2010-2018 seasons. This indicates that a one-point increase in the posted total leads to a 1.01 percent increase in the percentage of bets on the over. Bettors responded with a larger increase in percentage of bets on the over during the 2010-2018 seasons for each point higher the posted total was. The dummy variable is negative and significant indicating that the regression line rotated up during the 2010-2018 seasons. This means that the predicted percentage of bets on the over during the 2010-2018 seasons is lower than during the 2004-2009 seasons for most of the range of posted totals. However, given that the minimum posted total from 2010-2018 was 33 points, the model still predicts that more than 50 percent of bets occur on the over across all observations in the sample period. Also, since the concern here is the behavior of bettor preferences in the upper tail of the distribution, the results predict that a larger percentage of bets occur on the over in the 2010-2018 seasons for posted totals of 51.5 points and higher. This implies that bettors showed a *stronger* preference for betting the over for games with the highest predicted points scored in the latter sample.

The results in this section suggest that the movement of the highest posted totals towards the efficient point during the 2010-2018 seasons is not due to gamblers learning about a market inefficiency and changing their betting patterns. Gamblers like betting the over, and particularly enjoy betting the over on games with the highest posted total – a behavioral bias that has existed for a long time. Paul and Weinbach (2007, 2011) and Humphreys et al. (2013) argue that greater market competition among sportsbooks and the reduction in costs of gamblers acquiring information on historical may have caused sportsbooks to set totals simply as an unbiased forecast of the outcome of the game, rather than shading against bettor biases. Greater competition has resulted in increased betting limits on totals. A sportsbook shading lines away from the efficient point runs a greater risk of informed or "sharp" bettors taking advantage of the mispriced lines with higher limits. In addition, the speed and ease in acquiring and processing data may have increased the number of informed bettors in the market. Also, the reduced cost of moving funds to various online sportsbooks makes it easier for informed bettors to placing wagers.

SECTION IV: CONCLUSION

Paul and Weinbach (2002) identify a profitable betting strategy of betting the under in NFL games with the highest posted totals. This strategy won at a 58.7 percent rate during the 1979-2000 seasons. Common wisdom is that gambling markets are in general efficient, and that any inefficiencies will disappear over time as gamblers learn and adjust to the bias. The authors check to see if this profitable strategy holds out-of-sample by analyzing the totals market during the 2001-2018 NFL seasons. The authors find that this simple wagering strategy remained profitable for a considerable amount of time, winning at a 59.7 percent rate over the 2001-2009 seasons. This profitable strategy disappears during the 2010-2018 as the posted totals moved closer to their efficient point. Accounting for the increase in scoring during this time period did not improve the results. The authors check to see if changes bettor preferences can explain the change in pricing towards the efficient point. Using data on the percentage of bets placed on either side of the totals market from 2004-2018, The authors find that bettors prefer betting the

over with more than 50 percent of wagers being placed on all posted totals. This preference increases the higher posted total becomes. During the 2010-2018 sub-period, a larger percentage of bets occurred on the over for the highest posted totals. It appears that changes in sportsbook pricing behavior due to higher competition and reduced cost of acquiring information can best explain the movement towards the efficient price and not due to bettors learning about the bias.
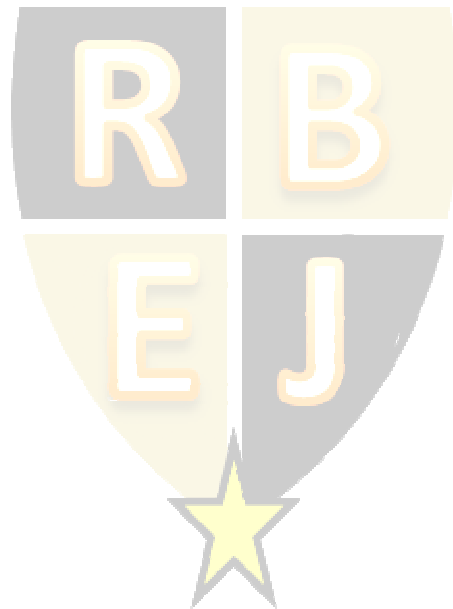
REFERENCES

Borghesi, R. (2008). Weather biases in the NFL totals market. *Applied Financial Economics*, 18, 947-953.

Boulier, B.L., Stekler, H. O., & Amundson, S. (2006). Testing the efficiency of the National Football League betting market. *Applied Economics*, 38, 279-284.

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

Fodor, A., DiFilippo, M., Kreiger, K., & Davis, J.L., (2013). Inefficient pricing from holder bias in NFL point spread markets. *Applied Financial Economics*, 15, 1407-1418.

Gray, P. K., & Gray, S. F. (1997). Testing market efficiency: Evidence from the NFL sports betting market. *Journal of Finance,* 52, 1725-1737.

Humphreys, B.R. (2011). The financial consequences of unbalanced betting on NFL games. *International Journal of Sports Finance,* 6, 60-71.

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

Kochman, L., & Goodwin, R. (2006). An under bias in the football betting market: Fact or fiction?: A note. *New York Economic Review,* 37, 32-36.

Levitt, S.D., (2004). Why are gambling markets organized so differently from financial markets? *The Economic Journal*, 114, 223-246.

Paul, R.J., & Weinbach, A.P. (2002). Market efficiency and a profitable betting rule: Evidence from totals on professional football. *Journal of Sports Economics,* 3, 256-263.

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

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

Sapra, S. G. (2008). Evidence of betting market intraseason efficiency and interseason overreaction to unexpected NFL team performance 1988-2006. *Journal of Sports Economics*, 9, 488-503.

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

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

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

Wever, S., & Aaland, D. (2012). Herd behavior and underdogs in the NFL. *Applied Economics Letters*, 19, 93-97.

APPENDIX

Table 1: Under Bets for All Games with Posted Total of 47.5 or More: 2001-2018

Season	Over-Under Record	Under Percentage	Contrarian Earnings $110
2001	9-10	52.6%	$10
2002	9-17	65.4%	$710
2003	7-10	58.8%	$230
2004	25-20	44.4%	-$750
2005	7-17	70.8%	$930
2006	2-12	85.7%	$980
2007	12-16	57.1%	$280
2008	14-17	54.8%	$160
2009	10-22	68.8%	$1100
2010	17-13	43.3%	-$570
2011	34-21	38.2%	-$1640
2012	32-34	51.5%	-$120
2013	43-38	46.9%	-$930
2014	43-43	50.0%	-$430
2015	27-33	55.0%	$330
2016	49-36	42.4%	-$1790
2017	23-28	54.9%	$270
2018	47-57	54.8%	$530
2001-2009	95-141	59.7%†	$3650
2010-2018	245-218	47.1%	-$4350

Notes: Under Percentage represents the percentage of games that went under the posted total.

†P-values of likelihood ratio and one-sided binomial tests of no profitability ($p = 0.524$) are 0.023 and 0.014, respectively.

Table 2: Total Scoring and Posted Total 2001-2018

	Total Scoring		Posted Total	
	2001-2009	2010-2018	2001-2009	2010-2018
Mean	42.7	45.2	41.6	44.8
Median	42	44	41.5	44.5
Standard Deviation	14.2	13.9	4.8	4.3
Q3	51	54	44.5	47.5
Q1	33	36	37.9	42
IQR	18	18	6.6	5.5

Notes: Summary statistics of total points scored and posted total in the sub periods 2001-2009 and 2010-2018.

Table 3: Under Bets for All Games with Posted Total of 51.5 or More: 2010-2018

	51.5 or more		
Season	Over-Under Record	Under Percentage	Contrarian Earnings $110
2010	0-0	-	-
2011	8-7	46.7%	-$180
2012	9-9	50%	-$90
2013	12-12	50%	-$120
2014	13-8	38.1%	-$630
2015	7-6	46.1%	-$170
2016	11-8	42.1%	-$410
2017	4-11	73.3%	$660
2018	23-21	47.7%	-$430
2010-2018	87-82	48.5%	-$1370

Notes: Proportion represents the proportion of games that went under the posted total. In 2010, the highest posted total was 51 points resulting in zero games meeting the 51.5 threshold.

Table 4: Percentage of Over Wagers 2004-2018

	2004-2009	2010-2018
Mean	64.4	64
Median	66	65

Standard Deviation	11.2	11.3
Q3	73	72
Q1	57	56
IQR	16	16
Min	25	24
Max	94	90

Notes: Percentage of wagers placed on the over for sub-periods 2004-2009 and 2010-2018.

Table 5: Results of Regression of Percentage Bet on Over and Posted Total 2004-2018

Variable	Coefficient	Standard Error	P-value
Intercept	31.99	2.33	<0.001
Posted Total	0.78	0.056	<0.001
Time Period Dummy	-13.2	3.28	<0.001
Dummy Interaction	0.23	0.076	<0.001

Note: Results above are for a regression of percentage of bets on the over on an intercept, the posted total (Posted Total), a dummy variable for the years 2010-2018 (Time Period), and a dummy-interaction term with the posted total. The time period is limited to data availability and covers the 2004-2018 NFL seasons.