DEPARTMENT OF
ECONOMICS
UNIVERSITY OF ALBERTA

Working Paper No. 2010-05

Prices, Point Spreads and Profits: Evidence from the National Football League

Brad R. Humphreys
University of Alberta

February 2010

Prices, Point Spreads and Profits: Evidence from the National Football League

Brad R. Humphreys*
University of Alberta
Department of Economics

February 2010

Abstract

Previous research on point spread betting assumed that bookmakers attract an equal volume of bets on either side of games in order to maximize profits. This paper examines the viability of this assumption from a theoretical and empirical perspective. The model of bookmaker behavior developed predicts that expected returns are not necessarily maximized when the volume of bets on each side of a game are equal. Analysis of a unique data set containing information on point spreads, game outcomes, and betting volume for the 2005-2008 NFL seasons reveals widespread imbalances in bet volumes. Simulations indicate that this imbalanced betting generated positive profits, including profits larger than would have been made if the betting volume was balanced on all games.

JEL Codes: L83, G11, G12

Introduction

The idea that bookmakers set point spreads to balance the volume of bets on each side of a game is ubiquitous in the literature on point spread betting in sports. Nearly every paper published in this area over the past 30 years includes a discussion of an implicit model of sports book behavior that includes this feature. For example, Woodland and Woodland (1991) wrote the in the *Journal of Political Economy*

> In most situations, the bookie has no desire to participate as an active gambler. Rather, he establishes an odds or spread line to balance the wagers so that his commission is independent of the final outcome of the contest. For the equilibrium spread line, this is equivalent to equalizing the total amount of money wagered on each team. For the odds or "money" line, the equilibrium weighting is proportional to the odds that are offered. This eliminates all risk for the bookie. (pages 638-639.)

*The Alberta Gaming Research Institute provided for financial support for this research. Address: 8-14 HM Tory, Edmonton, Alberta, Canada T6G 2H4; phone: (780) 492-5143; fax: (780) 492-3300; Email: *brad.humphreys@ualberta.ca*. Thanks to seminar participants at the University of Kentucky, Andy Weinbach and Dennis Coates for useful comments.

The model implicit in such statements, which I call the “balanced book” model, seems to generate clear predictions, even though it has never been formally written down to my knowledge. It predicts that, were we to observe the volume of bets on different sides of bets, the volume would be relatively balanced on either side, and that the sports book has an incentive to make sure this outcome occurs; and it predicts that observed changes in point spreads should be explained by imbalances in betting volume, as the model posits point spreads as the mechanism sports books use to balance betting that has become unbalanced as orders are taken in the market. This model has considerable support in the empirical literature. However, this support takes the form of a large number of tests confirming that point spreads are generally optimal predictors of actual score differences in games in many different settings. Sauer (1998) surveys this evidence and discusses its importance for understanding outcomes in betting markets.

Some recent research has emerged challenging the validity of the “balanced book” model. Levitt (2004) documented the outcome of a season long prediction contest for National Football League (NFL) games. While this contest did not resemble sports betting markets many respects, the contest generated detailed data on bettor behavior and clearly revealed that the volume of bets was not balanced on a majority of the games that were picked by contestants. In addition, Levitt (2004) developed a theoretical framework to explain the observed imbalances in betting on NFL games in this contest. Paul and Weinbach (2007) analyze detailed betting volume data from an on-line sports book, sportsbook.com, and found evidence of unbalanced volume on bets placed on NFL games in the 2006 season. Paul and Weinbach (2008) found evidence of unbalanced betting volumes on National Basketball Association (NBA) games in the 2004 - 2006 seasons. Taken together, the evidence in these papers suggests that the “balanced book” model may not describe actual outcomes in sports betting markets.

In this paper, I investigate the implications of unbalanced betting volume on the returns to sports books. I extend Levitt’s (2004) framework to an expected return maximizing model in which sports books choose the fraction of bets placed on the favored team by setting the point spread on games. The model shows that a “ balanced book” is a possible outcome, but not necessarily the outcome that maximizes expected returns for the sports book. It also generated predications about the expected return maximizing volume of betting on the favored team. I use a novel data set that includes information on betting volume on both sides of bets on all NFL games played in the 2005, 2006, 2007 and 2008 seasons. These data contain evidence of significant imbalances in bet volumes in 9 out of 10 NFL games, as well as evidence that the point spreads set on these games, like most others, are unbiased minimum variance predictors of actual game outcomes. The results of financial simulations indicate that the unbalanced betting experienced by these sports books was profitable over the course of these three seasons, and in one of the three seasons generated returns in excess of those that would have been generated by a perfectly balanced book. The results call into question the ability of the widely used “balanced book” model to describe basic outcomes in sports betting markets.

A Simple Model of Sports Book Behavior

While the behavior of bookmakers has been widely discussed in the literature, few have bothered to write down a formal model. Two notable exceptions exist. Levitt (2004) developed a framework for interpreting the results from a contest that involved picking winning NFL teams against the spread in the 2001 season. 285 handicappers participated in the contest. The choice variable in this framework was the probability that a team won the game, which does not closely resemble sports book behavior, since a sports book can only set a point spread and take bets that are made at this point spread. Cain, Law and Lindley (2000) developed a complete model of odds setting and bettor behavior in the market for bets on horse races. The model developed by Cain, Law and Lindley (2000) explicitly considers bettors behavior and the interaction between bettors and bookies, making it a particularly useful exercise. Hurley and McDonough (1995) also develop a model of bettor-bookie interaction in the context of odds betting on horse racing. This model also examines races with only two horses, but the pari-mutuel nature of horse race betting differs significantly from point spread betting in terms of payoffs and commission charged to bettors, making it difficult to apply this model to sports betting. While both models focus on races with only two horses, a situation similar to sports betting, horse race betting is either pari-mutual or odds betting, while much of the sports betting in North America is based on point spreads. This is an important difference when modeling the expected returns to bookies and the decisions made by bettors, and neither model offers much insight into behavior in point spread betting markets.

To motivate the model of sports book behavior, first consider the simple case where a sports book accepts wagers on a single game played by two teams, team 1 and team 2. Let H represent the total amount of dollars wagered on this bet, f_1 the fraction of dollars wagered on team 1, and $f_2 = (1 - f_1)$ the fraction of dollars wagered on team 2. In this setup, H can be normalized to one

$$H = f_1 + f_2 = f_1 + (1 - f_1) = 1$$

which allows the analysis to be carried out in terms of units bet or fractions bet on either team. Sports books operate by charging a fee or commission on losing bets only. Let v be the commission or "vig" charged on losing bets. Since each game has only two possible outcomes, the unconditional net gain or loss (R) on each game is

$$\begin{array}{ll} \text{Bets on Team 1 Win} & R = f_2(1+v) - f_1 \\ \text{Bets on Team 2 Win} & R = f_1(1+v) - f_2. \end{array} \tag{1}$$

If team 1 wins, then the sports book keeps all money wagered on team 2, plus the commission, and pays off those bets placed on team 1. If team 2 wins, then the sports book keeps all money wagered on team 1 plus the commission and pays off bets placed on team 2. The amount of profit or loss on a game depends on the amount that is wagered on each team. The gain or loss on a game can be expressed in terms of the fraction bet on team 1 by substitution

$$\begin{array}{ll} \text{Bets on Team 1 Win} & R = f_2(1+v) - f_1 \\ & \quad = (1 - f_1)(1+v) - f_1 \\ & \quad = 1 - f_1 + v - f_1 v - f_1 \\ & \quad = 1 - 2f_1 - f_1 v + v \\ \text{Bets on Team 2 win} & R = f_1(1+v) - f_2 \\ & \quad = f_1(1+v) - (1 - f_1) \\ & \quad = f_1 + f_1 v - 1 + f_1 \\ & \quad = 2f_1 + f_1 v - 1 \end{array} \tag{2}$$

Note that no matter which bets win, the book maker's profits increase with the commission. If bets on team 1 win, then gains on this game fall as the fraction of the bets on team 1 increase; if bets on team 2 win, then gains on this game rise as the fraction of the bets on team 1 increases.

A break even condition on bets taken on this game can be derived from equation (2) by setting the profit equation equal to zero and solving

$$
\begin{array}{ll}
\text{Bets on Team 1 Win} & R = 1 - 2f_1 - f_1 v + v = 0 \\
& 1 - 2f_1 - f_1 v + v = 0 \\
& 2f_1 + f_1 v = 1 + v \\
& f_1(2+v) = 1 + v \\
& f_1 = \frac{1+v}{2+v} \\
\text{Bets on Team 2 Win} & R = 2f_1 + f_1 v - 1 = 0 \\
& 2f_1 + f_1 v = 1 \\
& f_1 = \frac{1}{2+v}.
\end{array}
\tag{3}
$$

Note that $\frac{1}{2+v} < \frac{1+v}{2+v}$. These two terms constitute upper and lower bounds for profitability of a bet from the perspective of the sports book, no matter what the outcome of the game. So long as $\frac{1}{2+v} < f_1 < \frac{1+v}{2+v}$ the sports book makes a profit on bets on this game. These two expressions are familiar to anyone who has read the sports betting literature because they are used for the calculation of the fraction of bets that must be won by any betting strategy to earn a profit for the bettor. If $v = 0.1$, then $\frac{1}{2+v} = 0.476$ and $\frac{1+v}{2+v} = 0.524$, implying that so long as the fraction of the bets on team 1 is between these two bounds, the sports book makes a profit on the bet no matter what the outcome of the bet. The sports book might be able to make larger profits than this, conditional on the outcome, bettors expectations, or other factors, but this profit is unconditional.

Finally, as v increases, the bounds on the certain profit condition expands. By equating the expressions for the profit earned under each outcome in (2), when $v = 0.5$ the same profits are earned for either outcome for any value of f_1. From (3), $v = 0.5$ corresponds to unconditional profit bounds of $0.40 < f_1 < 0.6$. Put another way, if the sports book could charge a commission of 50%, that sports book would make a certain profit on bets on the game no matter what outcome, so long as the fraction of the bets on one team was between 40% and 60% of the total.

A "Balanced Book"

The break even condition can be used to illustrate the "balanced book" outcome frequently mentioned in the literature. In the context of this simple model, a "balanced book" refers to the case where a sports book sets the point spread on a game to balance the volume of betting on either side of the bet, so that $f_1 = f_2 = f = 0.5$. Under the balanced book condition, profits are

$$
\begin{array}{ll}
\text{Bets on Team 1 Win} & R = f(1+v) - f = fv = \frac{v}{2} \\
\text{Bets on Team 2 Win} & R = f(1+v) - f = fv = \frac{v}{2}
\end{array}
\tag{4}
$$

proportionate to the commission charged v no matter what the outcome of the wagers turns out to be. This outcome involves no risk on the part of the sports book; if the volume of bets on each team is equal, the sports book earns a profit proportional to the commission charged no matter what the outcome. This result motivates much of the empirical research on point spread betting in sports. Most research on point spread betting assumes that point spreads are set to achieve this outcome. However, an expanded model highlights some of the problems with this approach.

Adding Outcome Uncertainty

Levitt (2004) developed an expanded expression for the expected profit of a sports book, R, that includes a term for the probability that a bet on a given team wins, an extension to the unconditional analysis above. In Levitt's (2004) model, the probability that a bet on team 1 wins was assumed to be the choice variable for sports books, and this probability is manipulated by changing the point spread. Here, I take a different approach to modeling sports book behavior. Rather than formulate the model in terms of the probability that a bet on a given team wins, I explicitly model the probability that a bet on a given team wins. In particular, suppose that the probability that a bet placed on team 1 wins is π_1 and the probability that a bet on team 2 wins is $\pi_2 = (1 - \pi_1)$. In this case, the expected profit earned by the sports book is

$$E[R] = [(1 - \pi_1)f_1 + (1 - \pi_2)f_2](1 + v) - (\pi_1 f_1 + \pi_2 f_2). \tag{5}$$

The first term on the left hand side of equation (5) is how much the sports book keeps on losing bets. The second term is how much the sports book pays out on winning bets. This can be written in terms of team 1 outcomes by substituting $\pi_2 = 1 - \pi_1$ and $f_2 = 1 - f_1$. This expression, when written in terms of π_1 and f_1 can be simplified to

$$E[R] = (2 + v)[f_1 + \pi_1 - 2\pi_1 f_1] - 1. \tag{6}$$

For simplicity, assume that team 1 is the favored team. In this expanded expected return function, the expected return on any bet depends on both the amount bet on team 1, the "vig" and the probability that team 1 wins the game. Levitt (2004) assumed that the probability that a bet on team 1 wins, π_1, is the choice variable for the sports book, and derived an expression for the return maximizing probability. However, sports books do not directly control this probability. Sports books set a point spread and take bets at this point spread. The probability that a bet on Team 1 wins depends on the point spread, the relative strengths of the two teams, and random events that take place during the game. I assume that a relationship exists between the fraction of the money bet on team 1 and the probability that a bet on team 1 wins, and that sports books affect the fraction of bets on team 1 by setting the point spread. Formally,

$$\pi_1 = \sigma f_1 \tag{7}$$

This relationship can be motivated by interpreting π_1 as the objective probability that a bet on team 1 wins and f_1 as bettors' subjective probability that a bet on Team 1 will win. f_1 depends on bettors' preferences and expectations, the relative strengths of the two teams, random events that take place during games, and the point spread set by sports books. Sports books can affect f_1 by changing the point spread, and empirical evidence exists that f_1 increases with point spreads (Paul and Weinbach, 2007). π_1 depends on only the point spread, the relative strengths of the two teams, and random events that take place during the course of the game. σ captured the effect of bettors' expectations and preferences. If $\sigma = 1$, then $\pi_1 = f_1$ and bettors subjective expectation that that a bet on Team 1 will win is equal to the objective expectation. If $\sigma \neq 1$, then bettors expectations and preferences will distort their subjective expectation that a bet on Team 1 will win. There are a number of reasons to think that $\sigma \neq 1$. Conlisk's (1993) model of the utility of gambling predicts that bettors do not take into account the financial implications of a gamble, but instead derive utility from the act of gambling. Bettors who place bets on Team 1 for utility maximizing reasons may not take into account factors like the relative strengths of the two teams, or the point spread. These bettors would bet on Team 1 because they like the team's colors, or perhaps because Team 1 was the favored team, and they derive enjoyment from betting on and rooting for the favored team. Also, bettors who use heuristics to make decisions about gambling, like those described by Tversky and Kahneman (1973), might also exhibit biased decisions, leading their subjective probability that a bet on Team 1 will win to differ from the objective probability that this bet would win. If $\sigma > 1$, then $\pi_1 > f_1$ bettors subjective probability that a bet on Team 1 will win exceeds the objective probability; if $\sigma > 1$, then $\pi_1 > f_1$ bettors subjective probability that a bet on Team 1 will win is lower than the objective probability

Substituting equation (7) into the expression for the expected return to the sports book for each bet, equation (6), gives an expression for expected returns on a bet in terms of the choice variable, f_1, the fraction of money bet on team 1

$$E[R] = (2 + v)[f_1 + \sigma f_1 - 2\sigma f_1 \cdot f_1] - 1. \tag{8}$$

According to this expected return function, the expected return on bets first rises and then falls with f_1. Taking the derivative of this expression with respect to changes in f_1, setting equal to zero, and solving for f_1 gives an expression for the fraction of bets on team 1 that maximizes the expected return on a bet

$$f_1^* = \frac{1 + \sigma}{4\sigma}. \tag{9}$$

This expression defines the fraction of bets on team 1 that the sports book should target to maximize the expected return on bets taken on a game. Clearly, a balanced book only maximizes expected returns if bettors are "unbiased" in that $\sigma = 1$. However, if $\sigma \neq 1$, and biased bettors exist, then a balanced book does not maximize expected returns for the sports book. The sports book "shades" the point spread to systematically take advantage of the biased bettors in the market and increase expected returns above what they would expect to earn by balancing the volume of bets on either side of each game.

In the literature, this incentive to move point spreads away from the level that equalizes betting has been attributed to the presence of heterogeneous bettors in the market. Consider a market with informed and uninformed bettors. The uninformed bettors are insensitive to the point spread and bet on a given team for reasons like loyalty or non-monetary motivations. Levitt (2004) shows that participants in the betting contest he analyzed tended to bet on the favored team, suggesting that they placed bets in this way. Informed bettors make decisions based on the point spread, and their expectation of the actual outcome of the game. Uninformed bettors will bet on their preferred team at any point spread, while informed bettors will bet either side, depending on the point spread. In this case, book makers can increase their profits by changing the point spread in a way to decrease the probability that bets placed by uninformed bettors win. Levitt (2004) shows that the probability that a bets placed on the favored team had a less than 50% chance of winning in the betting contest he analyzed, and interprets this as evidence that sports books shade point spreads. Neither Levitt's (2004) model or this model explicitly consider how bettors make decisions, limiting both model's ability to explore this point.

This model of the expected returns to sports books shows that the "balanced book" model frequently discussed in the literature maximizes the expected returns earned by sports books only when all of the bettors in the market have an unbiased subjective probability that a bet on one of the teams in the game will win. When some biased bettors place bets on a game, and their subjective probability that a bet on the favored team will win does not equal the objective probability that a bet on the favored team wins, a sports book can make higher expected returns by accepting more bets on one side of the game, and shading the point spread to take advantage of these biased, or uninformed bettors. Previous research assumed that a balanced book was a relatively easy outcome for sports books to generate, or at least that sports books attempted to attract an equal volume of bets on either side of a proposition. Expanding a commonly used model of sports book behavior to the case where the probability that a bet on the favored team depends on the fraction of bets made on that team, and the sports book can manipulate this fraction by setting the point spread, and bettors switch their bets to the other team in response to changes in the point spread, generates a prediction that expected returns may not be maximized when the book is balanced. The next section examines the relationship between point spreads, betting volume, and actual gains and losses in the market for point spread bets on NFL games to assess how often the volume of betting is equal on either side of propositions in this market.

Empirical Analysis

Much of the previous research on point spread betting focused on examining the efficiency of point spread betting markets. Because point spreads and game outcomes were readily available, but betting volumes were not, researchers assumed that the point spread was set in a way to equalize the volume of betting on either side of a game, and focused attention on the ability of point spreads to predict game outcomes. Recently, economists have gained access to betting volume data from point spread betting markets and have begun to examine point spreads and betting volumes. Examples of this emerging literature include Levitt (2004), and Paul and Weinbach (2007, 2008). The model developed above, which builds on the theoretical framework in Levitt (2004), demonstrates the conditions under which an unbalanced book, an outcome where the the volume of bets is not equally distributed, can still generate the largest expected return for the sports book. In this section, I investigate a unique data set that contains information on actual bet volumes from four on-line sports books to see if evidence supporting the predictions of the model exist in these data.

Data Description

The balanced book model motivated much of the previous research on point spread betting. Despite the extensive reliance on this model to motivate empirical research on sports betting markets, little evidence exists that the balanced book model describes actual outcomes in sports betting markets. The primary reason for this lack of evidence is a lack of data on the volume of bets made on either side in individual games. While point spreads and game outcomes are readily observable, until now, researchers have not had access to data on betting volume. Because of this lack of data, researchers have proceeded under the assumption that point spreads are set by sports so that an equal amount of money is bet on either side of all games, and proceeded to tests of the efficiency of this market, or tests for inefficiencies in the form of profitable betting strategies.

Recently, data on betting volumes have become available to researchers, primarily from on-line sports books. Sports Insights, an online sports betting information service, recently began making betting data, including information on betting volume, available for a fee. The data analyzed here were purchased from Sports Insights. Sports Insights has agreements to obtain betting volume data from four large on-line sports books: BetUS, Carib Sports, Sportbet, and Sportsbook.com. The data files that Sports Insights makes available include the opening and closing point spreads, the actual score of the game, and the percentage of bets reported on each side of a proposition for all regular season games played in the National Football League (NFL) in the 2005 - 2008 seasons. The data collected by Sports Insights represents an average across the four participating sports books. The betting volume is not available for all sports books and is not available for each game played over the course of the season. In addition, the total dollars bet on each game is not known. Table 1 shows summary statistics for key variables in this data set.

The NFL regular season runs from September until early January each year. A number of "pre-season" games are played in August and early September, but these games do not count toward the league championship. The data analyzed here include both pre-season and regular season games. Most NFL regular season games are played on Sunday afternoon and evening. In addition to Sunday games, one (and occasionally two) games are played on Monday night, and some other games are

Table 1: Summary Statistics, NFL Betting 2005-2007

	2005		2006		2007		2008	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Score Difference	3.78	14.0	0.85	14.4	2.86	15.4	2.56	15.3
Opening Point Spread	2.51	5.6	2.77	5.6	2.46	6.8	2.70	6.1
Closing Point Spread	2.70	6.1	2.80	5.9	2.42	6.7	2.67	6.13
Home team bet win %	0.49		0.47		0.50		0.44	
Favored team bet win %	0.55		0.42		0.51		0.49	
Fraction of bets on favorite	61.60	13.2	60.33	14.4	62.34	12.3	61.5	12.8
Games	256		256		256		256	

played on Thursdays and Saturdays later in the season. Each team plays 16 regular season games spread over 17 weeks. A small number of teams with the best records during the regular season advance to the postseason knock-out tournament that culminates in the Super Bowl. Betting on NFL games takes place on a rigorous schedule. Sports books issue an opening point spread on each game early in the week for the entire slate of games scheduled to take place over the next week. The opening line is made public on Sunday evening or Monday morning. Throughout the week, information about the status of injured players is made public, and the sports books observe the order flow in the market. Point spreads are changed on some games, either in response to new information about players or weather conditions, or in response to observed betting volumes on the games. The final point spread is the point spread that is posted immediately prior to the start of each game, when betting ends.

The first row on Table 1 is the average actual score difference, expressed as home team score minus visiting team score, for all NFL games in the data set in each season. The second two rows are the average opening and closing point spread set by the four on-line sports books on each game in each of the three seasons. Note that the actual score difference exhibits considerably more variation than either of the point spreads. Sauer (1998) reported a similar pattern in data on betting on National Basketball Association (NBA) games in the 1980s, and many others have reported that actual outcomes are much more variable than point spreads. On average, the point spread changed by roughly one point from the opening line to the closing line an all three seasons. However, in a significant number of games, 28% of them, the point spread did not change over the course of the week. The next two rows show the win percentage of bets placed on the home team and the favored team in each game. The overall average winning percentage for bets on the home team (0.48) and bets on the favored team (0.49) are less than 0.50, suggesting that sports books may shade the point spread against these bets. However, these winning percentages show considerable variation across seasons, indicating the presence of an important random component in these outcomes. Note that two of these season average winning percentages, betting on home teams in 2008, and betting on favorites in 2005, have average winning percentages outside the absence of profit bounds (0.476, 0.524) described above.

The last row reveals some interesting information about betting volumes. The data set contains information on the volume of bets placed on either side of the propositions for each game. This fraction will not be equal to the fraction of money bet on each side when the average value of the bets on the two sides are different. However, anecdotal evidence suggests that the volume of bets

Table 2: Distribution of Bet Volume and Dollars Bet

Variable	2005	2006	2007	2008
Average % of bets on favorite	61.6	60.3	62.3	61.5
Median % of bets on favorite	63.0	62.5	64.0	63.5
Skewness	-0.36	-0.52	-0.54	-0.51
Kurtosis	2.31	2.48	3.24	2.82
% of games where $47.6 \leq$ % of bets on favorite ≤ 52.4	7.8	10.5	6.6	7.8
% of games where % of bets on favorite > 60.0	59.0	60.0	64.1	62.9
% of games where % of bets on favorite > 75.0	21.4	19.4	14.1	13.3

on each side is equal to the volume of money bet on each side. Clearly, the volume of bets on either side are not balanced very often in these data. The average fraction of bets are not equal to 50% in any season, and the standard deviations are relatively large, indicating substantial variation in the volume of bets. Bettors like to bet on favorites in the NFL. In each season, more than 60% of the bets placed were on the favored team. Again, the fact that a majority of the bets were placed on the favorite, and the win percentage of bets on the favorite was less than 50% suggests shading of the point spread by bookmakers, potentially to take advantage of uninformed bettors.

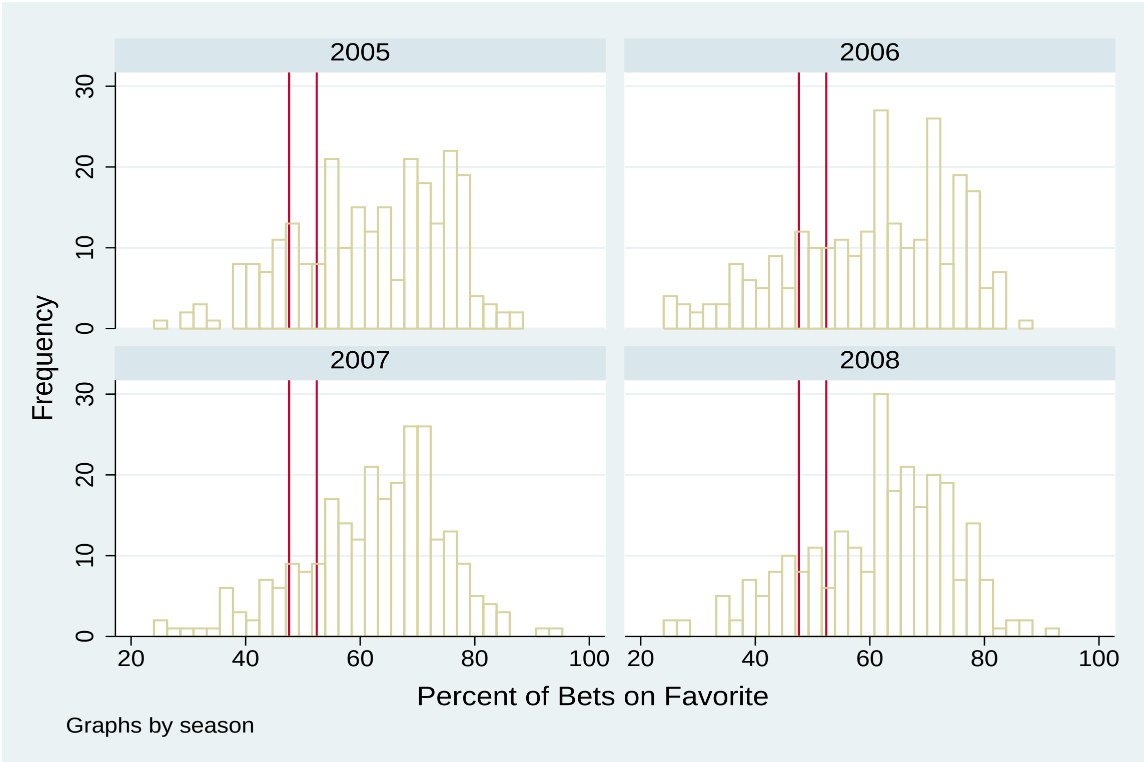
Distribution of Betting Volume

The disparity in the volume of bets placed on either side of games revealed on Table 1 does not fit with the typical "balanced book" model described in the literature. The volume of bets made is skewed toward favorites, the team that is expected to win the game, in all four NFL seasons. A closer look at the data on the fraction of bets made on either side of propositions shows a large number of games with unbalanced betting. The large standard deviations on the bet volume data shown on Table 1 suggest that the betting volume on individual games might be quite different from a balanced book.

Table 2 takes a closer look at the distribution of betting volume data across individual games. From the break even condition, equation (3), if the fraction of bets on the favored team falls between 47.6% and 52.4%, then the sports book makes a profit on the betting no matter which team wins the game. If the fraction of bets on the favored team falls outside this range, the then the sports book takes a position on the game, and can either gain more or lose more, depending on the outcome of the game. From Table 2, the distribution of the bets placed on the favored team is quite skewed. Sports books consistently take positions on games, and these positions are mostly on the underdog. The fifth row on Table 2 shows the percent of games in which the observed fraction of bets on the favorite fell inside the certain profit range in the 2005-2007 NFL seasons. In more than 90% of the games in these three NFL seasons, the observed fraction of bets on the favorite fell outside the certain profit range. In other words, the four on-line sports books represented in this data set took a position, on average, on 9 out of 10 NFL games that they took bets on. Either these books were exceptionally bad at setting point spreads to to equalize betting on either side of the game, or achieving a balanced book was not the goal of sports books taking bets on NFL games.

Figure 1 shows the distribution of the fraction of bets on the favored team in each game in each season. The red vertical lines on Figure 1 show the boundaries of the certain profit region. Figure 1

Figure 1: Distribution of Fraction of Bets on Favored Team



shows a large amount of variation in the fraction of bets on the favorite. Again, Figure 1 indicates that sports books took large positions on games, and that a majority of bettors prefer to bet on the favored team. This tendency for sports bettors to over bet favored teams has been documented by Woodland and Woodland (1994) in Major League Baseball and by Paul and Weinbach (2005) in the National Basketball Association.

The distribution of the bets on the favored team in point spread betting on the NFL falls well outside the certain profit range for almost all games, indicating that sports books take positions on games frequently. Most previous research has assumed that sports books attempt to set point spreads to balance the volume of bets on either side of the game. If this were the case, we would expect to see many more instances of the betting volume falling in the certain profit range. Previous research also indicates that point spreads are unbiased and minimum variance estimators of actual game outcomes. One reason for the unbalanced book outcomes observed above could be that point spreads were not efficient during these three seasons for some reason.

Market Efficiency Tests

One important characteristic used to evaluate sports betting markets is the efficiency of the market. Efficiency in sports betting markets is typically defined as the absence of profit making opportunities; that is, in efficient sports betting markets bettors are unable to make positive profits in the long run. Efficiency in sports betting markets implies that, given symmetry of the distribution of point score differences, the point spread set by sports books on a game is an unbiased, minimum variance estimator of the difference in points scored in the game. In practical terms, tests of

Table 3: Market Efficiency Tests

	2005	2006	2007	2008
Intercept	0.636	-1.58	0.186	-0.155
P-Value	0.472	0.096	0.837	0.871
Opening Point Spread	1.254	0.879	1.090	1.003
P-Value	0.001	0.001	0.001	0.001
R^2	0.231	0.116	0.228	0.162
Observations	256	256	256	256
F-stat, $\alpha = 0$, $\beta = 1$	2.81	2.78	0.37	0.01
P-value	0.062	0.058	0.689	0.986
Intercept	0.531	-1.50	0.137	-0.219
P-Value	0.545	0.112	0.878	0.818
Latest Point Spread	1.202	0.838	1.126	1.039
P-Value	0.001	0.001	0.001	0.001
R^2	0.243	0.118	0.242	0.173
Observations	256	256	256	256
F-stat, $\alpha = 0$, $\beta = 1$	2.06	3.28	0.65	0.05
P-value	0.129	0.039	0.524	0.954

efficiency in sports betting markets are based on a regression model

$$DP_i = \alpha + \beta PS_i + e_i \tag{10}$$

where DP_i is the difference in points scored by the two teams involved in game i, PS_i is the point spread set by sports books on game i, and e_i is an unobservable random variable assumed to be distributed with mean zero and constant variance that captures all other factors that affect the difference in points scored. In this regression model, tests of efficiency are based on the joint hypothesis test based of the null

$$H_o : \alpha = 0 \text{ and } \beta = 1.$$

By convention, the points scored variable is expressed as visitor's points scored minus home team's points scored, and the point spread is expressed as negative numbers when the home team is favored and positive numbers when the home team is the underdog. The distribution of the points scored variable is relatively symmetric, the mean is -2.42 and the median is 3, so regression based efficiency tests appear to be appropriate. A histogram of this variable can be found in the appendix. In this case, the opening and closing lines are observed, so efficiency tests can be performed for both the opening line and the closing line.

Table 3 shows the results of estimating equation 10 using data from the 2005, 2006, 2007 and 2008 seasons separately. The key statistics on this table are the F-statistics on the test of the joint hypothesis that the intercept is equal to zero and the slope parameter is equal to one. This is the conventional test of betting market efficiency; if the null is accepted, then the point spread is an unbiased minimum variance estimate of the difference in points scored. This rejection implies the absence of profit opportunities for bettors in this market. This null hypothesis is accepted at conventional significance levels for all seasons for both the opening point spread and the closing point spread. Sauer (1998) presents a detailed discussion of efficiency tests in this setting. Both the opening line and closing line are good predictors of the actual point score in NFL games in these three seasons. This result is consistent with other tests of NFL point spread betting markets found in the literature. The point spreads set by the four on-line sports books represented in these data indicate that the market for point spread betting on NFL games was efficient in these three seasons. Only the closing line in the 2006 season shows evidence that the point spread may not be a good predictor of the game outcome.

Losses, Gains, and Profits

The most important evidence supporting the model of sports book behavior developed above is based on the actual returns earned by the sports book, given the observed point spreads, game outcomes, and distribution of bets on either side of proposition is in this market. The data set analyzed here contains enough data to conduct financial simulations of the profitability of sports books. I do not have access to enough data to calculate the exact amount of profit earned by sports books, for four reasons. First, I do not have data from specific sports books. The data are averages across four different sports books. If there is a significant amount of heterogeneity in point spreads set, bets taken, and the volume of bets on each side of a proposition across sports books, then the average data will not reflect this heterogeneity. Second, I do not have access to data on the timing of individual bets. All I observe is the opening line and the closing line on each game and the final volume of bets on each side. Because the point spread changes in about 75% of the games, and changes by an average of about 1 point, I do not know the exact amount of money wagered on each side at each point spread that was available during the week that the sports books were taking bets on the games. This is important because point spread bets pay off based on the point spread that was posted at the time the bet was made, not based on the last point spread posted. Third, I do not know how much money was wagered on each game; I only have access to information on the fraction of bets, and the fraction of money bet on either side of each game. Fourth, I do not have access to the volume of money bet on each side in the games, only the fraction of bets. Because of these limitations, I use simulations to estimate the gains, losses, and profits earned by sports books on point spread bets in NFL games taken over these three seasons.

The simulations are straightforward. For each game, I compare the opening and final point spread to the actual game outcome to determine which side of the bet won, and which side lost. The winning bets go in the books as losses for the sports books, since they must pay the bettors who made these bets. The losing bets go in the books as gains to the sports books, plus the 10% commission charged on losing bets. That means for each $100 wagered on a losing bet, the sports book collects $110 from the bettor. In addition, I assume the the average size of bets on the favorite is equal to the average size of bets on the underdog. Under this assumption, the fraction of bets made on each side is equal to the fraction of dollars bet on each side. Using the fraction of bets placed on each side on each game, I calculate the gains and losses on each game, and add this up over both days and the entire season. Strumpf (2003) carried out similar simulations using data for several illegal bookmakers in the New York City area in the 1990s.

The first set of simulations assumes an equal number of dollars was bet on each game in the season. For simplicity, I assume that 100 total "units" were bet on each game. While this assumption probably does not match reality – the total amount bet on games probably varies depending on the teams involved – it is a convenient baseline for comparing the simulation results. Table 4 shows selected summary statistics for the baseline simulations. The top panel assumes that all bets are made at the opening point spread; the bottom panel assumes that all bets are made at the final point spread. The actual distribution of bets lies somewhere between these two points,

Table 4: Average Daily Loss, Gain, and Returns 2005-2008

	2005 Season		2006 Season		2007 Season		2008 Season	
	100 Units bet, all bets at opening spread							
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Loss	-50.1	17.7	-48.4	17.6	-50.3	17.5	-49.4	17.1
Gain	54.8	19.4	56.6	19.4	54.6	19.3	55.6	18.8
Return	4.7	37.2	8.4	37.0	4.2	36.8	6.2	36.0
	100 units bet, all bets at closing spread							
Loss	-50.6	17.7	-48.7	17.6	-50.0	17.5	-49.7	17.1
Gain	54.3	19.4	56.4	19.4	55.0	19.3	55.3	18.8
Return	3.7	37.2	7.7	37.1	5.0	36.8	5.5	36.0
Games	245		249		251		250	

unless the line does not change over the course of the week.

Several interesting features emerge from the baseline simulations. First, despite the lack of a balanced book for nearly all the games, the sports books make a positive return on average each day, and over the course of the season, no matter which point spread is used. Because of the assumption that 100 units are bet on each game, averages reported on the "Return" row can also be interpreted as the percent return on all bets taken. So, for example, in the 2005 season the average return on all bets taken, assuming that they were made at the latest point spread, was 3.7, or 3.7%. Second, notice that the average return varies quite a bit over the three seasons simulated.

Another interesting feature is that returns are lower on bets made at the latest point spread compared to bets made at the opening point spread. A number of papers have analyzed changes in point spreads, including Gandar, Zuber, O'Brien and Russo (1988), and Avery and Chevalier (1999). The focus of this line of research has been to explain why point spreads change over the course of the week in the NFL, and the extent to which informed traders or bettor sentiment, as captured by observable variables associated with past team performance, explains observed changes in point spreads.

The results of these simulations indicate that, no matter what the reason for moving point spreads, the effect of these changes appears to increase the returns earned by sports books on bets. Changing the point spread is a profit maximizing activity based on these simulations.

Also, note that the variability of losses is smaller than the variability of gains, and that the variability of profits is largest of all. Operating a sports book is a risky business, because profits are highly variable. The minimum and maximum values on Table 4 underscore just how risky the sports book business can be. Assuming that all bets are made at the latest point spread, the largest daily loss in each of the three seasons was between 64% and 73% of the average bet on each game. All of these maximum losses took place on Sunday, when there can be as many as 15 games played, so it was possible to have sizable losses even on days when many games were played. This is consistent with results in Strumpf (2002).

How do these returns compare to what would have been earned if the book was balanced on all games? Under the assumption of equal dollars bet on each game, this comparison is simple. From equation (4), if an equal amount of money was wagered on each side of every game ($f_1 = 0.5$), then the average return would be $v/2 = 5$ per game or per day. By setting point spreads in a way to produce an unbalanced book on 9 of 10 games, the sports books earned a return higher than this in the 2006 season, a return equal to this in the 2007 season, and a lower return in 2005. Again, I do not have access to enough data to estimate the variance of returns in this setting, so it is unclear how likely a sports book is to earn a return in excess of the certain return of 5% by taking positions on games. All I can conclude from the simulations is that it is possible to earn larger returns by operating an unbalanced book.

Discussion and Conclusions

The results above indicate that sports books regularly take large positions on NFL games. These results stand in contrast to the "balanced book" model that predicts an equal amount of betting on each side of a game. The emerging evidence from research on betting volumes indicates that sports books are willing to take positions on games, even when the point spreads set on these games are unbiased minimum variance predictors of the game outcomes in many different settings. The results here extend this research by showing that unbalanced betting on games generates profits for the sports books, and that these profits can exceed what would have been earned if the betting volume was perfectly balanced on all games. Since the "balanced book" model of sports book behavior cannot explain observed outcomes in sports betting markets, researchers should focus on developing models that can explain actual outcomes in these markets. Clearly, an improved model should take into account the presence of informed and uninformed bettors in the market and strategic interaction between sports books and informed bettors.

While the results above are interesting, they lack a complete explanation. The result that changes in point spreads affect profits is intriguing, and deserves additional attention. Several possible mechanisms could explain this result. Both line shading, where the sports book strategically moves the point spread to take advantage of known bettor preferences (for example, the tendency of bettors to bet on favored teams at any odds, or the tendency for bettors to bet on the home team), and incomplete information could explain the increase in profits associated with changes in the point spread found here. While I show that imbalanced betting volume and point spread changes are profitable for sports books, I have not fully explained why they are profitable.

Both the theoretical and empirical analysis can be extended in several useful directions. First, the model needs to be expanded to explicitly include point spreads, the natural choice variable for sports books. The point spread set by a sports books affects both the fraction of bets on either side of a proposition and the probability that a bet on the favored team wins. This extension will require quite a bit of modeling, as it makes both the variables that currently appear in the model endogenous; formulating a useful model that includes these two features will require the incorporation of additional elements of sports betting markets. Second, the model needs to be expanded to include a richer depiction of the choices made by bettors in this market. These extensions include formal modeling of the bettors' participation decisions and the introduction of bettor heterogeneity to capture the actions of informed and uninformed bettors. Cain, Law and Lindley (2000) provide one approach for this extension. Explicitly modeling bettor heterogeneity will also permit a full examination of shading of point spreads by sports books in this market, a widely documented phenomenon (Strumpf, 2002).

Several clear extensions to the empirical analysis exist. First, the financial simulations need to be expanded to include unequal betting volume on games. The simplest extension would assume that the volume bet on games is proportional to the size of the markets that teams play in, or proportional to past success by the teams involved. It is possible that variation in betting volume could change the returns to the sports book significantly, given the large variability of returns in the equal volume simulations performed here. A second extension is to perform a similar analysis in different settings. New data are becoming available on betting volume in other point spread betting markets like college football, and professional and college basketball. These sports have the advantage of many more games in each season than the NFL. However, betting on these sports probably exhibits much more variation in total dollars bet on games, which will place a premium on correcting for variation in the amount bet on each game when assessing the total returns to the sports book.

References

Avery C. & Chevalier, J. (1999). Identifying investor sentiment from price paths: The case of football betting, *Journal of Business*, 72(4), 493-521.

Cain, M., Law, D. & Lindley, D. (2000). The construction of a simple book, *Journal of Risk and Uncertainty*, 20(2), 119-140.

Conlisk, J. (1993). The utility of gambling, *Journal of Risk and Uncertainty*, 6(3), 255-275.

Hurley, W. & McDonough, L. (1995). A note on the Hayek hypothesis and the favorite-longshot bias in parimutuel betting, *American Economic Review*, 85(4), 949-955.

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

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

Paul, R. J. & Weinbach, A. P. (2005). Bettor misperceptions in the NBA, *Journal of Sports Economics*, 6(4), 390-400.

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

Paul, R. J. & Weinbach, A. P. (2008). Price setting in the NBA gambling market: Tests of the Levitt model of sportsbook behavior. *International Journal of Sports Finance*, 3(3), 2-18.

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

Strumpf, K. (2002). Illegal sports bookmakers, Mimeo, University of North Carolina. Department of Economics.

Tversky, A. and D. Kahneman (1974) Judgment under uncertainty: Heuristics and biases, *Science*, New Series, 185(4157), 1124-1131.

Woodland, B. M. & Woodland, L. M. (1991). The effects of risk aversion on wagering: Point spread versus odds, *Journal of Political Economy*, 99(3), 638-653.

Woodland, B. M. & Woodland, L. M. (1994). Market efficiency and the favorite-longshot bias: The baseball betting market, *Journal of Finance*, 49(1), 269-279.

Appendix

Derivation of the expected return maximizing fraction of bets on team 1. The first line is equation (6), the expected return on a bet. Equation (7) is substituted into the equation and the optimum f_1 is derived.

$$
\begin{aligned}
E[R] &= (2+v)(f_1+\pi_1-2\pi_1 f_1)-1 \\
\pi_1 &= \pi(f_1) \\
E[R] &= (2+v)[f_1+\pi(f_1)-2\pi(f_1)f_1]-1 \\
\frac{\partial E[R]}{\partial f_1} &= (2+v)[1+\pi'(f_1)-2(\pi(f_1)+\pi'(f_1)f_1)] \\
\frac{\partial E[R]}{\partial f_1} &= (2+v)[1+\pi'(f_1)-2\pi(f_1)-2\pi'(f_1)f_1)] \\
\frac{\partial E[R]}{\partial f_1} &= (2+v)[1-2\pi(f_1)-2\pi'(f_1)f_1+\pi'(f_1)]=0 \\
\frac{\partial E[R]}{\partial f_1} &= (2+v)-(2+v)2\pi(f_1)-2(2+v)\pi'(f_1)f_1+(2+v)\pi'(f_1)=0 \\
-2(2+v)\pi'(f_1)f_1^* &= (2+v)2\pi(f_1)-(2+v)-(2+v)\pi'(f_1) \\
2(2+v)\pi'(f_1)f_1^* &= (2+v)(1-2\pi(f_1)+\pi'(f_1)) \\
f_1^* &= \frac{(2+v)(1-2\pi(f_1))+\pi'(f_1))}{2(2+v)\pi'(f_1)} \\
f_1^* &= \frac{1}{2}\cdot\frac{1-2\pi(f_1)+\pi'(f_1))}{\pi'(f_1)}
\end{aligned}
\tag{11}
$$

$$
\begin{aligned}
E[R] &= (2+v)(f_1+\pi_1-2\pi_1 f_1)-1 \\
\pi_1 &= \sigma f_1 \\
E[R] &= (2+v)[f_1+\sigma f_1-2\sigma f_1\cdot f_1]-1 \\
E[R] &= (2+v)[f_1+\sigma f_1-2\sigma f_1^2]-1 \\
\frac{\partial E[R]}{\partial f_1} &= (2+v)[1+\sigma-4\sigma f_1] \\
\frac{\partial E[R]}{\partial f_1} &= (2+v)+(2+v)\sigma-4(2+v)\sigma f_1 \\
4(2+v)\sigma f_1^* &= (2+v)+(2+v)\sigma \\
f_1^* &= \frac{1+\sigma}{4\sigma}
\end{aligned}
\tag{12}
$$

Department of Economics, University of Alberta
Working Paper Series

http://www.economics.ualberta.ca/working_papers.cfm

2010-04: State-dependent congestion pricing with reference-dependent preferences - **Lindsey**
2010-03: Nonlinear Pricing on Private Roads with Congestion and Toll Collection Costs – Wang, **Lindsey**, Yang
2010-02: Think Globally, Act Locally? Stock vs Flow Regulation of a Fossil Fuel – Amigues, **Chakravorty,** Moreaux
2010-01: Oil and Gas in the Canadian Federation – **Plourde**
2009-27: Consumer Behaviour in Lotto Markets: The Double Hurdle Approach and Zeros in Gambling Survey Dada – **Humphreys, Lee, Soebbing**
2009-26: Constructing Consumer Sentiment Index for U.S. Using Google Searches – Della Penna, **Huang**
2009-25: The perceived framework of a classical statistic: Is the non-invariance of a Wald statistic much ado about null thing? - **Dastoor**
2009-24: Tit-for-Tat Strategies in Repeated Prisoner's Dilemma Games: Evidence from NCAA Football – **Humphreys, Ruseski**
2009-23: Modeling Internal Decision Making Process: An Explanation of Conflicting Empirical Results on Behavior of Nonprofit and For-Profit Hospitals – **Ruseski,** Carroll
2009-22: Monetary and Implicit Incentives of Patent Examiners – **Langinier,** Marcoul
2009-21: Search of Prior Art and Revelation of Information by Patent Applicants – **Langinier,** Marcoul
2009-20: Fuel versus Food – **Chakravorty,** Hubert, Nøstbakken
2009-19: Can Nuclear Power Supply Clean Energy in the Long Run? A Model with Endogenous Substitution of Resources – **Chakravorty**, Magné, Moreaux
2009-18 Too Many Municipalities? – **Dahlby**
2009-17 The Marginal Cost of Public Funds and the Flypaper Effect – **Dahlby**
2009-16 The Optimal Taxation Approach to Intergovernmental Grants – **Dahlby**
2009-15 Adverse Selection and Risk Aversion in Capital Markets – Braido, da Costa, **Dahlby**
2009-14 A Median Voter Model of the Vertical Fiscal Gap – **Dahlby,** Rodden, **Wilson**
2009-13 A New Look at Copper Markets: A Regime-Switching Jump Model – Chan, **Young**
2009-12 Tort Reform, Disputes and Belief Formation – **Landeo**
2009-11 The Role of the Real Exchange Rate Adjustment in Expanding Service Employment in China – **Xu**, Xiaoyi
2009-10 "Twin Peaks" in Energy Prices: A Hotelling Model with Pollution and Learning – **Chakravorty**, Leach, Moreaux
2009-09 The Economics of Participation and Time Spent in Physical Activity – **Humphreys, Ruseski**
2009-08 Water Allocation Under Distribution Losses: Comparing Alternative Institutions – **Chakravorty**, Hochman, Umetsu, Zilberman

Please see above working papers link for earlier papers