Efficiency Implications of the Intra-Week Evolution of Market Prices: An Analysis of Sequential NFL Betting Lines in New York City

Thomas W. Miller, Jr.*
Professor of Finance
John Cook School of Business
Saint Louis University
millertw@slu.edu

David E. Rapach
Associate Professor of Economics
John Cook School of Business
Saint Louis University
rapachde@slu.edu

First version: November 12, 2008
This version: October 20, 2009

*Corresponding author. Please send correspondence to Thomas W. Miller, Jr., Department of Finance, John Cook School of Business, Saint Louis University, 3674 Lindell Boulevard, St. Louis, MO 63108-3397; phone: 314-977-3830; fax: 314-977-1478; e-mail: millertw@slu.edu.

We thank Doug Young for his late-blooming inspiration. In addition, we acknowledge very helpful comments from Bill Dare, Brad Humphreys, Carolyn Moore, Chris Neely and from seminar participants at the 2009 International Conference on Gambling and Risk Taking, Stateline, NV. The usual disclaimer applies. Matt Harlow and Jared Jones, M.D. provided tireless and accurate research assistance. To generate the results reported in this paper, we used MATLAB 7.1 and SAS 9.2.

Disclaimer: In this study, the authors describe a series of sometimes illegal betting transactions. The identities of the participants and any specific betting information presented in this paper have been available to the general public for over 35 years. This length of time exceeds any State of New York statutes of limitations regarding laws that may or may not have been violated by any person described in this study. Nevertheless, the authors do not condone, encourage, or applaud illegal endeavors.

Efficiency Implications of the Intra-Week Evolution of Market Prices: An Analysis of Sequential NFL Betting Lines in New York City

Abstract

We study the intra-week evolution of price efficiency in the 1972 NFL betting market in New York City. Our unique data set provides weekly observations for three sequential betting lines: (i) an "Outlaw Line" set by a single agent at the beginning of the week; (ii) Tuesday's "Opening Line" shaped by bets from a group of 8–10 agents; and (iii) a game-time "Closing Line" formed by subsequent bets placed by the wider public. We analyze the information content and efficiency of these three betting lines as more agents participate in the market and as more information becomes available. Forecast encompassing tests clearly show that information content increases as the betting line evolves from the Outlaw to the Opening to the Closing. Using a sequential regression framework, however, we uncover significant evidence of pricing inefficiencies related to various measures of sentiment. Betting strategies based on exploiting these inefficiencies generate economically and statistically significant profits. In addition, actual bets made by a number of "professional" gamblers appear profitable, pointing to the existence of superior analysts in the market. Overall, our results show that while betting lines evolved in a manner consistent with a degree of efficient information processing, significant inefficiencies emerged in the price-setting process.

JEL classifications: C25, C53, G14, G17

Keywords: Outlaw Line; National Football League; Market efficiency; Forecast encompassing; Sentiment; Superior analysts

Efficiency Implications of the Intra-Week Evolution of Market Prices: An Analysis of Sequential NFL Betting Lines in New York City

1 Introduction

Bettors and brokers say convictions are opinions backed by money. Few markets allow researchers to measure convictions as readily as the National Football League (NFL) betting market. Accordingly, the NFL betting market has long served as a laboratory for testing market efficiency in a relatively direct manner.[1]

A vexing aspect of testing market efficiency is that the true value of the underlying asset remains unknown. Therefore, an appealing aspect of testing market efficiency in the NFL betting market is that the range of payoffs is known with certainty in advance. The NFL betting market is also an appealing setting to test market efficiency because its market structure is simple and payoffs occur soon after betting ends.

In this paper, we employ a unique data set that allows for an extensive analysis of the intra-week evolution of NFL betting lines and the implications for market efficiency. Our data set consists of three sequential NFL betting lines culled from *The National Football Lottery*, in which Mr. Larry Merchant chronicles his weekly betting activities in New York City during the 1972 NFL season. Of primary interest to us, Mr. Merchant presents three sequential betting lines for each week of the 1972 NFL season: the Outlaw Line, the Opening Line, and the Closing Line.

The Outlaw Line is set by a single agent at the beginning of the betting week. Soon thereafter, a small group of 8–10 "professional" bettors shapes the Outlaw Line into the Opening Line. The market is then opened to all bettors—professional and public alike—and bets made during the remainder of the week move the Opening Line to the Closing Line. In our data set, the Closing Line is the game-time betting line. Movements from the Opening Line to the Closing Line reflect both the collective views and sentiments of a much wider body of market participants as well as additional information that becomes available during the week. In our study, the Opening and Closing Lines are (illegally) provided to Mr. Merchant by New York "bookies." It is important

[1] In fact, Pankoff (1968) published the initial study of NFL betting market efficiency soon after the seminal works on market efficiency by Fama (1965) and Samuelson (1965). Other sports betting markets have also been studied, such as basketball, baseball, and horse racing. See Sauer (1998) for an extensive survey of the sports betting literature through the late 1990s.

to note that these agents stood ready to take the other side of any bet made by Mr. Merchant, so that the betting lines represent actual transaction prices.[2] To our knowledge, Levitt (2004), who analyzes points spreads for an online handicapping contest for 2001 NFL season games, is the only other study that uses actual transaction prices. However, the nature of the contest means that the bets analyzed by Levitt (2004) are not wages in the traditional sense.

We begin our analysis by comparing the information content in these three sequential betting lines. Previous studies typically estimate Mincer and Zarnowitz (1969) regressions to test for biases in the individual betting lines.[3] This approach, however, has limited power in uncovering information differences across forecasts. This fact motivates our use of the Harvey, Leybourne, and Newbold (1998) forecast encompassing test, which, to our knowledge, has not been previously employed to compare betting lines. This method allows us to test directly whether a particular betting line (i.e., forecast) contains information above and beyond the information contained in another line.

We find that encompassing tests signal significant differences in information content among betting lines and provide interesting insight on the evolution of information in the betting market. More specifically, we find that the Opening Line contains information beyond that contained in the Outlaw Line, but the converse does not hold. Similarly, the Closing Line contains information beyond that in the Opening Line, but the converse is not true. Our results thus indicate that the information content of betting lines increases as more agents participate in the betting market and as agents' potential relevant information set grows. This pattern of increasing information content is expected in a reasonably well-functioning market.

We further scrutinize market efficiency by investigating the role of sentiment in explaining fluctuations in the three sequential betting lines. We perform this inquiry using a sequence of regression models where the dependent variable is, in turn, the Outlaw Line, changes in the Outlaw Line to the Opening Line, and changes in the Opening Line to the Closing Line. The regressors are variables constructed using publicly available information that exists before the setting of any of the three betting lines. Because Mr. Merchant claims that power rankings play an important role in setting the Outlaw Line, the first regressor is a predicted point spread based on power rankings constructed using a methodology from Dana and Knetter (1994). The remaining regressors are

[2]In advance, Mr. Merchant and his illegal bookmakers agreed on bet sizes and the frequency with which the winner would pay the loser. In no instance did a bookmaker refuse a bet.

[3]Representative studies that estimate Mincer-Zarnowitz regressions to analyze NFL betting lines include Pankoff (1968), Gandar, Zuber, O'Brien, and Russo (1988), Golec and Tamarkin (1991), and Boulier, Stekler, and Amundson (2006).

eight sentiment measures similar to those used by Avery and Chevalier (1999).

Our results show that power ranking and a number of the sentiment variables are significant determinants of the Outlaw Line, and that these variables jointly explain nearly 75% of the variation in the Outlaw Line. These results confirm Mr. Merchant's claim of the importance of power rankings in setting the Outlaw Line. These results also indicate that the lone agent setting the Outlaw Line is sensitive to measures of sentiment in the NFL betting market. Furthermore, we find that power ranking and some of the sentiment variables are significant determinants of subsequent movements in betting lines. This latter finding is interesting because the information contained in the power rankings and sentiment variables is publicly available before the time the Outlaw Line is set. That is, this "stale" information continues to affect prices as the number of market participants grows during the betting week.[4]

Like Avery and Chevalier (1999), we estimate ordered logit models to examine whether the power rankings and sentiment variables are significantly related to the actual outcome of betting into the three lines. If the market efficiently processes all relevant information contained in the power ranking and sentiment variables available at the start of the week, none of this information should be significantly related to actual outcomes when betting into the three lines, especially the Opening and Closing Lines. We find, however, that the betting lines are inefficiently slanted against prestige teams identified *a priori* by Mr. Merchant and toward teams that display a "hot hand." This finding suggests that strategies based on betting with a group of prestige teams and against hot-hand teams are potentially profitable. Indeed, we show that such simple trading strategies earn economically and statistically significant profits during the 1972 season.

Finally, our unique data set also allows us to study the payoffs of bets made by a number of informed traders. In the *The National Football Lottery*, Mr. Merchant records his complete set of bets for the 1972 season, as well as those of four other gamblers. One of the gamblers is identified by Mr. Merchant as a "noise" bettor, while he views the others as informed bettors. Our analysis finds that Mr. Merchant and two of the three other informed bettors made handsome (and statistically significant) profits. The remaining informed bettor realized moderate (and statistically insignificant) profits. In contrast, the noise bettor suffered losses. We interpret the evidence of profitable trading by a number of informed bettors as strong evidence of the existence of superior analysts, because these traders were identified *a priori* by Mr. Merchant.

Overall, our results paint a nuanced picture of price formation in the 1972 NFL betting market.

[4]This finding is similar in spirit to the partial adjustment process for IPOs documented by Bradley and Jordan (2002).

The information content of market prices clearly increases during the week, as the market becomes thicker with traders and the potential relevant information set grows. Despite the expected general increase in the information content of market prices over the course of the week, information is not processed in a completely efficient manner. As a result, pricing errors exist. In addition, a set of *a priori* identified informed bettors earn economically and statistically significant profits, which, strictly speaking, represents a rejection of semi-strong form market efficiency.

The remainder of the paper is organized as follows. We present a brief review of research on the NFL betting market in Section 2. We describe the 1972 NFL betting market and our point spread data in Section 3, and we report forecast encompassing test results in Section 4. Our results for a sequence of regression models that use power rankings and sentiment variables to explain the Outlaw, Opening, and Closing Lines appear in Section 5, along with the results of our examination of the profitability of trading strategies based on prestige and hot-hand effects. Our investigation of the performance of professional bettors is in Section 6, and our conclusions are in Section 7.

2 Summary of Previous Research

2.1 Mechanics of the Point Spread Betting Market

The mechanics of betting on an NFL game are simple. A straight bet is the most common football bet. A straight bet is a contingent claim whose expiration value depends on the *ex post* difference in the actual points scored versus an *ex ante* point spread, commonly called "the betting line," or simply "the line." The betting line is a price set by a "bookmaker" or "bookie." With point-spread betting, one team is the favorite and the other is the underdog. In a straight bet, a bettor wins when the bettor's chosen team "covers the spread." That is, the bettor's team must win, or not lose, by a predetermined margin. Settlements are typically figured at odds of 11 to 10. To break even, the bettor must win 11 bets for every 10 bets lost, a winning percentage of 11/21, or 52.4%. These settlement odds represent a transaction fee imposed on bettors, colorfully known as the "vigorish," "vig," or "juice."

An example easily illustrates the betting line mechanics. Suppose that the Oakland Raiders are playing the Denver Broncos. Further suppose that the Raiders are a "four-point favorite." Bettors express the betting line in two ways: Broncos $+4$ or Raiders -4. If a (misguided) bettor places a \$500 bet on the Broncos, the bettor receives \$500 if: (i) the Broncos win (or tie); or (ii) the Broncos lose by fewer than four points. If the Raiders win by more than four points, the bettor loses \$550. If the point difference in the game is exactly four points, the bet is "pushed," and no money changes hands.[5]

When bookmakers act as brokers, their goal is to have an equal amount of money bet with customers on both sides of the line, thereby earning a 5% transaction fee on the total amount wagered. Bookmakers can also act as dealers by taking positions, but this, of course, exposes them to greater risk.[6] In general, betting lines express market participants' expectations of actual game outcomes, so they can be viewed as market-generated forecasts. These forecasts can be subjected to statistical and economic tests to assess the rationality of the forecasts and market efficiency.

2.2 Standard Tests of NFL Betting Market Efficiency

A minimal requirement for market efficiency is that the betting line be an unbiased estimator of the actual point difference. Researchers commonly test this implication using the Mincer and Zarnowitz (1969) regression model:

$$S_n = \alpha + \beta X_n + \varepsilon_n, \tag{1}$$

where S_n is the actual point difference, X_n is a betting line, ε_n is a an error term with an expected value of zero, and $n = 1, \ldots, N$ indexes individual games. Under the null hypothesis of forecast unbiasedness, $\alpha = 0$ and $\beta = 1$, so that $E(S_n) = E(X_n)$. To use the regression model (1) for the NFL betting market, the betting line and actual point difference need to be expressed in a consistent manner. This can be done, for example, by expressing both S_n and X_n as the home team score minus the visiting team score. Using this definition, suppose the betting line for game n is four ($X_n = 4$). This means that a bet placed on the home team wins (loses, pushes) if $S_n \geq 5$ ($S_n \leq 3$, $S_n = 4$).

By using betting lines published in newspapers for 1956–1965 NFL regular-season games, Pankoff (1968) was the first to estimate (1) for NFL betting lines. He could not reject the $\alpha = 0$, $\beta = 1$ null hypothesis for any of the samples at conventional significance levels. Zuber, Gandar, and Bowers (1985) and Sauer, Brajer, Ferris, and Marr (1988) generally fail to reject the $\alpha = 0$,

[5] Unlike in legal Las Vegas Sports books today, a bettor in 1972 did not have to give the bookie $500 in advance to place a bet. Also, betting lines are sometimes expressed in half points, which are called "hooks" by bettors. Hooks have the effect of reducing the number of pushes.

[6] We do not have the requisite data to determine the extent to which the bookies are taking positions. Davies and Abram (2001) observe, "Although the ideal situation for the bookie is for his books to be balanced at game time, that is seldom the case. In most instances at kickoff or tipoff time the bookmaker has some of his own money in play on every game; he is in effect betting against his customers." Again, without data, we cannot be sure to what extent this is the case for the 1972 NFL betting market. Levitt (2004) presents evidence from a handicapping contest during the 2001 NFL season that bookmakers take sizable positions. Using data on line movements and betting volume, Humphreys, Paul, and Weinbach (2009) recently present evidence that the "balanced book" model does not hold with respect to the betting market for 2007 NCAA Division IA men's basketball games.

$\beta = 1$ null using data from the 1983 NFL season, while Gandar, Zuber, O'Brien, and Russo (1988) also fail to reject this null for the 1980–1985 seasons.[7] Using data for the 1994–2000 seasons, Boulier, Stekler, and Amundson (2006) obtain similar results.

The $\alpha = 0$, $\beta = 1$ restrictions in the context of (1) represent a relatively weak form of market efficiency. Therefore, testing these restrictions does not necessarily constitute an informative test of market efficiency.[8] A more stringent and potentially more informative test of market efficiency augments (1) with additional regressors known at the time of forecast formation:

$$S_n = \alpha + \beta X_n + Z_n'\gamma + \varepsilon_n, \tag{2}$$

where Z_n is an m-vector of variables based on information available at the time the betting line X_n is set. In addition to the $\alpha = 0$, $\beta = 1$ restrictions, efficiency requires $\gamma = 0$.[9] Pankoff (1968) notes that including variables such as rushing and passing yardage, scoring, and winning percentages in Z_n does not significantly improve the predictive ability of the regression model. Sauer, Brajer, Ferris, and Marr (1988) and Gandar, Zuber, O'Brien, and Russo (1988) obtain similar results using comparable variables in Z_n, as do Boulier, Stekler, and Amundson (2006) using more recent data.

2.3 Extended Tests of NFL Betting Market Efficiency

Overall, studies employing regression tests based on (1) and/or (2) generally fail to reject market efficiency. It could be claimed, however, that these tests are simply not powerful enough to detect market inefficiencies. Two approaches have been pursued in response to this claim.

First, several studies estimate NFL betting models with a qualitative dependent variable. These models are relevant because the outcome of a bet depends only on whether the bettor's team covers the spread, and not the number of points by which the team covers the spread. Gray and Gray (1997) and Avery and Chevalier (1999) estimate ordered probit models that reflect the nature of the outcomes of betting on NFL games. For example, Avery and Chevalier (1999) use a dependent variable that takes a value of 1 if the home team covered the spread, 0 if the bet tied, and -1 if the

[7] Zuber, Gandar, and Bowers (1985) also fail to reject the $\alpha = 0$, $\beta = 0$ null hypothesis for many individual weeks of the 1983 season, an obvious violation of rationality. Sauer, Brajer, Ferris, and Marr (1988), however, argue that this null is rejected when the data from individual weeks are pooled, resulting in a more powerful test.

[8] Furthermore, Golec and Tamarkin (1991) and Dare and MacDonald (1996) elucidate potential misspecification biases in (1) that could lead to biasing the results in favor of efficiency in the form of the $\alpha = 0$, $\beta = 1$ restrictions. Dare and Holland (2004) and Borghesi (2007) recently build on this theme. These studies tend to find more evidence of inefficiencies in the form of home and/or favorite-team biases. Misspecification biases in (1) can be addressed to a certain extent using a suitably modified versions of (2).

[9] Some researchers impose the $\beta = 1$ restriction before estimating (2) and thus estimate the model, $S_n - X_n = \alpha + Z_n'\gamma + \varepsilon_n$, for which the relevant null hypothesis is $\alpha = 0$, $\gamma = 0$.

visiting team covered the spread. Under the maintained hypothesis of an efficient market, no variable based on information available when the betting line is set should predict the betting outcome. Gray and Gray (1997) and Avery and Chevalier (1999), however, both show that variables based on information unquestionably available before the setting of the betting line are significantly related to betting outcomes. Of special interest is the finding that bettors apparently display a hot-hand betting bias. That is, bettors collectively place too much emphasis on the recent performance of teams that covered point spreads.

Second, a number of studies examine the profitability of a variety of betting strategies based on information available when a betting line is set. The semi-strong form of market efficiency requires that such exploitable betting strategies should not be consistently profitable, at least after accounting for transaction costs. A number of studies analyze the profitability of various betting strategies.[10] The results from these studies are decidedly mixed. Some studies claim to uncover profitable trading strategies indicating market inefficiencies, while others fail to identify profitable strategies or question the robustness of purported profitable strategies.

3 Data Description

In 1973, Holt, Rinehart, and Winston published *The National Football Lottery*, by Larry Merchant. In his book, Mr. Merchant chronicles his far-flung exploits centered around his weekly betting activities during the 1972 NFL season. Mr. Merchant placed weekly bets on NFL games with his three bookies based in New York City.[11]

3.1 Betting Lines

Of particular interest for our study is the fact that Mr. Merchant provides weekly data for three sequential betting lines, two of which were obtained from bookies who stood ready to take the other side of the bet from Mr. Merchant. Following is a description of the three betting lines.

Outlaw Line. This is the early, or test, line put out privately on Sunday or Monday by

[10] Representative studies include Pankoff (1968); Vergin and Scriabin (1978); Tryfos, Casey, Cook, Leger, and Pylypiak (1984); Zuber, Gandar, and Bowers (1985); Sauer, Brajer, Ferris, and Marr (1988); Gandar, Zuber, O'Brien, and Russo (1988); Gray and Gray (1997); Avery and Chevalier (1999); Dare and Holland (2004); Boulier, Stekler, and Amundson (2006); and Borghesi (2007).

[11] Mr. Merchant had a main bookie named "Doctor" who took the bulk of Mr. Merchant's betting action. To shop for points and gather information, Mr. Merchant also placed bets with "Billy" and "Mr. Rhodes." According to Mr. Merchant, "If you're serious about betting pro football, you should have at least three bookies." To validate his published betting results, Mr. Merchant mailed his bets in advance to an attorney each week.

an individual linemaker (i.e., bookie). Mr. Merchant reports that the Outlaw Line was set by Mr. Bob Martin. According to Mr. Merchant, "Martin conjures a line every Monday from a witch's stew of records, power ratings, horned toads and intuition." Mr. Merchant's sources describe Mr. Martin as "The Main Man of Bookmaking" and "the Last of the Super Bookies."[12]

Opening Line. This line serves as the initial public line generally available to bettors on Tuesday. According to Mr. Merchant, the Outlaw Line set by Mr. Martin is "subsequently exposed to 'eight or ten' [gambling] pros [on Monday]." Mr. Merchant reports that this betting process "shakes the line down to the hard numbers that [form the Opening Line] posted for all comers on Tuesday."[13] Like most bettors at the time, Mr. Merchant can bet into the Opening Line. (Mr. Merchant is not one of the select bettors allowed to bet into the Outlaw Line.) By phone, Mr. Merchant's main bookie, Doctor, provides the Opening Line.

Closing Line. This is the final or game-time line. After the posting of the Opening Line on Tuesday, bettors can place bets throughout the rest of the week, right up until game time. The point spread can move as money flows to one team or the other. Mr. Merchant reports the Closing Line available from bookies just before kickoffs on games that were played on Saturday, Sunday, or Monday night.

We focus on the analysis of these three intra-week NFL betting market prices. Most researchers focus on a single line, although Avery and Chevalier (1999) and Gandar, Zuber, O'Brien, and Russo (1988) study both the Opening and Closing Lines and movements from the Opening to Closing Lines. To our knowledge, however, no previous study analyzes the Outlaw Line, including its information content and evolution into the Opening Line. Our data set thus allows us to analyze the intra-week price-setting process in the NFL betting market along a new dimension.

Another unique aspect of our betting line data is that the betting lines recorded by Mr. Merchant are actual transaction prices, acquired by Mr. Merchant directly from bookies standing ready to take the other side of the bet. Apart from Levitt (2004), existing studies of NFL betting markets use lines gathered from newspapers and/or gambling publications. While lines from these sources are likely to be close to actual transaction prices, we do not know that they provide observations on

[12] We surmise that Mr. Merchant obtained weekly Outline Law data from a professional gambler living in Las Vegas. Mr. Merchant traveled to Las Vegas during the 1972 season with the expressed purpose of meeting with this gambler and tracking down the source of the Outlaw Line, who turned out to be Mr. Bob Martin.

[13] The resulting Opening Line is the basis, with regional adjustments, for the betting lines "from Seattle to Miami, San Diego to Boston, Memphis to Saint Joe," according to Mr. Merchant.

legitimate transaction prices.[14] Our use of actual transaction price data should sharpen statistical and economic tests of NFL betting market efficiency.

Table 1 reports summary statistics for actual point differences and Outlaw, Opening, and Closing Lines for 1972 regular-season games. There were 26 teams in the NFL in 1972, each of which played 14 regular-season games, resulting in 182 regular-season games.

The mean for the actual point difference is 1.2, indicating that the home team won, on average, by just over one point during the 1972 season. The actual point difference, however, is quite volatile, with a standard deviation of 16.6. The means for the three betting lines are close to the mean for the actual point differences, while the standard deviations of the betting lines are approximately half that of the actual point difference. These relationships between the actual point difference and betting line means and standard deviations in Table 1 indicate that the three betting lines are at least "reasonable" forecasts of actual game outcomes.[15] Overall, the summary statistics for the actual point difference, Opening Line, and Closing Line generally accord with those reported elsewhere in the literature, such as Gandar, Zuber, O'Brien, and Russo (1988) for the 1980–1985 NFL seasons.

Table 1 also contains summary statistics for the magnitudes of the movements from the Outlaw to Opening Lines and Opening to Closing Lines. The mean magnitudes are 0.80 and 0.90 points, respectively. The standard deviations of the two line changes are also about one point. The average size and volatility of the Opening to Closing Line movements are, once again, similar to those reported in Gandar, Zuber, O'Brien, and Russo (1988).

3.2 Gamblers

Another unique aspect of our data set is that Mr. Merchant records his complete weekly betting activities, as well as the activities of four other gamblers—three of whom Mr. Merchant regards as experts/professional gamblers. Mr. Merchant regards the remaining gambler as a noise bettor. The five gamblers for whom we have a complete set of bets for the 1972 season are the following:[16]

Fast Eddie. Fast Eddie told Mr. Merchant that he bets $200 a game. According to Mr. Mer-

[14] Our Closing Line data are also actual game-time betting lines. In contrast, due to the publication timing of newspapers, betting lines gathered from newspapers are not necessarily game-time lines.

[15] With respect to volatility, forecasts obviously should not be more volatile than the actual values themselves. The fact that the betting line standard deviations are not greater than the actual point difference standard deviation shows that this condition is not violated.

[16] Mr. Merchant does not record the precise amount bet by each gambler, but for each game he records whether the gambler bet on the game and the nature of the bet (i.e., whether the bet was placed on the home team or visiting team to cover the spread).

chant, Fast Eddie earned his nickname because he "never knows who's playing until the weekend and then he bets on instinct." Fast Eddie can be classified as a noise bettor.

Mr. Rich. Mr. Rich tells Mr. Merchant that he bets on NFL games as a "hobby." According to Mr. Merchant, Mr. Rich bets up to $500 a game, and "his wife and fifteen-year old son help him handle the phone on Sunday as he shops among half a dozen bookies [for the best lines]." Mr. Merchant clearly regards Mr. Rich as an expert gambler. Indeed, Mr. Rich reveals that he can bet into the Outlaw Line, and, as Mr. Merchant observes, "hobbyists don't bet into it [i.e., the Outlaw Line]."

Greenwich Village Fats. According to Mr. Merchant, "Fats is a prosperous businessman—he has to be to support an old betting habit. Fats says he now bets $100 a game. I am especially interested in Fats because he is betting with information from a professional gambler who bets $250,000 a weekend, for himself and for pros in other cities." The identity of this professional gambler remains unrevealed, but his nickname is "The Mover."

Mr. Lester Banker. Mr. Lester "Lem" Banker is a well-known professional gambler. Mr. Banker moved to Las Vegas in 1957 and soon became one of its most famous denizens. In 1972, Mr. Banker was writing a twice-weekly column on betting for the *Las Vegas Review-Journal*.

Mr. Larry Merchant. When he wrote *The National Football Lottery*, Mr. Merchant was on a leave of absence from his job as a sports columnist for the *New York Post*.[17] Before he began writing the book, Mr. Merchant reports that he would bet $100 a game. In the book, however, Mr. Merchant reports that he bets $500–$2,000 per game.

We note that betting on NFL games outside of Las Vegas was an illegal activity in 1972. In fact, depending on the individual case, promoting gambling, possession of gambling records, or possession of gambling devices could be classified as either a class A misdemeanor or class E felony under the State of New York Criminal Law (Ch. 27, Articles 225 and 415). Nevertheless, the NFL betting market in New York City was by all accounts sizable and very active in 1972. Mr. Merchant consistently refers to the popularity of betting on NFL games in New York City.[18] Overall, the 1972 NFL betting market in New York City appeared thick with traders and information, and transaction costs were equal to those in Las Vegas.

[17] Mr. Merchant played football at Lafayette High School in Brooklyn and the University of Oklahoma. In 1978, Mr. Merchant joined HBO Sports as an analyst and became a well-known member of its boxing coverage team.

[18] Indeed, Mr. Merchant displays a libertarian streak and rails against authority figures who decry the immorality of betting, appealing to the failed experiment of Prohibition.

4 Forecast Encompassing Tests

The notion of forecast encompassing was developed by Chong and Hendry (1986) and Fair and Shiller (1990), among others. This method provides a means to compare the information content across forecasts. To our knowledge, forecast encompassing tests have not previously been employed to compare the information content across NFL betting lines.[19]

The intuition of forecast encompassing tests is as follows. Consider forming an optimal combination forecast as a convex combination of betting lines i and j:

$$X_n^* = (1-\lambda)X_{i,n} + \lambda X_{j,n}, \tag{3}$$

where $0 \leq \lambda \leq 1$. If $\lambda = 0$, then betting line i encompasses betting line j, because line j does not contain any information—beyond that already contained in line i—when forming an optimal composite forecast. In contrast, if $\lambda > 0$, then line i does not encompass line j, because line j does contain information useful for forming an optimal composite forecast (again, beyond the information already contained in line i). In the context of the forecast encompassing method, we conclude that a betting line is superior in information content if it "forecast encompasses" a competing betting line, while the converse is not true.

Harvey, Leybourne, and Newbold (1998, HLN) develop a statistic to test the null hypothesis that forecast i encompasses forecast j ($H_0 : \lambda = 0$) against the (one-sided, upper-tail) alternative hypothesis that forecast i does not encompass forecast j ($\lambda > 0$). Define $z_{i,j,n} = (u_{i,n} - u_{j,n})u_{i,n}$, where $u_{i,n} = S_n - X_{i,n}$ and $u_{j,n} = S_n - X_{j,n}$.[20] Letting $\bar{z}_{i,j} = (1/N)\sum_{i=1}^{N} z_{i,j,n}$, the modified version of the HLN statistic is given by

$$MHLN_{i,j} = \frac{N-1}{N}[V(\bar{z}_{i,j})^{-0.5}]\bar{z}_{i,j}, \tag{4}$$

where $V(\bar{z}_{i,j}) = (1/N)[(1/N)\sum_{n=1}^{N}(z_{i,j,n} - \bar{z}_{i,j})^2]$. The $MHLN_{i,j}$ statistic has a standard normal limiting distribution under the null of forecast encompassing. HLN recommend basing inference on the student-t distribution with $N-1$ degrees of freedom to improve finite-sample performance.

For each possible encompassing test among the three betting lines, Table 2 reports the estimate of λ in (3) and p-value corresponding to the $MHLN$ statistic given by (4). Each entry in the table

[19] We estimated conventional Mincer-Zarnowitz regressions for the three betting lines. We are unable to reject the null hypothesis that $\alpha = 0$, $\beta = 1$ at any reasonable significant level in (1) for all three betting lines. We also tested for differences in mean square forecast error (MSFE) across the three betting lines using the popular Diebold and Mariano (1995) test. This test does not indicate significant differences in MSFE across the three betting lines at conventional significance levels. The complete Mincer-Zarnowitz and Diebold-Mariano results are available upon request from the authors.

[20] We express S_n, $X_{i,n}$, and $X_{j,n}$ as the home team score minus the visiting team score.

corresponds to the null hypothesis that the betting line given in the row heading encompasses the betting line given in the column heading.

Reading across the "Outlaw Betting Line" row in Table 2, we see that the Outlaw Line does not forecast encompass the Opening and Closing Lines (at the 10% and 5% significance levels, respectively). In contrast, the entries in the "Outlaw Betting Line" column indicate that the Opening and Closing Lines encompass the Outlaw Line. Taken together, these results imply that the Opening and Closing Lines are superior in their information content to the Outlaw Line.

The lower entry in the "Closing Betting Line" column in Table 2 shows that the Opening Line does not encompass the Closing Line (at the 10% level), while the rightmost entry in the "Closing Betting Line" row indicates that the Closing Line encompasses the Opening Line. Therefore, according to the forecast encompassing tests, the Closing Line is superior in information content to the Opening Line.

The forecast encompassing test results in Table 2 point to a clear increase in the information content as we move from the Outlaw to the Opening to the Closing Lines. Information content thus appears to increase during the betting week as more agents participate in the betting market and as the potential relevant information set for agents grows. Accordingly, we have evidence that the 1972 NFL betting market processed information in a manner consistent with a reasonably well-functioning market.

5 Modeling the Evolution of Betting Lines

To investigate the evolution of the intra-week betting lines in more detail, we estimate two sets of regression models. The first is a set of three linear regression models with the Outlaw Line, changes from the Outlaw Line to the Opening Line, and changes from the Opening Line to the Closing Line serving, in turn, as the regressand. The second set comprises three ordered logit models for betting outcomes based on betting, in turn, into the Outlaw, Opening, and Closing Lines.

The regressors for both sets of these models are predicted point spreads based on power rankings and eight variables designed to measure sentiment. Our consideration of predicted point spreads based on power rankings is motivated by Mr. Merchant's contention that, as noted in Section 3.1, the setting of the Outline Line by Mr. Bob Martin is influenced by some type of power ratings. The sentiment variables are similar to those used by Avery and Chevalier (1999). Avery and Chevalier (1999) select these variables because of their similarity to potential behavioral strategies followed by stock market investors.[21]

5.1 Construction of Explanatory Variables

5.1.1 Power Rankings

The predicted point spread based on power rankings uses the "Stationary Parameters Model" in Dana and Knetter (1994).[22] We use this model to generate a set of point spread predictions for week-t games by first estimating the following system:

$$S_t = \mu\iota + D_t\theta + e_t, \tag{5}$$

where S_t is a 13×1 vector of actual point differences for week t; ι is a 13×1 vector of ones; D_t is a 13×25 matrix of binary variables, with element $d_{j,k,t} = 1$ ($d_{j,k,t} = -1$) if team k is the home (visiting) team in game j and zero otherwise; and e_t is a 13×1 vector of i.i.d. disturbance terms, each with variance σ_e^2. The parameter μ is the home-team advantage, and the 25×1 vector θ represents a set of team power rankings. Because an intercept is included in (5), we exclude one team by lot (the Atlanta Falcons) from θ. The elements of θ are interpreted as power rankings because they take the form of the expected point spread between a given team and the Atlanta Falcons for a game played at a neutral site.

We form simulated real-time predictions of weekly point spreads based on (5) using the following recursive procedure. For week $t = 1$, we estimate (5) via OLS using actual point differences for the last four (of six) weeks of the preseason, $t = -3, \ldots, 0$, generating $\hat{\mu}_0$ and $\hat{\theta}_0$. Predicted point spreads for the first week of regular-season games are then given by $\hat{S}_1 = \hat{\mu}_0\iota + D_1\hat{\theta}_0$. To compute a set of predicted point spreads for the second week of the regular season, we estimate (5) using data for $t = -3, \ldots, 1$, yielding the predicted vector, $\hat{S}_2 = \hat{\mu}_1\iota + D_2\hat{\theta}_1$. We proceed in this manner to produce a complete set of weekly regular-season predicted point spreads based on power rankings: $\{\hat{S}_t\}_{t=1}^{14}$.[23]

[21] Avery and Chevalier (1999) discuss in more detail how these sentiment variables relate to potential sentiment biases studied in the behavioral finance literature. See Shleifer (2000), Hirshleifer (2001), and Barberis and Thaler (2003) for extensive surveys of the behavioral finance literature.

[22] The authors acknowledge that Dr. Doug Young developed, but did not publish, a similar model at Montana State University in the late 1970s.

[23] For brevity, the complete set of power rankings is not reported. It is available upon request from the authors. We note that weekly power rankings appear quite plausible. For example, the Miami Dolphins have the highest power ranking for the final week of the regular season. Miami went on the win the Super Bowl, capping the first (and only) undefeated season in NFL history. Other playoff teams also have relatively high power rankings. We also experimented with different initial samples for the recursive estimation procedure; for example, we used actual point differences from the 1971 season to generate $\hat{\mu}_0$ and $\hat{\theta}_0$. The results are qualitatively unchanged.

5.1.2 Sentiment Variables

Like the predicted point spread based on power rankings, all of the sentiment variables are based on information available at the beginning of the betting week.

The first four sentiment variables constitute "prestige" variables. The prestige variables are designed to identify popular teams, resulting, for example, from extensive media coverage. The first prestige variable, PRESTIGE4, corresponds to four teams identified by Mr. Merchant as prestige teams at the start of the 1972 NFL season: the Green Bay Packers, Miami Dolphins, Pittsburgh Steelers, and Washington Redskins. Mr. Merchant relates that Mr. Bob Martin, the setter of the Outlaw Line, viewed these teams as popular with the public. We assign each of these four teams a value of 1 and all of the remaining teams a value of 0. PRESTIGE4 is the difference in the binary variables for the home and visiting teams.

Following Avery and Chevalier (1999), we create two variables corresponding to teams that made the playoffs or finished in last place in the previous season. Under the assumption that teams that performed well during the previous season are more popular, playoff teams should be popular and last-place finishers should be unpopular. We assign a value of 1 to each team that made the playoffs (finished last in its division) during the 1971 season and 0 to all remaining teams; PLAYOFF (LASTPLACE) is the difference between the home and visiting team binary variables.

Mr. Merchant places his bets into New York betting lines, and he notes the popularity of the New York teams in the New York betting market. The final prestige variable allows for the possibility that New York teams are especially popular bets with their numerous devout fans. We assign a value of 1 to the New York Giants and the New York Jets and 0 to all remaining teams; NYTEAMS is the difference between the binary variables for the home and visiting teams.[24]

The next three sentiment variables relate to potential hot-hand biases, which occur when the public bets heavily on teams that have recently performed well. We use three hot-hand variables defined by Avery and Chevalier (1999). The first, WINSTREAK, is based on recent winning and/or losing steaks. More specifically, for each team in a given week, we compute the number of games with the same outcome as the previous week. The measure is positive for teams that won the previous week and negative for teams that lost. WINSTREAK is the difference between

[24]Mr. Merchant acquaints us with a particularly rabid New York Giant fan, Mr. Danny Levezzo. Before the 1972 season, Mr. Levezzo declared that he would bet only on the New York Giants. According to Mr. Merchant, "Bookies and friends, knowing this, always give him the worst of the spread." In week 11, all over New York, the Giants were 14–14.5 point favorites to beat the Eagles. Mr. Levezzo, however, was offered a 15.5 spread favoring the Giants. Mr. Merchant reports, "So, he [Mr. Levezzo] shrewdly took the Eagles [and the points]." Sadly for Mr. Levezzo, the Giants won the game by 52 points.

this measure for the home and visiting teams. The next two hot-hand variables, LAG1 and LAG2, reflect how teams have performed recently relative to the betting line. LAGk is equal to: [(home team actual point difference k weeks ago) minus (Closing Line k weeks ago)] minus [(visiting team actual point difference k weeks ago) minus (Closing Line k weeks ago)].[25]

The final sentiment variable allows for the possibility that bettors read special meaning into previous games between the same pair of teams. Following Avery and Chevalier (1999), PREVMATCH has a value of 1 (-1) if the home (visiting) team won the previous matchup during the 1972 season and 0 if the teams did not play each other previously in the season.

5.2 Linear Regression Models for Betting Lines

The second through fourth columns of Table 3 report results for three linear regression models. Each model includes nine explanatory variables: the predicted point spread based on power rankings, the four prestige variables (PRESTIGE4, PLAYOFF, LASTPLACE, NYTEAMS), three hot-hand variables (WINSTREAK, LAG1, LAG2), and PREVMATCH.

The second column of Table 3 reports results for a linear regression model with the Outlaw Line serving as the dependent variable. This regression allows us to explore how the power rankings and sentiment variables affect the setting of the initial betting line by a single agent, Mr. Bob Martin. The R^2 statistic shows that these variables explain a substantial 74% of the variation in the Outlaw Line. The predicted point spread based only on the power rankings is an important determinant of the Outlaw Line, with a t-statistic near 12. The estimated coefficient indicates that Mr. Martin increases the line by approximately half a point for each point predicted by the relative power rankings.

Four (three) of the eight sentiment variables are also significant determinants of the Outlaw Line at the 10% (5%) level. PRESTIGE4 enters with a negative coefficient, meaning that Mr. Martin systematically lowers (increases) the spread when Green Bay, Miami, Pittsburgh, or Washington is the home (visiting) team, after accounting for their respective power rankings. PLAYOFF, WINSTREAK, and PREVMATCH enter with significantly positive coefficients. The estimated coefficient on PLAYOFF shows that the Outlaw Line increases (decreases) by almost two points when the home (visiting) team made the playoffs during the previous season. If the home (visiting) team won the previous matchup between the teams earlier in the 1972 season, the estimated coefficient on PREVMATCH predicts that Mr. Martin increases (decreases) the Outlaw Line by more

[25] Camerer (1989) and Brown and Sauer (1993) test for a hot-hand bias in point spreads for National Basketball Association games.

than a point.

It is important to realize that if Mr. Martin responds to sentiment variables in setting the Outlaw Line, this fact does not necessarily imply that the Outlaw Line is set inefficiently. The sentiment variables can reflect some "informed" sentiment. That is, these variables could contain information about the "fundamental" value of a team that is not accounted for by the power rankings alone. As emphasized by Avery and Chevalier (1999), we need to examine whether the sentiment variables are significantly related to betting outcomes to address pricing inefficiencies resulting from behavioral biases. We investigate this issue in Section 5.3.

The third column of Table 3 reports regression results when the difference between the Opening and Outlaw Lines serves as the regressand. LAG1 enters the model significantly. This result suggests that the group of 8–10 "special" agents that move the Outlaw Line to the Opening Line focuses on how the teams performed the previous week with respect to covering the spread. More specifically, these bettors drive the line up (down) slightly when the home (visiting) team covered the spread by more the previous week. PLAYOFF is nearly significant at conventional levels. Its estimated coefficient shows that the line increases (decreases) by approximately a quarter of a point when the home (visiting) team made the playoffs in 1971.

The power rankings and sentiment variables collectively explain 10% of the movements from the Outlaw to Opening Lines. It is not surprising that the R^2 statistic in the third column is smaller than that in the second column, because all of the regressors are based on information already available when the Outlaw Line is set. Movements in the Outlaw to Opening Lines presumably reflect differences in the collective views of the special agents from the views of Mr. Martin, as well as any additional information that becomes available in the limited time between the setting of the two early lines.

Results for the final linear regression model are reported in the fourth column of Table 3. In this model, the movement from the Opening to Closing Lines is the dependent variable. Power ranking is a significant determinant of the movement from the Opening to Closing Lines. The magnitude of the estimated coefficient, however, indicates that the public drives the line up only an additional 0.06 points for each point predicted by the power rankings. The public thus demonstrates a statistically significant, but economically minor, disagreement with Mr. Martin's and the special agents' use of the power rankings. Note that LAG2 is nearly significant at conventional levels in the fourth column of Table 3. The positive estimated coefficient on LAG2 reveals that the public believes that Mr. Martin and the special agents underestimate the performance of the home team in covering the spread two weeks ago.

The regressors jointly explain 19% of the movements from the Opening to Closing Lines. The fact that this number is smaller than the R^2 statistic in the second column is again not surprising, because this model uses stale information. Many of movements from the Opening to Closing Lines presumably reflect information that becomes available during the course of the betting week.

Taking the results in the second through fourth columns of Table 3 together, we have the following picture. The set of nine explanatory variables based on information available at the beginning of the betting week explain a substantial 74% of the variations in the Outlaw Line set by Mr. Martin at the start of the week. A group of four prestige teams, 1971 playoff teams, relative winning steak, the outcome of a previous matchup, and especially power ranking are all significant individual determinants of the setting of the Outlaw Line. When 8–10 special agents bet into the Outlaw Line to form the Opening Line, they significantly increase the betting line when the home team covered the spread the previous week by a wider margin than the visiting team. They also appear to drive up the betting line when the home team made the playoffs during the 1971 season. Given that this information is available at the start of the week, this movement reflects differences in the opinions of Mr. Martin and the special agents. Finally, when the wider public bets during the remainder of the week as the Closing Line is formed, it increases the betting line for home teams with higher power rankings—albeit by a limited amount—relative to the interpretation of the power rankings by Mr. Martin and the special agents. Compared to these agents, the public also seems to place greater emphasis on whether the home team covered the spread two weeks ago by a wider margin than the visiting team.

5.3 Ordered Logit Models for Betting Outcomes

In the spirit of Avery and Chevalier (1999), we estimate a series of ordered (or ordinal) logit models to analyze the efficiency implications of the linear regression results reported in the second through fourth columns of Table 3. The ordered logit model is used to analyze a qualitative dependent variable that has a natural ordering. In our study, the qualitative dependent variable reflects three outcomes: the visiting team covers the spread, a push, or the home team covers the spread. Treating the outcome as a qualitative variable is appropriate, because betting outcomes depend only on whether the actual point difference is below, equal to, or above the betting line. That is, betting outcomes do not depend on how far above (below) the actual point difference is in relation to the betting line when the home (visiting) team covers the spread.

The ordered logit model treats the log of the odds ratio as a linear function of observable variables. More specifically, let Y_n equal -1 if $S_n < X_n$ (visiting team covers the spread), 0 if $S_n = X_n$ (push), and 1 if $S_n > X_n$ (home team covers the spread), where S_n and X_n denote the actual point difference and a betting line, respectively. Given these three possible outcomes, the following pair of log odds ratios are defined:

$$\pi(Y_n = -1) = \log\{P(Y_n = -1)/[1 - P(Y_n = -1)]\} = \log[\exp(\delta_1 - \psi' x_n)] = \delta_1 - \psi' x_n, \quad (6)$$

$$\pi(Y_n = 0) = \log\{P(Y_n \leq 0)/[1 - P(Y_n \leq 0)]\} = \log[\exp(\delta_2 - \psi' x_n)] = \delta_2 - \psi' x_n, \quad (7)$$

where x_n is a 9×1 vector of observations for the power ranking and eight sentiment variables for game n, ψ is a 9×1 vector of slope coefficients, and δ_1 and δ_2 are thresholds.[26] The parameter vector, $(\psi', \delta_1, \delta_2)'$, can be estimated via maximum likelihood.

The fifth column of Table 3 reports ordered logit parameter estimates when X_n is the Outlaw Line. If a line is set efficiently, then any information available at the time the line is set should have no systematic effect on betting outcomes. Interpreting the second and fifth columns of Table 3 together, we can thus investigate whether Mr. Martin sets the Outlaw Line efficiently with respect to the power ranking and sentiment variables.

The results reported in the fifth column of Table 3 suggest that, with one exception, Mr. Martin efficiently adjusts the Outlaw Line in light of the power ranking and sentiment variables. As previously discussed, the results in the second column show that the Outline Line responds significantly to the power ranking, PRESTIGE4, PLAYOFF, WINSTREAK, and PREVMATCH. However, the coefficients on the power ranking, PLAYOFF, WINSTREAK, and PREVMATCH are all insignificant in the fifth column. This result implies that Mr. Martin efficiently incorporates relevant information from these variables when he sets the Outlaw Line. Note that while PLAYOFF, WINSTREAK, and PREVMATCH are classified as sentiment variables, Mr. Martin apparently captures the informed sentiment in these variables.

The significant explanatory variable in the fifth column of Table 3 is PRESTIGE4. The estimated coefficient is positive, meaning that the home (visiting) team is significantly more likely to cover the spread when the home (visiting) team is one of the four prestige teams. The Outlaw Line is thus set "too low" ("too high") when one of the four prestige teams is the home (visiting) team. As we noted in Section 5.1.2, Mr. Merchant relates that Mr. Martin viewed these four teams as especially popular with the public during the 1972 season.

[26]Corresponding to the log odds ratios, the probabilities for each outcome are given as follows: $P(Y_n = -1) = [1 + \exp(\psi' x_n - \delta_1)]^{-1}$, $P(Y_n = 0) = [1 + \exp(\psi' x_n - \delta_2)]^{-1} - [1 + \exp(\psi' x_n - \delta_1)]^{-1}$, $P(Y_n = 1) = 1 - [1 + \exp(\psi' x_n - \delta_2)]^{-1}$.

A potential explanation for the significant coefficient on PRESTIGE4 is that Mr. Martin anticipated that bettors would be "irrationally exuberant" (to borrow a phrase) about these teams. Therefore, he set the line at a level he believed would entice bettors to bet on these teams, thereby taking a position himself. This belief, however, appears to have been misplaced, creating an "undervaluation" of these teams and an inefficiency in the Outlaw Line. Nevertheless, this is the only pricing inefficiency apparent in the Outlaw Line, given the set of explanatory variables.

The sixth column of Table 3 contains the ordered logit estimation results when X_n is the Opening Line. Recall from the linear regression results that the only explanatory variable that significantly affects movements from the Outlaw to Opening Lines is LAG1, which *increases* the betting line. Interestingly, LAG1 is significantly *negative* in the sixth column of Table 3. This result indicates that the special agents forming the Opening Line move the Outlaw line in a way that creates an Opening Line pricing inefficiency. Home (visiting) teams that cover the spread by a relatively greater margin during the previous week are significantly less (more) likely to cover the spread this week. That is, by placing too much emphasis on recent performance, the special agents create a hot-hand, or momentum, bias in the Opening Line.

Like the logit results for the Outlaw Line in Table 3, PRESTIGE4 is significant when X_n is the Opening Line. In the linear regression results, PRESTIGE4 is not a significant determinant of movements from the Outlaw to Opening Lines. This result implies that the special agents fail to move the line in a way that corrects the pricing inefficiency relating to PRESTIGE4 in the Outlaw Line.

While pricing inefficiencies corresponding to PRESTIGE4 and LAG1 are evident in the Opening Line in Table 3, we must be mindful of the encompassing test results in Section 4. These results show that the Opening Line has a general information advantage over the Outlaw Line. The ordered logit results in the sixth column of Table 3 identify particular inefficiencies in the Opening Line, but there is also apparently some useful information utilized by the special agents that is ignored by, or not available to, Mr. Martin. This information could be available at the start of the betting week but not contained in our nine explanatory variables. Or, it could be information that becomes available in the relatively short window of time between the setting of the Outlaw and Opening Lines.

Ordered logit estimation results with X_n representing the Closing Line are reported in the final column of Table 3. Three variables enter significantly in this ordered logit model: PRESTIGE4, LAG1, and LAG2.[27] Recalling the results in the fourth column of Table 3, neither PRESTIGE4 nor LAG1 is a significant determinant of movements from the Opening to Closing Lines. The fact that these two variables are significant in the last column implies that the wider betting public fails to correct the inefficiencies relating to these variables in the Opening Line.

Furthermore, the significance of LAG2 in the last column indicates that the wider public introduces an additional pricing inefficiency. Recall from the fourth column that LAG2 is positively related (and "nearly" statistically significant) to movements from the Opening to Closing Lines. The significantly negative estimated coefficient on LAG2 in the last column implies that this line movement is inefficient. More specifically, the wider public places too much emphasis on the relative performance of the teams that covered the spread two weeks ago, another example of a hot-hand bias.[28]

We are again mindful of the encompassing test results from Section 4. Those results indicate that the Closing Line is superior in its overall information content to the Opening Line. Despite the apparent pricing inefficiencies not corrected—as well as created—by the wider public as they move the Opening Line to the Closing Line, some information that becomes available during the week relating to fundamental value appears to be utilized by the wider public.

Overall, the ordered logit results in the last three columns of Table 3 point to pricing inefficiencies that emerge during the betting week relating to three sentiment variables. Mr. Martin appears to discount the price of four teams that he believes are especially popular with the public during the 1972 season. In addition, 8–10 special agents and the wider public exhibit a hot-hand, or momentum, bias. That is, they place too much emphasis on the relative performance of teams in covering the spread during the previous two weeks.[29]

5.4 Profitability of Betting Strategies

We analyze the profitability of trading strategies suggested by the ordered logit model results shown in Table 3. We emphasize that we view this exercise as a check on the economic significance of the

[27] Avery and Chevalier (1999) also find that LAG1 is significantly related to betting outcomes for the Closing Line during the 1990–1994 NFL seasons.

[28] Recall that the results in the fourth column of Table 3 also show that power ranking is a significant determinant of movements from the Opening to Closing Lines. As discussed in Section 5.2, however, the estimated coefficient is economically small and thus does not create a pricing inefficiency according to the ordered logit results in the final column of Table 3.

[29] To test for potential home and favorite-team biases, we also estimated ordered logit models that include the home-favorite binary variable suggested by Dare and Holland (2004). This variable is not significant in any of the logit models, and its presence has little influences on the other estimated coefficients.

statistically significant results for PRESTIGE4, LAG1, and LAG2 in Table 3. This exercise focuses on whether straightforward betting strategies based on PRESTIGE4, LAG1, and LAG2 generate economically and statistically significant profits during the 1972 season. It does not determine whether these strategies could have been identified *a priori* and whether they are profitable during previous and subsequent seasons.

We analyze the economic and statistical significance of betting strategies using the method developed by Tryfos, Casey, Cook, Leger, and Pylypiak (1984). Let N_W, N_L, and N_O denote the number of wins, losses, and pushes, respectively, for a betting strategy that bets on B games during the season. Also, let $P_W = N_W/B$ and $P_L = N_L/B$. The average profit of a strategy is given by

$$\bar{r} = P_W - (1.1)P_L. \tag{8}$$

Also, define

$$S_{\bar{r}}^2 = (1/B)\{[P_W + (1.21)P_L] - [P_W - (1.1)P_L]^2\}. \tag{9}$$

Under a standard set of assumptions, $(\bar{r} - \mu)/S_{\bar{r}} \sim N(0,1)$ for large B, where μ is the true average profit corresponding to a particular strategy. Then, we can easily test the null hypothesis that a betting strategy is unprofitable ($\mu \leq 0$) against the alternative that it is profitable ($\mu > 0$) by comparing $\bar{r}/S_{\bar{r}}$ to the upper-tail critical values from the standard normal distribution. For each strategy considered, Table 4 reports the number of bets placed, number of wins, number of losses, proportion of wins, average profit (in percent), and Z-statistic (given by $\bar{r}/S_{\bar{r}}$).

Panel A of Table 4 reports results for a simple betting strategy based on PRESTIGE4. The ordered logit results in Table 3 indicate that the four prestige teams identified *a priori* by Mr. Martin are undervalued. In line with this, the strategy bets on the home (visiting) team when PRESTIGE4 equals 1 (-1). No bet is placed when PRESTIGE equals 0. The results in Panel A of Table 5 show that this strategy is highly profitable whether we bet into the Opening or Closing Line, with average profits of 41.85% and 38.15%, respectively. The corresponding Z-statistics are both significant at the 1% level.

LAG1 and LAG2 are the other significant sentiment variables in the ordered logit models in Table 3. The results point to a hot-hand bias, so that home (visiting) teams with a high (low) LAG1 or LAG2 are overvalued. We implement trading strategies based on LAG1 and LAG2 using thresholds. More specifically, we bet on the visiting team if LAG1 or LAG2 is greater than a prespecified positive threshold value. We bet on the home team if LAG1 or LAG2 is less than -1 times the threshold value. No bet is placed if the absolute value of LAG1 or LAG2 is less than the threshold. A higher threshold thus reduces the number of bets placed, meaning that a stronger potential mispricing signal is required to initiate a bet.

Panel B of Table 4 shows that trading strategies based on LAG1 are only slightly profitable or unprofitable when the threshold equals 0 or 5. This is true for betting into the Opening or Closing Line. The corresponding Z-statistics clearly indicate that the profitability of the strategies is insignificant. In contrast, trading strategies based on LAG1 are economically sizable and statistically significant when we use thresholds of 10 and 15. Increasing the threshold obviously reduces the number of bets placed, but the winning percentage increases by enough to make the strategies profitable. The average profit is 13.63% (17.14%) for a threshold of 10 when betting into the Opening (Closing) Line. It increases to 19.67% (21.50%) for a threshold of 15. The Z-statistics corresponding to LAG1 strategies for thresholds of 10 and 15 are statistically significant at conventional levels.

According to Panel C of Table 4, trading strategies based on LAG2 are only marginally profitable or unprofitable for all of the thresholds. The corresponding Z-statistics clearly show that the profitability of the strategies is not statistically significant. Overall, the results in Table 4 provide evidence of economically and statistically significant profits accruing to trading strategies based on two sentiment variables, PRESTIGE4 and LAG1 (for higher thresholds). This suggests that PRESTIGE4 and LAG1 are economically important in the ordered logit models in Section 5.3. While LAG2 is also statistically significant in the ordered logit models, it does not appear economically significant in Table 4.[30]

6 Performance of Identified Gamblers

In this section, we analyze the profitability of the betting strategies pursued by the five individual gamblers described in Section 3.2. Mr. Merchant views one of the gamblers—Fast Eddie—as a noise bettor. The other four gamblers—Mr. Rich, Greenwich Village Fats, Mr. Lem Banker, and

[30]Following Gandar, Zuber, O'Brien, and Russo (1988), we also analyzed trading strategies based on movements from the Opening to the Closing Lines. Using thresholds of 0, 0.5, 1.0, and 1.5, we considered "same-direction" strategies where where a bet is placed on the home (visiting) team if the Closing Line minus the Opening Line is greater (less) than the threshold (-1 times the threshold). Similarly, we considered "opposite-direction" strategies where where a bet is placed on the visiting (home) team if the Closing Line minus the Opening Line is greater (less) than the threshold (-1 times the threshold). None of these strategies generated economically or statistically significant profits. A same-direction strategy that bets into the Opening Line based on movements from the Outlaw Line to the Opening Line generated economically sizable profits of 13.97% and 20.54% for thresholds of 0.5 and 1.0, respectively. The corresponding Z-statistics were slightly above 0.10. In another betting market, Johnson, Jones, and Tang (2006) recently present evidence that movements in horse racing odds present potential profitable trading opportunities.

Mr. Merchant himself—are regarded by Mr. Merchant as informed bettors.[31]

Pankoff (1968) was the first to mention that semi-strong form efficiency for the NFL betting market could also be tested by examining the performance of individual bettors. However, neither Pankoff (1968) nor any other subsequent study of NFL betting market efficiency has had the requisite data on individual gamblers' betting activities to test for semi-strong form efficiency by testing for superior analysts. Our data on the complete bets made by these gamblers during the 1972 season allows us to test for the existence of superior analysts. If most of the gamblers identified *a priori* by Mr. Merchant as informed bettors realize economically and statistically significant profits during the 1972 season, while the noise bettor fails to earn such profits, this plausibly represents a rejection of semi-strong market efficiency for the 1972 NFL betting market.[32]

We have data on the nature of each bet placed during the 1972 season for each gambler, namely, whether the gambler bet on a particular game, and when he did, whether he bet on the home or visiting team. We analyze the profitability of bets made by individual gamblers again using the method in Tryfos, Casey, Cook, Leger, and Pylypiak (1984).[33] Results for the individual gamblers' strategies are reported in Table 5. Because Mr. Merchant does not indicate the precise line that each gambler bets into, we calculate results for betting into both the Opening and Closing Lines.

Table 5 shows that the number of bets placed ranges from 51 (Mr. Rich, 28% of all games) to 109 (Mr. Banker, 60% of all games). The results for Fast Eddie show that his strategy was unprofitable, whether betting into the Opening Line (average profit of -5.00%) or Closing Line (average profit of -13.50%). Before the start of the 1972 season, Mr. Merchant clearly identifies Fast Eddie as an uninformed bettor. Fast Eddie's performance during the season gives us little reason to doubt Mr. Merchant's assessment.

In contrast to Fast Eddie, the four gamblers identified by Mr. Merchant as informed bettors before the start of the 1972 season earn sizable average profits. Betting into either line, the winning percentages are all well above the 52.4% break-even level. Betting into the Opening Line, the

[31] It is less clear to what extent Mr. Merchant views Greenwich Village Fats as an expert bettor. First, Mr. Merchant reveals that Greenwich Village Fats was previously addicted to gambling. Second, as mentioned in Section 3.2, Mr. Merchant is primarily interested in Greenwich Village Fats bets during the season because of his potential use of information passed on to him by "The Mover."

[32] Two qualifications should be noted at this point. First, with respect to testing semi-strong form market efficiency, it is implicitly assumed that the bettors are utilizing publicly available information and are not privy to "insider" information. Second, insofar as information is costly to obtain and process, we would expect superior analysts to earn "abnormal" returns as compensation for these costs (Grossman and Stiglitz, 1980); while, strictly speaking, the existence of superior analysts represents a rejection of semi-strong form market efficiency, it does not necessarily contradict efficiency from this broader perspective.

[33] We do not have data on the amount wagered for each game. The Tryfos et al. (1984) method implicitly assumes that the same amount is wagered for each game.

average profit for the four informed traders ranges from 11.01% for Greenwich Village Fats to 25.08% for Mr. Merchant. Betting into the Closing Line, the average profit ranges from 9.75% for Greenwich Village Fats to 21.76% for Mr. Rich. The Z-statistics indicate that the average profits earned by Mr. Rich, Mr. Banker, and Mr. Merchant are all significant at the 10% level or better betting into either line.

In our view, the fact that all four of these bettors, identified *a priori* as informed, earn economically sizable profits provides strong evidence for the existence of superior analysts in the 1972 NFL betting market. This fact, coupled with the statistical significance of three of the bettors' profits, makes it difficult to attribute the outcome to chance alone. Our results appear to differ from those in Levitt (2004), who fails to find significant evidence of superior analysts among bettors in an online handicapping contest covering the 2001 NFL season. These contrasting results could be due to differences in the size and scope of the betting markets analyzed in the present paper and Levitt (2004).

As a final exercise, we estimate ordered logit models for the nature of the bets placed by each gambler based on the same set of explanatory variables from Section 5. The qualitative dependent variable equals 1 if the gambler bets on the home team to cover the spread, 0 if no bet is placed, and -1 if the gambler bets on the visiting team to cover the spread. The results are reported in Table 6.

The nature of the bets for Fast Eddie and Mr. Rich only responds significantly to one of the explanatory variables (power ranking for Fast Eddie and NYTEAMS for Mr. Rich). The results in Section 5 indicate that neither of these variables are related to pricing inefficiencies in the betting lines. This fact means that these gamblers are not significantly exploiting the pricing inefficiencies documented above. This result is not surprising in the case of Fast Eddie, a likely noise bettor. However, because Mr. Rich earns sizable profits and is identified as an informed bettor, he appears to have superior information about fundamental value and/or recognizes inefficiencies stemming from sentiment not captured in our set of explanatory variables.

The fourth and fifth columns of Table 6 show that the nature of the bets for Greenwich Village Fats and Mr. Banker responds significantly to at least two of the three explanatory variables associated with pricing inefficiencies documented in Section 5. Interestingly, both Greenwich Village Fats and Mr. Banker are significantly more likely to bet on the visiting team to cover the spread as LAG1 increases. This finding suggests that they are attempting to exploit the pricing inefficiency associated with LAG1 in the final two columns of Table 3. Mr. Banker also responds significantly to LAG2 and in a manner consistent with exploiting the pricing inefficiency corresponding to LAG2 in Table 3.

Both Greenwich Village Fats and Mr. Banker are significantly less likely to bet on the home (visiting) team when the home (visiting) team is one of the four prestige teams. In contrast to their responses to LAG1 and LAG2, their responses to PRESTIGE4 reinforce the pricing inefficiency corresponding to PRESTIGE4 in Table 3. Presumably, these gamblers would have fared better if they ignored this variable or bet in the opposite direction. Nevertheless, Greenwich Village Fats and (especially) Mr. Banker earn economically significant profits during the 1972 season. They apparently took advantage of the mispricings associated with LAG1 and LAG2 (and possibly other sentiment variables not included in our models) and/or utilized superior information concerning fundamental value to earn substantial profits.

The last column of Table 6 indicates that the nature of the bets for Mr. Merchant does not respond significantly to any of the power ranking or sentiment variables. Thus, the sizable profits he earns apparently result from exploiting pricing inefficiencies corresponding to sentiment variables not included in our set of explanatory variables and/or superior information pertaining to fundamental value. Mr. Merchant does not provide a detailed model of his betting patterns. However, the sense one gets from reading his book is that Mr. Merchant does not rely on sentiment. Rather, based on his his experience as a player and sportswriter, Mr. Merchant strongly believes that he "knows football." Accordingly, he appears to rely extensively on analyses of corresponding strengths and weaknesses of the competing teams.[34]

7 Conclusion

The NFL betting market has been a popular laboratory for testing market efficiency because the range of payoffs is known with certainty in advance and payoffs occur soon after betting ends. We extend the NFL betting market literature using a unique data set culled from Mr. Larry Merchant's book, *The National Football Lottery*. This unique data set provides observations on three sequential betting lines for the 1972 NFL season: (i) an Outlaw Line set by a single agent, Mr. Bob Martin, at the start of the week; (ii) Tuesday's Opening Line shaped by bets from a group of 8–10 agents; and (iii) a game-time Closing Line formed by subsequent bets placed by the wider gambling public.

Nearly all studies of the NFL betting market focus on the Closing Line. However, Avery

[34]For example, while a particular team has a relatively good pass defense, it can be weak against the run. If its opponent this week has a good running back and solid offensive line, Mr. Merchant would overweight this information in his analysis.

and Chevalier (1999) and Gandar, Zuber, O'Brien, and Russo (1988) study both the Opening and Closing Lines. Our study is the first to analyze the Outlaw Line in conjunction with the Opening and Closing Lines. This fact allows us to examine the intra-week evolution of market prices from initial prices set by a single agent to final prices influenced by the wider betting public, comprised of recreational and professional bettors.

In addition, the data we use are from bookmakers standing ready to take either side of the bet. Consequently, the betting lines in our data set are true market prices. Most betting line data published in newspapers and other sources purport to be "representative" betting lines. That is, it is possible that these published betting lines are not necessarily lines into which bettors can bet.

We analyze the efficiency implications of the intra-week evolution of betting lines using forecast encompassing tests, linear regression models of line setting based on power rankings and sentiment variables, and ordered logit models of betting outcomes. Forecast encompassing tests, which we introduce to the NFL betting market literature, point to an increase in the general information content from the Outlaw Line to the Opening Line and to the Closing Line. This pattern is consistent with a degree of efficient information processing in the 1972 NFL betting market.

Our linear regression and ordered logit model results, however, identify significant pricing inefficiencies associated with prestige and momentum. More specifically, in setting the Outlaw Line, Mr. Martin undervalued four prestige teams that he viewed as especially popular with the public. This undervaluation is not corrected in the setting of the Opening and Closing Lines during the remainder of the betting week. The group of 8–10 agents who shape the Outlaw Line into the Opening Line overvalued teams that covered the spread by wider margins the previous week. This overvaluation persisted into the Closing Line. Furthermore, in influencing the Closing Line, the wider betting public introduced an additional overvaluation of teams that covered the spread two weeks ago. A number of straightforward betting strategies based on these inefficiencies also appear to generate economically and statistically meaningful profits, according to the method in Tryfos, Casey, Cook, Leger, and Pylypiak (1984).

Our data set also allows us to test for the existence of superior analysts. Mr. Merchant provides information on the complete set of bets for the 1972 season made by five gamblers, four of which he clearly identifies *a priori* as informed bettors (including himself), as well as another he regards as an uninformed bettor. Interestingly, using the method in Tryfos, Casey, Cook, Leger, and Pylypiak (1984), we find that all four of these informed bettors earned an economically sizable average profit, three of which are statistically significant. In contrast, the uninformed bettor fared poorly.

These results can be interpreted as evidence supporting the existence of superior analysts in the 1972 NFL betting market.

Our results provide a nuanced view of the 1972 NFL betting market. As more agents participate in the market and the information set grows during the betting week, the information content of prices increases. Furthermore, the single agent who sets prices at the start of the betting week does so in a manner that precludes profitable betting opportunities relating to power rankings and a number of sentiment variables. These results are consistent with a reasonable degree of efficiency in the price-setting process. Nevertheless, the Outlaw Line undervalued a group of four prestige teams, and pricing inefficiencies relating to hot-hand, or momentum, biases emerged during the course of the betting week. In addition, the existence of superior analysts suggests that all relevant information is not optimally reflected in market prices.

The most pressing question raised by our paper is the extent to which our results are characteristic of betting markets more generally. The 1972 NFL betting market in New York City appears sizable, suggesting that our results are not an artifact of a relatively inactive market with thin trading and are thus applicable to large, active markets. Nevertheless, it would be highly informative to perform similar analyses for other betting markets. Of course, the major problem is data availability. We focus on the betting market for the 1972 NFL season, as it is the only season for which data are available in *The National Football Lottery*. Unfortunately, similar rich data sets are not available for other seasons. Insofar as comparable data become available for other NFL seasons, we can better determine the generality of our results. This analysis is especially interesting in light of the legalization of gambling in certain parts of the country since 1972, as well as the advent and tremendous growth of internet gambling. Our extensive analysis of the Outlaw, Opening, and Closing Lines and individual gambler performance for the 1972 season provides a benchmark for analyzing similar data and price setting in other NFL seasons when data availability permits.

References

Avery, C. and J. Chevalier (1999). “Identifying Investor Sentiment from Price Paths: The Case of Football Betting.” *Journal of Business* 72, 493–521.

Barberis, N.C. and R.H. Thaler (2003). “A Survey of Behavioral Finance.” In G. Constantinides, R. Stulz, and M. Harris (Eds.), *Handbook of the Economics of Finance*, Amsterdam: North Holland.

Boulier, B.L., H.O. Stekler, and S. Amundson (2006). “Testing the Efficiency of the National Football League Betting Market.” *Applied Economics* 38, 279–284.

Bradley, D.J. and B.D. Jordan (2002). “Partial Adjustment to Public Information and IPO Underpricing.” *Journal of Financial and Quantitative Analysis* 37, 595–615.

Brown, W.O. and R.D. Sauer (1993). “Does the Basketball Market Believe in the Hot Hand? Comment.” *American Economic Review* 83, 1193–1209.

Camerer, C. (1989). “Does the Basketball Market Believe in the ‘Hot Hand’?” *American Economic Review* 79, 1257–1261.

Chong, Y.Y. and D.F. Hendry (1986). “Econometric Evaluation of Linear Macro-Econometric Models.” *Review of Economic Studies* 53, 671–690.

Dana, J.D. and M.M. Knetter (1994). “Learning and Efficiency in a Gambling Market.” *Management Science* 40, 1317–1328.

Dare, W.H. and S.S. MacDonald (1996). “A Generalized Model for Testing the Home and Favorite Team Advantage in Point Spread Markets.” *Journal of Financial Economics* 40, 295–318.

Davies, R.O. and R.G. Abram (2001). *Betting the Line: Sports Wagering in American Life*. Columbus, Ohio: The Ohio State University Press.

Diebold, F.X. and R. Mariano (1995). “Comparing Predictive Accuracy.” *Journal of Business and Economic Statistics* 13, 256–265.

Fair, R.C. and R.J. Shiller (1990). “Comparing Information in Forecasts from Econometric Models.” *American Economic Review* 80, 375–389.

Fama, E. (1965). “The Behavior of Stock Market Prices.” *Journal of Business* 38, 34–105.

Gandar, J.M., R. Zuber, T. OBrien, and B. Russo (1988). “Testing Rationality in the Point Spread Betting Market.” *Journal of Finance* 43, 995–1007.

Golec, J. and M. Tamarkin (1991). “The Degree of Inefficiency in the Football Betting Market.” *Journal of Financial Economics* 30, 311–323.

Gray, P.K. and S.F. Gray (1997). “Testing Market Efficiency: Evidence from the NFL Sports

Betting Market." *Journal of Finance* 52, 1725–1737.

Grossman, S. and J.E. Stiglitz (1980). "On the Impossibility of Informationally Efficient Markets." *American Economic Review* 70, 123–136.

Harvey, D.I., S.J. Leybourne, and P. Newbold (1998). "Tests for Forecast Encompassing." *Journal of Business and Economic Statistics* 16, 254–259.

Hirshleifer, D.A. (2001). "Investor Psychology and Asset Pricing." *Journal of Finance* 56, 1533–1598.

Humphreys, B.R., R.J. Paul, and A. Weinbach (2009). "Understanding Price Movements in Point Spread Betting Markets: Evidence from NCAA Basketball." Manuscript, University of Alberta.

Johnson, J.E.V., O. Jones, and L. Tang (2006). "Exploring Decision Makers' Use of Price Information in a Speculative Market." *Management Science* 52, 897–908.

Levitt, S.D. (2004). "Why are Gambling Markets Organized So Differently from Financial Markets?" *Economic Journal* 114, 223–246.

Merchant, L. (1973). *The National Football Lottery*. New York: Holt, Rinehart, and Winston.

Mincer, J. and V. Zarnowitz (1969). "The Evaluation of Economic Forecasts." In J. Mincer (Ed.), *Economic Forecasts and Expectations*, New York: National Bureau of Economic Research.

Pankoff, L.D. (1968). "Market Efficiency and Football Betting." *Journal of Business* 41, 203–214.

Samuelson, P. (1965). "Proof That Properly Anticipated Prices Fluctuate Randomly." *Industrial Management Review* 6, 41–49.

Sauer, R.D. (1998). "The Economics of Wagering Markets." *Journal of Economic Literature* 36, 2021–2064.

Sauer, R.D., V. Brajer, S.P. Ferris, and M.W. Marr (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.

Shleifer, A. (2000). *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford: Oxford University Press.

Tryfos, P., S. Casey, S. Cook, G. Leger, and B. Pylypiak (1984). "The Profitability of Wagering on NFL Games." *Management Science* 30, 123–132.

Vergin, R.C. and M. Scriabin (1978). "Winning Strategies for Wagering on National Football League Games." *Management Science* 24, 809–818.

Zuber, R., J.M. Gandar, and B. Bowers (1985). "Beating the Spread: Testing the Efficiency of the Gambling Market for National Football League Games." *Journal of Political Economy* 93, 800–806.

Table 1
Summary statistics, 1972 NFL season

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Actual point difference	Outlaw Line	Opening Line	Closing Line	|Opening – Outlaw|	|Closing – Opening|
Mean	1.23	1.93	1.75	1.47	0.80	0.90
Standard deviation	16.57	8.05	8.34	8.91	0.85	1.16
Minimum	−44	−17.0	−17.0	−16.0	0	0
Maximum	52	18.0	18.0	19.0	4.5	12.5

Notes: The number of available observations is 182 for each variable for the 1972 NFL season.

Table 2
Estimated combining weights in optimal composite forecasts and Harvey, Leybourne, and Newbold (1998) test results, 1972 NFL season

(1)	(2)	(3)	(4)
	X_j		
X_i	Outlaw Betting Line	Opening Betting Line	Closing Betting Line
Outlaw Betting Line		1.21 [0.08†]	1.44 [0.03*]
Opening Betting Line	−0.21 [0.60]		1.78 [0.08†]
Closing Betting Line	−0.44 [0.76]	−0.78 [0.80]	

Notes: The table reports estimates of λ for the regression model, $S_n = (1-\lambda)X_{i,n} + \lambda X_{j,n} + \varepsilon_n$. Statistical significance is based on the modified Harvey, Leybourne, and Newbold (1998) statistic given in (4) in the text, which corresponds to a test of the null hypothesis that X_i encompasses X_j ($\lambda = 0$) against the (one-sided, upper-tail) alternative hypothesis that X_i does not encompass X_j ($\lambda > 0$); p-value is given in brackets. †, *, and ** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 3
Linear and ordered logit regression results based on power rankings and sentiment variables, 1972 NFL season

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Regressand:	Outlaw Line	Opening Line − Outlaw Line	Closing Line − Opening Line	Outcome, bet into Outlaw Line	Outcome, bet into Opening Line	Outcome, bet into Closing Line
Regressor	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)
Power ranking	0.506 (11.91**)	0.001 (0.10)	0.065 (4.83**)	−0.010 (−0.46)	−0.013 (−0.61)	−0.017 (−0.80)
PRESTIGE4	−1.517 (−2.16*)	0.201 (1.06)	−0.140 (−0.63)	1.402 (3.59**)	1.371 (3.55**)	1.268 (3.39**)
PLAYOFF	1.866 (3.28**)	0.245 (1.60)	0.157 (0.87)	0.453 (1.55)	0.459 (1.57)	0.256 (0.90)
LASTPLACE	−0.368 (−0.65)	0.058 (0.38)	0.201 (1.12)	0.382 (1.35)	0.352 (1.23)	0.309 (1.10)
NYTEAMS	0.782 (0.87)	−0.117 (−0.49)	0.033 (0.12)	0.230 (0.51)	0.269 (0.60)	0.290 (0.66)
WINSTREAK	0.366 (3.61**)	0.002 (0.08)	−0.023 (−0.73)	−0.057 (−1.11)	−0.037 (−0.75)	−0.017 (−0.36)
LAG1	0.008 (0.36)	0.015 (2.70**)	−0.008 (−1.16)	−0.016 (−1.49)	−0.019 (−1.78†)	−0.022 (−2.06*)
LAG2	0.008 (0.46)	−0.004 (−0.80)	0.009 (1.58)	−0.012 (−1.32)	−0.013 (−1.47)	−0.016 (−1.74†)
PREVMATCH	1.212 (1.74†)	0.008 (0.04)	−0.178 (−0.81)	−0.123 (−0.35)	−0.167 (−0.47)	−0.075 (−0.22)
Constant(1)	0.679 (2.04*)	−0.169 (−1.89†)	−0.465 (−4.42**)	0.161 (0.96)	0.120 (0.71)	−0.029 (−0.18)
Constant(2)				0.237 (1.40)	0.221 (1.31)	0.144 (0.85)
R^2	0.74	0.10	0.19			

Notes: The results in the second through fourth columns are for linear regression models with the regressands given at the top of the table. The results in the fifth through seven columns are for ordered logit models where the qualitative dependent variable represents the outcome of the game relative to the betting line given at the top of the table; the dependent variable equals −1 (0, 1) if the actual point difference is less than (equal to, greater than) the betting line. The regressors given in the first column are defined in the text. †, *, and ** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 4
Profitability of trading strategies, 1972 NFL season

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Bet into Opening Line						Bet into Closing Line					
Threshold	Bets	Wins	Losses	Wins/ (Wins+Losses)	Average profit (%)	Z-statistic	Bets	Wins	Losses	Wins/ (Wins+Losses)	Average profit (%)	Z-statistic
A. Trading strategy based on PRESTIGE4												
PRESTIGE4$\neq 0$	54	38	14	0.73	41.85	3.35 [0.00**]	54	36	14	0.72	38.15	3.07 [0.00**]
B. Trading strategy based on LAG1												
$\lvert$LAG1$\rvert > 0$	168	85	79	0.52	−1.13	−0.14 [0.56]	168	84	77	0.52	−0.42	−0.05 [0.52]
$\lvert$LAG1$\rvert > 5$	130	68	58	0.54	3.23	0.36 [0.36]	130	68	57	0.54	4.08	0.45 [0.33]
$\lvert$LAG1$\rvert > 10$	91	52	36	0.59	13.63	1.28 [0.10†]	91	53	34	0.61	17.14	1.63 [0.05*]
$\lvert$LAG1$\rvert > 15$	60	36	22	0.62	19.67	1.52 [0.06†]	60	36	21	0.63	21.50	1.68 [0.05*]
C. Trading strategy based on LAG2												
$\lvert$LAG2$\rvert > 0$	152	76	73	0.51	−2.83	−0.34 [0.63]	152	77	68	0.53	1.45	0.17 [0.43]
$\lvert$LAG2$\rvert > 5$	125	62	60	0.51	−3.20	−0.34 [0.63]	125	62	57	0.52	−0.56	−0.06 [0.52]
$\lvert$LAG2$\rvert > 10$	90	47	41	0.53	2.11	0.19 [0.42]	90	48	40	0.55	4.44	0.41 [0.34]
$\lvert$LAG1$\rvert > 15$	69	37	31	0.54	4.20	0.34 [0.37]	69	38	30	0.56	7.25	0.58 [0.28]

Notes: The strategy in Panel A bets on the home (visiting) team if the PRESTIGE variable takes a value of 1 (−1). The strategy in Panel B bets on the visiting team if the LAG1 variable is greater than the positive threshold value indicated in the first column; the strategy bets on the home team if the LAG1 variable is less than −1 times the threshold value. The strategy is Panel C is the same as that in Panel B except that it is based on the LAG2 variable. Z-statistic is the Tryfos, Casey, Cook, Leger, and Pylypiak (1984) statistic corresponding to the null hypothesis that the trading strategy is unprofitable against the (one-sided, upper-tail) alternative hypothesis that it is profitable; p-value is given in brackets. †, *, and ** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 5
Profitability of professional gamblers' trading strategies, 1972 NFL season

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Bet into Opening Line						Bet into Closing Line					
Gambler	Bets	Wins	Losses	Wins/ (Wins+Losses)	Average profit (%)	Z-statistic	Bets	Wins	Losses	Wins/ (Wins+Losses)	Average profit (%)	Z-statistic
Fast Eddie	60	30	30	0.50	−5.00	−0.37 [0.64]	60	26	31	0.46	−13.50	−1.03 [0.85]
Mr. Rich	51	32	19	0.63	21.76	1.53 [0.06†]	51	32	19	0.63	21.76	1.53 [0.06†]
Green. Vill. Fats	79	45	33	0.58	11.01	0.95 [0.17]	79	44	33	0.57	9.75	0.84 [0.20]
Mr. Lem Banker	109	62	44	0.58	12.48	1.28 [0.10†]	109	61	42	0.59	13.58	1.41 [0.08†]
Mr. Larry Merchant	63	40	22	0.65	25.08	2.00 [0.02*]	63	38	23	0.62	20.16	1.60 [0.06†]

Notes: Z-statistic is the Tryfos, Casey, Cook, Leger, and Pylypiak (1984) statistic corresponding to the null hypothesis that the trading strategy is unprofitable against the (one-sided, upper-tail) alternative hypothesis that it is profitable; p-value is given in brackets. †, *, and ** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 6
Ordered logit regression results for professional gamblers based on power rankings and sentiment variables, 1972 NFL season

(1)	(2)	(3)	(4)	(5)	(6)
	Fast Eddie	Mr. Rich	Greenwich Village Fats	Mr. Lem Banker	Mr. Larry Merchant
Regressor	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)
Power ranking	0.045 (2.11*)	0.004 (0.20)	0.063 (3.07**)	−0.004 (−0.21)	−0.004 (−0.21)
PRESTIGE4	0.170 (0.47)	0.393 (1.05)	−0.560 (−1.65†)	−0.902 (−2.63**)	0.544 (1.52)
PLAYOFF	0.223 (0.76)	−0.134 (−0.44)	−0.479 (−1.75†)	0.256 (0.97)	0.433 (1.52)
LASTPLACE	0.083 (0.29)	−0.208 (−0.70)	0.292 (1.10)	−0.197 (−0.78)	−0.142 (−0.50)
NYTEAMS	−0.154 (−0.31)	1.114 (2.25*)	−0.735 (−1.75†)	0.486 (1.15)	−0.738 (−1.61)
WINSTREAK	0.005 (0.10)	−0.014 (−0.27)	0.042 (0.88)	0.023 (0.45)	−0.014 (−0.29)
LAG1	0.015 (1.41)	0.008 (0.72)	−0.024 (−2.39**)	−0.023 (−2.28*)	−0.007 (−0.63)
LAG2	−0.008 (−0.87)	−0.006 (−0.67)	−0.002 (−0.19)	−0.030 (−3.56**)	−0.007 (−0.75)
PREVMATCH	0.062 (0.18)	−0.536 (−1.44)	−0.127 (−0.40)	−0.326 (−1.02)	0.120 (0.35)
Constant(1)	−1.388 (−6.91**)	−1.733 (−7.92**)	−1.389 (−7.00**)	−1.302 (−6.68)	−1.093 (−5.88**)
Constant(2)	2.221 (8.79**)	2.132 (8.61**)	1.389 (7.03**)	0.694 (3.94)	2.620 (8.90**)

Notes: The table report results for ordered logit models where the qualitative dependent variable equals −1 (0, 1) if the gambler given at the top of the table bets on the visiting (no, home) team. The regressors given in the first column are defined in the text. †, *, and ** indicate significance at the 10%, 5%, and 1% levels, respectively.