Does the Hot Hand Drive the Market? Evidence from Betting Markets

Michael Sinkey*and Trevon Logan†

December 2011

Abstract

This paper investigates how market makers respond to behavioral strategies and the implications of these responses for market efficiency. In particular, we ask whether market makers rationally price out certain strategies at the expense of leaving other strategies profitable, resulting in potential market inefficiency. We answer this question by testing for betting market efficiency in an amateur sport, American college football, using data from over 11,000 games from 1985 to 2003. We find that the market is inefficient; favorites are statistically overpriced while home teams are statistically underpriced. We show that the magnitude of this bias is large enough to generate both economic and statistical inefficiency in this betting market. Furthermore, we provide suggestive evidence for the cause of this inefficiency: betting houses deliberately inflate the betting lines in order to counteract bettors' "hot hand" beliefs. While eliminating the "hot hand" bias is efficient for a betting house, tempering the "hot hand" results in consistently profitable simple betting strategies.

JEL Codes: D03, D22, D40, D83, G10, G12

Key Words: behavioral finance, market efficiency, "hot hand"

*Corresponding Author. Department of Economics, University of West Georgia, 1601 N. Maple Street, Carrollton, GA 30117 e-mail: msinkey@westga.edu

†Ohio State University and NBER, e-mail: logan.155@osu.edu

Andrew Weinbach graciously shared data sources. Nathan Fritz-Joseph , Kyle Kain, Michael Kovach, Justin Sawyer, and Christopher Zukauckas provided excellent research assistance. We thank participants in the OSU Micro Workshop, Benjamin Anderson, Rodney Andrews, Vicki Bogan, Lisa D. Cook, John Gray, P.J. Healy, Pok-Sang Lam, Peter McGee, Jim Peck, Brandon Restrepo, Brian Roe, and Bruce Weinberg for helpful comments and conversations. The usual disclaimer applies.

1 Introduction

How do prediction markets function when responding to behavioral strategies? We study the behavior of a prediction market–the American college football betting market–where a market maker, the betting house, sets the price for the entire market. This paper examines how a market maker prices behavioral strategies when there may be several different sources of bias. In a betting market, the market maker has one choice variable, the market price, with which to control an entire array of betting strategies. How does a market maker price different available strategies when she is not able to price discriminate? Does balancing leave open opportunities for bettors to profit, and are those profits large and consistent enough to cause market inefficiency?

In any betting market bettors have a number of different strategies available, but how betting houses balance these strategies in their pricing is unknown. Betting houses would presumably minimize their exposure to the most prevalent forms of bias, but this could open them to exposure to less prevalent forms that bettors could exploit, which mirrors more familiar "contrarian strategies" (Grinblatt and Keloharju 2000 [19], Lakonishok et al. 1994 [25]). This usually leaves us in the situation of searching for sophisticated strategies that a handful of investors could exploit for profit (White 2000 [37], Sullivan et al. 1999 [35], Conrad et al. 2003 [9]).

We examine the pricing of a narrow set of simple, transparent betting strategies that unsophisticated bettors could adopt (such as always betting on the home team) to ask how betting houses minimize their exposure to prevalent forms of bias. While one could argue that a small number of especially sophisticated bettors could always strategize for profit, we concentrate here on strategies that an unsophisticated lay bettor could use. Presumably, these are the strategies that the betting house is most sensitive to since the majority of bettors are unsophisticated.

We use data from the universe of college football betting outcomes from 1985 to 2003, over 11,000 games, to test for market efficiency in the college football betting. Specifically, our direct test for market efficiency estimates whether or not bettors can consistently make profit by betting clearly identifiable, simple strategies. We find robust evidence that, for some of these strategies, the market is inefficient, in contrast with other work on college football, including Dare and McDonald 1996 [10] and Fair and Oster 2007. [14] In particular, we find that favorites are consistently overpriced. We estimate that, for example, a bet of $1,000 against a favored team in prominent games would yield $1,117 in expectation, a gross return of nearly twelve percent and a net return of two percent after accounting for the transaction cost associated with placing the bet.

More importantly, we analyze how this betting market functions to understand the source of this inefficiency. The key piece of information used in a betting market is the point spread, also known as the betting line, which is the predicted margin of victory in a given game. Bettors place bets that a team will "beat the spread" (exceed the predicted margin of victory) or not. Existing analysis of market function postulates that previous performance against the line should not be predictive of future performance if the market is efficient (Fama 1991, among others). We find that betting lines are not independent from game to game. Betting lines have memory– we present robust evidence that they are functions of previous betting market results. In particular, we find that betting lines are systematically greater for teams who beat the betting line the previous week. We also find that the magnitude of these increases is significantly greater in this market than in other betting markets, such as professional basketball [5] and professional football [6]. Given this serial correlation, we next test to see whether or not teams that exceed the betting line are likely to do so in the following week. We find that teams who exceed the betting line in one week are no more likely to do so in the following week. Thus bettors who believe that teams are more likely to exceed the betting line in subsequent weeks believe so erroneously, as "hot" teams are priced efficiently. We find no evidence that would suggest that a bettor could examine past performance against the spread and construct a strategy based on this data to make profit.

We then ask why some strategies (like the hot hand) are priced efficiently while others are priced inefficiently. As a rule, betting houses are particularly sensitive to any bias among bettors-their profit motive is to have an equal amount of money on either side of the bet to minimize their exposure to the risk of more bets being placed on a winning bet than on a losing bet. Our analysis suggests that betting houses are particularly sensitive to potential "hot hand" bias among bettors. If a sizable number of bettors believe in the hot hand-specifically, that betting on teams based on recent performance against the spread is profitable-the profit motive of betting houses will cause them to eliminate any possible profitability associated with this strategy by increasing the betting lines of "hot" teams in order to avoid this risk, leaving the betting houses vulnerable to other possible betting strategies. We support our conclusion with both qualitative and empirical evidence. Qualitatively, we use narrative evidence from a variety of sources to document how previous performance against the betting line is commonly used by bettors to predict current performance.

We find strong evidence that links the inefficient pricing found for a few simple strategies to the efficient pricing found for "hot" teams. Empirically, we show how the mispricing of games varies by whether or not they contain "hot" teams. One consequence of the adjustment for the "hot hand" is that it makes other conditional strategies, such as betting on home underdogs, consistently profitable. Both of these pieces of evidence lend credence to our conclusion that counteracting the "hot hand" creates profitable strategies in this betting market. As such, our study expands the market efficiency literature by providing evidence suggesting that market makers intentionally leave some strategies mispriced to account for other strategies, yielding a potentially rational explanation for observed inefficiencies.

2 The College Football Betting Market

In this paper we examine the pricing of behavioral strategies by using data from the betting market for American college football. Sports betting markets are well-suited for examining the pricing of different strategies since the predicted event will be realized with certainty and strategies are fairly transparent. For this reason, researchers have focused on sports betting markets as fertile ground for tests of market efficiency (Sauer et al. 1988 [34], Zuber et al. 1985 [39], Camerer 1989 [7], Woodland and Woodland 1994 [38], Gil and Levitt 2007 [16]).The existing literature on the efficiency of betting markets has focused on professional sports (Gray and Gray 1997 [18], Sauer et al. 1988 [34], Gandar et. al 1988 [15], Zuber et al. 1985 [39], Woodland and Woodland 1994 [38], Brown and Sauer 1993 [6], Bröder and Scheibehenne, 2007 [4]). Relatively few studies have focused on amateur sports which, despite their amateur status, have large betting markets.[1] More importantly, we expand these previous studies by examining why market inefficiency might occur. We examine how profitable strategies may arise as a consequence of pricing other strategies with the same choice variable, the betting line. Unlike many previous works, we do not treat the profits of strategies as being determined independently from each other, since betting houses use the same betting line for all strategies.

Among sports betting markets, amateur betting markets share characteristics with professional betting markets, but also contain economically interesting differences. College football games have different properties from professional football games. College coaches are not required to provide weekly injury reports; whether or not key players are playing may be obscured from both the opponent and the betting market. College football teams also have new players enter and exit more frequently than the NFL. It may be more difficult to discern the quality of a team with younger (freshman or sophomore) players because of few observations. Even if the status of key players is known, their substitutes are more likely to be unknown quantities. There may also be fewer sources of public information in college sports; while every professional sports team receives substantial newspaper and television coverage, relatively few college programs are subject to such intense scrutiny. Indeed, many major college sports powers are outside of major media markets, which can act to limit the availability of information about teams [2].

Despite the size and economically interesting features of amateur betting markets, little is known about their economic properties. There are few existing studies of the college football betting market, and the majority have suggested that college football betting is efficient (Dare and McDonald 1996 [10], Golec and Tamarkin 1991 [17], and Fair and Oster 2007[14]). Only Paul et. al 2003 [31] have found any evidence for inefficiency. A drawback to all of these existing studies of college football is that they employ methodological approaches which are indirect tests of market efficiency, as we describe below.

Sports betting markets are particularly tractable environments with which to examine how market makers price behavioral strategies. They hinge on a relatively simple metric, the betting

[1] For example, it is estimated that as much as $12 billion is bet on the NCAA Final Four Basketball tournament each Spring (Matuszewki 2009 [29]).

[2] For example, every single NFL (National Football League) city has all four major TV affiliates and a daily newspaper with their own dedicated sports reporters, and many cities have multiple newspapers.

line, which is the estimate of how large the margin of victory will be in a given sporting event. Bettors simply place wagers on whether or not a given team will win (lose) by more (less) than the predicted margin of victory (the betting line). In order to win, a bettor need only be on the correct side of the betting line. That is, bettors who bet on the favorite win if the team predicted to win wins by a margin greater than the predicted margin of victory, and bettors who bet on the underdog win if the team predicted to lose loses by less than the predicted margin of victory. Sports betting markets are also quite large and many bettors are repeat participants- if a consistently profitable strategy were available bettors are likely to exploit it. A priori, there should be limited room for expected profit in sports betting markets. Most importantly, outcomes are observed, and we can directly assess whether or not a given bet or betting strategy was profitable.

Betting houses facilitate the betting process in both professional and amateur sports in the same way, by setting the betting line that bettors wager on. They derive their profit by taking a fixed percentage of all bets placed, known as the vigorish. Intuitively, betting houses do not risk losing money if exactly half of the total amount of money bet is on one side of the betting line and the other half on the other side, regardless of the outcome. The money that is bet on the losing side of the outcome is used by to pay off the winning wagers. If this does not occur, the betting house incurs some risk.

Although betting houses could profit handsomely if there were more losing bets than winning bets, such a strategy would involve substantial risk to the betting house.[3] For this reason, betting houses are primarily interested in setting betting lines that will guarantee equal betting on either side of the betting line [4] , and the betting literature usually assumes this to be the betting house's goal (see Gray and Gray 1997 [18] and others for use of this assumption). The betting line can be better understood as the wager-weighted median of the distribution of bettor beliefs about the outcome of the game in question as opposed to the expected value of the outcome. This profit motive on the part of the betting house, and its implication for what the betting line should represent, forms the basis of the tests of market efficiency. Intuitively, bettors will be indifferent about a particular betting line when it contains all of the available information that bettors would use to place their wager. If information is available that would induce bettors to be on a particular side of a bet, then the betting line has not been properly set.

3 Conceptual Framework

We begin by establishing a set of probabilities that would allow a bettor to make profit in the betting market. We then describe how betting houses aggregate bettor beliefs about these probabilities

[3] Levitt (2004) [26] finds that betting houses, under certain betting formats, may take positions with respect to game outcomes in order to maximize profit using NFL data. Krieger, Fodor and Stevenson (2011) [24] find that this strategy is substantially less profitable using later data.

[4] Kilby et al. 2002 [23] and Roxborough 1991 [33] point out in guides for sports book management that a bookmaker's primary objective is, in fact, to minimize risk.

to set betting lines. Combining the probabilities with bettor beliefs allow us to define market efficiency.

With some probability, p, the bet will be successful, and if the event has only two outcomes, the remaining probability, $1-p$, captures the instance when the bet is not successful. The revenue of a bettor is multiplied by the bet size, B [5]. Also, every bet has a fixed percent that is given to the betting house, c, the vigorish. The threshold for any particular bet is thus a probability of success that exceeds the sum of the probability of failure and the transaction cost. For a market to be efficient, it must be the case that, for all strategies,

$$pB - cB \geq (1-p)B \tag{1}$$

and

$$(1-p)B - cB \geq pB \tag{2}$$

The left hand side of (1) represents the profit made by betting on a team to beat the betting line and the bet being successful. The left hand side of (2) represents the profit made by betting against a team to beat the betting line and the bet being successful. The right hand sides of both equations represent the loss realized from an unsuccessful bet. Factoring out B and rearranging terms yields the key relationship: $p \in \left\{\frac{1-c}{2}, \frac{1+c}{2}\right\}$. A bettor will place a bet on a team to beat the spread if she believes that the team has a greater than $\frac{(1+c)}{2}$ percent chance of beating the betting line. Conversely, a bettor will bet against a team beating the spread if she believes that the team has less than a $\frac{(1-c)}{2}$ percent chance of beating the betting line.

In this paper we test to see whether a particular strategy can exceed this threshold. If $c = 0$, there were no vigorish, then a risk-neutral bettor would be indifferent between betting that a team would beat (or, conversely, lose to) the line if the probability that a team would beat the spread was equivalent to the probability that the same team would lose to it-this would represent a coin flip. Otherwise, the bettor would bet on the outcome that was more likely. Note that since all participants in the betting market pay the vigorish, they are inherently risk-loving. Betting houses aggregate this betting behavior to make profit.

We define a market as efficient only when there are no profitable betting strategies that can be employed by bettors to make profit in expectation, which is a common definition in the betting market literature. That is, while betting a particular strategy in any given game may result in a win or loss, betting a simple strategy (such as always betting on home teams) is not expected to yield a consistent profit in the long run.

Using this measure as our test does not require the actual margin of victory to be any particular

[5] Betting houses may cap the size of the bet; for example, Kilby et al. 2002 [23] suggests that gambling houses should cap the size of the bet at $2,000 for college football, although betting houses may choose to pursue higher limits if they feel that bettors are particularly uninformed. However, this would not prevent bettors from having others bet for them. Thus, we model bet size to be arbitrarily large, although in principle the bet size is capped. We return to this idea later, when we discuss potential limits to arbitrage in this market.

distance from the betting line. The only criterion for profit is whether or not a particular game or team characteristic is statistically related to beating the betting line, which is a binary outcome. Thus, we can test to see whether or not particular characteristics beat the line more often than not (or particular characteristics lose to the line more often than not), which allows us to directly assess whether or not betting based on a particular game characteristic could make profit in expectation.

4 Data

We exploit two sources of data to test for market efficiency in this betting market. One source is final betting lines taken from sports handicapper Jim Feist's workbook, as used in Paul et al. 2003 [31]. This data source consists of final betting lines from all Division 1-A games from 1985-2003 for which betting lines exist.[6] From this data we obtain information on game location, game results, and the betting lines themselves. As this data is a comprehensive set of the universe of college football betting lines, we refer to this in our analysis as the "total data."

We also combine the betting line data with a richer set of game characteristics first presented in Logan 2010 [28]. The Logan data contains the team's record before and after the game and detailed information on the teams' opponents, including opponent's contemporaneous and season winning percentages and their poll ranking in both the AP and Coaches Polls. There are limits to the Logan data, however, as it covers only 25 of the most popular teams in college football. On the other hand, popular teams are likely to have the largest and most robust betting markets. These teams are listed in the appendix. Two key strengths of the Logan data are that it allows us to both construct measures such as opponent strength, and to consider poll data, or public signals of quality. These additional characteristics allow us to consider a richer set of strategies than is available in the raw betting lines themselves. We refer to this matched data as the "Logan sample."

Table 1 presents the summary statistics for both the total data and the Logan sample. We have two observations for each game in the sample; we exclude games where teams play teams that are not in Division 1-A [7]. In the total data, the average margin of victory is less than half a point. The average betting line is close to zero, and teams are more likely to be favored when playing at home than away. Because the Logan sample consists of 25 of the most successful programs, these teams are favored more than 70 % of the time, and yet beat the betting line only slightly more often (52%). The average betting line in the Logan sample is 7.81 points, which means that teams in

[6] We note that while there may be events that would influence betting market results after the final betting line, these events are neither known to bettors nor betting houses at the time of the bet, and thus these events do not impact market efficiency. Put another way, we only examine betting house pricing in terms of what betting houses are able to control for *ex-ante*, not what they cannot control, such as within-game injuries, for *ex-post*.

[7] These opponents are substantially weaker and are almost always away underdogs. These data were excluded from our data set because of the fact that betting houses do not provide lines for these teams unless they are playing major opponents, indicating the markets for betting on these teams individually are thin. Because betting houses do not pay attention to these teams except in the light of major opponents, we do not include results with these data in our analysis. However, our results are robust to the inclusion of these teams.

the Logan sample are favored to win by slightly more than a touchdown. By game location, teams in the Logan sample are much more likely to be favored. Consistent with the data being for the most successful college football teams, the average ranking of a team in the Logan sample is 10th in both the AP and Coaches polls. On average, the teams in the Logan sample play opponents who win one more game than they lose in a given season, which we would expect given the conference structure under which the most prominent teams play.

5 Empirical Results

5.1 Testing for Betting Strategies

Various tests of sports market betting efficiency exist. One potential method is to regress the actual outcome of a game on the associated betting line and a list of meaningful covariates, and then to test whether or not the betting line is statistically different from the actual outcome. Variations of this method have been used by Zuber et. al 1985 [39], Gandar et. al 1988 [15], Sauer et. al 1988 [34], Golec and Tanarkin 1991 [17], Dare and McDonald 1996 [10], Dare and Holland 2004 [11], and others on NFL data. It has also been used by Fair and Oster 2007 [14] on college football data for a limited number of years. Most of these studies have found that the betting line is not statistically different from the margin of victory on average, and use this to infer that betting markets are efficient. [8]

While this technique is a test of whether or not the betting line is a predictor of the actual margin of victory, it is not a direct test of whether or not the betting market is efficient. The technique described above characterizes efficiency as how close the actual margin of victory is to the betting line, but this is not important to market participants (either the betting houses or bettors), as shown earlier. Betting houses are concerned with getting even money on either side of the bet, and bettors are concerned with whether or not the actual margin of victory will be greater or less than predicted. These standard techniques do not address these questions. At best, the results of standard tests of market efficiency are indirect tests of market efficiency.

We test for efficiency in the market by considering the strategic considerations of bettors, who are only interested in whether or not a team beats the betting line, and the strategies bettors could use when making a bet. We denote whether a team beats the line as Y_i, which takes the value 1 if team i beats the line and 0 otherwise. [9] Strategies, we argue, would be based on game characteristics, which are independent sources of variation between games. For example, whether or not a team is playing at home, or whether or not a team is favored by the betting line itself are both pieces of information that individuals could use as a rule of thumb with which to bet on the outcome of a game. If favored teams beat the line sufficiently often, it would be profitable to

[8] We show the results of the traditional test for our data in the appendix.

[9] If a team neither loses to nor beats the betting line, it is dropped from the analysis, as all bets related to that game are returned.

bet on favored teams, and we denote this as a possible betting strategy. If the betting market is efficient, then no strategy based on readily observable game characteristics should yield positive expected profit. These sorts of patterns, we argue, should readily be discerned by bettors, and betting houses should move to eliminate the possibility for arbitrage.

To capture whether or not these strategies meaningfully increase the possibility of winning a bet, we estimate the marginal effects from a probit model of the form:

$$Pr(Y_i = 1) = \Phi(X_i^{'}\beta) \tag{3}$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution, and X_i is a vector of characteristics that pertain to specific strategies (see Gray and Gray 1997 [18] for an application of this method to professional football). These characteristics include a team's rank in the AP poll, the strength of the opponent, and the week of the season that the game is played. This approach allows us to uncover whether betting houses, and therefore bettors, misprice the likelihood of a team beating the line based on public information about a given game.

We test simple betting strategies in order to fully characterize and test for efficiency. These strategies are easily constructed and are more likely to be used by bettors, and focusing on only these strategies allows us to avoid data mining. There may be many different rules of thumb that bettors would employ in order to try and make profit consistently in betting markets. Different strategies have different expected outcomes. For example, an individual might choose to bet only on home teams. Home teams may have an advantage because their fans may be able to cheer very loud, affecting the other team's ability to use timeouts and call plays, or the visiting team may be unfamiliar with the playing surface, wind conditions, temperature, or other variables that may impact the result of a game. An individual might bet on home teams if he thought these factors appreciably improved a team's probability of winning a bet.

Table 2 reports the results of the probit specification, where we test whether or not particular strategies are consistently profitable. We find that there is evidence, both in the total data and in Logan sample, for favorites being overpriced relative to the line. In the total data, favorites are 1.86 percent less likely to beat the line. Home teams are 2.05 percent more likely to beat the line, which suggests that home teams are underpriced and that favorites are overpriced relative to the line. In the Logan sample, favorites are 6.38 percent less likely to beat the line, which suggest that favorites who happen to be teams with strong traditions are significantly overpriced. In the Logan sample, betting against favored teams with strong traditions would allow for large profit if the strategy was consistently applied. Whether or not a team is favored significantly impacts that team's associated probability of beating the betting line in every model. Favorites are consistently less likely to beat the betting line in every specification, from 1.86 percent to 2.54 percent in the total data to 1.82 to 6.84 percent in the Logan sample. Bettors systematically overprice favorites.

We find that betting $1,000 against favorites in the Logan data would, in expectation, yield

$1,017 after accounting for the vigorish, nearly a two percent return. Betting houses significantly overprice the favored team when the favored team has a strong tradition, which may indicate that betting houses profit from utility gains that bettors may receive from betting on their favorite team.

We consider the profitability of a richer set of specific strategies in Table 3, which tests the marginal effects of interaction strategies using both the total data and Logan sample. In particular, we test to see whether or not betting on a team that possesses a particular joint characteristic (e.g., a favorite playing at home) is more profitable than betting randomly. We find that in both the total data and the Logan sample pricing of home underdogs is statistically inefficient. This finding is consistent with previous findings for professional football, but has not been documented in college football. In the total data, home underdogs are 1.92 percent more likely to beat the spread, and this is significant at the 95 percent level. Furthermore, in the Logan sample, home underdogs are 10.8 percent more likely to beat the spread after controlling for week of season, the opponent's win-loss record at the time of the game, and AP poll rank. While the Logan results may be a function of a small number of home underdogs, the results suggest the presence of a profit opportunity for this set of teams.

As further evidence, we also check that the data is consistent with the inverse of the strategies. Since betting on home underdogs implies betting against away favorites (excluding games played at neutral sites), we test to see if betting against away favorites is a profitable strategy. We find that betting against away favorites is a profitable strategy in all specifications. In the total data, away favorites are 1.92 percent less likely to beat the spread. We find this to be strong evidence for the existence of profitable strategies in this betting market. [10]

6 The Mechanism of Inefficiency

The evidence presented in the previous section establishes that the college football betting market is inefficient (in terms of positive expected profit) when pricing both home teams and favorite teams. While finding inefficiency is interesting, the mechanism for inefficiency is more economically interesting, since the mechanism may explain inefficiencies in other markets. Here, we analyze and identify a potential source of those inefficiencies: memory in the betting lines.

Memory in asset pricing has been documented in other economic settings (see Lo 1991 [27] and Hirshleifer 2001 [20]). Jegadeesh and Titman 1993 [21] find that strategies derived from buying stocks that have performed well in recent periods and selling stocks that have performed poorly

[10] Dare and Holland 2004 suggest that these effects should be exactly opposite to each other, and that this should therefore be a model restriction. Although our results are consistent with their intuition, we do not impose equality of coefficients as a model restriction. First, there may be unobserved heterogeneity associated with away underdogs but not with home favorites, particularly because point spreads are not set in all college football games. 1-AA teams that are away underdogs may perform differently in games compared to their 1-A counterparts since they appear in games with point spread relatively infrequently. In these cases, imposing equality between the two effects may eliminate this interesting variation.

generate significant positive returns. Their findings are robust to measures for systematic risk and suggest the presence of both overreaction and underreaction in asset pricing. Studies have documented that market participants may believe that data clustering represents future trends rather than statistical anomalies (see Hirshleifer 2001 [20], Barberis and Thaler 2002 [2], and Durham, Heitzel, and Martin 2005 [12]), which may lead to strategies such as momentum trading. Past documented increases in an asset's price may lead to beliefs that future prices of the asset may increase, and may indeed lead to price changes in an asset. These beliefs may be correct and past performance may represent strong asset fundamentals. However, these beliefs may also lead to biased estimates of future prices, as investors or market participants may recall or use only favorable subsets of available information to make predictions of future prices. Relying on biases may be a sign of coarse thinking, as described in Ellison and Holden 2008 [13], Mullainathan, Schwartzstein, and Shleifer 2008 [30] and others.

We examine how strong asset fundamentals may lead to biases in bettor beliefs. In particular, if bettors perceive the likelihood of winning a bet to be higher based on whether or not a team previously exceeded a betting line, then betting houses have incentive to derive pricing rules that account for this perception. Bettors may well be correct in thinking that a team is stronger and that a team's performance will improve over time; however, a team's improved quality may be accounted for in the new betting lines. If lines move appreciably higher and bettors fail to account for this movement in their bets, then their analysis may represent coarse thinking in that they fail to account for betting house pricing changes in their beliefs. Looking for memory in betting lines thus allows us to identify whether or not bettors' beliefs about the likelihood of winning change due to past events. We are also able to identify both how a betting house responds to these belief changes, and the consequences for the betting houses with respect to pricing other strategies efficiently.

We proceed in three parts: first, we describe how betting houses set lines in the presence of memory, and we demonstrate empirically that betting lines exhibit memory. We then document evidence of bettor belief changes in the college football betting market. Finally, we test to see whether or not these belief changes contribute to the aggregate inefficiency found with pricing of home and favorite teams.

6.1 Does Past Performance Predict the Betting Line in College Football?

We first analyze how betting houses set betting lines. We test to see whether betting lines have memory in Table 4 by estimating the following equation:

$$bettingline_t = \gamma_0 + \gamma_1 beatspread_{t-1} + \gamma_2 beatspread_{t-2} + \gamma_3 beatspread_{t-3} + \gamma_4 WINSTREAK_t + \epsilon_t \tag{4}$$

If beating the betting line is not important to a betting house at time t, we would expect the coefficients on beat spread to be close to zero and statistically insignificant, while if there is evidence that beating the betting line matters to betting houses, we would expect the coefficients to be positive, indicating that betting houses make it more difficult for a team to beat the betting line in successive weeks. In this specification we also control for strong performance of a team by including a dummy variable, $WINSTREAK_t$, equal to one if a team has won in its previous three games and zero otherwise.[11]

We find strong evidence that beating the spread is important to betting houses in columns 1-3 and column 5 of Table 4: beating the betting line adds over 2 points to the betting line for a team, after accounting for both previous games that a team has beat the spread and whether or not a team is playing at home. Additionally, betting lines have memory from weeks prior to the previous week. Column 3 of Table 4 indicates that if a team beats the spread in the previous three weeks, then the betting house adds over 7.5 points to the betting line in the current week. Even after including a control for previous strong performance ($WINSTREAK_t$), we find that beating the spread just in the previous week adds a little over a half a point to the current week's betting line outside of a team's strong perforamnce, and that beating the spread two and three weeks ago adds nearly a point for each week to the betting line.

Furthermore, the magnitude of this result is significantly larger than magnitudes found in other markets. Brown and Sauer 1993 find that in professional basketball, betting houses add roughly 2/3 of a point to the point spread for a four game winning streak against the betting line; by contrast, we find that a three game streak adds roughly seven points to the betting line without performance controls, and two-and-a-half points after performance close. Our results mirror Avery and Chevalier 1999 [1] who find that betting houses add points to the final line from the closing line if a team beats the spread in previous weeks for NFL games. However, we also find that betting houses add points to the final line compared to the previous performance of a team based on winning streaks, which somewhat differs from Avery and Chevalier 1999 [1], who find that betting houses reduce the line during a week based on winning streaks.

We also test to see if the magnitude by which a team beats the betting line is important to betting houses. Beating the betting line by a small number of points may appear random to a betting house, while beating the line by many points may represent an increase in team quality and perceived strength. That is, we estimate:

$$bettingline_t = \delta_0 + \delta_1 magnitude_{t-1} + \delta_2 magnitude_{t-2} + \delta_3 magnitude_{t-3} + \upsilon_t \tag{5}$$

We report the results in columns 4-6 of Table 4. We find that teams who beat the betting line by many points are favored more heavily in subsequent games, and that this is more important to betting houses than whether or not a team beats the betting line. In particular, every additional point that a team beats the betting line by adds between .082 and .108 points to a subsequent betting line without performance controls, and adds between .033 and .048 points to a subsequent

[11]We tried other definitions of WINSTREAK with similar results.

betting line with performance controls.

The results from Table 4 reject the idea that betting lines have no memory from week to week, and, from the perspective of the betting house, reject the hypothesis that bettor beliefs do not contain serial correlation. If bettor beliefs did not change from week to week based on previous performances against the spread, then the betting house would have no reason to change the line, because that would cause it to take on unnecessary risk. Furthermore, our estimates indicate that increases in the betting line based on betting line performances are not just proxies for improved performance of a team. We find evidence that, even after controlling for winning streaks of teams, betting lines account for strong performance against the spread by increasing the betting line.

6.2 Profitability of Betting on Hot Teams

Although betting houses account for previous performance against the spread in betting lines, a well-documented behavioral betting strategy is betting on teams who beat the spread in consecutive weeks. This strategy evolves over time and changes from week to week. Thus an individual could not construct a strategy based on this characteristic in advance of the season, but would have to tabulate how this strategy evolved weekly. As noted previously, a strategy derived from this characteristic might be to bet on teams that perform well relative to the spread in previous weeks. This strategy is known as the "hot hand effect," and is well documented in the literature (see Camerer 1989 [7], Brown and Sauer 1993 [5], and Avery and Chevalier [1] for a discussion). The "hot hand" is a notion that captures the idea that an individual's probability of succeeding at an event changes over time in a deterministic way, such as the individual improving at a task, like shooting a basketball during a game, or even winning games in general. From our search of the narrative record, the hot hand strategy is very popular among bettors.[12] We are examining the "hot hand" effect solely in the context of consecutive performance against the spread.

We find that betting lines are functions of previous outcomes, but do those previous outcomes predict performance against the spread? To examine the profitability of betting on hot teams, we estimate (3) with betting on hot teams as a strategy available to bettors. We report the results in Table 5, which reports the profitability of betting on a team who beat the betting line in previous weeks. Column 1 of Table 5 indicates that whether or not a team beats the betting line in the previous week does not improve the probability of a team beating the betting line in the current

[12] We document this claim from our search of the narrative record, with some key examples presented here:

> "Following a team's winning streak is one of the best way of making money in sports gambling, as everybody loves to ride the hot team!" [13]

> "You need to find out which teams are blazing hot and seemingly can't lose to anybody and the teams that are cold as ice which look like they couldn't beat themselves. It's basic common knowledge and handicapping strategy to try and ride a winner until she bucks you and to stay away from dead beats." [14]

> "I am leaning towards Boise and the points. Right now, the line is sitting at 9.5. Boise has whipped Fresno 4 straight and I think the streak may even continue." [15]

week. Columns 2 and 3 show that this result is robust to consecutive weeks of beating the spread. In all of our specifications, beating the spread is not statistically different from zero. Using previous spread results as criteria for current betting is an unprofitable strategy in the college football market. Our results show that deriving a strategy based on whether a team beat a spread would be no better or worse than flipping a coin.

The "hot hand" strategy is well-accounted for by betting houses. Betting houses increase the spread even more prominently for teams that beat the spread multiple weeks in a row. Table 5 indicates that betting on hot teams is not a profitable strategy for bettors, as the probability of winning a bet falls well within the efficient range given by equations (1) and (2). In fact, betting on hot teams is literally no better than flipping a coin, as demonstrated in specifications 1-3 of Table 5. Bettor beliefs about the likelihood of hot teams continuing to beat the spread are empirically misplaced. Intuitively, perhaps why bettors have these seemingly erroneous beliefs is that teams win more games after beating the spread, but do not fare better against the spread. A simple intuition is that the team is indeed improving in quality, but the betting house is fully accounting for this improvement with the point spread, as evidenced in Table 4, where betting houses add sizably to the point spread if the team is on a winning streak. Bettors may be unable to disentangle winning the actual game from winning against the point spread, which may resemble coarse thinking. Bettors may weigh past examples of "hot teams" more heavily in their decision-making process with regard to their betting, while forgetting teams that beat the spread one week and lose to the spread the next week.

6.3 Market Function and Market Inefficiency

Our key argument in this work is that in order for hot hand pricing to be the driving force behind the observed inefficiencies in pricing of simple strategies, counteracting the hot hand must cause inefficiencies with respect to other strategies. This is because the betting house only has one variable with which to control an entire vector of strategies, and the strategies are not perfectly correlated with each other. Pricing one strategy more carefully necessarily leaves open other strategies, but it is not clear that the strategies left open are simple, easily derived strategies. To test this, we analyze how betting houses price simple strategies, such as betting on home or favorite teams, conditional on the betting houses' accounting for the "hot hand." Table 4 indicates that teams are priced differently based on how they perform against the betting line the week before. This differential pricing provides an additional source of variation both for bettors (bettors can construct a list of teams who performed well or poorly relative to the betting line and devise a strategy based on this information) and for the betting house, which can observe performance compared to the spread and price games based on this rule and other characteristics.

In principle, betting houses have no incentive to price simple strategies differently from week to week, because bettors do not believe that these characteristics should be priced differently.

However, these strategies might nonetheless be priced differently in different weeks depending on a team's performance against the betting line. If bettors believe that a team's performance against the spread is the foremost predictor of future outcomes against the line, then a betting house may price the hot hand strategy more closely in order to minimize risk, at the expense of other strategies. By estimating our original model in the light of hot hand pricing, we can isolate whether or not simple strategies are priced differently after accounting for the "hot hand." Additionally, we might be able to learn what drives the magnitude of the inefficiencies that we find in the total sample, where we ignore the presence of hot hand pricing.

To analyze the effect of hot hand pricing on simple strategies, we estimate (3) in the light of hot hand pricing. We define a team as being "hot" if it has won against the betting line in multiple weeks, and we define a team as being "cold" if it has lost to the betting line in two consecutive weeks. These definitions are the smallest possible streaks that bettors can view and infer when deciding to bet on a game.[16] We then estimate our baseline model with a control variable for whether or not a team is "hot" or "cold" and interaction terms between whether a team is hot and the simple strategies we examine in Tables 2 and 3: whether or not a team is favored, is playing at home, is a home underdog, home favorite, away underdog, or away favorite.

If the hot hand does not influence pricing and profitability of strategies, then we would expect that these strategies would not be priced differently because a team is hot. The marginal effects of betting on a team when it is hot and playing at home, for example, would not be statistically different from betting on home teams in general, nor would it be different from betting on teams that are "cold" and playing at home. If, however, the pricing is statistically different, then the inefficiency for a particular strategy, such as betting on home underdog teams, might arise only when it occurs in conjunction with the "hot hand".

Table 6 reports the estimates of our model for teams that have beat the spread in consecutive weeks. In each specification, we include a control for whether or not a team is "hot" and cluster our standard errors by team to capture any within-team pricing correlation. Column 1 shows that hot favored teams are not statistically more or less likely to beat the spread. However, Column 2 shows that hot home teams are statistically more likely to beat the spread- 3.80 percent more likely, compared with 2.05 percent in our total sample. Home teams are nearly twice as likely to beat the spread in instances where the betting house has moved the line up to account for the "hot hand" effect. These results tell us the source of the underpricing that we find in the market for home teams. Home teams are systematically underpriced by nearly double the magnitude of the estimates in Table 2, which does not account for hot hand pricing. The magnitude of the marginal increase in probability is driven specifically by teams priced differently due to a time-varying characteristic.

Columns 3-6 of Table 6 report the marginal effects associated with interaction strategies in the light of hot hand pricing. Column 4 of Table 6 presents evidence indicating that the source of the

[16] We explore other definitions of "hot" and "cold" in the Appendix, where we conduct subsample analysis of strategy profitability.

widely-found "home underdog" bias is hot hand pricing. We find evidence that home underdogs are four percent more likely to beat the spread if "hot," and that this is statistically different from zero. Home underdogs are also nearly twice as likely to be underpriced when "hot," indicating that a betting house's desire to increase the betting line to price out performance against the spread does not adequately account for the strength of the home underdog. Away favorites are three percent less likely to beat the spread if "hot," and this strategy also statistically different from zero, demonstrating consistency across the inverse of the strategy. Home underdogs are substantially more likely to to be underpriced relative to home favorites when hot, while away favorites are substantially overpriced relative to away underdogs.

Table 7 reports the estimates of our model for teams that lost to the spread in two consecutive weeks. These teams may be considered to be "cold." In all of our specifications, every estimate is insignificant. We find no evidence that "cold teams" are priced inefficiently in any of our estimates. Furthermore, our estimates for hot home underdogs, hot home teams, and hot away favorites are statistically different from the estimates for cold home underdogs, cold home teams, and cold away favorites.

Taken together, Tables 6 and 7 parse out how inefficiency is created in the market. We found that home teams are underpriced and favorites are overpriced in our total data. Table 6 indicates that home teams are particularly underpriced for teams that the betting house prices differently because they are "hot." We find evidence that the underpricing of home underdogs is generated from the pricing of hot teams. In other words, the market inefficiency of under-pricing home strategies is driven by the betting house's desire to closely price the hot hand effect. We point out that, for teams included when the definition of "cold" is expanded to include teams that have only lost to the spread in the previous week, favorites are overpriced. We report evidence supporting these results in the Appendix.

Interpreting the home and favorite marginal effects from the total sample ignores the fact that betting houses price games differently based on behavioral strategies. Because this underlying pricing drives the magnitudes of both effects, the original point estimates do not fully characterize the inefficiency of the market. Accounting for whether or not a team is priced differently based on performance against the spread yields substantially different marginal effects. Anecdotally, the bettors that bet on hot teams may represent a large fraction of the total bets, which would drive betting houses to price hot teams so precisely that other strategies become profitable, although these strategies are unused or used infrequently. If, for example, betting houses track bettors' bets over time, then they may be able to reconstruct and tabulate bettor strategies, and may choose to price the most frequently-used strategies more closely.

Tables 6 and 7 also reveal interesting strategies available to bettors in this market. A bettor could gather a list of all the teams that beat the spread and only bet on home teams, or gather a list of all the teams that beat the spread in previous weeks and bet against those teams that are away favorites, or for teams that are home underdogs. These conditional strategies are more profitable than unconditionally betting for home teams and against favorite teams.

7 Conclusion

We find two key features of prediction market pricing: one is that the price does not accurately represent all of the information available in the market, thus there is room for profitable betting. A second feature is that prediction market makers price popular behavioral strategies much more accurately than more fundamental strategies. In the market of college football betting, betting houses seek to eliminate the behavioral strategy of betting on "hot teams." This is done at the expense of pricing other strategies more closely; we provide evidence that the college football betting market exhibits market inefficiency with the pricing of simple strategies. We test for market efficiency directly by using the predicted success of specific betting strategies. We find evidence for statistical inefficiency when pricing home teams, favored teams, and teams that play weak opponents. Moreover, we find evidence for profitable betting against favored teams with strong traditions. We find that betting houses systematically overprice and underprice particular subsets of games. These inefficiencies could be exploited by bettors to make significant profit in the betting market.

More interestingly, we find that betting houses' responses to the behavioral biases of bettors are a source of this inefficiency. In this market, betting houses seek to eliminate profitable opportunities that may arise from a team entering a winning streak relative to the betting line by increasing the threshold that this team has to overcome in order to beat the spread the next week. By doing so, the betting house removes one potentially profitable (a widely popular) strategy from bettors, but this leaves open other, less well-known, strategies. Because the betting line is only a one-dimensional instrument, betting houses may prefer to price a more commonly played strategy more aggressively than uncommon ones, since eliminating one profitable strategy leaves another strategy priced less closely. If bettors do not use profitable betting strategies, then the inefficiencies we see in betting house pricing do not affect the betting house's profitability; in fact, the betting house may rationally maximize its profit by leaving these strategies available and pricing other strategies more closely. Additionally, betting houses are more likely to have private information regarding bettor beliefs, since they are able to track betting patterns by the same bettors and may be able to construct bettor strategies from observed bets. Because of this asymmetry between our data and betting house data, we conjecture that observed inefficiencies may not be indicative of irrational or erroneous pricing by betting houses, but rather erroneous beliefs about strategy profitability from bettors.

Further research should investigate the link between market function and inefficiency. If, in other markets, market makers respond to behavioral strategies very strongly, then trading strategies that are unrelated to those strategies may become profitable. This may have strong implications for pricing in asset markets where behavioral strategies are prominent (see Hirshleifer 2001 [20], Jegadeesh and Titman 2001 [22], Durham et al. 2005 [12]). Indeed, the success of some investors may be due to their ability to uncover the strategies that become profitable when market makers act to take away prominent behavioral sources of potential arbitrage.

Table 1: Summary Statistics

	Total Data					Logan Sample				
Variables	N	Mean	Std. Dev.	Min.	Max.	N	Mean	Std. Dev.	Min.	Max.
Home	22674	0.500	0.500	0	1	7144	0.521	0.500	0	1
Favorite	20877	0.498	0.500	0	1	4796	0.711	0.453	0	1
Beat Spread	22239	0.492	0.500	0	1	4821	0.516	0.500	0	1
Margin of Victory	22674	-0.023	21.946	-81	81	7161	8.802	20.599	-77	81
Betting Line	22440	-0.013	14.850	-59	59	4897	7.819	13.906	-42	55
Home Underdog	20877	0.183	0.387	0	1	4796	0.093	0.290	0	1
Home Favorite	20877	0.316	0.465	0	1	4796	0.420	0.494	0	1
Away Favorite	20877	0.183	0.386	0	1	4796	0.238	0.426	0	1
Away Underdog	20877	0.318	0.466	0	1	4796	0.158	0.365	0	1
Rank in AP Poll (before game)						5451	10.543	6.561	1	25
Rank in AP Poll (after game)						5330	10.460	6.523	1	25
Opponent Strength						6669	0.882	3.151	-11	12

See data appendix for variable definitions.
The means for home, favorite, home favorite, away favorite, home underdog, and away underdog represent the percentages of those teams in our data. We do not include games in which the opponent is a 1-AA team, since these teams are not normally priced by betting houses. Our results are robust to the inclusion of these teams and games.

Table 2: Marginal Effects in Betting Line Models

	Total Data			Logan Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Home	.021***		.010	.006		.016
	(.007)		(.007)	(.014)		(.014)
Favorite		-.019***	-.022***		-.064***	-.068***
		(.007)	(.008)		(.016)	(.020)
Observations	22239	20877	20877	4805	4720	4704
Pseudo R-Squared	.001	.000	.001	.003	.002	.026

This table reports the increases or decreases in the probability of a winning bet from betting on teams using strategies as described in equation (3), and we report the coefficients as the marginal effects of the probit regressions. The dependent variable is a dummy variable set equal to one if a team beat the spread and 0 otherwise. All standard errors are robust and are clustered by team. Significance Levels: $^{***}p < .01, ^{**}p < .05, ^{*}p < .1$

Table 3: Marginal Effects of Interaction Strategies

	Total Data				Logan Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Home Favorite	-.008				-.018			
	(.008)				(.016)			
Home Underdog		.019**				.108**		
		(.009)				(.048)		
Away Favorite			-.019**				-.006	
			(.009)				(.016)	
Away Underdog				.009				.024
				(.007)				(.037)
Opponent Strength Controls?					X	X	X	X
Week of Season Controls?					X	X	X	X
AP Poll Controls?					X	X	X	X
Observations	20877	20877	20877	20877	2691	2705	2691	2705
Pseudo R-Squared	.000	.000	.000	.000	.001	.002	.001	.001

This table reports the increases or decreases in the probability of a winning bet from betting on teams using interaction strategies as described in equation (3), and we report the coefficients as the marginal effects of the probit regressions. The dependent variable is a dummy variable set equal to one if a team beat the spread and 0 otherwise. All standard errors are robust and are clustered by team. Significance Levels: $^{***}p < .01, ^{**}p < .05, ^{*}p < .1$

Table 4: Betting House Responses to Behavioral Strategies

	Hot Hand Pricing					Margin Over Spread Pricing			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Beat Spread (t-1)	2.218***	2.135***	2.066 ***		.634***				
	(.208)	(.218)	(.231)		(.228)				
Beat Spread (t-2)		2.638***	2.546***		.947***				
		(.218)	(.231)		(.228)				
Beat Spread (t-3)			2.650***		.941***				
			(.231)		(.229)				
Winning Streak				10.855***	9.937***				9.419***
				(.726)	(.758)				(.772)
Margin Above Spread (t-1)						.091***	0.087***	.083***	.033***
						(.007)	(.007)	(.007)	(.007)
Margin Above Spread (t-2)							.108***	.102***	.048***
							(.007)	(.007)	(.007)
Margin Above Spread (t-3)								.099***	.041***
								(.007)	(.008)
Constant	-1.022	-2.222***	-3.467***	-3.117***	-2.371***	.066	.118	.0839	-1.783***
	(.642)	(.643)	(.667)	(.622)	(.523)	(.631)	(.625)	(.626)	(.537)
Observations	19712	17403	15218	18571	15208	19712	17403	15218	15208
R-Squared	.006	.014	.022	.079	.088	.010	.023	.034	.091

This table presents the results associated with OLS estimates of 4 and 5. The dependent variable in all specifications is the final betting line for the game. Specifications (1)-(5) present results associated with equation (4), where "Beat Spread" is a dummy variable that takes the value of one if a team beats the spread and 0 otherwise for time t-1 (one week prior), t-2 (two weeks prior) and t-3 (three weeks prior). Winning Streak is a dummy variable that takes the value of one if a team wins three weeks in a row and is 0 otherwise. Specifications (6)-(9) present results associated with equation (5), where "Margin Above Spread" is the difference between the margin of victory and the betting line for one week prior, two weeks prior, and three weeks prior. Standard errors are robust and are clustered by team to capture any within-team pricing correlation. Significance Levels: $^{***}p < .01, ^{**}p < .05, ^{*}p < .1$

Table 5: Marginal Effects of Betting on "Hot" Teams

	(1)	(2)	(3)
Beat Spread (t-1)	.001	.003	.004
	(.007)	(.008)	(.008)
Beat Spread (t-2)		.007	.010
		(.007)	(.008)
Beat Spread (t-3)			.001
			(.008)
Observations	19712	17403	15218
Pseudo R-Squared	.000	.000	.000

This table reports the increases or decreases in the probability of a winning bet from betting on teams that overcame in the betting line in previous weeks. We report the coefficients as the marginal effects of the probit regressions. The dependent variable is a dummy variable set equal to one if a team beat the spread and 0 otherwise. All standard errors are robust and are clustered by team. Significance Levels: $^{***}p < .01,^{**} p < .05,^{*} p < .1$

Table 6: Marginal Effects of Interaction Strategies for Hot Teams

	(1)	(2)	(3)	(4)	(5)	(6)
Hot	-.015	.009	-.003	-.009	.004	-.001
	(.011)	(.012)	(.007)	(.010)	(.009)	(.009)
Hot*Favorite	-.020					
	(.016)					
Hot*Home		.038***				
		(.014)				
Hot*Home Favorite			.001			
			(.017)			
Hot*Home Underdog				.040*		
				(.021)		
Hot*Away Favorite					-.030*	
					(.017)	
Hot*Away Underdog						-.002
						(.017)
Observations	17666	17666	17666	17666	17666	17666
Pseudo R-Squared	.001	.001	.000	.000	.000	.000

This table reports the increases or decreases in the probability of a winning bet on different strategies conditional on these teams overcoming the betting line in previous weeks. All specifications include a control for whether or not the team was "hot" at the time of the game, which is defined as beating the spread in consecutive weeks. We report the coefficients as the marginal effects of the probit regressions. The dependent variable is a dummy variable set equal to one if a team beat the spread and 0 otherwise. All standard errors are robust and are clustered by team. Significance Levels: $^{***}p < .01,^{**} p < .05,^{*} p < .1$

Table 7: Marginal Effects of Interaction Strategies for Cold Teams

	(1)	(2)	(3)	(4)	(5)	(6)
Cold	.010	.001	.009	.003	.001	-.007
	(.012)	(.012)	(.010)	(.010)	(.009)	(.010)
Cold*Favorite	-.019					
	(.018)					
Cold*Home		-.017				
		(.015)				
Cold*Home Favorite			-.025			
			(.020)			
Cold*Home Underdog				-.007		
				(.020)		
Cold*Away Favorite					.002	
					(.020)	
Cold*Away Underdog						.026
						(.017)
Observations	17712	18755	17712	17712	17712	17712
Pseudo R-Squared	.001	.001	.001	.000	.000	.001

This table reports the increases or decreases in the probability of a winning bet on different strategies conditional on these teams overcoming the betting line in previous weeks. All specifications include a control for whether or not the team was "cold" at the time of the game, which is defined as losing to the spread in consecutive weeks. We report the coefficients as the marginal effects of the probit regressions. The dependent variable is a dummy variable set equal to one if a team beat the spread and 0 otherwise. All standard errors are robust and are clustered by team. Significance Levels: $^{***}p < .01, ^{**}p < .05, ^{*}p < .1$

References

[1] Christopher Avery and Judith Chevalier. Identifying Investor Sentiment From Price Paths: The Case of Football Betting. *Journal of Business*, 72:493–521, 1999.

[2] Nicolas Barberis and Richard Thaler. A Survey of Behavioral Finance. 2002. NBER Working Paper 9222.

[3] Jimmy Boyd. College Football Trends: NCAA Football Betting Trends, September 2009. [online] `http://www.locksmithsportspicks.com/college-football-trends/`.

[4] Arndt Broder and Benjamin Scheibehenne. Predicting Wimbledon 2005 Tennis Results by Mere Name Recognition. *International Journal of Forecasting*, 23(3):415–426, 2007.

[5] William Brown and Raymond Sauer. Does the Basketball Market Believe in the Hot Hand? Comment. *American Economic Review*, 83:1377–1386, 1993.

[6] William Brown and Raymond Sauer. Fundamentals or Noise? Evidence from the Professional Basketball Betting Market. *Journal of Finance*, 48:1193–1209, 1993.

[7] Colin Camerer. Does the Basketball Market Believe in the Hot Hand? *American Economic Review*, 79:1377–1386, 1989.

[8] A. Colin Cameron, Jonah Gelbach, and Douglas Miller. Robust Inference with Multi-Way Clustering. 2006. NBER Technical WP 327.

[9] Jennifer Conrad, Michael Cooper, and Gautum Kaul. Value versus Glamour. *Journal of Finance*, 58(5):1969–1995, 2003.

[10] William Dare and Scott McDonald. A Generalized Model for Testing the Home and Favorite Team Advantage in Point Spread Markets. *Journal of Financial Economics*, 40:295–318, 1996.

[11] William H. Dare and A. Steven Holland. Efficiency in the NFL Betting Market: Modifying and Consolidating Research Methods. *Applied Economics*, 36(1):9–15, 2004.

[12] Gregory Durham, Michael Hertzel, and J. Spencer Martin. The Market Impact of Trends and Sequences in Performance: New Evidence. *Journal of Finance*, 60:2551–2569, 2005.

[13] Glenn Ellison and Richard Holden. A Theory of Rule Development. 2008. mimeo.

[14] Ray C. Fair and John F. Oster. College Football Rankings and Market Efficiency. *Journal of Sports Economics*, 8(1):3–18, 2007.

[15] John Gandar, Richard Zuber, Thomas O'Brien, and Ben Russo. Testing Rationality in the Point Spread Betting Market. *Journal of Finance*, 43(4):995–1007, 1988.

[16] Ricard Gil and Steven Levitt. Testing the Efficiency of Markets in the 2002 World Cup. *Journal of Prediction Markets*, 1(3):255–270, 2007.

[17] Joseph Golec and Maury Tamarkin. The Degree of Inefficiency in the Football Betting Market. *Journal of Financial Economics*, 30:311–323, 1991.

[18] Philip K. Gray and Stephen F. Gray. Testing Market Efficiency: Evidence From The NFL Sports Betting Market. *Journal of Finance*, 52(4):1725–1737.

[19] Mark Grinblatt and Matti Keloharju. The Investment Behavior and Performance of Various Investor Types: A Study of Finland's Unique Data Set. *Journal of Financial Economics*, 55(1):43–67, 2000.

[20] David Hirshleifer. Investor Psychology and Asset Pricing. *Journal of Finance*, 56(4):1533–1597, 2001.

[21] Narasamhan Jegadeesh and Sheridan Titman. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48(1):65–91, 1993.

[22] Narasamhan Jegadeesh and Sheridan Titman. Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *Journal of Finance*, 56(2):699–720, 2001.

[23] Jim Kilby, Jim Fox, and Anthony Lucas. *Casino Operations Management, Second Edition*. Wiley, 2002.

[24] Keven Krieger, Andy Fodor, and Greg Stevenson. The Sensitivity of Findings of Expected Bookmaker Profitability. *Journal of Sports Economics*. forthcoming.

[25] Josef Lakonishok, Andrei Shliefer, and Robert Vishny. Contrarian Investment, Extrapolation, and Risk. *Journal of Finance*, 49(5), 1994.

[26] Steven Levitt. Why are Gambling Markets Organized so Differently From Financial Markets. *Economic Journal*, 114:223–246, 2004.

[27] Andrew Lo. Long Term Memory in Stock Prices. *Econometrica*, 59(5):1279–1313, 1991.

[28] Trevon Logan. Econometric Tests of American College Football's Conventional Wisdom. *Applied Economics*, 2010. forthcoming.

[29] Erik Matuszewski. Vegas Betting Cops Help NCAA Find Suspicious Wagers, March 2009. [online] `http://www.bloomberg.com/apps/news?pid=20601079&sid=aDptJsOc5pGo&refer=home`.

[30] Sendhil Mullainathan, Joshua Schwartzstein, and Andrei Shliefer. Coarse Thinking and Persuasion. 2005. mimeo, Harvard University.

[31] Rodney Paul, Andrew Weinbach, and Chris Weinbach. Fair Bets and Profitability in College Football Gambling. *Journal of Economics and Finance*, pages 236–242, 2003.

[32] Ron Raymond. Rice Owls vs. Oklahoma State Cowboys- College Football Betting, September 2009. [online] `http://www.atsdatabase.com/blog/blog1.php/2009/09/15/rice-owls-vs-oklahoma-state-cowboys-coll`.

[33] Michael Roxborough and Michael Rhoden. *Race and Sports Book Management: A Guide for the Legal Bookmaker*. Michael Roxborough and Michael Rhoden, 1991.

[34] Raymond Sauer, Vic Brajer, Stephen Farris, and M. Wayne Farr. Hold Your Bets: Another Look at the Efficiency of Gambling for National Football League Games. *Journal of Political Economy*, 96:206–213, 1988.

[35] Ryan Sullivan, Allan Timmerman, and Halbert White. Data Snooping, Technical Trading Rule Performance, and the Bootstrap. *Journal of Finance*, 54(5):1647–1691, 1999.

[36] Against the Line. Boise at Fresno: College Football Message Boards, September 2009. [online] `http://www.againsttheline.com/board/showthread.php?t=1146&highlight=streak`.

[37] Halbert White. A Reality Check for Data Snooping. *Econometrica*, 68(5):1097–1126, 2000.

[38] Linda Woodland and Bill Woodland. Market Efficiency and the Favorite-Longshot Bias: The Baseball Betting Market. *Journal of Finance*, pages 269–279, 1994.

[39] Richard Zuber, John Gandar, and Benny Bowers. Beating the Spread: Testing the Efficiency of the Gambling Market for National Football League Games. *Journal of Political Economy*, 93:800–806, 1985.

Appendix

A Data Appendix

We use betting line data from the years 1985-2003 for 119 Division I-A teams. If a team was originally in Division I-AA and entered Division I-A, then we only include data after the team entered Division I-A. Data was taken from Jim Feist's betting workbook and was used in Paul et. al, 2003. For a given team, we have data on the date of the game, the opponent, the betting line, the location of the game, and the week of the season. In addition, for the Logan sample, we have data on the poll rank before and after the game and the win/loss record of the opponent for the season for a sample of 25 of the most prominent teams over that period, listed in Appendix Table 1. The Logan sample methodology is described in Logan (2009).

In addition to the existing data, we define the following variables:

- Home favorites are defined as teams that are both playing at home and favored for a particular game. We assign a team one if it meets both criteria and zero otherwise. We define away favorites, away underdogs, and home underdogs in the same manner.
- "Beat Spread" is a variable set equal to one if a team beats the spread. If a team does not beat the spread, it is set equal to zero. We do not assign values for pushes.
- "Margin Above Spread" refers to the difference between the actual margin of victory and the margin of victory predicted by the betting line. For example, if a team is favored to win by 10, but wins by 3, then we define the margin above spread to be -7. If the team is favored to win by 10, but wins by 17, then we define the margin above spread to be 7.
- "Opponent Strength" is defined to be the number of wins minus the number of losses for a team's opponent at the time the game is played.
- "Week of Season" is the poll week of the season. The first poll corresponds to preseason rankings, thus we start with week 0, which corresponds to the first poll before any games are played.

A.1 Standard Test of Market Efficiency

Appendix Table 2 provides estimation results for the standard model for efficiency in the literature:

$$MOV_i = \beta_0 + \beta_1 LINE_i + \epsilon_i \tag{A.1}$$

where MOV_i is the margin of victory (or margin of defeat) for team i and $LINE_i$ is the number of points that team i is favored to win (or lose) by the betting line. The traditional test for efficiency is that the constant, β_0 is zero and the coefficient on the betting line, β_1 is statistically indistinguishable from 1. The prevailing logic in the literature is that the betting line contains all of the information needed to predict the margin of victory, on average. If this is true, the betting line will accurately and consistently predict the margin of victory. Our results for this particular test mirror those found in this literature, and are detailed in the appendix.

We reproduce the standard test of market efficiency using both the total sample and the Logan sample. Using both sources, we find corroborating evidence for these previous results. Columns I and II of Appendix Table 2 show that the coefficient on betting lines is statistically indistinguishable from 1. In the total sample, however, playing at home adds one additional point to margin of victory that is not captured by the betting line, which suggests that while the betting line may predict margin of victory on average it excludes information that influences the actual margin of victory. In the Logan sample the predictive power of the line is slightly weaker, and it diminishes significantly once additional variables are added to the specification. For example, column V indicates that opponent strength, playing at home, and AP rank are all important indicators of the actual margin of victory. Also, the inclusion of these additional game characteristics causes the betting line to be statistically different from zero. [1]

One could argue that the results of standard tests, where we show that the predictive power of the betting line declines when additional game characteristics are added to the specification, is simply due to colinearity. If the betting line captures these game characteristics including them in the regression results in biased estimates. But note that such an argument also implies that, conditional on the betting line, these game characteristics would have no predictive power, as they have been captured by the betting line.

A.2 Robustness

We separate our robustness checks into three categories. First, we broaden our definitions of "hot" and "cold" to include teams that have performed well or poorly relative to the betting line only in the previous week. Bettors may believe that even a strong performance against the betting line in just the previous week represents a larger trend, and likewise may believe that poor performance against the betting line represents a trend as well. As described in the text, betting houses have incentive to strategically price these games differently because of bettor beliefs. We examine pricing of simple strategies in the light of this expanded definition of "hot" and "cold."

We next analyze robustness related to time and robustness related to potential violations of independence within clusters of our standard errors. We test to see whether or not time plays a role in our estimation; perhaps our results are driven either by variation in betting house behavior between years or by variation in betting house behavior during a given year. If either of these arguments is true, then the probabilities we estimate for the total sample do not capture an unconditional increase or decrease in the profitability of a particular strategy. Rather, our results would then capture either an anomaly or the fact that betting houses are initially imprecise in their pricing but improve over time.

We finally examine the sensitivity of our overall results to potential error correlation caused by within-group similarities. We test for within-group autocorrelation by estimating a logit model with multi-way clustering on year, team, and date. If errors are correlated at any of these group levels, then standard errors from our probit estimates will be neither i.i.d., nor homoskedastic, and will, in general, be too small.

[1] If the betting line includes all of the information that is used by bettors, the inclusion of these characteristics would cause a multicolinearity problem.

A.2.1 Subsample Analysis of Hot and Cold Team Pricing

Appendix Tables 3 and 4 report estimates of our basic model utilizing the additional variation provided from hot hand pricing. To estimate how the hot hand influences pricing, we split our sample into four subsamples: teams that beat the spread last week (teams that are relatively "hot"), teams that beat the spread in two consecutive weeks (teams that are "hotter"), teams that lost to the spread last week (teams that are relatively "cold"), and teams that lost to the spread in either of the last two weeks (teams that are relatively "colder"). These definitions are alternative definitions to those provided in the main text. We then estimate our baseline model in Appendix Tables 3 and 4 for those different subsamples. If the hot hand does not influence pricing, then we would expect that these strategies would not be priced differently from our total sample: the marginal effects should be the same. If, however, the pricing is statistically different, then the inefficiency for a particular strategy, such as betting on home teams, might arise only when it occurs in conjunction with the "hot hand".

The first three columns of Appendix Table 3 report the estimates of our model using data only for the teams that beat the spread last week. We characterize these teams as being "hot", because the most recent result for bettors to remember is a victory over the spread. Of those teams, home teams are more likely to beat the spread- 3.92 percent more likely, compared with 2.05 percent in our total sample. Home teams are nearly twice as likely to beat the spread in instances where the betting house has moved the line up to account for the "hot hand" effect. If those "hot" teams are favored, they are neither more likely nor less likely to beat the spread. Specifically, we find that favorites are 1.45 percent less likely to beat the betting line, but that this result is not statistically different from zero. These results tell us the source of the underpricing that we find in the market for home teams. Home teams are systematically underpriced by nearly double the magnitude of that found in the total sample. The home teams that are underpriced in our sample statistically share the feature of having beat the betting line the week before– the magnitude of the marginal increase in probability is driven specifically by teams priced differently due to a time-varying characteristic.

Columns 4-6 of Appendix Table 3 provide additional evidence to suggest that betting houses misprice home teams because they are accounting for the "hot hand." In these columns, we report our basic model specifications for teams that beat the spread in two consecutive weeks, implying that there is a more pronounced "streak" for bettors to focus on. Intuitively, bettors might view these teams as being "hotter" than other teams because they have had consecutively strong performances against the betting line. Table 4 indicates that betting houses add around two points to the betting line for these teams after accounting for strong performance from winning streaks. We find that home teams are 3.83 percent more likely to beat the spread for this subsample-betting houses continue to misprice home teams by accounting for the hot hand effect. Once again, we find no evidence that the entire set of favorites is overpriced in this subsample as the marginal effects are statistically insignificant.[2]

We also find evidence indicating that the source of the widely-found "home underdog" bias is hot hand pricing. Column 3 of Appendix Table 3 finds that home underdogs are substantially more likely to to be underpriced relative to home favorites in our sample of hot teams. This evidence is mildly supported by the results contained in Column 6, which indicates that home underdogs are again underpriced, although this result is only marginally significant.

[2] We provide further evidence of these effects in the appendix, where we demonstrate two results: that standard tests of market efficiency are unchanged by subsample analysis, and that home teams that beat the spread last week are more profitable among home teams. These can be viewed as robustness checks for our main results.

The first three columns of Appendix Table 4 include only teams that did not beat the spread in previous weeks. As these teams have lost to the spread last week, they are not considered "hot;" rather, they may be considered to be "cold." Of those teams, teams that are favored are 2.6 percent less likely to beat the spread, compared with 1.86 percent in our total sample. Again, the magnitude of the marginal effect found in the total sample is explained by pricing of teams that did not beat the spread in the previous week. While we find that the line is systematically lower for these teams, if the team is favored, the line is still systematically high enough to generate a three percent overpricing of favorites. The magnitude of the favorite overpricing found in the total sample is again driven by a subset of teams priced differently because of a behavioral strategy. For teams that did not beat the spread before, home has no predictive power of whether or not a team beat the spread for the subset of teams that did not beat the spread before, but favorite does, and it does with a larger magnitude. Columns 4-6 of Appendix Table 4 report results for teams that lost to the spread in either week (or, potentially, in two consecutive weeks). We find that these teams, if favored, continue to be overpriced.

A.2.2 Time and Season Fixed Effects

Our results could be biased due to particular years being heavily mispriced. For example, favorites could have been mispriced more heavily in a given year; the results from that year could in turn lead to both the magnitude and the significance of the inefficiencies we find in the market. We test for this possibility by including year fixed effects in the model, which allow us to control for any sensitivities related to year. If a particular year drives the results of our model, then including that year as a fixed effect would mitigate the inefficiency result that we find.

We report our results for the richer set of covariates in the Logan data; we find that the coefficients on both favorite and opponent strength retain statistical significance. These results are reported in Appendix Table 5. We do not find any evidence that betting houses priced teams significantly less efficiently in a given year.

We also test to see whether or not our results are sensitive to the week of the season that the game was played in. Perhaps betting houses become better at pricing teams over the course of a season after the quality of a team is more precisely revealed. If this is true, then games that were priced at the beginning of the season would be priced less efficiently, and the results that we find for the total sample might be driven by the inefficiencies found at the beginning of the year. To test this hypothesis, we include week of season fixed effects in our basic model for the Logan sample. We find that including week of season fixed effects does not significantly alter our estimated marginal effects. We do not find any evidence of a week of season bias driving our results.

A.2.3 Clustering of Standard Errors

We examine whether or not our results are robust to multi-way clustering. For example, team and season year should likely not be considered independent clusters in the data, and financial returns are commonly clustered by date. To account for the potential of multi-way clustering we reestimate our model to account for clustering by season year, team, and date of game in a logit model using procedures outlined in Cameron et al. 2006 [8]. If errors are also correlated across years or for particular dates, then the standard errors from our main estimates may be biased as the assumption of independence is violated within clusters of the independent variables; more specifically, we will not have any independent variation within large groups of our independent variables. Thus our standard errors may be too small and we may reject the null too often.

We report the results in Appendix Table 6. We find that favorites are still overpriced in both the total sample and in the Logan sample, and this result is significant at a level just under the 95 percent level in the total sample and is significant at the 99 percent level in the Logan sample. Favorites are still significantly overpriced among teams with strong traditions. We find that home teams are still underpriced, but this result is only marginally significant; however, we find that profitability of interaction strategies, such as betting on home underdogs, are still significant.

Appendix Table 1: Teams Used in Sample

Alabama	Miami, Florida
Arkansas	Michigan
Auburn	Nebraska
Boston College	Notre Dame
Brigham Young (BYU)	Ohio State
California	Oklahoma
California, Los Angeles (UCLA)	Penn State
Colorado	Southern California (USC)
Florida	Stanford
Florida State	Tennessee
Georgia	Texas
Iowa	Texas A & M
Louisiana State (LSU)	

Appendix Table 2: Standard Tests for Market Efficiency

	Total Data		Logan Sample					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	-.011	-0.858***	1.334***	.434	.913**	.781	2.017***	3.839**
	(.108)	(.198)	(.267)	(.333)	(.392)	(.564)	(.731)	(1.414)
Betting Line	1.005***	.987***	.941***	.920***	.893***	.921***	.865***	.816***
	(.008)	(.008)	(.017)	(.017)	(.023)	(.018)	(.027)	(.048)
Home		1.704***		1.696***	2.052***	2.090***	2.912***	2.922 ***
		(.276)		(.477)	(.493)	(.477)	(.623)	(.685)
Opponent Strength					-.237***			-.289**
					(.089)			(.117)
Week of Season						-.0345		-.044
						(.043)		(.051)
Rank in AP Poll (before game)							-.0858*	-.123*
							(.046)	(.051)
Team Clustered Standard Errors		X						X
Observations	22239	22239	4894	4878	4585	4859	3014	2787
R-Squared	.463	.464	.401	.404	.411	.405	.363	.363

Dependent Variable: margin of victory.
Robust standard errors in parentheses.
Significance Levels: $^{***}p < .01,^{**}p < .05,^{*}p < .1$
Specifications (1)-(2) include the total data.
Specifications (3)-(8) include data from the Logan sample.

Appendix Table 3: Marginal Effects in Betting Line Models: Hot Teams Only

	Beat Spread Last Week			Beat Spread Consecutive Weeks		
	(1)	(2)	(3)	(4)	(5)	(6)
Home	.039***		.029***	.038**		.029*
	(.010)		(.009)	(.015)		(.015)
Favorite		-.014	-.022**		-.020	-.027
		(.010)	(.011)		(.016)	(.017)
Observations	9723	9217	9217	4241	4031	4031
Pseudo R-Squared	.001	.000	.001	.001	.000	.001

This table reports the increases or decreases in the probability of a winning bet from betting on teams using strategies as described in equation (3), and we report the coefficients as the marginal effects of the probit regressions. The first three columns (specifications (1)-(3)) are estimates for teams that were priced differently by betting houses due to past positive performance against the spread in the previous week. The last three columns (specifications (4)-(6)) are estimates for teams that were priced differently by betting houses due to consecutive positive performances against the betting line. The dependent variable is a dummy variable set equal to one if a team beat the spread and 0 otherwise. All standard errors are robust and are clustered by team. Significance Levels: $^{***}p < .01, ^{**}p < .05, ^{*}p < .1$

Appendix Table 4: Marginal Effects in Betting Line Models: Cold Teams Only

	Lost to Spread Last Week			Lost to Spread Either Week		
	(1)	(2)	(3)	(4)	(5)	(6)
	(.010)		(.011)	(.008)		(.009)
Favorite		-.026**	-.024**		-.017**	-.018 *
		(.010)	(.011)		(.009	(.011)
Observations	9989	9333	9333	13162	12404	12404
Pseudo R-Squared	.000	.001	.001	.000	.001	.001

This table reports the increases or decreases in the probability of a winning bet from betting on teams using strategies as described in equation (3), and we report the coefficients as the marginal effects of the probit regressions. The first three columns (specifications (1)-(3)) are estimates for teams that were priced differently by betting houses due to negative performance against the spread in the previous week. The last three columns (specifications (4)-(6)) are estimates for teams that were priced differently by betting houses due to negative performances against the betting line in either week. The dependent variable is a dummy variable set equal to one if a team beat the spread and 0 otherwise. All standard errors are robust and are clustered by team. Significance Levels: $^{***}p < .01, ^{**}p < .05, ^{*}p < .1$

Appendix Table 5: Sensitivity Checks- Logan Sample

	(1)	(2)	(3)
Home	.016	.015	.014
	(.015)	(.016)	(.020)
Favorite	-.069***	-.069***	-.069***
	(.017)	(.017)	(.017)
Year fixed effects?	X		X
Week of season fixed effects?		X	X
Observations	4697	4685	4683
Pseudo R-Squared	.004	.004	.005

Dependent Variable: $Y_i = 1$ if team i beat the spread and 0 otherwise.
Robust standard errors in parentheses.
Significance Levels: $^{***}p < .01, ^{**}p < .05, ^{*}p < .1$
We report the coefficients as the marginal effects of the probit regressions.
For dichotomous variables, such as home or favorite, the effects represent the change in probability from 0 to 1.

Appendix Table 6: Marginal Effects in Betting Line Models

	Total Data			Logan Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Home	.087*		.039	.018		.006
	(.053)		(.050)	(.085)		(.064)
Favorite		-.075*	-.866**		-.253***	-.273***
		(.039)	(.042)		(.067)	(.082)
Observations	22239	20877	20877	4812	4711	4695
Pseudo R-Squared	.000	.000	.001	.002	.000	.003

Dependent Variable: $Y_i = 1$ if team i beat the spread and 0 otherwise.
Significance Levels: $^{***}p < .01, ^{**}p < .05, ^{*}p < .1$
We report the coefficients as the marginal effects from a logistic regression.
For dichotomous variables, such as home or favorite, the effects represent the change in probability from 0 to 1.
Standard errors are clustered by team, date of game, and season year.

On the one hand, if you bet that inflation will be high and go with a real annuity, you run the risk of seriously overpaying if inflation turns out to be low. On the other hand, if you bet that inflation will be low and therefore choose a nominal annuity, you run the risk of inflation being much higher and receiving significantly more worthless payouts as your retirement years advance. How likely is it that annual inflation will average just 2.12% for the next 30 years?Â That in turn suggests inflation over the next several decades may be much higher than the market currently expects. Brown conjectured that many retirees are not putting enough weight on the risk of higher inflation in coming decades.