Asset Pricing and Sports Betting

Tobias J. Moskowitz*

ABSTRACT

Two unique features of sports betting markets provide an informative laboratory to test behavioral theories of cross-sectional asset pricing anomalies: 1) the bets are idiosyncratic, having no systematic risk exposure; 2) the contracts have a known and short termination date where uncertainty is resolved that allows mispricing to be detected. Analyzing more than one hundred thousand contracts spanning almost three decades across four major professional sports (NBA, NFL, MLB, and NHL), there is strong evidence of momentum and weaker evidence of value effects that move prices from the open to the close of betting, which are then completely reversed by the game outcome. These findings are consistent with delayed overreaction theories of asset pricing, and are inconsistent with underreaction or rational pricing. In addition, a novel implication of overreaction uncovered in sports betting markets is shown to also predict returns in financial markets, where momentum is stronger and value is weaker when information is more uncertain. Despite evidence of mispricing, the magnitudes of momentum and value effects in sports betting markets are much smaller than those in financial markets, and are not large enough to overcome transactions costs, which prevent them from being arbitraged away.

*Yale University, AQR Capital, and NBER. I have benefited from the suggestions and comments of Nick Barberis, Antonio Bernardo, John Cochrane, Lauren Cohen, Josh Coval, Eugene Fama, Ken French, Xavier Gabaix, Bryan Kelly, Steve Levitt, Gregor Matvos, Lasse Pedersen, Amit Seru, Jesse Shapiro, John Shim, Richard Thaler, and Rob Vishny, as well as seminar participants at the SIFR conference in Stockholm, Sweden and the FRA meetings in Las Vegas. I also thank John Shim and Brian Weller for outstanding research assistance. Moskowitz thanks the Center for Research in Security Prices and the Initiative on Global Markets at the University of Chicago, Booth School of Business for initial financial support and Yale SOM for additional financial support. Moskowitz is also a principal with AQR Capital, who has nothing to do with and no interest in sports betting markets.

Correspondence to: Tobias J. Moskowitz, Yale School of Management, 165 Whitney Ave., New Haven, CT 06511. E-mail: tobias.moskowitz@yale.edu.

The asset pricing literature is replete with predictors of financial market security returns, yet much debate remains on their interpretation. Risk-based theories of rational asset pricing, behavioral theories of mispricing and limits to arbitrage, and statistical explanations such as data mining provide three distinct views of these patterns with vastly different implications for understanding asset pricing's role in the broader economy. Indeed, security characteristics that describe expected returns have become the focal point for discussions of market efficiency, for risk sharing, resource allocation, and investment decisions, where debate centers on whether these variables represent compensation for bearing risk in an informationally efficient market, predictable mispricing in an informationally inefficient market (due perhaps to investor biases and market frictions), or a statistical fluke.

Progress on the efficient markets question is mired by the joint hypothesis problem (Fama (1970)) that any test of efficiency is inherently a test of the underlying equilibrium asset pricing model, leading to a host of rational and behavioral theories for the same return predictors. Rational theories link return premia to aggregate systematic risks (e.g., macroeconomic shocks or proxies for state variables representing the changing investment opportunity set and marginal utility of investors), while behavioral theories link returns to investor cognitive errors and biases.

Capital market security returns provide a particularly difficult empirical laboratory to distinguish between these broad views of asset pricing since the researcher cannot directly observe marginal utility or investor preferences, and where both rational *and* behavioral forces could simultaneously be at work.[1]

To circumvent the joint hypothesis problem, I propose an alternative asset pricing laboratory – sports betting markets. The idea is simple. Assuming asset pricing models should apply to all markets (which is more appealing than asset-specific models), there are two key features of sports betting markets that provide a direct test of behavioral asset pricing distinct from and not confounded by any rational asset pricing framework: 1) sports bets are idiosyncratic and have no relation to any risk premia in the economy; 2) sports contracts have a short and known termination date where uncertainty is resolved, revealing a true value that is independent from betting activity that allows mispricing to be detected.

For the first feature – the idiosyncratic nature of the bets – the critical point is to examine the *cross-section* of sports betting contracts – comparing betting lines across games at the same time and even across different bets on the same game. While aggregate risk preferences and changing risk premia might affect the entire betting market as a whole, they have no bearing on the cross-section of games or the cross-section of contracts

[1] Complicating matters further is the role played by institutional, market, funding, trading, delegation, and regulation constraints that may also affect prices and interact with rational and behavioral forces to exaggerate or mitigate return patterns.

on the same game.[2] Hence, rational asset pricing theories have nothing to say about return predictability for these contracts. On the other hand, sports betting contracts should be subject to the same behavioral biases that are claimed to drive the anomalous returns in financial security markets. The behavioral models focus on beliefs or preferences that deviate from rational expectations and neoclassical theory regarding generic risky gambles (see Barberis and Thaler (2003)). Evidence from experimental psychology that provides the backbone for these theories comes from finite risky, generic bets. Hence, these theories pertain as much to idiosyncratic sports bets as they do to capital market securities. The cross-section of idiosyncratic sports bets, therefore, provides a unique asset pricing laboratory for behavioral theory.

The second key feature of sports betting contracts is that they have a known, and very short, termination date, where uncertainty is resolved by outcomes (e.g., the game score) that are independent of investor behavior.[3] These contracts provide a true terminal value for each security; something rarely seen in financial markets. Moreover, that terminal value depends solely on the outcome of the game, which is independent of betting activity, bettor sentiment, or preferences, something we can never know for sure in financial markets. The exogenous terminal value allows for the identification of mispricing, providing a stronger test of the behavioral models, which assume prices deviate from fundamental values due to cognitive biases or erroneous beliefs. The alternative hypothesis that these markets are efficient implies that information moves prices and there is no mispricing. Hence, mispricing implies return predictability while rational pricing implies no return predictability (since there is no risk premium embedded in these contacts). The combination of both features: 1) no risk premia and 2) an exogenous finite terminal value makes sports betting contracts unique and useful for isolating tests of behavioral asset pricing theories.[4]

The direction of any pricing correction at the terminal date also helps distinguish among competing behavioral theories. For example, overreaction models (Daniel, Hirshleifer, and Subrahmanyam (1998))

[2]For example, changing risk aversion and/or risk premia might affect betting behavior and prices for the entire NFL betting market as a whole – how much is bet, the willingness to bet, and perhaps betting odds in aggregate – but should have no impact on the betting behavior and prices of the Dallas vs. New York game relative to the Washington vs. Philadelphia game happening at the same time. Moreover, sports betting contracts are in zero-net supply and it is rare that one side of the market is being bet by individuals with the bookmaker taking the other side (in fact, spreads are typically set so that both sides are roughly even), providing another reason why aggregate risk premia would not be expected in these markets.

[3]Unless one believes in conspiracy theories and rampant game fixing by paying players to perform differently than they otherwise would in order to affect betting outcomes, there should be no relation between betting behavior and game outcomes. While there are some infamous cases where game fixing is claimed to have happened – the 1919 Chicago Black Sox in the World Series, the Dixie Classic scandal of 1961, the CCNY Point Shaving Scandal in 1950-51, and the Boston College basketball point shaving scandal of 1978-79 – such cases are extremely rare, have typically involved obscure and illiquid games, and have never actually been proven. For the games analyzed in this paper, game fixing related to betting behavior should not be a concern given the depth of the sports markets analyzed, the attention and scrutiny paid to these contests, and the stakes and salaries of professional athletes over the sample period, which would make fixing games extremely costly. Finally, for any of this to matter for the interpretation of the results in this paper, it would have to be correlated with the cross-sectional return predictors – momentum, value, and size – which seems unlikely.

[4]While other assets also have finite terminal dates, such as fixed income and derivative contracts, they also carry potentially significant risk premia, and may have terminal values that are affected by investor preferences or behavior.

imply a return reversal from the revelation of the true price, while underreaction models (Barberis, Shleifer, and Vishny (1998) and Hong and Stein (1999)) imply a return continuation. These additional implications of behavioral models are very difficult to test in financial markets because there is typically no known terminal date or revelation of true value for financial securities.

To draw connections to the broader asset pricing literature, this study examines cross-sectional predictors of returns found in financial markets applied to the cross-section of sports betting contracts. Specifically, I focus on the three characteristics that have received the most attention: size, value, and momentum.[5] One objective, therefore, is to derive analogous measures for size, value, and momentum in sports contracts. Momentum, which is typically measured by past performance or returns, is relatively straightforward. For value, a variety of "fundamental"-to-price ratios, long-run reversals, and relative valuation measures are used, and size is measured by local market and team size.[6]

Using data from the largest Las Vegas and online sports gambling books across four U.S. professional sports leagues: the NBA, NFL, MLB, and NHL, covering more than one hundred thousand contracts and spanning almost three decades, I find that price movements from the open to the close of betting react to momentum and (to a lesser extent) value measures in a manner consistent with evidence from financial markets. Size has no return predictability. These price movements are fully reversed, however, by the game outcome, when the true terminal value is revealed. The evidence suggests that bettors follow momentum and value signals (e.g., chasing past performance and "cheap" contracts) that push prices away from fundamentals, which then get reversed when the true price is revealed. The results are most consistent with overreaction models (e.g., Daniel, Hirshleifer, and Subrahmanyam (1998)).

These patterns are robust across a variety of specifications and measures and are found in each of the four different sports and within each sport across three separate betting contract types (point differential, who wins, and total number of points scored by both teams), providing a total of 12 different samples. The

[5]There is a host of evidence that size, value, and momentum explain the cross-section of returns over many markets and time periods. For recent syntheses of this evidence and its application to other markets, see Fama and French (2012) and Asness, Moskowitz, and Pedersen (2013). The behavioral and risk-based asset pricing models also focus predominantly on these three characteristics: Daniel, Hirshleifer, and Subrahmanyam (1998), Barberis, Shleifer, and Vishny (1998), and Hong and Stein (1999) for behavioral models and Gomes, Kogan, and Zhang (2003), Zhang (2005), Belo (2010), Berk, Green, and Naik (1999), Johnson (2002), Sagi and Seasholes (2007), Hansen, Heaton, and Li (2008), and Lettau and Maggiori (2013) for risk-based explanations.

[6]There are other robust cross-sectional predictors of returns in financial markets that include liquidity risk (Pastor and Stambaugh (2003) and Acharya and Pedersen (2005)), carry (Koijen, Moskowitz, Pedersen, and Vrugt (2013)), profitability (Novy-Marx (2011)), and defensive or low risk strategies (such as Frazzini and Pedersen's (2012) betting against beta strategy or quality measures from Asness, Frazzini, and Pedersen (2013)) that are not analyzed here for several reasons. The first being to keep the analysis manageable and focus on the cross-sectional characteristics receiving the most attention from both the behavioral and rational asset pricing theories. The second being that many of these other variables are not applicable to sports betting contracts. For example, carry (as defined by Koijen, Moskowitz, Pedersen, and Vrugt (2013) to be the return an investor receives if prices do not change) is literally zero for all sports betting contracts and defensive or low risk strategies such as betting against beta cannot be examined either since beta is zero across all contracts due to their purely idiosyncratic nature.

remarkably consistent patterns across these dozen independent samples make the results very unlikely to be driven by chance, providing a wealth of out-of-sample tests for behavioral theories of value and momentum.

An additional implication of overreaction is that continuation is stronger when there is greater uncertainty about valuations (Daniel, Hirshleifer, and Subrahnayam (1998, 1999), Rabin (2002), Rabin and Vayanos (2010)). Consistent with this idea, stronger momentum effects and weaker value effects are found when there is more uncertainty, such as near the beginning of a season, when the quality of teams is uncertain and for games not heavily involved in parlays, which are bets on multiple contracts where the payoff requires that all the bets win, so games more involved in parlays should be those investors feel more certain about.

Using these insights from sports betting markets, I flip the analysis around and apply this novel idea to financial security returns to establish a further link to capital markets. Firm valuation is more certain immediately following an earnings announcement. Splitting the sample of firms into those who recently announced earnings versus those whose last earnings announcement was several months ago, I find stronger price momentum and weaker value returns for firms with stale earnings (more uncertainty). Using dispersion in analyst forecasts of earnings, I find firms with wide analyst dispersion of opinion have stronger momentum premia and weaker value premia than firms with tight consensus of opinion, consistent with uncertainty strengthening momentum and weakening value. These results match those in sports betting markets and are consistent with overreaction theories.

While the results support the view that momentum and value effects may be related to mispricing due to overreaction, this raises two questions. First, what prevents arbitrageurs from eliminating these mispricings? Using the actual costs of betting on the sample of contracts, the returns to momentum and value are easily wiped out by trading costs in sports betting markets, preventing would-be arbitrageurs from eliminating these patterns in prices.

Second, how generalizable are the results to return premia more generally in financial markets? While sports betting markets isolate tests of behavioral theories from risk-based theories, other differences between sports and financial markets could also matter for generalizing the results. For example, if investor preferences and/or arbitrage activity are vastly different in the two markets then any connection may be tenuous. There are reasons to be both aggressive and cautious in generalizing the results. On the aggressive side, bettors prefer to make rather than lose money,[7] and the experimental psychology evidence motivating the behavioral

[7] In addition, the majority of sports betting volume is comprised of investors who use this market professionally and not simply for entertainment, such as professional gamblers and institutional traders. These include sports betting hedge funds – see Centaur Galileo, a UK-based sports-betting hedge fund that was launched in 2010 but subsequently closed in January 2012. Peta (2013) discusses the industry of professional gambling and the use of financial tools from Wall Street in the sports betting market, including launching his own sports betting hedge fund.

theories comes from generic risky gambles, and hence should apply equally to sports betting contracts as they do to financial securities. Finding that the same predictors in financial markets (momentum and value) also explain returns in sports betting markets, and that both vary with uncertainty in both markets, provides a more direct link that implies either that behavioral biases are (at least partially) responsible for the same cross-sectional return patterns in both markets or that this is just a remarkable coincidence. An alternative is to offer different explanations for the same patterns in different markets, but this seems less satisfying.

On the cautious side, the magnitudes of value and momentum effects in sports betting markets are about one-fifth (per unit of risk) the size of those found in financial markets, perhaps suggesting that the majority of the return premia in financial markets may be coming from other sources. In addition, while significant co-variation in value and momentum returns across securities, markets, and even asset types (Asness, Moskowitz, and Pedersen (2013)) is a feature present in financial markets, there is no covariance structure for value or momentum returns in sports betting contracts. While the lack of common risk in sports betting markets is a virtue allowing behavioral forces to be isolated, it also suggests that these forces are insufficient to drive the common variation we see in momentum and value returns more generally in financial markets. Hence, if a significant part of the return premia to value and momentum in general comes from common sources, the sports betting laboratory has little to say on the majority of these risk premia more generally. Put differently, on the cautious side, the empirical setting in betting markets may identify a behavioral explanation for value and momentum that can only account for a small fraction of the premia found in financial markets.

The rest of the paper is organized as follows. Section I motivates why sports betting markets are a useful laboratory for asset pricing, provides a primer on sports betting, and sets up a theoretical framework for the analysis. Section II describes the data and presents some summary statistics. Section III presents results of cross-sectional asset pricing tests of momentum, value, and size in sports betting markets, and Section V compares the results to those from financial markets, including a novel test in financial markets generated from insights in the sports betting market. Section VI concludes.

I. Motivation, Primer, and Theory

This section discusses why the sports betting market may be a useful laboratory to test asset pricing theory, provides a brief primer on how these markets work, and develops a theoretical framework for the analysis.

A. A Useful Asset Pricing Laboratory

Both financial markets and sports betting markets contain investors with heterogenous beliefs and information who seek to profit from their trades and, like derivatives or the delegated active management industry, the sports betting market is a zero-sum game.(Levitt (2004) discusses the similarities and differences between financial and sports betting markets.) However, there are two key features of sports betting markets that make it a useful laboratory to test behavioral asset pricing theory while abstracting from any risk-based pricing effects. The first is that the cross-section of gambles in these markets are completely idiosyncratic, having no relation to any systematic risk. The second key feature of these markets is that the contracts have a known, and very short termination date, where uncertainty is resolved by outcomes that are independent of investor behavior. The exogenous terminal value allows for the identification of mispricing, providing a stronger test of behavioral models, where any mispricing due to investor behavior will be visibly corrected by the game outcome. The direction of price correction allows for tests of competing behavioral models such as over- and underreaction as discussed below.

Identifying price correction is difficult in financial markets because there is typically no known terminal date for financial securities, and no known or observed true terminal value. In addition, time-varying discount rates may confound mispricing or its correction in financial markets. Sports contracts, being purely idiosyncratic and having very short horizons, eliminate this possible confounding influence.

While several papers study the efficiency of sports betting markets, with the evidence somewhat mixed,[8] this paper is chiefly interested in linking cross-sectional predictors of returns in financial markets with sports betting markets to provide a cleaner test of behavioral asset pricing theory, where aggregate risk and capital market institutional forces have no influence. In addition, this study uniquely takes insights from the sports betting markets through the lens of behavioral asset pricing models and applies a novel test to financial markets. The tests are novel to both the sports betting and asset pricing literatures.

B. Sports Betting Primer

The sports betting market by some counts accomodates more than $500 billion in wagers annually, though no one really knows the exact amounts because sports gambling is illegal in every state except Nevada and hence much of it conducted off shore or under the table.[9] Three separate betting contracts are examined

[8] Golec and Tamarkin (1991), Gray and Gray (1997), Avery and Chevalier (1999), Kuypers (2000), Lee and Smith (2004), Sauer, Brajer, Ferris, and Marr (1988), Woodland and Woodland (1994), and Zuber, Gandar, and Bowers (1985) examine the efficiency of sports betting markets in professional football (NFL) and baseball (MLB).

[9] According to the 1999 Gambling Impact Study, an estimated $80 billion to $380 billion was illegally bet each year on sporting events in the United States. This estimate dwarfed the $2.5 billion legally bet each year in Nevada (Weinberg (2003)).

for each game in each sport: the Spread, Moneyline, and Over/Under contract. Each contract's payoffs are determined by the total number of points scored by each of the two teams, where P_k is the number of points scored by team k.

B.1. Spread contract

The Spread (S) contract is a bet on a team winning by at least a certain number of points known as the "spread." For example, if Chicago is a 3.5 point favorite over New York, the spread is quoted as -3.5, which means that Chicago must win by four points or more for a bet on Chicago to pay off. The spread for betting on New York would be quoted as $+3.5$, meaning that New York must either win or lose by less than four points in order for the bet to pay off. Spreads are set to make betting on either team roughly a 50-50 proposition or to balance the total amount bet on each team, which are not necessarily the same thing, but often are very close (see Levitt (2004)). The typical bet is \$110 to win \$100. So, the payoffs for a \$110 bet on team A over team B on a spread contract of $-N$ points are:

$$\text{Payoff}^S = \begin{cases} 210, & \text{if } (P_A - P_B) > N \quad (\text{“cover”}) \\ 110, & \text{if } (P_A - P_B) = N \quad (\text{“push”}) \\ 0, & \text{if } (P_A - P_B) < N \quad (\text{“fail”}) \end{cases} \tag{1}$$

where "cover, push, and fail" are terms used to define winning the bet, tying, and losing the bet, respectively. For half-point spreads, ties or pushes are impossible since teams can only score in full point increments. The \$10 difference between the amount bet and the amount that can be won is known as the "juice" or "vigorish" or simply the "vig," and is the commission that sportsbooks collect for taking the bet.

B.2. Moneyline contract

The Moneyline (ML) contract is simply a bet on which team wins. Instead of providing points to even the odds on both sides of the bet paying off as in the Spread contract, the Moneyline instead adjusts the dollars paid out depending on which team is bet. For example, if a bet of \$100 on Chicago (the favored team) over New York is listed as -165, then the bettor risks \$165 to win \$100 if Chicago wins. Betting on New York (the underdog) the Moneyline might be $+155$, which means risking \$100 to win \$155 if New York wins. Again, the \$10 difference is commission paid to the sportsbook. The payoffs for a \$100 bet on team A over team B on a Moneyline contract listed at $-\$M$ are as follows:

$$\text{Payoff}^{ML} = \begin{cases} M + 100, & \text{if } (P_A - P_B) > 0 \quad (\text{“win”}) \\ max(M, 100), & \text{if } (P_A - P_B) = 0 \quad (\text{“tie”}) \\ 0, & \text{if } (P_A - P_B) < 0 \quad (\text{“lose”}) \end{cases} \tag{2}$$

where M is either > 100 or < -100 depending on whether team A is favored or team B is favored to win.[10]

[10] Woodland and Woodland (1991) argue that use of point spreads rather than odds that only depend on who wins (such as the Moneyline) maximizes bookmaker profits when facing risk averse bettors. Consistent with this logic, anecdotal evidence

B.3. Over/Under contract

The Over/under contract (O/U), is a contingent claim on the total number of points scored, rather than who wins or loses. Similar to the Spread contract, the O/U contract is a bet of \$110 to win \$100. Sportsbooks set a "total", which is the predicted total number of points the teams will score combined. Bets are then placed on whether the actual outcome of the game will fall under or over this total. If the total for the Chicago versus New York game is 70 points, then if the two teams combine for more than 70 points the "over" bet wins and the "under" bet loses, and vice versa. The payoffs for a \$110 bet on the over of team A playing team B with an O/U contract listed at T total points are as follows:

$$\text{Payoff}^{O/U} = \begin{cases} 210, & \text{if } (P_A + P_B) > T \quad (\text{“over”}) \\ 110, & \text{if } (P_A + P_B) = T \quad (\text{“push”}) \\ 0, & \text{if } (P_A + P_B) < T \quad (\text{“under”}) \end{cases} \tag{3}$$

For the "under" bet on the same game the payoffs for the first and third states are flipped.

B.4. Bookmaking

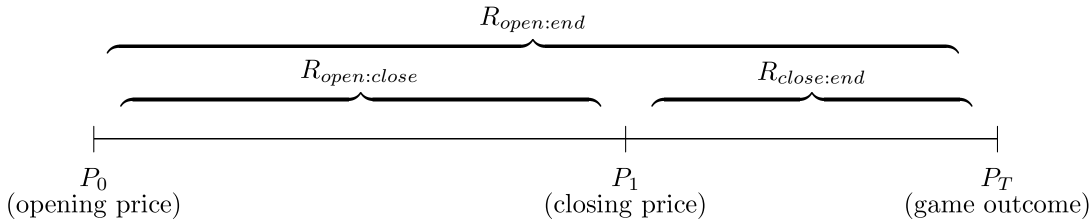
Bookmakers set an initial "line" or "price" on each contract, which are the "opening" lines. Bookmakers set prices to maximize their risk-adjusted profits by either equalizing the dollar bets on both sides of the contract or equalizing the probabilities of the two teams winning the bet, so that they receive the vig with no risk exposure, which are not necessarily the same thing. If bookmakers are better on average than gamblers at predicting game outcomes or can predict betting volume, they may also choose to take some risk exposure to earn even higher profits.[11] I use the empirical distribution of actual betting prices to back out estimated probabilities in order to compute returns, which accounts for whatever bookmakers are doing in setting prices.

Once the opening price is set, betting continues until the start of the game. As betting volume flows, the line can change if the bookmaker tries to balance the money being bet on either side of the contract. For instance, if bettors think the bookmaker has mispriced the contract initially or if new information is released – like an injury to a key player – prices may move after the opening of the contract up until the game starts when a final closing price is posted. Line movements are typically small and infrequent. Bettors receive the price at the time they make their bet, so if the line later changes they still retain the original line at the time they placed the bet. For some contracts (e.g., the NFL) the time between open and close is typically six days, while for others (e.g., the NBA) it may only be a few hours.

suggests that retail and casual bettors prefer Spread bets more than Moneyline bets, which tend to be more dominated by professional or institutional gamblers.

[11] Levitt (2004) finds that NFL bookmakers predominantly do the former, though sometimes also do the latter. In some cases they are good at predicting betting volume and strategically take advantage of investor biases such as over-betting favorites or local teams. However, bookmakers are careful not to distort prices so much as to make a simple betting strategy, like always betting on the underdog, become profitable.

The figure below illustrates the timeline of prices on each betting contract and the three periods over which returns are calculated.



To compute returns, each contract line is converted into a price by estimating the probability, $\pi(S_t)$, of a payoff occurring and calculating the expected value of the contract based on the probability and value of each payoff state for that contract, where the payoff probability depends on the time t betting line (Spread, Moneyline, or Over/under value). To calculate prices and returns, the probabilities of payoff states for each contract for both opening and closing lines are calculated. Two approaches are adopted for computing probabilities. The first is a theoretical approach, where contract prices are set such that the odds of winning and losing each bet are equal. The second approach empirically estimates the probabilities of the payoff states from the data, where a logit, probit, and non-parametric kernel density estimator are used to extract empirical probabilities of the payoff states. This is similar in spirit to the literature that tries to recover real probabilities from risk-neutral probabilities from option contracts (e.g., Ross (2013), Andersen, Fusari, and Todorov (2014), Borovicka, Hansen, and Scheinkman (2014), Jensen, Lando, and Pedersen (2015)). However, the difference here is that sports betting contracts are not confounded by risk premia, hence extraction of the true probabilities from prices is simpler.

Returns for all contracts $c \in \{S, m, O/U\}$ over the three horizons above are then simply,

$$\mathrm{R}^c_{open:end} = \frac{P^c_T}{P^c_0}, \qquad \mathrm{R}^c_{close:end} = \frac{P^c_T}{P^c_1}, \qquad \mathrm{R}^c_{open:close} = \frac{P^c_1}{P^c_0}. \tag{4}$$

The returns are estimated for each contract type (Spread, Moneyline, and O/U) for every game in each sport. Appendix A details how these returns and the payoff probabilities are calculated.

C. A Theoretical Framework

Prices can move from the open to the close for information or non-information reasons, and might respond rationally or irrationally to information. If prices move for information reasons—e.g., a key player is injured after the open but before the game starts—and if the market reacts rationally to the news, then the closing price (which reflects the news) will be a better predictor of the game outcome than the opening price (which did not contain the news). If priced rationally, there will be no return predictability from the close to the end of the game, since the closing price equals the expectation of the terminal value, $P_1 = E[P_T]$. Movement from the opening to the close will therefore not have any predictive value for the return from the close to the end of the game. Since the return from the open to the game outcome is the sum of the returns from the open to the close plus the return from the close to the end, it will also equal the return from open to close in expectation ($E[R_{open:end}] = E[R_{open:close}] + E[R_{close:end}]$, where $E[R_{close:end}] = 0$ if priced rationally). More formally, running the regression

$$R^j_{close:end} \quad = \quad \alpha + \beta_1 R^j_{open:close} + \epsilon^j \tag{5}$$

where returns for contract j are as defined in the previous section, the rational response to information hypothesis predicts,[12]

> **Prediction 1:** If prices move ($P_0 \neq P_1$) for information reasons and markets respond rationally to the news, then $\beta_1 = 0$.

Alternatively, prices could move from the open to the close for purely non-information reasons, such as investor sentiment or noise. In this scenario the closing price will be wrong and will be corrected once the game ends and reveals the true price. Hence, closing prices will be poorer predictors of game outcomes than opening prices, implying predictability in returns from the open-to-close on final payoffs. Moreover, the open-to-close return should negatively predict the close-to-end return as prices revert to the truth at the terminal date, and if there was no information content in the price movement, then prices will fully revert back to the original price at the open. Under this scenario, equation (5) predicts

> **Prediction 2:** If prices move ($P_0 \neq P_1$) for non-information reasons, then $\beta_1 = -1$.

Another possibility is that prices may move for information reasons, but that the market reacts irrationally to the news. For example, markets may underreact or overreact to the information in an injury report or weather forecast. Under- and overreaction are two of the leading behavioral mechanisms proposed in the asset pricing literature (Daniel, Hirshliefer, and Subrahmanyam (1998), Barberis, Shleifer, and Vishny (1998), and Hong and Stein (1999)). In this case, closing prices would still be wrong and would therefore imply predictability of the close-to-end return by the open-to-close return. However, the sign of this return predictability depends on the nature of the misreaction to news on the part of investors. For example, if markets overreact to the news, then the open-to-close return will negatively predict the close-to-end return,

[12] Alternatively, one could run the regression $R^j_{open:end} = \alpha + \beta_0 R^j_{open:close} + \epsilon^j$ and test if $\beta_0 = 1$. Since $R_{open:end} = R_{open:close} + R_{close:end}$, $\beta_0 = 1 + \beta_1$.

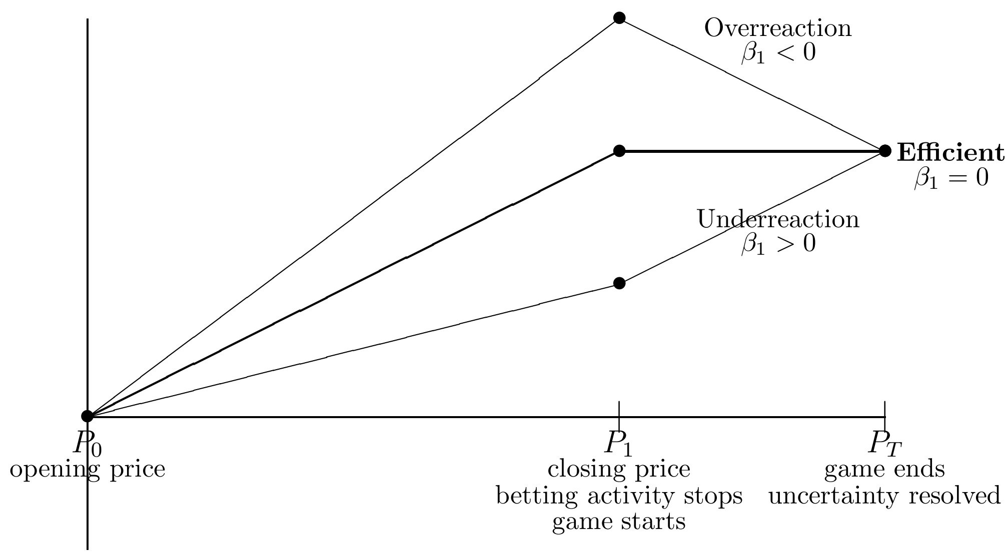
but if the market underreacts to the news, then the open-to-close return will positively predict the close-to-end return. More formally,

> **Prediction 3:** If prices move ($P_0 \neq P_1$) for information reasons but markets respond irrationally to the news, then
>
> (a) $\beta_1 > 0$ if underreaction
>
> (b) $\beta_1 < 0$ if overreaction.

All three hypotheses—rational information, non-information, and irrational information response—make distinct predictions for the regression coefficients from equation (5).

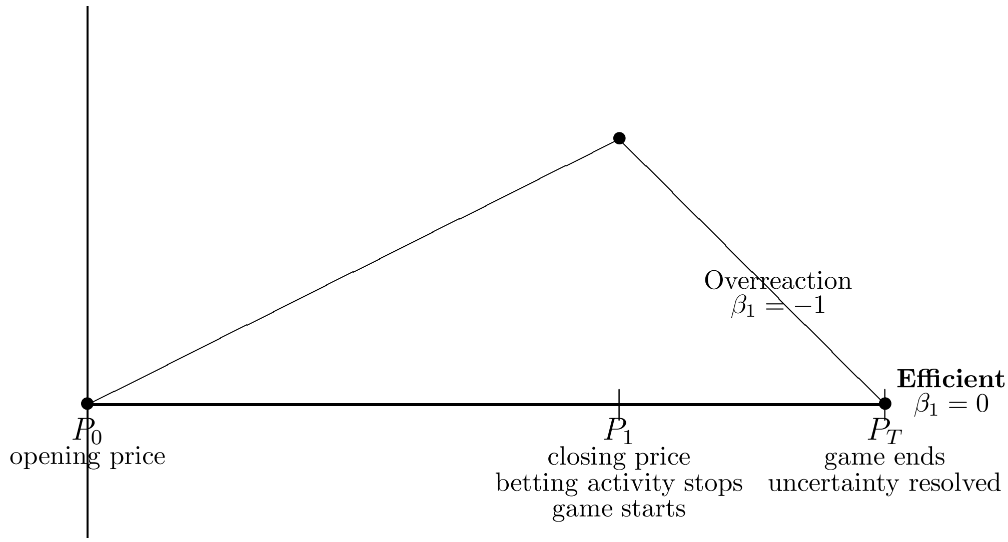
The figures below summarize the implications of Predictions 1 through 3. In the first case, assume that the line movement contains some information, where the market can respond efficiently or inefficiently to the news, with the latter resulting in either over- or underreaction.

Case 1: Response to news



If the market efficiently responds to news, then the closing price is efficient and in expectation equals the terminal value at date T (highlighted in **bold**). If the market misreacts to news, then the betting line could overshoot the true value (overreaction) or underreact to the news. Since the true terminal value at date T remains the same, overreaction implies the return from P_1 to P_T will be negative or opposite that from P_0 to P_1, while in the case of underreaction the returns will be of the same sign.

Case 2: No information (pure noise or total overreaction)



If prices move for purely non-information reasons, then there should be a full price reversal by the end of the game. Since the efficient price response would have been no price change, where $P_0 = P_1 = P_T$, if prices move from P_0 to P_1, then since P_T is exogenous, the return from time = 0 to 1 will exactly be offset by an opposite signed return from time = 1 to T of equal magnitude. In this way, non-information price moves are a special case or extreme form of overreaction.

These tests are unique to the literature and can shed light on asset pricing theory more generally. The general idea of using differences in the ability of the opening versus final point spreads to predict game outcomes to tease out information from sentiment effects is also explored in Gandar et al. (1988, 1998) and Avery and Chevalier (1999), who examine the NBA and NFL, respectively. However, the tests here are novel in several respects. First, they provide a stronger set of tests of the rational and behavioral theories by examining return patterns over different horizons of the contract, whereas the previous papers look only at the close-to-end return. Second, they can differentiate among behavioral theories, namely over- versus underreaction. Third, redefining betting lines in terms of financial returns allows for more powerful tests and more information. For example, Gandar et. al (1988, 1998) and Avery and Chevalier (1999) only look at whether closing lines are better at predicting game outcomes, but neither paper can examine return reversals since they only look at point spreads. In addition, couching the bets in terms of financial returns (per dollar invested) allows comparisons to financial returns in capital markets, which is the goal of this paper. Finally, linking the cross-sectional characteristics from the financial markets literature – value, momentum, and size – to the cross-sectional return patterns in betting markets, is unique to both literatures and is the primary motivation in this paper.

II. Data and Summary Statistics

This study examines the most comprehensive betting data to date, which includes multiple betting contracts on the same game across four different sports, many of which have not previously been explored in depth.

A. Data

Data on sports betting contracts are obtained from two sources: Covers.com (via SportsDirectInc.com) and SportsInsights.com, two leading online betting resources.

The first data set comes from Covers.com via SportsDirect, Inc., who provides a host of historical data on sports betting contract prices, spreads, and outcomes, as well as historical team and game information. The data come from the largest sportsbooks in both Nevada (where it is legal in the U.S.) and outside of the U.S. and pertain only to Spread contracts for four different professional sports:

1. The National Football League (NFL) from 1985 to 2013.
2. The National Basketball Association (NBA) from 1999 to 2013.
3. The National Hockey League (NHL) from 1995 to 2013.
4. Major League Baseball (MLB) from 2004 to 2013.

For MLB and the NHL, Spread contracts exhibit no cross-sectional variation, merely reporting an identical -1.5 point spread for all favorites, but where payouts adjust for the probability of winning, which is the same as the Moneyline contract. These are simply known as "run" lines and "puck" lines in MLB and the NHL, respectively.

The second data set comes from SportsInsights.com and has a shorter time-series, but contains a larger cross-section of betting contracts. The data begin in 2005 and end in May 2013 for all four sports. However, in addition to the Spread contract, the SportsInsights data also contains information on the Moneyline and Over/under contracts. Opening and closing betting lines are provided on all three betting contracts for each game. In addition, information on betting volume (the total number of bets, not dollars) is provided from three sportsbooks per game, which come from Pinnacle, 5Dimes, and BetCRIS who are collectively considered the "market setting" sportsbooks that dictate pricing in the U.S. market. The betting lines are those from the Las Vegas legalized sportsbooks and online betting sportsbooks, where all bookmakers offer nearly identical lines on a given game.[13]

[13]In sports betting parlance, market setting means that other sportsbooks would "move on air" meaning that if one of these three big sportsbooks moved their line, other sportsbooks would follow even without taking any significant bets on the game.

The data from both sources include all games from the regular season, pre-season, and playoffs/post-season. All games include the team names, start and end time of game, final score, and the opening and closing betting lines across all contracts on each game. Both data sets also include a host of team and game information and statistics, which are supplemented with information obtained from ESPN.com, Baseball Reference.com, Basketball Reference.com, Football Reference.com, Hockey Reference.com, as well as the official sites of MLB, the NFL, the NBA, and the NHL.

For each data set, I first check duplicate games for accurate scores and remove lines that represent either just first half, second half, or other duplicate entries that show the same teams playing on the same date (except for a few "double headers" in MLB which had to be hand-checked). I also remove pre-season games, all-star games, rookie-sophomore games, and any other exhibition-type game that does not count toward the regular season record or playoffs. I then merge the two data sets and check for accuracies for the same game contained in both data sets over the overlapping sample periods. When a discrepancy arises in final score (which is less than 0.1% of the time), I verify by hand the actual score from a third source and use the information from the data set that matched the third source. When a discrepancy arises in the point spread (less than 1% of the time), I throw out that game.[14]

Table I summarizes the data on betting contracts across sports. Reported are the sample periods for each sport, number of games, and total number of betting contracts. The NBA covers 18,681 games from 1999 to 2013 and 38,939 betting contracts on those games. The NFL contains 7,035 games and 10,775 betting contracts over the period 1985 to 2013. MLB contains 23,986 games and 47,964 betting contracts from 2005 to 2013. The NHL has 9,890 games and 19,764 betting contracts from 2005 to 2013. Overall, the data contain 59,592 games and 117,442 betting contracts. Table I also reports the distribution of closing lines/prices for each betting contract in each sport. The mean, standard deviation, and 1st, 10th, 25th, 50th, 75th, 90th, and 99th percentiles are reported. The last row in each panel reports the implied probability of the home team winning for each Moneyline value at each reported distributional percentile. The distribution of implied win probabilities is widest for the NBA and NFL and tightest in MLB and the NHL.

B. Return Distributions

Figure 1 aggregates all of the betting returns into a portfolio by placing an equal-dollar bet in every contract and every game. Specifically, every month for each contract type (Spread, Moneyline, and Over/under) an equal-weighted portfolio of the betting returns on all games played that month is created by systematically

[14] Results are robust to taking an average of the spreads or to simply using the SportsInsights spread.

betting on the home team and the favored team for the Spread and Moneyline contracts, and on the over for the Over/under contract (so that the investment is on only one side of the contract). The total return over that month across all of these bets for each contract type and across all sports is computed, equal-weighted across sports. Since no sport has a season lasting a full 12 months, no month contains all four sports with most months containing two and sometimes three sports. Repeating this every month over the sample period produces a time-series of aggregate sports contract returns.

Figure 1 plots the cumulative returns to aggregate sports betting for each of the three contract types along with the cumulative returns on the U.S. stock market (Center for Research in Security Prices value-weighted index) over the same sample period. The sports betting aggregate returns are flat and slightly negative (due to the vig), indicating that systematically betting on the home team, the favored team, or the over is not profitable (i.e., that markets are efficient with respect to these attributes). More importantly, time-series variation in the monthly sports betting returns do not appear to move at all with the stock market. The monthly correlation between the stock market index return and the aggregate return to betting on the home, favorite, or over is 0.06, -0.01, and 0.03, respectively, none of which are significantly different from zero. Across all sports and games, sports betting returns appear to be independent from financial market returns, which is not surprising given they are idiosyncratic bets.

Figure 2 plots the cross-sectional distribution of returns across betting contracts. For brevity, results are reported for the NBA only, but Figure B1 in Appendix B reports the cross-sectional distribution of returns for the other sports. Panel A shows returns from open to close for the Spread, Moneyline, and Over/under contracts and Panel B shows returns from close to end. As the figures show, the open-to-end and close-to-end returns have similar distributions that have a mass at -1 (losing the bet) and significant right skewness for positive payoffs. The open-to-close returns are centered at zero since the majority of the time the betting line does not move, but there is also significant variation in line movements across games.

Looking across contracts, the distribution of Moneyline returns has much fatter tails than Spread or Over/under contracts, which makes sense since Moneyline contracts have embedded leverage in them because they adjust payoffs rather than probabilities. The Spread and Over/under contracts are very similarly distributed. A table of summary statistics on the returns of each contract type is included at the bottom of Figure 2. The Moneyline has the lowest mean, but fattest tails, and the Spread and Over/under contracts are very similarly distributed with a slight negative mean for the Spread contract and a slight positive mean for the Over/under contract, neither of which is reliably different from zero. Return distributions for the other sports are provided in Appendix B and show similar results.

Table II reports return correlations for each of the three betting contracts for each game. The non-bolded numbers in Table II are the correlations *across* contract types. The correlation in returns between the Spread and Moneyline contracts is about 0.69 on average for open-to-end and close-to-end returns, which makes sense since the Spread and ML contracts both bet on a particular team winning, though the former contract adjusts for points while the latter adjusts the payoffs, which is why the correlations are not one. For the O/U contracts, however, the correlation of returns with both the Spread and ML contracts are zero at all return horizons. Essentially, the O/U contract provides an uncorrelated bet on the *same* game, which provides a set of independent return observations which can be used to test for cross-sectional return predictability for the same game.

Within each contract type, the correlations of returns at different horizons are highlighted in **bold**. Not surprisingly, the open-to-end return is very highly correlated with the close-to-end return: 0.96 for Spread, 0.99 for Moneyline, and 0.94 for O/U contracts, which simply reflects the fact that lines do not move much between the open and the close. However, when the lines do move between open and close, the return from open-to-close is slightly positively correlated with the open-to-end return and is negatively correlated with the close-to-end return.

C. Hypothetical "Point Returns"

Since the payoffs to the betting contracts are discrete, betting outcomes may truncate useful information. For example, suppose there are two games facing the exact same point spread of -3.5, but in one game the favored team wins by 4 points, while in the other game the favored team wins by 20 points. In both cases, the spread contract pays off the same amount. However, treating these two games equally throws out potentially useful information since one team barely beat its spread, while the other exceeded its spread by a wide margin. In order to extract more information from these betting contracts, I also compute hypothetical returns from points scored rather than simply the discrete dollar outcomes of the contracts. Specifically, I compute hypothetical "point" returns by replacing the dollar payoffs with the actual points scored. These "point returns" use more information from the distribution of game outcomes and can be applied to both Spread and O/U contracts. For Moneyline contracts the concept of point returns does not apply since those contracts are simply bets on who wins or loses, where the payoffs have already been adjusted to reflect the likelihood of who wins.

The correlation between the dollar returns and the hypothetical point returns is 0.79 for open-to-end and close-to-end returns, indicating that the returns are highly correlated, but that there is also additional information in the point returns. For the open-to-close returns, the correlation between dollar returns and point returns is around 0.28. Results are reported in Table B1 in Appendix B.

III. Cross-Sectional Asset Pricing Tests

I start with a general test of separating information versus sentiment-based price movements by first estimating regression equation (5) to test Predictions 1 through 3 generally in these markets, and then focus on applying these tests to the cross-sectional characteristics of value, momentum, and size.

A. Testing General Price Movements

Panel A of Table III reports results for the full sample of bets for each sport separately. The first row reports results for the NBA for all three betting contracts, which shows a consistently strong and highly significant negative coefficient for β_1 for all three betting contracts, indicating that the close-to-end return is strongly negatively related to the open-to-close price movement. These results reject Prediction 1—the rational informationally efficient hypothesis. The regression coefficients for all three contracts are also statistically different from -1, thus rejecting Prediction 2 – the pure noise hypothesis. The results are most consistent with Prediction 3b, the overreaction to information hypothesis. The magnitude of the coefficient, which is around -0.50 for the Spread and Over/under contracts, suggests that about half of the total price movement from open to close is reversed at game outcome, suggesting that the market overreacts on average by about 50 percent to price-relevant news (or that half of the betting line movements contain no information).

The next three rows of Panel A of Table III report results for the NFL, MLB, and NHL betting contracts. Results for the NFL contracts are very similar to those for the NBA, with negative coefficients of around -0.50. For MLB and the NHL, where only Moneyline and Over/under contracts are available, the results are also similar.

Panel B of Table III repeats the regressions excluding the betting lines where there was no price movement. The results are nearly identical, suggesting that the conditional expectation of close-to-end returns conditional on no price movement from open to close is also zero (e.g., $E[\frac{P_T}{P_1}|\frac{P_1}{P_0} = 0] = 0$).

B. Microstructure Effects?

The results in Table III point to overreaction in betting markets according to Predictions 1 – 3. However, an alternative hypothesis for the negative relation between open-to-close and close-to-end returns that is outside of the theories in Section II is microstructure effects, which could confound the asset pricing tests for investor behavior. For instance, if bettors are better informed than bookmakers, then like the market makers in Stoll (1978), the bookmakers may set closing prices that do not reflect the average betting price in order to shield their exposure to adverse selection. In this case, they may move prices more aggressively with volume, which could induce negative correlation between the opening return and the closing return.

To test this hypothesis, Panel C of Table III reestimates equation (14) for the highest and lowest betting volume games across all sports. Specifically, sorting contracts in each sport into the top and bottom third of betting volume (number of contracts), β_1 is estimated separately for the high and low volume contracts. If microstructure is driving the negative coeffiecient on β_1, then there should be a stronger negative coefficient for the most heavily bet games since this is where bookmakers face the greatest risk of big losses. As Panel C shows, however, the coefficients are, if anything, more negative for the low volume games, though the last row of Panel C shows that the differences between high and low volume coefficient estimates are statistically indistinguishable from zero. The results are inconsistent with a microstructure story for $\beta_1 < 0$.

Another test of the microstructure story is to directly examines cases where bettors may be better informed. The volume data also contains information on the number of contracts involved in "parlays." A parlay bet is a portfolio of bets (across games or across type of contract), where the bet only pays off if all contracts win, where the payout is levered. For example, a two, three, or four game parlay pays out 13/5, 6/1, or 10/1. Since parlay bets only pay out if all bets in the parlay win, bettors should include the games in a parlay they are most confident about. Hence, contracts involved more heavily in parlays should be those where bettors are more informed (or at least think they are). According to the microstructure story, these will also be the games bookmakers (who also see the parlay volume) would adjust prices more aggressively as a result.

Panel D of Table III sorts games by parlay volume relative to total volume and reestimates β_1 for the highest and lowest third of games involved in parlays (as a percentage of total volume). The difference in beta estimates is negligible and no different from zero, inconsistent with microstructure effects driving the negative relation between open-to-close and close-to-end returns.

The results in Table III across different betting contracts and different sports appear most consistent with overreaction. However, the question in this paper is not whether sports betting markets in general are prone to investor sentiment, but rather whether the same cross-sectional return predictors found in financial markets have import for sports betting markets. When lines move from open to close is it because investors chase returns by following momentum, value, or size-related signals? And, does such price movement reverse at the game outcome? Or, are the characteristics that predict returns in financial markets unrelated to return predictability of sports betting contracts? These questions are investigated next.

C. Cross-Sectional Return Characteristics

The primary goal of this paper is to investigate whether the cross-sectional characteristics of momentum, value, and size (all defined below) are related to return predictability in sports betting markets. I first examine whether movement from the open to the close is related to these characteristics by running the following regression:

$$\tilde{R}_{i,0:1} \quad = \quad \alpha_1 + \beta_1 Char_i + \tilde{\epsilon}_{i,0:1}, \tag{6}$$

where $Char_i = \{\mathrm{M}om, \mathrm{V}al, \mathrm{S}ize\}_i$ is the characteristic of contract i.

Following equation (5) and the logic of the previous subsection, close-to-end returns are regressed on the same characteristic, where the sign of the coefficient indicates what theories most likely explain any relation between the characteristic and open-to-close returns. The regression is:

$$\tilde{R}_{i,1:T} \quad = \quad \alpha_T + \beta_T Char_i + \tilde{\epsilon}_{i,1:T}. \tag{7}$$

In effect, the characteristics momentum, value, and size are being used as instruments for betting line movements, where equation (6) is the first stage and equation (7) is the second stage regression, represented in reduced form.

The characteristics could affect opening betting lines, closing lines, or movement from open to close and hence affect return over the three horizons considered. The pattern by which these returns are affected by each characteristic helps distinguish among various theories, summarized by the following hypotheses:

1. **H1: No Relevance.** The characterstic is not related to either information or sentiment $\Rightarrow \beta_1 = \beta_T = 0$.

2. **H2: Information Efficiency.** The characteristic is related to information that is priced efficiently $\Rightarrow \beta_1 = \beta_T = 0$.

3. **H3: Non-information/Noise.** The characteristic is related to non-information or pure noise that erroneously moves prices $\Rightarrow \beta_1 \neq 0$, $\beta_T = -\beta_1$.

4. **H4: Information Inefficiency/Sentiment.** The characteristic is related to information, which moves prices ($\beta_1 \neq 0$), but the market responds inefficiently to that information. There are two types of misreaction:

   (a) **Underreaction** $\Rightarrow \beta_1 \times \beta_T > 0$

   (b) **Overreaction** $\Rightarrow \beta_1 \times \beta_T < 0$.

Hypotheses 1 and 2 cannot be distinguished because both imply no relation between returns and the cross-sectional characteristics over any horizon. Under H1 the characteristic either has no information value or is not an attribute bettors care about. Under H2 the characteristic has relevant information content, but it is priced efficiently so there is no predictability in returns.

Hypotheses 3 and 4 are behavioral models. As Barberis and Thaler (2003) summarize, the behavioral models deviate from rational expectations in one of two fundamental ways: differences in beliefs or differences in preferences. Under H3, prices move for non-information reasons, such as preferences for a certain team due to geographic proximity, the color jersey they play in, etc. Alternatively, investors may use signals that are pure noise, but erroneously believe they have information content. In either case, if prices move for non-information reasons, prices are inefficient and returns will be predictable, where movement from the open to the close will reverse sign from the close to the end of the game. If non-informative signals are related to momentum, value, or size, then those patterns will show up in regression equations (6) and (7), where movement in prices from open to close will be completely reversed from the close to the game outcome (e.g., $\beta_T = -\beta_1$).

H4 is more about differences in beliefs, where the characteristics may have information content, but the market misreacts to that information – under- or overreacting – causing mispricing. Behavioral asset pricing models are often motivated by, and experimental evidence from psychology suggests, individuals underreact to mundane pieces of information and overreact to dramatic news (see Kahenman and Tversky (1979) generally, and for references to financial applications see Barberis and Thaler (2003)). Two of the most prominent behavioral theories for momentum and value in financial markets focus on these different aspects of information misreaction. Barberis, Shleifer, and Vishny (1998) and Hong and Stein (1999) derive a model of underreaction to explain momentum and subsequent correction to capture the value effect. Daniel, Hirshleifer, and Subrahmanyam (1998) have a model of delayed overreaction that generates momentum in the short-term and eventual reversals in the long-term related to valuation ratios.

Regression equations (6) and (7) help distinguish between these competing behavioral theories if the characteristics are associated with moving betting prices. For example, underreaction implies that there is return predictability from open-to-close and from close-to-end that is of consistent sign. Underreaction implies markets slowly respond to the same information, where prices are updated in the same direction. In this case, the total return from open-to-end is greater than the return from close-to-end as the closing price will be closer to the true price.

Overreaction, on the other hand, implies that prices deviate further from the truth as investors continue to overreact to news. In this case, if prices are initially too high, for example, because investors overreacted to good news, delayed overreaction will cause prices to move even higher by the close, which implies that the return from open-to-close will be positive and the return from close-to-end will be negative; the opposite of the underreaction implication. The absolute magnitude of the close-to-end return will be greater than the open-to-end return as a result, which is again opposite to that predicted by underreaction.

Of course, a combination of effects is also possible. Opening prices may be inefficient and closing prices efficient, or vice versa. Looking at opening price, open-to-close, and closing price returns tests whether opening or closing prices are affected differently by the same characteristics and allows inferences for how prices are determined in this market through the hypotheses above.

D. Measuring Cross-Sectional Characteristics

The goal is to derive cross-sectional characteristics of the sports betting contracts that are analogous to those that describe the cross-section of returns in financial markets. The measures are illustrated for the NBA, but similar measures are used for other sports.

D.1. Momentum

The easiest characteristic to match to financial markets is momentum, since it is typically measured based on past performance, which is also easily defined for sports betting contracts. The literature uses the past 6 to 12 month return on a security to measure momentum, known as "price" momentum. Other measures of momentum related to earnings (such as earnings surprises and earnings momentum, see Chan, Jegadeesh, and Lakonishok (1996), Novy-Marx (2015)) are termed "fundamental" momentum.

Analogous "price" and "fundamental" momentum measures for sports betting contracts are constructed from past performance. Given the short maturity of the contracts, the horizon relevant for momentum is likely different than in financial markets. Theory provides little guidance on what horizon is appropriate for momentum in general.[15] Therefore, I construct a number of past return measures over various horizons, much like Jegadeesh and Titman (1993) did in their original study. Specifically, lagged measures of performance based on: wins, point differential, dollar returns on the same team and contract type, and point returns on the same team and contract type over the past 1, 2, . . ., 8 games are constructed. Eight games is roughly 10% of the NBA season (each team plays 82 games in the regular season).

[15] The theoretical literature has thus far been silent on the question of horizon for momentum. Empirically, momentum is found for past returns less than 12 months, both in the cross-section (Jegadeesh and Titman (1993) and Asness (1994)) and in the time series (Moskowitz, Ooi, and Pedersen (2012)), with subsequent reversals occurring two to three years after portfolio formation.

A series of out-of-sample tests are also conducted to assess the robustness of these measures and to guard against data mining that might overfit the sample. These out-of-sample tests include applying the same measures analyzed in one sport to all other sports as well as finding measures from one time period and applying it to other time periods – for example, using even-numbered years to test momentum measures of various horizons, and then examining those measures out of sample in odd-numbered years. Results are stable along two dimensions. First, variables that appeared stronger in even years were also stronger out of sample in odd years. Second, (and shown below) the horizons over which momentum are measured generate similar results. These selected measures for the NBA are then applied out of sample to other sports, where, as will be shown, the results are also consistent.

The first two momentum measures based on past wins and point differential are at the team level and the last two measures based on dollar and point returns are at the betting contract level (Spread, Moneyline, O/U), with the former being "fundamental" momentum and the latter "price" momentum.[16] A momentum index of all measures is also created by taking a principal component-weighted average of the momentum measures from their correlation matrix (since the variables are in different units and scales).

Since every game is a contest between two teams, the momentum measure for the betting contract is the *difference* between the team momentum measures. For example, the momentum measure on a Spread contract bet on team A versus team B using past dollar returns over the last N games is,

$$MOM_t^{\$ret,N}(\text{A vs. B}) = \sum_{g=1}^{N} \text{R}_{t-g,close:end}^{S,\$}(\text{A}) - \sum_{g=1}^{N} \text{R}_{t-g,close:end}^{S,\$}(\text{B}), \quad (8)$$

where $\text{R}_{t-g}^{S,\$}(\text{A})$ is the dollar return from betting on team A in a Spread contract in the most recent g^{th} game that team A has played prior to time t. Equation (8) can be easily estimated for other momentum measures by substituting point returns, wins, or point differentials in place of the past dollar returns. In addition, equation (8) can be estimated for other sports betting contracts such as the Moneyline, where the past Spread contract returns are replaced with the past Moneyline contract returns.

For Over/under contracts, which is a bet on the total number of points scored rather than who wins and by how much, the *sum* of the two team's momentum measures is taken rather than their difference. However, not all momentum measures make as much sense here since the O/U contract is a bet on total points scored and not who wins, momentum measures based on past wins should be much weaker predictors for O/U contracts than for Spread or ML contracts.

[16] All lagged measures within a season only pertain to games within that season. Games from the previous season are not used to construct past game performance measures since there is a significant time lag between seasons where teams can change significantly. Hence, for the NBA, the first set of betting contracts each season starts with game nine when the momentum measure requires an eight game lag.

Panel A of Table IV reports the correlations of the various momentum variables across games for different lags. The measures are all highly correlated, with the "fundamental" team-level momentum measures (win percentage and net points) having a 0.81–0.88 correlation with each other, the "price" contract-level momentum measuures (dollar and point returns) having a 0.78–0.79 correlation to each other, and the correlation between fundamental and price momentum being around 0.65.[17] The last row of Panel A of Table IV reports the correlation of the momentum index (Mom_{index}) with each group of momentum measures averaged across all lags, which is highly correlated with each momentum measure.

D.2. Value

While analogous momentum measures in sports betting markets are straightforward and intuitive, value is a more difficult characteristic to match to financial markets. A general measure of value applied to diverse asset classes is one that looks "cheap" following Asness, Moskowitz, and Pedersen (2013). In equities, value is often measured by the ratio of book value of equity-to-market value of equity or some other ratio of "fundamental" value to market value of the firm (e.g., E/P, D/P, CF/P; see Fama and French (1992, 1993, 1996, 2012), Lakonishok, Shleifer, and Vishny (1994), Asness, Moskowitz, and Pedersen (2013), and Israel and Moskowitz (2013).) Another value measure used that is highly correlated with fundamental-to-market value ratios is the negative of the long-term past return on the asset, following DeBondt and Thaler (1985, 1987), Fama and French (1996), and Asness, Moskowitz, and Pedersen (2013), who apply value measures to equities, futures, fixed income, currencies, and commodities, and show that long-short equity strategies sorted on the negative of past five year returns are 0.86 correlated to strategies formed on book-to-market equity in both the U.S. and globally across a dozen developed equity markets.

Using these concepts, a number of value measures motivated by the finance literature are constructed which can be grouped into four categories:

1. **Long-term past performance,** measured over the previous one, two, and three seasons.
2. **Team Fundamental-to-market ratios,** book values of the team franchise (e.g., book value of equity, ticket revenue, total revenue, player payroll) divided by the current price on the Spread contract.

   The idea is to use a measure of team fundamental value scaled by the current market price, in this case the spread on the game itself. For example, differences in player payroll between the two teams divided by the spread captures a notion of "cheapness," where if the labor market for athletic talent is fairly efficient, then two teams facing different payrolls should in principal face different probabilities of winning the game. Since payroll is a slow-moving long-term measure of value/team quality, and the spread provides the market's immediate assessment of how likely the team will win, then a game

[17] These correlations among different momentum measures are actually very similar to those used in financial markets, where, for instance, Chan, Jegadeesh, and Lakonishok (1996) find that price and earnings momentum are about 0.60 correlated.

that has big differences in payroll and little differences in spread (or opposite signed differences in the spread), will look "cheap," or a value bet.

3. **Game Fundamental-to-market ratio,** following the same theme, the ratio of the fundamental value of the game itself relative to its price. The sports analytics community has derived a number of measures of team quality or strength for use in predicting wins. The most popular is known as the Pythagorean win expectation formula. Appendix C provides details and intuition for this formula, which has been applied to all sports. The formula provides an expected win percentage for each team, which is used as a relative strength measure by taking the difference between the measures for each team and then dividing that difference by the current betting line or contract price, $E(P)/P$, which is a value measure.

4. **Residual measures.** Finally, another way to measure value is to directly try to identify surprisingly cheap or expensive contracts by calculating the expected Spread (or Moneyline or O/U total) based on observable information and use deviations of the actual Spread from its expectation as a value indicator. Contracts with positive residuals are expensive and those with negative residuals are cheap relative to observable information.

A principal component-weighted average of the value measures within each category and across all categories from their correlation matrix is used to compute a value index.

Panel B of Table IV reports the correlations across the value measures. The long-term past return measures are highly correlated with each other, as are the team fundamental-to-price ratios, with correlations ranging from 0.52 to 0.97. The game fundamental-to-price ratios are weakly positively correlated to the team measures as well. However, the correlations between the long-term performance measures and the fundamental ratio measures are negligible. The regression residual is uncorrelated to any of the other value measures, which is not surprising since it controls for long-term observable variables (which the other value measures are based on) and hence largely captures short-term or unobservable information related to deviations in price. The value index has correlations of about -0.30 with long-term past performance and correlations of about 0.45 to fundamental-to-price ratios. These correlations are consistent with what is found in financial markets, where long-term past performance is negatively correlated to fundamental-to-price ratios. (For equities in the U.S., the average correlation between long-term past returns and book-to-market equity ratios is -0.40, for instance.)

The value measure for each game is the *difference* between the team value measures for the Spread and Moneyline contracts and the *sum* of the two measures for the Over/under contracts.

D.3. Size

Size is simply measured as the annual franchise value, ticket revenue, total revenue, and player payroll. These measures are highly correlated with the size of the local market in which the team resides as well as the popularity of the team, which is a function of many things including long-term historical performance.

These are very slow moving variables whose cross-sectional ranking does not change much over time (e.g., the New York teams are always "larger" than the teams in Pittsburgh). Panel C of Table IV shows that the size measures are all highly correlated with each other, with payroll being the lowest around 0.55.

Finally, Panel D of Table IV reports the correlations of the momentum, value, and size indices. Consistent with the measures used in financial markets, the value and momentum measures are negatively correlated, though not as strongly as they are in financial markets (Asness, Moskowitz, and Pedersen (2013)), and so are the size and value measures (Fama and French (1992, 1993)). Size and momentum are slightly positively correlated. The correlations across characteristics are consistent with those used in financial markets.

D.4. Cross-validation

While it is nearly impossible to get consensus on what the right momentum, value, and size measures are (and this is true in general) the above measures seem to reasonably match those examined in financial markets. The use of many measures allows for an assessment of the robustness of the results and using an average across measures helps reduce noise (see Israel and Moskowitz (2013)). Additionally, the many out-of-sample tests performed should alleviate any data mining or overfitting concerns.

I also asked two of the leading scholars from both sides of the efficient markets debate: Eugene Fama (proponent of the rational risk-based view and winner of the 2013 Nobel Prize in Economic Sciences in part for his work on efficient markets) and Richard Thaler (pioneer of behavioral finance and former co-director of the Behavioral Finance working group at the NBER) to weigh in on the plausibility of these measures *before* they (or I) saw the results.[18] The following are quotes from each when asked whether they thought these were reasonable measures of momentum, value, and size analogous to those used in financial markets:

> **Fama:** "Most of these make sense to me. . . . I like past team record longer-term for value, shorter-term for momentum. But the rest seemed ok."
>
> **Thaler:** "Momentum is easier. For value, since that's my [*referring to long-term past performance*] measure with DeBondt, I guess I have to like that one. I also like the difference in power rankings or quality divided by contract price as a measure of what behavioralists think of value."

Armed with these measures for every game, regressions (6) and (7) are estimated to test whether each characteristic is related to the cross-section of returns in these markets. A full set of results for the NBA is reported and then an abbreviated set of results for the other sports are reported for brevity.

[18] This way nobody could complain ex post about the measures if the results failed to confirm their priors.

E. Momentum

Table V reports results from estimating equations (6) and (7) by regressing the returns from open-to-close and close-to-end, respectively, on the momentum measures.[19] The first row of Panel A of Table V reports results for estimates of β_1 for the open-to-close returns of Spread contracts. For the point differential momentum measures, there is a strong positive relation between momentum and the open-to-close return at a one game lag, and a weakly positive relation between momentum and returns at 4 and 8 game lags, with similar sized coefficients. Using the past dollar return momentum measures, there is an even stronger positive relation with open-to-close returns at all game lags, with the strongest being at a one game lag. These results indicate that movement in the betting line from the open to the close is positively related to recent past performance, consistent with momentum, and where fundamental and price momentum measures yield similar results. The last column reports results for the momentum index, which shows a positive and significant coefficient (t-statistic = 2.34).

The second row of Panel A of Table V reports results for estimates of β_T for the close-to-end returns on the momentum measures. The coefficients are predominantly negative, and marginally significant for the most recent past dollar returns. These results indicate a reversal of the momentum effect that moved prices from the open to the close. This result is consistent with a model of overreaction for momentum, where investors push prices too far from the open to the close based on past performance that then reverses at the game outcome. The third row reports a test of the difference between the two return responses, $\beta_1 - \beta_T$, which captures the magnitude of the initial price move plus the subsequent reversal from momentum. The magnitude of the movement is uniformly positive and in the majority of cases statistically significant, indicating momentum is associated with moving betting lines. The last row reports the sume of the two returns responses, $\beta_1 + \beta_T$, which captures the cumulative effect on total returns over the life of the contract − from the open to game outcome. The sum of the coefficients, representing the cumulative return effect is uniformly zero across all momentum measures, indicating that $\beta_1 = -\beta_T$ so the initial movement in betting prices from momentum is fully reversed by the game outcome. These results are consistent with bettors overreacting to recent past performance that appears to have little information content. The overreaction to recent past performance is consistent with the long literature on the hot hand fallacy in sports (Gilovich, Tversky, and Vallone (1981)), where regardless of whether there is or is not a hot hand in sports performance, *belief* in the predictive power of recent performance seems to greatly outweigh what little evidence might exist for the hot hand (Moskowitz

[19] The win percentage and point differential momentum measures are highly correlated and yield similar results, hence for brevity, I only report the point differential measures.

and Wertheim (2011)). At the team level, which is the focus here, there is no systematic evidence of recent past performance predicting future performance (Moskowitz and Wertheim (2011)).

Since the total return from open to end is zero, the evidence supports the opening price being set efficiently by bookmakers on average with respect to momentum/recent past performance, but bettors seem to chase past performance pushing the closing price in the direction of momentum, which gets reversed by the game outcome. Bettors seem to overreact to past performance that causes positive momentum from the open to the close, and then a full reversal from the close to the game outcome when the true terminal value is revealed. These patterns are consistent with the implicatios of the overreaction model for momentum of Daniel, Hirshleifer, and Subrahmanyam (1998).

The remaining stanza of results in Panel A of Table V repeats the regressions using hypothetical "point" returns instead of dollar returns, where the point returns may capture more information since their payoffs are not truncated. Consistsent with this notion, the results from point returns are stronger and match directionally the results using dollar returns: momentum has strong positive predictability for open-to-close returns that reverses for close-to-end returns, consistent with overreaction to recent past performance.

Panel B of Table V repeats the same regressions for the Moneyline contracts. For Moneyline contracts, only dollar returns are available. The results are mixed. There is weak positive predictability from open-to-close that is not significant (except the past four games using dollar returns) and the reversals are negligible. The Moneyline contract is said to be more popular among professional bettors than the Spread or O/U contracts, and hence may be less prone to momentum patterns, but the direction of results is at least consistent with Panel A.

Panel C of Table V looks at the same regression results for the Over/under contracts. Recall, that the returns of the Over/under contract are uncorrelated with those on the Spread contract (see Table II). Despite this zero correlation on average, the results for momentum on Over/under contracts are very consistent with those for the Spread contract in Panel A, providing an independent test. There is strong positive predictability from momentum on open-to-close returns using past point differentials, which makes sense since this contract is based on total points scored and points may be more salient to bettors, and a significant reversal from close-to-end that appears to fully reverse that price movement. In addition, while past win percentage delivers positive momentum for the Spread contract, which is a bet on who wins, it does not predict Over/under returns well, which is also consistent with this conjecture. Overall, Table V provides evidence across three different betting contracts generally consistent with a delayed overreaction story for momentum.

Table VI reports results on momentum for the other sports, estimated separately for the NFL (Panel

A), MLB (Panel B), and the NHL (Panel C) (there are no Spread contracts for MLB and the NHL). The results are consistent with those for the NBA in Table V – a generally positive effect on closing prices from momentum, followed by a reversal, where the total price movement is significant and the cumulative return effect from open to game end is zero, consistent with a full reversal. Like the NBA, results in the NFL for the Moneyline contract are weaker. However, in MLB and the NHL, where Spread contracts do not exist, there are stronger momentum effects in the Moneyline contract. Finding the same patterns in three other distinct sports and three different betting contract types that are independent across sports using the same momentum measures provides a compelling set of out-of-sample evidence.

While the coefficients are of consistent sign in every sport and contract type, their statistical significance is sometimes lacking, which could be due to low power. To increase the power of the tests, Panel D of Table VI reestimates the regressions pooling all four sports (including the NBA) into one regression, with sport fixed effects. The results show a statistically strong and significant positive momentum pattern from open to close and then an equally strong reversal from close to end. The t-statistics on the positive momentum coefficients from open to close are all large and significant, ranging from 3.66 for dollar returns to 10.19 for "point returns" and the negative reversals from close to end exhibit significant t-statistics ranging from -2.33 to -2.97. The differences between $\beta_1 - \beta_T$, representing the magnitude of the absolute price movements from open to close and then reversing from close to end, are large and significant with t-statistics of 2.66 for O/U contracts and 4.26 for Spread contracts based on dollar returns. For hypothetical point returns, the t-statistics are almost 7.0. The sum of the coefficients are also close to zero and generally indistinguishable from zero, indicating that the price movement based on momentum is fully reversed by the game outcome.

Figure 3 summarizes the results visually across all four sports by plotting the t-statistics of the betas on momentum (using the average across all momentum measures) over each return horizon, for each betting contract type, in each sport. The regressions are estimated for each momentum measure from Section III individually and then the t-statistics from these regressions are aggregated across all momentum measures and plotted over each return horizon for each sport separately in Panels A through D as well as for the NBA and NFL combined (Panel E) and across all four sports (Panel F) combined.

A consistent pattern emerges, especially for the Spread and Over/under contracts in every sport: the momentum returns exhibit a tent-like shape over the three horizons – near zero from open-to-end, significantly positive from open-to-close, and significantly negative from close-to-end, with the initial price movement from open-to-close from momentum being fully reversed by the game outcome. The patterns for the Moneyline contracts exhibit the same tent-like shape, but are much less pronounced. These patterns mirror those implied by the overreaction model in Section II, exemplified by the illustrative figures in that section that resemble the empirical patterns in Figure 3.

F. Value

Table VII reports regression results from estimating equations (6) and (7) for the value measures for the NBA. Panel A contains the results for Spread contracts, and Panels B and C report results for the Moneyline and Over/under contracts. Each value measure is signed so that it can be interpreted as a measure of cheapness (i.e., negative of past performance, negative of residual) and in this way have the opposite expected sign as the momentum variables.

The first row of Panel A of Table VII reports the value regression results for open-to-close returns. Almost all of the coefficients on the value measures are negative. The second row reports results for close-to-end returns, where the coefficients on value are predominantly positive. Cheap contracts get cheaper between the open and close, and the positive returns from close to end indicate that part of this price movement is inefficient. These patterns are consistent with bettors overreacting to news that generates a value effect, where there is a negative relation between future returns and fundamental-to-price ratios or negative long-term past returns. However, most of the coefficients are not reliably different from zero, save for the past one and two seasons of return performance.

A generous interpretation emphasizes that the point estimates tell a consistent story that matches overreaction and is consistent with the opposite-signed results for momentum, where the lack of significance could be due to low power and the difficulty and noise in measuring value. A less generous interpretation argues that value may have no predictive content in this market either because the general return premium for value in financial markets is related to systematic risk, which is absent in sports betting, or because value is a much slower moving variable than momentum and the market is therefore better able to price it.

The bottom two rows of Panel A of Table VII repeat the regressions using the hypothetical "point returns," which may contain more information. There is stronger predictability for value on the returns from the open to the close and from close to end, in a direction consistent with overreaction.

Estimating the regressions for the Moneyline (Panel B) and Over/under (Panel C) contracts, similar results are obtained: consistently negative coefficients on value measures from open-to-close and positive coefficients from close-to-end, all of which are supportive of an overreaction story, but the evidence is statistically weak.

Table VIII reports results for the value regressions among the other sports. Just the composite value index is used for brevity. The results are consistent with those shown in Table VII for the NBA. To increase the power of the tests, Panel D of Table VIII repeats the regression pooling all sports (NBA, NFL, MLB, and NHL). The results show a statistically strong negative coefficient on value for open-to-close returns, consistent with an overreaction story and consistent with the increased power of the tests coming from combining all sports. However, the coefficients from open to end, though positive, remain insignificant. Overall, the magnitude of price movements is significant, while the cumulative return is no different from zero, consistent with price movement and then correction from initial overreaction.

Figure 4 summarizes the results for value across all four sports by plotting the average t-statistics of the value betas estimated for open-to-end, open-to-close, and close-to-end returns for each betting contract type. A consistent pattern is evident, where the return to value is negative between the open and close and then fully rebounds positively between the close and game end for a total return of zero. These patterns are consistent with an overreaction story, where cheap or value contracts continue to get cheaper between the open and the close, becoming too cheap and then rebounding positively by game end. This picture is the mirror image of momentum, where value is negatively related to past performance, and hence the pictures for momentum and value tell the same story. (Though, recall the measures for value and momentum were only mildly negatively correlated.)

G. Size

Table IX reports results for regressions on size. There is no return predictability for size – statistically, economically, or even of consistent sign – for any contract over any horizon, regardless of the size measure used. Size, which is the slowest moving of all the variables examined, seems to be either efficiently priced in both opening and closing prices, or is irrelevant to sports bettors. Table X repeats the size regressions for the other sports and similarly finds nothing.

Table B4 in the appendix reports results for multivariate regressions that use momentum, value, and size simultaneously in the regressions, using the index measures of momentum, value, and size. The results are largely consistent with the univarirate regressions for each characteristic.

IV. Link to Financial Markets

Given the existence of momentum and value return premia in idiosyncratic sports betting contracts, a natural question arises as to how to compare these returns to analogous momentum and value return premia in financial markets? In addition, since the return patterns in sports betting markets are consistent with investor overreaction, does this help shed light on the patterns found in financial markets?

A. A Further Test of Overreaction

Theory suggests that overreaction effects are stronger when there is more uncertainty. For example, Daniel, Hirshleifer, and Subrahmanyam (1998) model overreaction from investor overconfidence that causes the investor to trade too aggressively on information that eventually gets corrected by future information the investor slowly updates to. Greater uncertainty implies a lack of precise information to update to, which allows the overreaction to continue unchecked. Another possible connection between uncertainty and overreaction comes from the "law of small numbers" of Rabin (2002) (and Rabin and Vayanos (2010), as well as initial motivation from Kahneman and Tversky (1971, 1974)). The idea is that people treat small samples as if they represent the true population. This "representativeness" bias leads to two possible judgement errors. In cases where the process (i.e., the probability of outcomes) is known, individuals fall prey to the gambler's fallacy, where they erroneously predict reversals in order to match population moments in their small sample. For example, a fair coin flipped heads previously will more likely be predicted to land tails on the next flip (see Chen, Moskowitz, and Shue (2016) for field evidence of the gambler's fallacy influencing judgement). The prediction of too many reversals can be viewed as a form of underreaction. Conversely, if the process is unknown (i.e., the true probabilities uncertain), then individuals tend to follow the hot hand fallacy, erroneously predicting continuation in a sequence of outcomes because they are trying to infer the true probabilities (Rabin (2002)). Hence, a sequence of heads from a series of coin flips with unknown probabilities is predicted to be followed by another heads. The prediction of too many streaks can be viewed as an overreaction to recent outcomes. If greater uncertainty implies less knowledge of the underlying process, then according to this framework it should lead to stronger overreaction.

To test the interaction between uncertainty and overreaction, I estimate the returns to momentum and value in sports betting markets as uncertainty varies. Per the discussion above, the predictions of overreaction theories imply that momentum will be stronger when uncertainty is high and value will be weaker, if overreaction is the driving force behind these phenomena. When uncertainty is high, bettors will overreact to recent news, pushing prices further in the same direction – e.g., momentum. When uncertainty is low, reversals or value will be more pronounced.

To implement these tests empirically, two measures of uncertainty are used. The first examines games early in the season, when, as discussed previously, there is more uncertainty about team quality. Panel A of Table XI reports the open-to-close and close-to-end returns to momentum and value for games early in each sports season (first 25 percent of games) versus the remainder of the season, as well as the difference between them. The last row of each set of results also reports the difference between open-to-close and close-to-end returns as well as the difference in differences from open-to-close and close-to-end from early versus late in the season.

The results are consistent with the predictions of overreaction models: momentum returns from open-to-close are much stronger early in the season when uncertainty is highest. Bettors seem to chase past performance more strongly when there is more uncertainty, pushing opening prices in the direction of past performance. Absent solid information about team quality early in the season, bettors may simply place too much weight on recent performance, as Rabin's (2002) agent who follows the law of small numbers does when the underlying process (in this case team quality) is uncertain. Or, as the representative investor in the Daniel, Hirshleifer, and Subrahmanyam (1998) model, absent any precise information on team quality, overconfidence places more weight on recent news (in this case recent performance) causing more overreaction. Furthermore, there is stronger evidence of mispricing, as there is a stronger reversal from close-to-end predicted by the initial price movement, consistent with overreaction. The differences between early versus late in the season are statistically significant as well. These patterns are also consistent across the three betting contracts: Spread, Moneyline, and Over/under, providing independent verification of these results in three different samples.

For value, the opposite patterns emerge: value returns appear weaker early in the season and are stronger later in the season. This suggests value or reversal effects are stronger when there is less uncertainty, consistent with the prediction of overreaction theories. Using a fundamental anchor of team quality (which the value measures attempt to capture) provides a benchmark that prompts bettors to identify contracts that look "cheap" or "expensive." In other words, later in the season, when team quality is better known, bettors expect performance to revert back to that quality benchmark, much like Rabin's (2002) agent when the underlying process is known. Or, as in Daniel, Hirshleifer, and Subrahmanyam (1998), the signal of team quality mitigates the overconfidence in recent performance.

The second measure of uncertainty exploits variation across contracts rather than time. Specifically, I use the parlay volume (divided by total betting volume) as a measure of uncertainty. As described previously, parlay bets only pay off if all bets involved in the parlay win. Hence, bets more heavily involved in parlays are those bettors feel more certain about. Dividing up the sample into the top and bottom third of contracts involved in parlays (per betting volume and within each sport), and recomputing the returns to momentum and value for high and low parlay games, Panel B of Table XI reports the results for high and low parlay bets and their difference. Low parlay implies higher uncertainty. Consistent with the predictions of overreaction models, low parlay contracts have stronger momentum effects and weaker value effects and high parlay games, where bettors believe outcomes are more certain, have less momentum and stronger value effects. These results are consistent across the three contract types and provide another unique test of overreaction and uncertainty. Once again, the stronger the initial price move from open-to-close due to either value or momentum, the stronger the subsequent correction from close-to-end, indicative of mispricing.

Figure 5 summarizes and illustrates the results from Table XI nicely, where the patterns from momentum and value returns are consistent with those from bettor overreaction, with the effects for momentum being much stronger early in the season (Panel A) and for low parlay games (Panel B) when uncertainty is greater, while value effects are stronger when there is less uncertainty. The graphs are plotted on the same scale to highlight the relative strength and magnitude of these patterns across the subsamples.

B. From Sports to Financial Markets

Using the new insights of overreaction and uncertainty found in sports betting markets for momentum and value effects, I apply the same idea to financial security market returns. This analysis is novel to the financial markets literature and also helps establish a tighter link between asset pricing in betting markets and asset pricing in capital markets.

Table XII reports results for tests of momentum and value returns in U.S. equities that try to mimic the tests in Table XI on uncertainty for sports betting contracts. Specifically, Table XII reports equity momentum and value returns when uncertainty of firm earnings is high versus low. Panel A reports results using the recency of earnings announcements as a proxy for uncertainty, where firms with high earnings uncertainty are those whose most recent earnings number is stale, while firms with low earnings uncertainty are those with very recent earnings announcements before the portfolio formation date. The idea is that earnings provides an important signal of value investors can anchor to, where more recent earnings are clearly more informative than earnings from some time ago. Earnings announcement dates are also set and announced in advance A(usually several weeks before the earnings announcement) and are highly stable and predictable (Frazzini and Lamont (2006)), making them less susceptible to other market influences.

Momentum and value returns are computed for U.S. stocks from the Center for Research in Security Prices over the July 1963 to December 2013 sample period for firms with recent and distant earnings announcements, where firms are first sorted into groups based on the recency of their latest earnings announcement. Specifically, at month t, all firms are sorted by how recent their last earnings announcement was and two groups are formed: "recent" firms who had an earnings announcement less than two weeks ago and "stale" firms who had an announcement at least 11 weeks ago. Since earnings occur quarterly for most firms (about every 12 weeks) and to cover enough firms, a two week window at the beginning and end of the earning cycle is chosen. Results are robust to other window choices and generally get weaker as the window expands (see below). Within each group stocks are then further sorted by momentum (past 12-month average returns, skipping the most recent month) or value (book-to-market ratio, BE/ME) and a portfolio long the top third of stocks and short the bottom third based on momentum or value sorts is created, where stocks are value weighted by their beginning of month market capitalization within each group. The returns to the long-short momentum (winners minus losers) and value (high BE/ME minus low BE/ME) portfolios within recent and stale earnings announcement groups are then reported in Panel A of Table XII along with their differences.

As Panel A of Table XII shows, price momentum is much stronger among the stale earnings firms, exhibiting an 11.81 percent annual return compared to only a 1.98 percent annual return for price momentum among firms whose earnings announcement occurred within the last two weeks. The difference of 9.83 percent between the two momentum portfolios among stale versus recent earnings firms is statistically significant. At first, this result may seem surprising given the evidence of earnings momentum (Bernard and Thomas (1989), Chan, Jegadeesh, and Lakonishok (1996), Novy-Marx (2015)), but a key difference here is that firms are simply sorted by whether they had an earnings announcement within the last two weeks or at least 11 weeks ago and not on the content of the earnings announcement (not sign, magnitude, or surprise of earnings), which is what the earnings momentum literature is about. Alphas with respect to the Fama and French four factors that include the market, size, value, and a momentum factor show an 11.44 difference in annual returns for momentum among stale versus recent earnings announcement firms.

The second row of Panel A of Table XII reports the results for value portfolios, which are opposite those of momentum. Value returns are strongest amongst the recent earnings announcers and weakest (statistically negligible) among the stale earnings announcers. The value premium is 9-10 percent higher per year among the recent announcers than it is among the stale announcers. These results are consistent with those from sports betting contracts in Table XI, where momentum is stronger among firms with more valuation uncertainty (proxied here by stale earnings information) and value is stronger among firms with less uncertainty (more recent earnings information). These results are consistent with an overreaction story for momentum and value premia in financial markets.

As a further test of the robustness of using earnings lag to identify uncertainty, Figure 6 plots the difference in momentum premia between stale versus recent earnings announcers for various windows of time used to define "stale" and "recent:" $>$ 11 versus $<$ 2 weeks since announcement, $>$ 8 versus $<$ 4 weeks since announcement, and > 7 versus < 6 weeks since announcement. As the figure shows, the momentum premium differences decline as the definition of recent versus stale becomes more coarse. This is reassuring that there is a direct link between staleness of earnings information and the strength of price momentum strategies. In addition, the figure plots the difference in returns of these momentum portfolios in years two to three *after* portfolio formation to test whether the subsequent reversals that often accompany momentum strategies are also stronger for the stale announcers. The long-term reversals of momentum strategies have often been characterized as a measure of price correction following momentum price movements (Jegadeesh and Titman (1993, 2001), Hong and Stein (1999), Hong, Lim, and Stein (2000)). Much like the terminal value of sports betting contracts provides a measure of mispricing, the long-term reversals are interpreted as price corrections where the long-term price is closer to the true price. The results for the long-term reversals are the mirror image of the momentum profits – when momentum is stronger, long-term reversals are stronger too, indicating a stronger correction occurs when there are bigger initial price movements coming from momentum. These results are remarkably consistent with the cleaner patterns found in sports betting markets, and the fact that their strength varies with uncertainty in the same way makes a compelling case for overreaction as a plausible story driving momentum premia in both markets.

Panel B of Table XII provides another measure of uncertainty across stocks using the dispersion in analyst forecasts of earnings, which is similar in spirit to the parlay volume measure used in sports betting markets. Following Diether, Malloy, and Scherbina (2002), dispersion in analyst forecasts is measured as the cross-sectional standard deviation across analysts of their fiscal year earnings-per-share forecast, scaled by the mean forecast for a given stock (requiring at least five analysts). Firms are sorted independently into five quintiles based on analyst dispersion and five quintiles based on momentum or value, where the intersection between each group produces 25 analyst dispersion $\times$ momentum categories and 25 analyst dispersion $\times$ value categories. The value weighted average return of each group is then computed monthly from July 1990 to December 2016. The returns to a strategy that is long the top quintile of momentum and short the bottom quintile of momentum is then computed within each analyst dispersion quintile and reported are the average annualized returns of the momentum profits within the lowest 20% and the highest 20% of analyst dispersion stocks, along with their t-statistics. The same set of results is reported for value strategy profits that go long the highest quintile of BE/ME stocks and short the lowest within the lowest and highest analyst dispersion firms.

High analyst dispersion of opinion implies higher uncertainty. As Panel B of Table XII shows, momentum is stronger among high analyst dispersion stocks, exhibiting a 4.68 percent higher return or 4.92 percent higher alpha than low analyst dispersion stocks. Conversely, value premia are stronger among low analyst dispersion stocks, exhibiting 6.30 percent per year stronger returns than among high dispersion stocks. These results are consistent with uncertainty exacerbating momentum returns and mitigating value returns, consistent with the implications of overreaction models and matching the patterns found in sports betting markets.[20]

C. Trading Strategy Profits and Costs

To compare the economic significance of the predictability of returns in sports betting markets to financial markets, Table XIII reports trading strategy profits from using momentum, value, and size in both markets.

C.1. Sports Betting Trading Strategies

Using the composite index measures for momentum, value, and size, a trading strategy is formed that invests positively (for the favorite) when predicted movements in the betting line are expected to be positive (i.e., continuation) and invests negatively (takes the other side of the bet) when line movements are predicted to be negative (i.e., reversals), where the dollars bet are in proportion to how strong the signal or characteristic predicting price movement is relative to the cross-sectional average. Specifically, the average trading strategy returns are computed as follows:

$$\bar{R}_{p,t} \quad = \quad \frac{1}{T}\sum_{t=1}^{T}\phi_t\sum_{j=1}^{N}(Char^i_{j,t} - \frac{1}{N}\sum_{j=1}^{N}Char^i_{j,t})\tilde{r}^i_{j,t} \quad \forall i \in (\text{S, ML, O/U}) \tag{9}$$

where $Char = \{$momentum, value, size$\}$ are signed to predict returns positively (hence, momentum and value enter with opposite sign to predict returns) for game j. For Spread and Moneyline contracts, a positive weight means betting for the favored team and a negative weight means betting on the underdog. For Over/under contracts a positive weight means betting on the over and a negative weight means betting on the under. The total bets summed across all games are rescaled to add up to one dollar at each point in time, which is what the scalar ϕ_t does.[21] All bets across all contracts are aggregated monthly. Results are reported separately for each characteristic by itself as well as a combination of all three characteristics used to predict

[20] Interacting the momentum measures with measures of investor salience – a feature that stands out or captures the focus of investors who may face limited attention (Barber and Odean (2008)) – is another test of overreaction. From experimental psychology (Kahneman and Tversky (1971)), humans tend to overreact to extreme or low probability events. To capture and test this idea, a dummy variable is created that identifies rare or extreme events that fans and the media may pay extra attention to: a player in the NBA scoring at least 50 points in the most recent game or recent eight games (only 104 occurrences in the 18,681 games examined), "triple doubles," team total points greater than 140, all of which yield similar results. For other sports similar salience measures are used, such as three or more touchdowns by any player in the NFL, a "hat trick," which is three goals or more in hockey, and multiple home runs by any player in MLB. Table B3 in the appendix reports the results, which for brevity, are only reported for the momentum index for both dollar and point returns. As the table shows, the interaction terms between momentum and the saliency dummy have the predicted sign – positive open to close and negative close to end, which indicates slightly more overreaction to past performance when one of these rare events also occurs, though the results are statistically insignificant. Chan (2003) examines momentum and reversals following news and no-news periods, but does not examine rare or extreme news separately from other news. To my knowledge, no one has tested this idea in financial markets.

[21] Similar results are obtained assuming $1 is bet equally across all contracts.

line movements called "multi-strategy." Trading strategy profits are computed for both open-to-close returns and for close-to-end returns.

Two average returns are computed: 1) gross returns and 2) net returns that account for transactions costs (the "vig"). The gross Sharpe ratio of the strategy is also reported. All returns and Sharpe ratios are reported as annualized numbers. Panel A reports results for the NBA only and Panel B reports results by combining all of the bets across all four sports into one portfolio, where each sport is given equal weight in the portfolio.

The first column of Panel A of Table XIII reports that annualized gross returns to betting on momentum, value, and size for open-to-close returns in the NBA are 60, 21, and -26 basis points, respectively. The positive predictability in returns from momentum and value is consistent with the previous findings from the regressions. The multi-strategy portfolio that combines all three characteristics yields 64 basis points on average per year. The second set of four rows report the gross returns from close to end (where the sign is reversed so that all trading strategies in expectation yield positive returns). Here, momentum produces 52 basis points and value 156 basis points, with the multi-strategy yielding 256 basis points.

While the gross returns are positive (except size), the net returns reported in the second column are significantly negative because of the vig, which at 10 percent on average easily wipes out any profits from trading on momentum and value, especially for open-to-close returns, since the vig is incurred twice (once at the open and then again at the close when unwinding the position). The third column reports the gross Sharpe ratio of each strategy, which ranges from -0.04 for size to 0.10 for the multi-strategy.

Panel B of Table XIII reports trading strategy profits using bets across all four sports. While the gross returns are pretty similar in magnitude to those from the NBA, the net returns and Sharpe ratios are typically bigger due to lower volatility from increased diversification of using many more uncorrelated bets across four different sports. Once again, however, none of these strategies are able to overcome trading costs. Given the size of the vig, attempting to arbitrage away the predictability of returns coming from momentum and value would prove unprofitable, and hence justifies why such predictability remains in the data. Moreover, this analysis further ignores price impact or "moving the line" from betting large quantities, which would raise trading costs further.

C.2. Financial Market Strategies

Panel C of Table XIII reports a comparison set of returns and Sharpe ratios for momentum, value, and size strategies in financial markets. The first set of results use the Fama and French long-short factors for size

(SMB), value (HML), and momentum (UMD) from Ken French's website for U.S. stocks for comparison, and the second set of results pertain to international factors for size, value, and momentum, constructed the same way as Fama and French's factors, from four markets: U.S., U.K., Europe (excluding the U.K.), and Japan, following Asness, Moskowitz, and Pedersen (2013) and Asness, Frazzini, Israel, Moskowitz, and Pedersen (2014). This provides a comparison of size, value, and momentum returns in a single equity market as well as across a diversified set of markets to compare to the diversified set of results applied across all four sports.

Since the sports betting contacts face no aggregate risk, to make cleaner comparisons I look at returns adjusted for the market by running a regression on the CRSP value-weighted index return in excess of the Treasury Bill rate for the U.S. protfolios or the MSCI World index in excess of the T-bill rate for the international portfolios. The market-adjusted or CAPM alphas of each strategy are then reported. In addition, I also estimate the net alpha of these strategies using the results from Frazzini, Israel, and Moskowitz (2014), who estimate total trading costs of these three strategies using proprietary live trading data from a large institutional money manager. The annualized information ratio (CAPM alpha divided by residual standard deviation) of these strategies is also reported gross of trading costs, for comparison to the Sharpe ratios of the sports betting strategies.

The first three columns of Panel C report the results for U.S. equities, which show a robust premium for momentum and value and a weak and essentially non-existent size premium (especially net of trading costs). Since the sports betting strategies and equity market strategies have different volatilities, the easiest comparison between them is the Sharpe ratio or information ratio. Comparing the NBA to the U.S. equity results, momentum is about one sixth as large in sports betting markets as it is in financial markets and value is about one fifth as large. The multi-strategy Sharpe ratio is about one fourth the size of its counterpart in financial markets. Comparing the Sharpe ratios diversified across all four sports to those in equity markets diversified across four geographies, the relative magnitudes are consistent – momentum, value, and their combination generate return premia per unit of risk about one fifth the size of those in financial markets.

Given these magnitudes, an interpretation of the data is that momentum and value effects, that are not coming from a systematic source of risk, produce returns per unit of risk about one fifth as large as what we observe in financial markets. From a behavioral view, one might argue that this evidence suggests that a significant, or at least non-trivial, portion of momentum and value patterns in markets could be coming from non-systematic sources, possibly related to investor behavior and sentiment. On the other hand, from a rational risk-based view, the magnitude of the premia generated from these non-systematic sources is small compared to the premia in financial markets, and hence the behavioral theories may only contribute a small part to these premia. Furthermore, given the large trading costs in sports markets, the magnitude of these effects may be even smaller in financial markets if arbitrage activity is less limited there.

D. Covariance Structure

Another key piece of evidence from financial markets on momentum and value (and size) returns is that there is significant covariance structure in the returns to these strategies (e.g., Fama and French (1993, 1996), Carhart (1997), Asness, Moskowitz, and Pedersen (2013), Lewellen, Nagel, and Shanken (2009) and Daniel and Titman (2010)). The significant covariance structure among these portfolios has often been interpreted as consistent with underlying common risk factors driving their returns. Behavioral models do not explicitly have a role for comovement and indeed this is one of the distinguishing features that Daniel and Titman (1997) try to exploit in testing rational versus behavioral paradigms. On the other hand, more recent models of investor behavior and institutional frictions derive a role for covariation among these portfolios through herding and benchmarking (e.g., Barberis and Shleifer (2004)) or correlated behavior and rebalancing (e.g., Vayanos and Wooley (2012)).

Given the idiosyncratic nature of sports betting contracts, it is interesting to examine whether there is any covariance structure among portfolios of these contracts sorted by the same characteristics. If covariance structure exists in these markets, it cannot be driven by aggregate risk sources and hence must come from other sources such as correlated investor behavior. If, however, no such comovement is present in these markets, then this could possibly be consistent with the common risk factor interpretation given to the comovement witnessed in financial markets.

Table XIV reports regression results of portfolios of contracts formed on momentum and value and regressed on a momentum and value "factor," similar in spirit to what is done in the financial markets literature. The test portfolios are formed by sorting all games in a given month (with at least 40 games) by their momentum (value) characteristic into five quintiles and then taking the equal-weighted average return of all games within each group. This provides a monthly return to quintile sorted portfolios based on momentum (value), whose returns are then regressed on the monthly returns of the momentum (value) factor, which is the high minus low quintile spread returns, Q5 − Q1.

Panel A of Table XIV reports results for test assets and factors formed from the *same* games such that there is overlap of the betting contracts that comprise the test portfolios and the factors. Panel B of Table XIV reports results for the same exercise where the test assets and factors are formed from non-overlapping games—e.g., one set of games is used to form the test assets and another set of games is used to form the factors. To compute the test portfolios and the factors independently each month the number of games are split randomly into two groups, with one used to form the test assets and the other used to form the factors, with no overlap in contracts on the left and right hand side of the regression. Reported are the coefficient estimates (β) on the factor, its t-statistic in parentheses, and the R^2 from each regression. The intercept is not reported for brevity.

Panel A, which reports results for the overlapping sample, where test assets and factors share the same contracts, shows strong covariation for the extreme portfolios (the low and high quintiles), but no covariation for quintiles 2 through 4. The covariation with the extreme portfolios are simply mechanical in that the factor itself is comprised of the same games that enter the extreme quintile portfolios. The lack of significance of any betas and the low R^2s for quintiles 2 through 4 indicate that there is no covariance structure present once this mechanical relation is no longer present. Confirming this conclusion, Panel B shows that there is no covariance structure at all across any of the quintiles, including the two extremes, when the test assets and factors contain non-overlapping games. Hence, there appears to be no significant covariance structure to either momentum or value-sorted portfolios among the sports betting contracts. The difference in results between Panels A and B highlights the dangers in using overlapping securities on both the left and right hand sides of the regression in conducting these tests generally. The lack of covariance structure is consistent with a behavioral interpretation of the momentum and value premia in these markets and is in contrast to the significant covariance structure found in financial markets for momentum and value strategies.[22]

V. Conclusion

A new testing ground for asset pricing anomalies is proposed: sports betting contracts. Two key aspects of these markets allow for distinguishing tests of behavioral asset pricing theories not confounded by rational risk-based theories. First, the contracts are idiosyncratic, implying that the cross-section of returns cannot be driven by aggregate or economy-wide risks. Second, the revelation of a true terminal value (e.g., the game outcome) in a short period of time that is independent of any betting activity provides an additional set of tests for detecting mispricing and testing competing theories.

Examining momentum, value, and size characteristics of these contracts, analogous to those in financial markets, I find that momentum exhibits significant predictability for returns, value exhibits significant but weaker predictability, and size exhibits no return predictability. The patterns of return predictability over

[22]I also examined whether there was any covariation in value or momentum sorted portfolios across sports. Not surprisingly, since there is no correlation within a sport, there is also no correlation across sports in terms of portfolio returns.

the life of the betting contracts – from opening to closing prices to game outcomes – matches those from models of investor overreaction. Moreover, the patterns are stronger in places and times where overreaction should be more prevalent. Applying these insights to financial market returns yields novel empirical results that match those in sports betting markets. The results suggest that at least part of the momentum and value patterns observed in capital markets could be related to similar investor behavior. The magnitude of return predictability in the sports betting market is about one-fifth that found in financial markets, where trading costs associated with sports betting contracts are too large to generate profitable trading strategies, possibly preventing arbitrage from eliminating the mispricing.

The results help shed light on behavioral asset pricing theories put forth to explain cross-sectional return patterns in capital markets while avoiding the joint hypothesis problem of specifying the stochastic discount factor that plagues such tests in financial markets. The revelation of a true terminal value also provides distinguishing tests of competing behavioral theories, where the evidence is most consistent with overreaction models. Further research may well use sports betting markets as a useful laboratory to investigate other patterns found in financial markets that may similarly help identify, or exclude, various asset pricing theories for their existence.

REFERENCES

Daniel, Kent D., and Sheridan Titman, 2006, Testing factor-model explanations of market anomalies, Working paper, University of Texas, Austin.

Fama, E. F., and K. R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, E. F., and K. R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

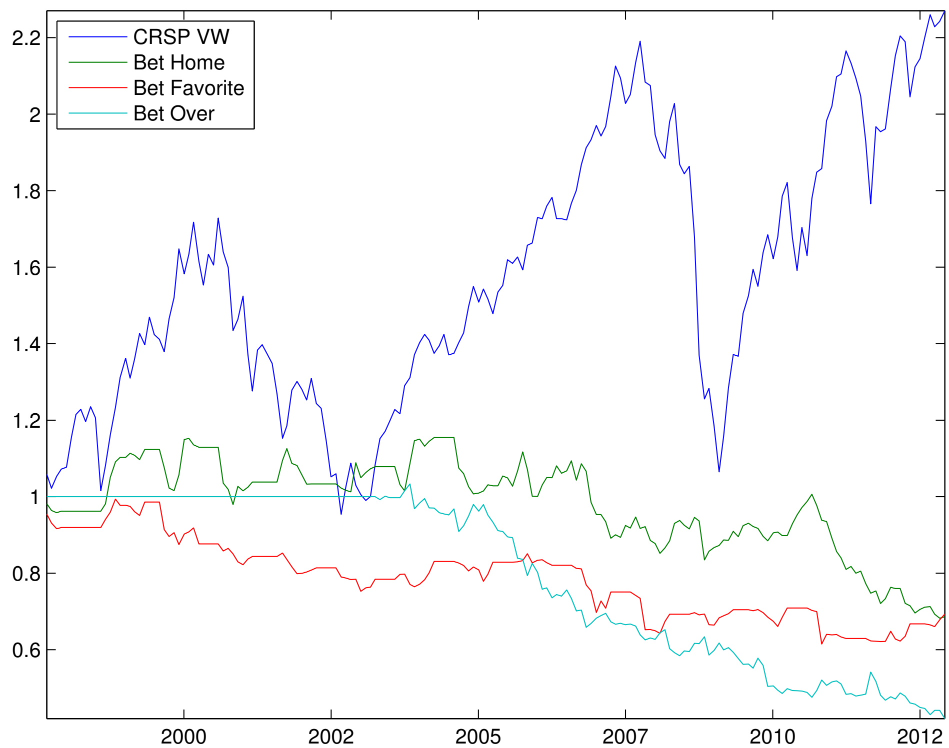
Fama, E. F., and K. R. French, 2007, The anatomy of value and growth stock returns, *Financial Analysts Journal* 63, 44-54.

Fama, E. F., and J. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Lewellen, Jonathan, Stefan Nagel, and Jay Shanken, 2006, A skeptical appraisal of asset-pricing tests, NBER working paper No. 12360.

Figures

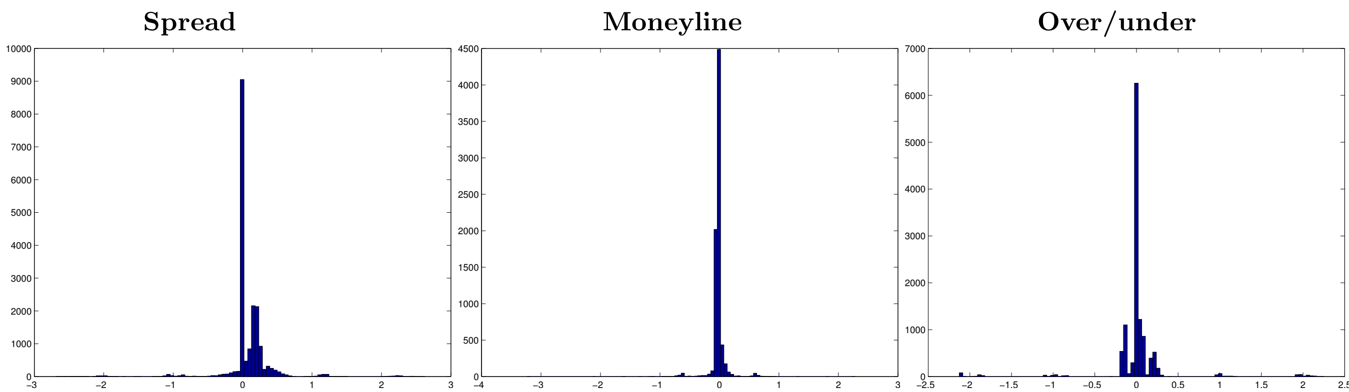
Figure 1: **Cumulative Returns to Sports Betting and the Stock Market.** The figure plots the cumulative returns to betting on the home team, the favorite team, and betting on the over across all sports—NBA, NFL, MLB, and NHL—and using all betting contracts—Spread, Moneyline, and Over/under. The cumulative returns to these bets versus the stock market (CRSP value-weighted index) are plotted weekly over time. Specifically, every week three portfolios of bets are formed by 1) betting on the home team using the Spread and Moneyline contracts, 2) betting on the favorite team using the Spread and Moneyline contracts, and 3) betting on the over using the Over/under contract. The portfolios are equal-weighted bets of one dollar in each game within each sport and then equal-weighted across sports, covering the NBA, NFL, MLB, and NHL. Since none of the sports have seasons that last a full year, and occur at different times of year, the majority of the time only two or three sports are covered. The weekly returns are aggregated monthly and their cumulative returns are plotted over time along with the cumulative return to the CRSP value-weighted U.S. stock index. All sports betting returns pertain to open-to-end returns, calculated following equation (14), using the probabilities estimated from the non-parametric specification for the spread and over/under contracts and using a probit model for the moneyline contracts. For ease of comparison, all series are scaled to the same ex post volatility that matches the sample volatility estimate of the stock market. Also reported is a correlation matrix of the return series. The sample period is November 1998 to March 2013.



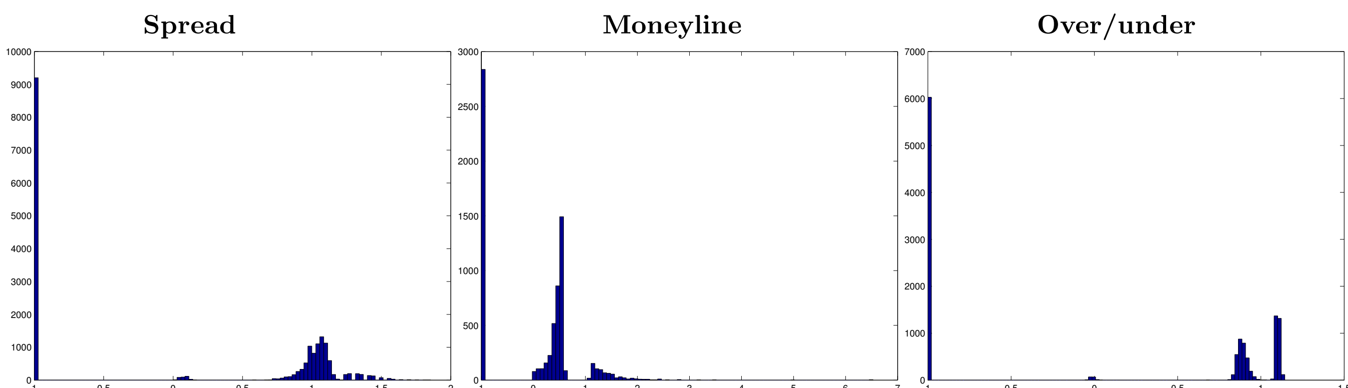
Correlation matrix of returns			
	Home	Favorite	Over
Stock Market	0.06	-0.01	0.03
Home		0.10	-0.01
Favorite			-0.01

Figure 2: **Return Distributions for NBA Contracts.** The figure plots the distribution of gross returns to all NBA game betting contracts. Panel A plots the open-to-close returns and Panel B plots the close-to-end returns for the spread, moneyline, and over/under betting contracts. Returns are calculated following equation 14, using the probabilities estimated from the non-parametric specification in the appendix. A table is included at the bottom of the three panels that reports the mean, standard deviation, skewness, and excess kurtosis of the net returns to each contract.

Panel A: Open-to-close returns



Panel B: Close-to-end returns



Summary statistics on returns (%)				
	mean	stdev	skew	ex-kurt
Spread contract				
$r_{open:close}$	-0.11	28.9	-0.37	32.90
$r_{close:end}$	-0.16	99.7	0.01	-1.98
Moneyline contract				
$r_{open:close}$	-1.60	13.5	-2.10	87.40
$r_{close:end}$	-2.09	89.7	0.63	2.13
Over/under contract				
$r_{open:close}$	0.07	35.9	0.07	24.06
$r_{close:end}$	0.10	100.2	0.02	-1.99

Table I: **Summary Statistics of Sports Betting Contracts Across Sports**

The table reports summary statistics for the sports betting contracts for each sport. Panel A reports statistics for the NBA, Panel B for the NFL, Panel C for MLB, and Panel D for the NHL. The number of seasons, total number of games, and total number of betting contracts are reported. Each game can have up to three betting contracts that depend on the game's outcome: 1) the Spread contract, which is a bet on whether a team wins by at least a certain amount of points, known as the "spread"; 2) the Moneyline contract, which is a bet on which team wins for different dollar amounts, specified as betting $|x|$ dollars to win \$100 if $x < 0$ or betting \$100 to win \$$x$ if $x > 0$; and 3) the Over/under contract, which is a bet on whether the total score (sum of both teams' points) is over or under the specified number. Reported below are summary statistics on the mean, standard deviation, 1st, 10th, 25th, 50th, 75th, 90th, and 99th percentiles of the Spread, Moneyline, and Over/under values in each sport for the closing lines for each contract assuming a bet is placed on the home team. MLB and the NHL do not have Spread contracts as every Spread contract is quoted at ±1.5, otherwise known as the "run" line (MLB) or "puck" line (NHL) that simply indicates which team is expected to win, with no cross-sectional variation in these contracts across games. The final row of each panel reports the betting-implied probability of the home team winning the game based on the implied probabilities from the Moneyline contracts at each reported distributional percentile.

	mean	stdev	$1^{st}\%$	$10^{th}\%$	$25^{th}\%$	$50^{th}\%$	$75^{th}\%$	$90^{th}\%$	$99^{th}\%$
Panel A: NBA, 1999 - 2013 18,681 games; 38,939 betting contracts									
Spread	-3.4	6.0	-15.0	-10.5	-7.5	-4.5	1.5	5.0	10.0
Moneyline	-220.0	438.8	-2200.0	-565.0	-315.0	-172.0	107.0	177.0	330.0
Over/Under	196.1	11.4	172.0	182.5	188.0	195.0	203.5	211.0	226.0
P(win)			*0.991*	*0.894*	*0.763*	*0.636*	*0.477*	*0.342*	*0.127*
Panel B: NFL, 1985 - 2013 7,035 games; 10,775 betting contracts									
Spread	-2.6	6.0	-15.5	-10.0	-7.0	-3.0	2.5	5.5	11.5
Moneyline	-160.6	208.0	-700.0	-370.0	-270.0	-174.0	-112.0	144.0	264.0
Over/Under	42.3	4.8	32.5	36.5	38.5	42.5	45.5	48.0	54.5
P(win)			*0.963*	*0.836*	*0.752*	*0.594*	*0.455*	*0.311*	*0.118*
Panel C: MLB, 2005 - 2013 23,986 games; 47,964 betting contracts									
Spread	-1.5	–	–	–	–	–	–	–	–
Moneyline	-69.7	128.9	-265.0	-187.0	-154.0	-124.0	104.0	130.0	173.0
Over/Under	8.7	1.1	6.5	7.5	8.0	9.0	9.5	10.0	11.5
P(win)			*0.726*	*0.652*	*0.606*	*0.550*	*0.510*	*0.435*	*0.366*
Panel D: NHL, 2005 - 2013 9,890 games; 19,764 betting contracts									
Spread	-1.5	–	–	–	–	–	–	–	–
Moneyline	-93.6	120.7	-280.0	-201.0	-165.0	-133.0	-105.0	120.0	158.4
Over/Under	5.6	0.4	5.0	5.0	5.5	5.5	6.0	6.0	6.5
P(win)			*0.737*	*0.668*	*0.622*	*0.571*	*0.512*	*0.455*	*0.387*

Table II: **Return Correlations**

The table reports return correlations for each of the three betting contracts for each game: the Spread contract (S), the Moneyline contract (m), and the Over/under contract (O/U). Three sets of returns are calculated for each contract: the return from the opening line to the outcome (open:end), the return from the closing line to the outcome (close:end), and the return from the opening line to the closing line (open:close). Returns are calculated following equations (10)-(14), using the probabilities estimated from the non-parametric specification. Highlighted in **bold** are the correlations within the same contract for different returns (open-to-end, close-to-end, and open-to-close). The non-bolded numbers are the correlations across contracts (Spread, Moneyline, Over/under).

	$R^S_{open:end}$	$R^S_{close:end}$	$R^S_{open:close}$	$R^m_{open:end}$	$R^m_{close:end}$	$R^m_{open:close}$	$R^{O/U}_{open:end}$	$R^{O/U}_{close:end}$	$R^{O/U}_{open:close}$
$R^S_{open:end}$	**1.00**	**0.96**	**0.15**	0.68	0.69	-0.02	-0.01	-0.01	0.01
$R^S_{close:end}$		**1.00**	**-0.14**	0.67	0.69	-0.03	-0.01	-0.01	0.01
$R^S_{open:close}$			**1.00**	0.03	0.02	0.04	0.00	0.00	0.01
$R^m_{open:end}$				**1.00**	**0.99**	**0.04**	0.01	0.01	0.00
$R^m_{close:end}$					**1.00**	**-0.11**	0.01	0.01	0.00
$R^m_{open:close}$						**1.00**	0.01	0.01	0.01
$R^{O/U}_{open:end}$							**1.00**	**0.94**	**0.18**
$R^{O/U}_{close:end}$								**1.00**	**-0.18**
$R^{O/U}_{open:close}$									**1.00**

Table III: **Testing Information versus Sentiment in Betting Price Movements**

The table reports estimates of regression equation (5)

$$R_{close:end} = \alpha + \beta_1 R_{open:close} + \epsilon.$$

Panel A reports results for the full sample of bets for each sport separately, for the NBA and NFL combined, and for all four sports simultaneously. Panel B reports estimates of the regression for the subsample of betting contracts where the price moved from open to close (e.g., excluding betting lines with no open-to-close price movement). Panel C reports regression estimates for the most heavily and least heavily bet games using betting volume (number of contracts) across all sports, where the sample is split into the top third and bottom third of betting volume games within each sport and the coefficients among the highest and lowest volume games are estimated. Panel D reports results across games involved in parlay bets – portfolios of bets where every bet in the portfolio must win for the bet to pay off and where the payout is highly levered – a proxy for investor confidence in the bet (due to either their information or pereived information) since every bet in a parlay must win. The sample is split into the top and bottom third of games within each sport based on parlay volume and the regression coefficients are estimated separately. Panels C and D report results measured across all sports.

	Spread	Moneyline	O/U	Spread	Moneyline	O/U
	PANEL A: FULL SAMPLE			PANEL B: PRICE MOVES ONLY		
NBA						
β_1	-0.49	-1.32	-0.50	-0.49	-1.43	-0.50
	(-19.30)	(-17.91)	(-20.14)	(-19.68)	(-18.11)	(-20.24)
NFL						
β_1	-0.34	-0.58	-0.46	-0.38	-0.52	-0.46
	(-6.32)	(-2.55)	(-8.31)	(-7.19)	(-1.90)	(-8.34)
MLB						
β_1		-0.95	-0.31		-0.95	-0.31
		(-10.51)	(-10.09)		(-7.11)	(-10.51)
NHL						
β_1		-0.26	-1.01		-0.19	-0.84
		(-2.37)	(-8.30)		(-1.72)	(-7.47)
NBA + NFL						
β_1	-0.46	-1.15	-0.49	-0.47	-1.28	-0.50
	(-20.16)	(-16.41)	(-21.79)	(-20.86)	(-16.89)	(-21.89)
All Sports						
β_1	-0.46	-1.04	-0.44	-0.47	-1.23	-0.48
	(-20.16)	(-21.25)	(-24.38)	(-20.86)	(-17.66)	(-27.02)

	Spread	Moneyline	O/U
PANEL C: BY BETTING VOLUME			
High			
β_1	-0.45	-1.14	-0.42
	(-8.41)	(-12.79)	(-11.67)
Low			
β_1	-0.48	-0.99	-0.49
	(-9.17)	(-11.04)	(-15.72)
Difference	0.03	-0.15	0.07
	(0.40)	(-1.19)	(1.47)
PANEL D: BY PARLAY BETTING (BETTOR CONFIDENCE)			
High			
β_1	-0.50	-1.43	-0.51
	(-8.75)	(-12.86)	(-13.36)
Low			
β_1	-0.44	-1.38	-0.46
	(-8.70)	(-14.59)	(-11.61)
Difference	-0.06	-0.05	-0.05
	(-1.05)	(-0.45)	(-1.31)

Table IV: **Momentum, Value, and Size Correlations**

The table reports correlations among various momentum, value, and size measures. Panel A reports correlations among various fundamental and price momentum measures as defined in Section III. Panel B reports correlations between various value measures based on long-term past returns, fundamental ratios, including a ratio of the expected contract price from an analytical model to the actual contract price, and a regression residual representing the error term from a model that predicts the expected betting contract price that includes home and away team dummies within each season, the record (winning percentage) of the team before each game, and the number of games played within the last week. Panel C reports correlations between various size measures, including the Forbes' estimates of franchise value, ticket revenue, total revenue, and player payroll expense. All measures are described in Section III. Also reported for each panel are the correlations of each variable or set of variables with a principal component-weighted average index of all momentum measures (Panel A), all value measures (Panel B) and all size measures (Panel C), where variables are weighted by the eigenvector associated with the largest eigenvalue of the correlation matrix among the variables. Panel D reports the correlations between the indices for momentum, value, and size.

Panel A: Momentum measures

	Fundamental momentum		Price momentum		Fundamental momentum		Price momentum	
	Win%	Net points	\$ returns	Point returns	Win%	Net points	\$ returns	Point returns
	Lag = 1 game				Lag = 2 games			
Win%	1.00	0.81	0.62	0.67	1.00	0.83	0.60	0.65
Net points		1.00	0.68	0.86		1.00	0.65	0.82
\$ returns			1.00	0.79			1.00	0.78
Point returns				1.00				1.00
	Lag = 4 games				Lag = 8 games			
Win%	1.00	0.85	0.57	0.61	1.00	0.88	0.52	0.56
Net points		1.00	0.61	0.76		1.00	0.56	0.70
\$ returns			1.00	0.78			1.00	0.78
Point returns				1.00				1.00
	Win% (1,2,4,8)		Net points (1,2,4,8)		\$returns (1,2,4,8)		Point returns (1,2,4,8)	
Mom_{index}	0.52		0.60		0.48		0.51	

Panel B: Value measures

	Long-term past returns($)			Team fundamental ratios				Game fundamental ratio	Regression
	1 season	2 seasons	3 seasons	Value/P	Tix/P	Rev/P	Payroll/P	$E(P)/P$	residual
1 season	1.00	0.72	0.60	0.02	0.02	0.02	0.02	-0.08	0.00
2 seasons		1.00	0.83	0.01	0.01	0.01	0.00	-0.08	-0.01
3 seasons			1.00	-0.01	-0.02	-0.01	-0.01	-0.08	-0.01
Value/P				1.00	0.87	0.97	0.52	0.17	-0.01
Tix/P					1.00	0.89	0.59	0.17	0.00
Rev/P						1.00	0.53	0.16	-0.01
Payroll/P							1.00	0.06	0.00
$E(P)/P$								1.00	0.01
Residual									1.00
Val_{index}	-0.30	-0.32	-0.30	0.43	0.45	0.46	0.50	0.56	0.52

Panel C: Size measures

	Value	Tix	Rev	Payroll
Franchise value	1.00	0.88	0.97	0.52
Ticket revenue		1.00	0.89	0.59
Total revenue			1.00	0.53
Player payroll				1.00
Size PC	0.90	0.88	0.95	0.71

Panel D: Momentum, Value, and Size

	Mom_{index}	Val_{index}	Size_{index}
Mom_{index}	1.00	-0.11	0.06
Val_{index}		1.00	-0.19
Size_{index}			1.00

Table V: **Momentum and the Cross-Section of NBA Betting Returns**

The table reports regression results of open-to-close returns and close-to-end returns from equations (6) and (7) on various fundamental-based and price-based momentum measures for Spread contracts (Panel A), Moneyline contracts (Panel B), and Over/under contracts (Panel C) in the NBA for 18,132 games from 1998 to 2013. The regressions equations are:

$$\begin{aligned}\tilde{R}_{i,0:1} &= \alpha_1 + \beta_1 Char_i + \tilde{\epsilon}_{i,0:1}\\ \tilde{R}_{i,1:T} &= \alpha_T + \beta_T Char_i + \tilde{\epsilon}_{i,1:T}.\end{aligned}$$

Results are reported for both actual dollar returns from the contracts as well as hypothetical returns based on points scored ("point returns"), except for the Moneyline contract where points-based returns are irrelevant. The momentum measures are the cumulative past point differentials and cumulative past dollar returns over the past N games. Also reported are tests for the difference between the two coefficients $\beta_1 - \beta_T$, which measures the cumulative absolute magnitude of the return movement from open to close and then from close to the game outcome to capture the total price movement due to momentum. In addition, we report tests on the sum of the two coefficients $\beta_1 + \beta_T$ to capture the total net return affect from open to game outcome that accounts for any offsetting return movement from open to close and then from close to game outcome.

	Past point differential			Past \$ returns			
Past games	1	4	8	1	4	8	Mom$_{index}$
Panel A: Spread contracts							
			Open-to-close dollar returns, $R^{\$}_{open:close}$				
$\beta_{1,Mom}$	0.03	0.04	0.04	0.04	0.06	0.06	0.06
	(2.82)	(1.86)	(1.52)	(3.40)	(2.19)	(1.46)	(2.34)
			Close-to-end dollar returns, $R^{\$}_{close:end}$				
$\beta_{T,Mom}$	-0.04	-0.02	0.04	-0.09	-0.11	-0.02	-0.07
	(-1.08)	(-0.27)	(0.42)	(-2.01)	(-1.16)	(-0.16)	(-1.82)
Difference	0.07	0.06	0.01	0.13	0.17	0.08	0.13
	(2.72)	(1.53)	(0.29)	(3.75)	(2.40)	(1.27)	(2.24)
Sum	-0.01	0.02	0.08	-0.05	-0.05	0.04	-0.01
	(-0.39)	(0.51)	(1.32)	(-1.44)	(-0.71)	(0.64)	(-0.17)
			Open-to-close "point returns", $R^{pts.}_{open:close}$				
$\beta_{1,Mom}$	0.41	1.01	1.14	0.48	1.33	1.84	1.03
	(8.71)	(12.02)	(10.55)	(8.92)	(11.76)	(10.92)	(10.07)
			Close-to-end "point returns", $R^{pts.}_{close:end}$				
$\beta_{T,Mom}$	-1.02	-1.21	-1.00	-1.27	-1.96	-1.52	-2.20
	(-2.31)	(-1.52)	(-0.99)	(-2.49)	(-1.83)	(-0.96)	(-2.29)
Difference	1.43	2.22	2.14	1.75	3.29	3.36	3.23
	(4.35)	(4.70)	(4.03)	(4.55)	(4.81)	(4.14)	(4.70)
Sum	-0.61	-0.20	0.14	-0.79	-0.63	0.32	-1.17
	(-1.85)	(-0.42)	(0.26)	(-2.05)	(-0.92)	(0.39)	(-1.70)

	Past point differential			Past \$ returns			
Past games	1	4	8	1	4	8	Mom_{index}
Panel B: Moneyline contracts							
	Open-to-close dollar returns, $R^{\$}_{open:close}$						
$\beta_{1,Mom}$	0.01	0.02	0.02	0.01	0.04	0.04	0.02
	(0.65)	(1.53)	(0.86)	(1.18)	(2.06)	(1.25)	(1.28)
	Close-to-end dollar returns, $R^{\$}_{close:end}$						
$\beta_{T,Mom}$	0.05	0.03	0.05	-0.04	-0.02	-0.01	0.02
	(1.03)	(0.98)	(1.76)	(-0.69)	(-0.14)	(-0.05)	(1.82)
Difference	-0.04	-0.01	-0.03	0.05	0.06	0.05	0.01
	(-0.93)	(-0.70)	(-1.05)	(1.05)	(0.99)	(0.79)	(0.07)
Sum	0.06	0.05	0.07	-0.03	0.02	0.03	0.04
	(1.39)	(1.12)	(1.61)	(-0.63)	(0.33)	(0.47)	(1.02)
Panel C: Over/under contracts							
	Open-to-close dollar returns, $R^{\$}_{open:close}$						
$\beta_{1,Mom}$	0.08	0.07	0.05	0.01	-0.01	0.01	0.15
	(4.37)	(2.52)	(1.37)	(0.54)	(-0.62)	(1.28)	(3.95)
	Close-to-end dollar returns, $R^{\$}_{close:end}$						
$\beta_{T,Mom}$	-0.02	-0.19	-0.18	0.01	0.01	0.01	-0.08
	(-0.46)	(-2.22)	(-1.77)	(0.31)	(0.51)	(0.07)	(-0.77)
Difference	0.10	0.26	0.23	0.00	-0.01	0.01	0.23
	(4.28)	(3.71)	(2.63)	(0.54)	(-0.51)	(1.28)	(3.77)
Sum	0.06	-0.12	-0.13	0.00	0.01	0.01	0.07
	(2.57)	(-1.71)	(-1.49)	(0.54)	(0.51)	(1.28)	(1.15)
	Open-to-close "point returns", $R^{pts.}_{open:close}$						
$\beta_{1,Mom}$	1.32	1.22	1.09	0.05	0.07	0.14	2.56
	(12.18)	(6.88)	(5.05)	(3.56)	(2.47)	(3.06)	(11.31)
	Close-to-end "point returns", $R^{pts.}_{close:end}$						
$\beta_{T,Mom}$	-1.11	-6.00	-6.50	0.10	0.04	-0.02	-3.36
	(-1.23)	(-4.06)	(-3.59)	(0.85)	(0.19)	(-0.04)	(-1.76)
Difference	2.43	7.22	7.59	-0.05	0.03	0.16	5.92
	(5.17)	(5.74)	(4.80)	(-0.60)	(0.31)	(1.57)	(5.02)
Sum	0.21	-4.78	-5.41	0.15	0.11	0.12	-0.80
	(0.45)	(-3.80)	(-3.42)	(1.80)	(1.14)	(1.17)	(-0.68)

Table VI: **Momentum and the Cross-Section of All Sports Betting Returns**

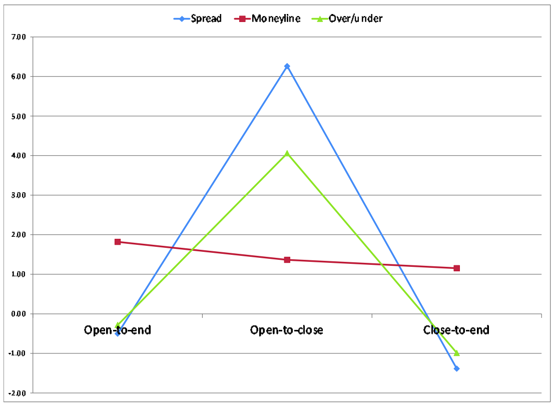
The table reports regression results from Table V for the NFL (Panel A), MLB (Panel B), NHL (Panel C), and for all sports combined (Panel D), including the NBA. The regressions are the open-to-close returns and close-to-end returns on the momentum index measure as described in equations (6) and (7) estimated separately for Spread contracts, Moneyline contracts, and Over/under contracts. Spread contract results only pertain to the NBA and NFL, since there are no Spread contracts for MLB and the NHL, while the Moneyline and Over/under contracts pertain to all sports. Results are reported for both actual dollar returns from the contracts as well as hypothetical "point returns" based on points scored, except for the Moneyline.

	Dollar returns			"Point returns"	
	Spread	Moneyline	O/U	Spread	O/U
	PANEL A: NFL				
	Open-to-close returns, $R_{open:close}$				
$\beta_{1,Mom}$	0.03	0.03	0.09	0.55	0.69
	(0.81)	(0.71)	(1.24)	(3.11)	(1.96)
	Close-to-end returns, $R_{close:end}$				
$\beta_{T,Mom}$	-0.17	0.10	-0.17	-0.94	-4.48
	(-1.89)	(0.45)	(-0.84)	(-0.75)	(-1.57)
Difference	0.20	-0.07	0.26	1.49	5.17
	(2.44)	(-0.39)	(1.65)	(1.74)	(2.06)
Sum	-0.14	0.13	-0.08	-0.39	-3.79
	(-1.71)	(0.72)	(-0.51)	(-0.46)	(-1.51)
	PANEL B: MLB				
	Open-to-close returns, $R_{open:close}$				
$\beta_{1,Mom}$	–	0.05	0.01	–	0.04
	–	(3.80)	(2.18)	–	(6.09)
	Close-to-end returns, $R_{close:end}$				
$\beta_{T,Mom}$	–	-0.26	-0.01	–	-0.01
	–	(-1.55)	(-0.27)	–	(-0.15)
Difference	–	0.31	0.02	–	0.05
	–	(2.17)	(0.96)	–	(2.67)
Sum	–	-0.21	0.00	–	0.03
	–	(-1.47)	(0.29)	–	(1.60)
	PANEL C: NHL				
	Open-to-close returns, $R_{open:close}$				
$\beta_{1,Mom}$	–	0.07	0.05	–	0.05
	–	(1.61)	(0.88)	–	(0.85)
	Close-to-end returns, $R_{close:end}$				
$\beta_{T,Mom}$	–	-0.25	-0.59	–	-1.42
	–	(-0.56)	(-0.81)	–	(-0.63)
Difference	–	0.32	0.64	–	1.47
	–	(2.17)	(1.69)	–	(1.48)
Sum	–	-0.18	-0.54	–	-1.37
	–	(-1.22)	(-1.42)	–	(-1.38)

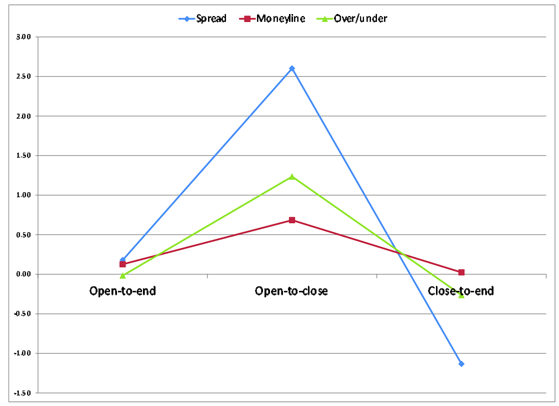
	Dollar returns			"Point returns"	
	Spread	Moneyline	O/U	Spread	O/U
	PANEL D: ALL SPORTS				
	Open-to-close returns, $R_{open:close}$				
$\beta_{1,Mom}$	0.04	0.05	0.05	0.48	0.25
	(3.66)	(5.14)	(3.75)	(9.69)	(10.19)
	Close-to-end returns, $R_{close:end}$				
$\beta_{T,Mom}$	-0.15	-0.21	-0.02	-0.79	-0.37
	(-2.80)	(-2.33)	(-0.34)	(-2.97)	(-2.76)
Difference	0.19	0.26	0.07	1.27	0.62
	(4.26)	(3.48)	(2.66)	(6.69)	(6.90)
Sum	-0.11	-0.16	0.03	-0.31	-0.12
	(-1.59)	(-2.14)	(1.14)	(-1.63)	(-1.33)

Figure 3: **Momentum Beta Patterns.** Plotted are the average t-statistics of the momentum betas estimated from equations (6) and (7) for open-to-end, open-to-close, and close-to-end returns for each betting contract type: Spread, Moneyline, and Over/under in each sport. The regressions are estimated for each momentum measure from Section III individually and then the t-statistics from these regressions are aggregated across all momentum measures and plotted below over each return horizon. The momentum betas are estimated and plotted separately for the NBA (Panel A), the NFL (Panel B), MLB (Panel C), the NHL (Panel D), the NBA and NFL combined (Panel E), and across all four sports (Panel F).

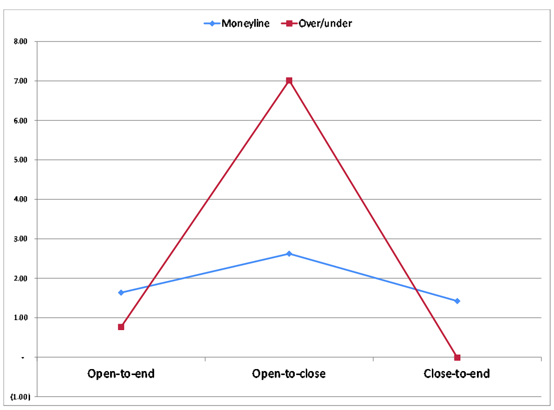
Panel A: NBA



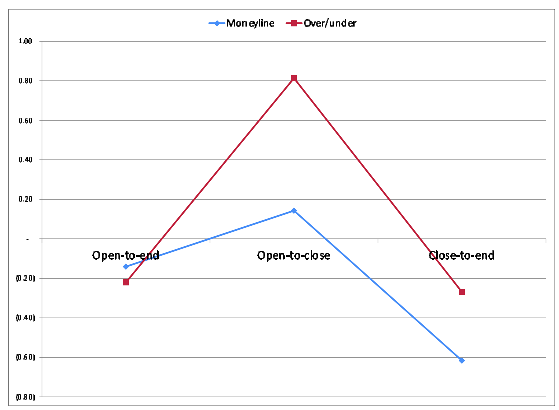
Panel B: NFL



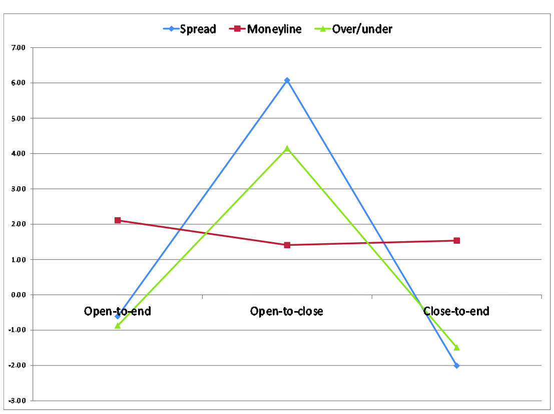
Panel C: MLB



Panel D: NHL



Panel E: NBA + NFL



Panel F: All sports

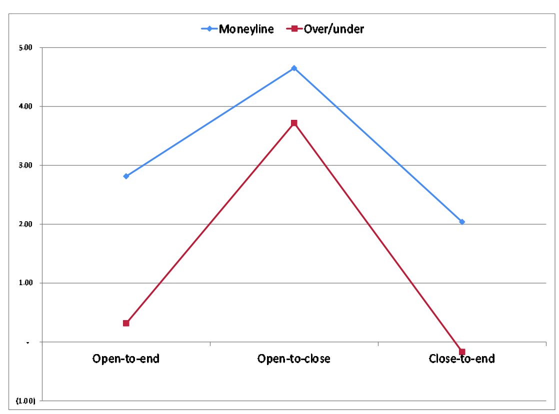


Table VII: **Value and the Cross-Section of NBA Betting Returns**

The table reports regression results of open-to-close returns and close-to-end returns from equations (6) and (7) on various value measures for Spread contracts (Panel A), Moneyline contracts (Panel B), and Over/under contracts (Panel C) in the NBA for 18,132 games from 1998 to 2013. The regression equations are:

$$\begin{aligned}\tilde{R}_{i,0:1} &= \alpha_1 + \beta_1 Char_i + \tilde{\epsilon}_{i,0:1} \\ \tilde{R}_{i,1:T} &= \alpha_T + \beta_T Char_i + \tilde{\epsilon}_{i,1:T}.\end{aligned}$$

Results are reported for both actual dollar returns from the contracts as well as hypothetical returns based on points scored, except for the Moneyline contract where points-based returns are irrelevant. The value measures are 1) the negative of the long-term return from betting on the team: negative of the cumulative returns from betting on the team over the past 1 and 2 seasons; 2) fundamental-to-price ratios: the franchise dollar value, ticket revenue, total revenue, and player payroll expense each scaled by the "price" (P) of the betting contract as proxied by the point spread; 3) a ratio of the expected contract price from an analytical model to the actual contract price, $E(P)/P$; and 4) a regression residual representing the error term from a model that predicts the expected betting contract price that includes home and away team dummies within each season, the record (winning percentage) of the team before each game, and the number of games played within the last week. Finally, the value index of all these measures is also used as a regressor. Also reported is a test for whether the predictability completely reverses, $\beta_{O:C} = -\beta_{C:E}$, where a "yes" indicates a failure to reject and a "no" indicates rejection (at the 5% significance level) of the null hypothesis that the betas are of equal magnitude but opposite sign.

Value measures	Long-term reversals: 1 season	2 seasons	Fundamental-to-price ratios: Value/P	Tix/P	Rev/P	Player/P	$E(P)/P$	Negative of residual	Val_{index}
Panel A: Spread contracts									
Open-to-close dollar returns, $R^{\$}_{open:close}$									
$\beta_{1,Val}$	-0.19	-0.22	0.20	-0.02	0.01	-0.03	-0.08	0.02	-0.06
	(-1.55)	(-1.23)	(0.29)	(-0.54)	(0.19)	(-0.57)	(-1.44)	(0.20)	(-0.37)
Close-to-end dollar returns, $R^{\$}_{close:end}$									
$\beta_{T,Val}$	0.82	1.40	2.11	0.05	0.07	0.17	0.16	0.47	0.44
	(2.01)	(2.37)	(0.90)	(0.32)	(1.02)	(1.09)	(0.88)	(1.34)	(0.82)
Difference	-1.01	-1.62	-1.91	-0.07	-0.07	-0.20	-0.24	-0.49	-0.50
	(-2.85)	(-3.03)	(-0.86)	(-0.56)	(-1.02)	(-1.43)	(-1.73)	(-1.44)	(-1.02)
Sum	0.63	1.18	2.31	0.03	0.07	0.14	0.08	0.45	0.38
	(1.78)	(2.21)	(1.05)	(0.24)	(1.02)	(1.00)	(0.58)	(1.32)	(0.77)
Open-to-close "point returns", $R^{pts.}_{open:close}$									
$\beta_{1,Val}$	-2.40	-2.84	-7.45	-0.69	-0.22	-0.65	-0.56	0.37	-1.23
	(-4.80)	(-3.93)	(-2.59)	(-3.73)	(-2.56)	(-3.44)	(-2.52)	(-0.84)	(-1.79)
Close-to-end returns, $R^{pts.}_{close:end}$									
$\beta_{T,Val}$	8.13	10.28	8.26	0.79	0.10	1.46	7.55	3.33	0.84
	(1.72)	(1.50)	(0.30)	(0.45)	(0.13)	(0.82)	(3.55)	(0.82)	(0.14)
Difference	-10.53	-13.12	-15.71	-1.48	-0.32	-2.11	-8.11	-2.96	-2.07
	(-2.80)	(-2.38)	(-1.00)	(-1.46)	(-1.04)	(-1.63)	(-4.07)	(-0.80)	(-0.71)
Sum	5.73	7.44	0.81	0.10	-0.12	0.81	6.99	3.70	-0.39
	(1.52)	(1.35)	(0.05)	(0.10)	(-0.39)	(0.62)	(3.50)	(1.00)	(-0.13)

Value measures	Long-term reversals 1 season	2 seasons	Fundamental-to-price ratios Value/P	Tix/P	Rev/P	Player/P	$E(P)/P$	Residual	Val_{index}
Panel B: Moneyline contracts									
Open-to-close dollar returns, $R^{\$}_{open:close}$									
$\beta_{1,Val}$	-0.13	-0.10	-2.32	-0.15	-0.07	-0.10	-9.44	-0.06	-0.12
	(-1.54)	(-0.84)	(-4.86)	(-4.99)	(-4.93)	(-2.97)	(-10.25)	(-0.81)	(-1.06)
Close-to-end dollar returns, $R^{\$}_{close:end}$									
$\beta_{T,Val}$	1.07	1.80	2.69	0.16	0.10	0.20	17.65	0.11	0.22
	(1.96)	(2.39)	(0.91)	(0.87)	(1.08)	(0.94)	(2.91)	(0.24)	(0.33)
Difference	-1.20	-1.90	-5.01	-0.31	-0.17	-0.30	-27.09	-0.17	-0.34
	(-2.42)	(-2.64)	(-2.76)	(-2.83)	(-2.82)	(-1.95)	(-6.33)	(-0.52)	(-0.73)
Sum	0.94	1.70	0.37	0.01	0.03	0.10	8.21	0.05	0.10
	(1.90)	(2.36)	(0.20)	(0.09)	(0.50)	(0.65)	(1.92)	(0.15)	(0.21)
Panel C: Over/under contracts									
Open-to-close dollar returns, $R^{\$}_{open:close}$									
$\beta_{1,Val}$	-0.14	-0.13	0.02	0.01	0.01	-0.01	-21.35	-0.12	-0.04
	(-3.49)	(-3.15)	(0.10)	(0.16)	(0.13)	(-0.04)	(-2.04)	(-1.73)	(-0.22)
Close-to-end dollar returns, $R^{\$}_{close:end}$									
$\beta_{T,Val}$	-0.12	-0.13	-0.98	-0.08	-0.03	-0.06	34.42	0.23	0.76
	(-1.11)	(-1.12)	(-2.14)	(-1.98)	(-2.19)	(-2.32)	(1.18)	(1.15)	(1.32)
Difference	-0.02	0.00	1.00	0.08	0.03	0.06	-55.77	-0.35	-0.80
	(-0.28)	(0.00)	(2.21)	(1.98)	(2.19)	(2.32)	(-2.53)	(-2.25)	(-1.44)
Sum	-0.26	-0.26	-0.96	-0.08	-0.03	-0.06	13.07	0.11	0.72
	(-3.65)	(-3.30)	(-2.12)	(-1.98)	(-2.19)	(-2.32)	(0.59)	(0.71)	(1.30)
Open-to-close "point returns", $R^{pts.}_{open:close}$									
$\beta_{1,Val}$	-2.49	-2.57	-3.42	-0.30	-0.10	-0.18	-11.54	-0.31	-3.40
	(-10.37)	(-10.30)	(-3.49)	(-3.41)	(-3.50)	(-3.40)	(-8.34)	(-0.74)	(-2.97)
Close-to-end "point returns", $R^{pts.}_{close:end}$									
$\beta_{T,Val}$	-0.22	-0.13	-10.99	-0.95	-0.35	-0.68	26.12	3.69	-11.16
	(-0.11)	(-0.06)	(-1.35)	(-1.30)	(-1.41)	(-1.58)	(5.09)	(1.03)	(-1.10)
Difference	-2.27	-2.44	7.57	0.65	0.25	0.50	-37.66	-4.00	7.76
	(-6.01)	(-7.24)	(1.18)	(1.13)	(1.25)	(1.42)	(-9.45)	(-1.20)	(0.97)
Sum	-2.71	-2.70	-14.41	-1.25	-0.45	-0.86	14.57	3.38	-14.56
	(-7.18)	(-8.02)	(-2.25)	(-2.17)	(-2.26)	(-2.44)	(3.66)	(1.02)	(-1.81)

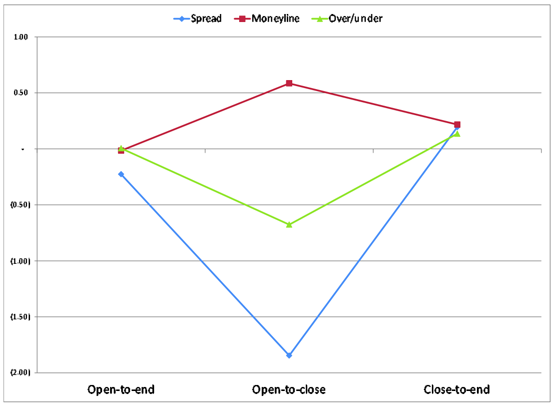
Table VIII: **Value and the Cross-Section of All Sports Betting Returns**

The table reports regression results from Table VII for the NFL (Panel A), MLB (Panel B), NHL (Panel C), and for all sports combined (Panel D), including the NBA. The regressions are the open-to-close returns and close-to-end returns on the value index measure as described in equations (6) and (7) estimated separately for Spread contracts, Moneyline contracts, and Over/under contracts. Spread contract results only pertain to the NBA and NFL, since there are no Spread contracts for MLB and the NHL, while the Moneyline and Over/under contracts pertain to all sports. Results are reported for both actual dollar returns from the contracts as well as hypothetical "point returns" based on points scored, except for the Moneyline.

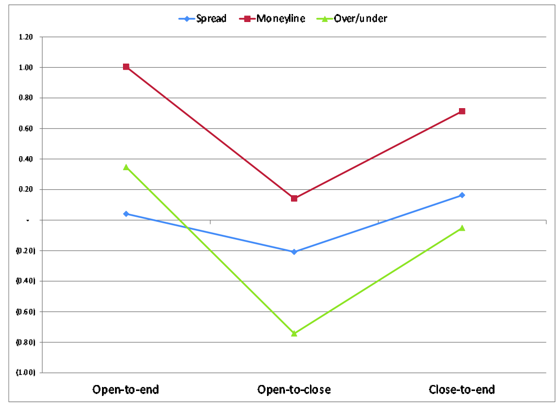
	Dollar returns			"Point returns"	
	Spread	Moneyline	O/U	Spread	O/U
			PANEL A: NFL		
		Open-to-close returns, $R_{open:close}$			
$\beta_{1,Val}$	0.04	-0.72	-1.54	-3.36	-2.29
	(0.19)	(-0.86)	(-1.58)	(-0.95)	(-0.45)
		Close-to-end returns, $R_{close:end}$			
$\beta_{T,Val}$	0.92	6.93	1.88	11.06	-12.95
	(0.37)	(0.35)	(1.99)	(0.33)	(-0.18)
			PANEL B: MLB		
		Open-to-close returns, $R_{open:close}$			
$\beta_{1,Val}$	–	-0.15	-0.17	–	-1.31
	–	(-2.40)	(-1.90)	–	(-2.47)
		Close-to-end returns, $R_{close:end}$			
$\beta_{T,Val}$	–	0.34	0.73	–	2.78
	–	(0.50)	(1.54)	–	(1.28)
			PANEL C: NHL		
		Open-to-close returns, $R_{open:close}$			
$\beta_{1,Val}$	–	-0.12	-0.46	–	-0.82
	–	(-1.35)	(-3.11)	–	(-3.30)
		Close-to-end returns, $R_{close:end}$			
$\beta_{T,Val}$	–	0.36	0.06	–	0.52
	–	(0.35)	(0.05)	–	(0.19)
			PANEL D: ALL SPORTS		
		Open-to-close returns, $R_{open:close}$			
$\beta_{1,Val}$	-0.09	-0.45	-0.13	-5.75	-2.10
	(-1.67)	(-3.18)	(-2.89)	(-3.96)	(-9.12)
		Close-to-end returns, $R_{close:end}$			
$\beta_{T,Val}$	0.17	1.07	0.11	16.58	0.58
	(0.96)	(0.89)	(0.60)	(1.21)	(0.31)
Difference	-0.26	-1.52	-0.24	-22.33	-2.68
	(-1.93)	(-0.61)	(-2.96)	(-1.07)	(-9.43)
Sum	0.08	0.62	-0.02	10.83	-1.52
	(0.60)	(0.25)	(-0.25)	(0.52)	(-5.35)

Figure 4: **Value Beta Patterns.** Plotted are the average t-statistics of the value betas estimated from equations (6) and (7) for open-to-end, open-to-close, and close-to-end returns for each betting contract type: Spread, Moneyline, and Over/under in each sport. The regressions are estimated for each value measure from Section III individually and then the t-statistics from these regressions are aggregated across all value measures and plotted below over each return horizon. The value betas are estimated and plotted separately for the NBA (Panel A), the NFL (Panel B), MLB (Panel C), the NHL (Panel D), the NBA and NFL combined (Panel E), and across all four sports (Panel F).

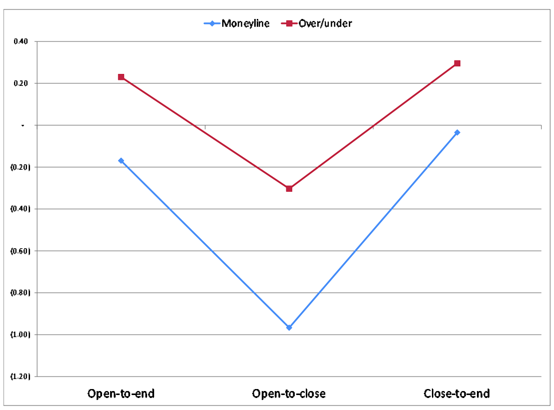
Panel A: NBA



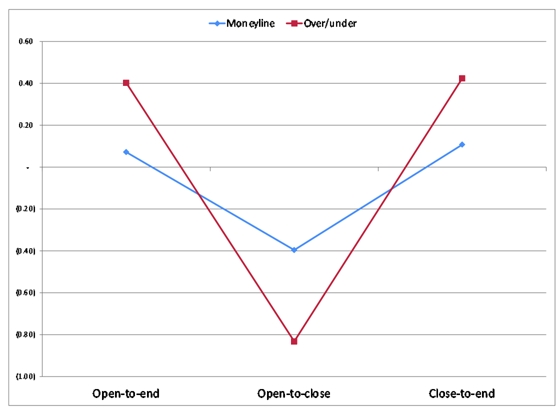
Panel B: NFL



Panel C: MLB



Panel D: NHL



Panel E: NBA + NFL

Panel F: All sports

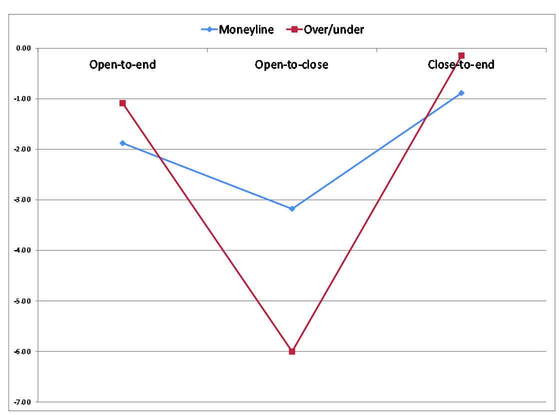


Table IX: **Size and the Cross-Section of NBA Betting Returns**

The table reports regression results of close-to-end returns and open-to-close returns on various size measures for Spread contracts (Panel A), Moneyline contracts (Panel B), and Over/under contracts (Panel C) in the NBA for 18,132 games from 1998 to 2013. Results are reported for both actual dollar returns from the contracts as well as hypothetical returns based on points scored, except for the Moneyline contract. Size is measured as the total franchise dollar value as estimated by Forbes, ticket revenue, total revenue, and player payroll expense of the team.

	Dollar returns				Point returns			
Size measures	Team value	Tix revenue	Total revenue	Player payroll	Team value	Tix revenue	Total revenue	Player payroll
Panel A: Spread contracts								
Open-to-close returns, $R_{open:close}$								
$\beta_{1,Size}$	0.15	0.01	0.01	0.01	-0.01	-0.01	-0.01	-0.01
	(0.66)	(0.63)	(0.50)	(0.97)	(-0.07)	(-0.36)	(-0.63)	(-1.72)
Close-to-end returns, $R_{close:end}$								
$\beta_{T,Size}$	-0.46	-0.02	-0.03	-0.07	-0.06	-0.01	-0.01	-0.01
	(-0.59)	(-0.38)	(-1.12)	(-1.40)	(-0.67)	(-1.44)	(-1.15)	(-2.11)
Panel B: Moneyline contracts								
Open-to-close returns, $R_{open:close}$								
$\beta_{1,Size}$	-0.10	-0.01	-0.01	0.00				
	(-0.58)	(-0.45)	(-0.66)	(0.09)				
Close-to-end returns, $R_{close:end}$								
$\beta_{T,Size}$	1.80	0.13	0.06	0.05				
	(1.72)	(1.93)	(1.71)	(0.68)				
Panel C: Over/under contracts								
Open-to-close returns, $R_{open:close}$								
$\beta_{1,Size}$	-0.27	-0.02	-0.01	-0.01	0.01	-0.01	0.00	0.00
	(-0.96)	(-1.24)	(-1.42)	(-0.53)	(0.76)	(-0.91)	(0.71)	(0.17)
Close-to-end returns, $R_{close:end}$								
$\beta_{T,Size}$	-0.35	-0.03	-0.01	-0.03	-0.01	0.01	0.00	-0.01
	(-0.44)	(-0.52)	(-0.13)	(-0.48)	(-0.06)	(0.01)	(0.17)	(-0.43)

Table X: **Size and the Cross-Section of All Sports Betting Returns**

The table reports regression results of close-to-end returns and open-to-close returns on various size measures for Spread contracts, Moneyline contracts, and Over/under contracts using the size index and estimated separately for the NFL (Panel A), MLB (Panel B), NHL (Panel C), and for all sports combined (Panel D), including the NBA. Spread contract results only pertain to the NBA and NFL, since there are no Spread contracts for MLB and the NHL, while the Moneyline and Over/under contracts pertain to all sports. The regressions are repeated here:

$$\tilde{R}_{i,O:C} = \alpha_{O:C} + \beta_{O:C,Char} Char_i + \tilde{\epsilon}_{i,O:C}$$
$$\tilde{R}_{i,C:E} = \alpha_{C:E} + \beta_{C:E,Char} Char_i + \tilde{\epsilon}_{i,C:E}.$$

Results are reported for both actual dollar returns from the contracts as well as hypothetical returns based on points scored, except for the Moneyline.

	Dollar returns			Point returns	
	Spread	Moneyline	O/U	Spread	O/U
Panel A: NFL					
	Open-to-close returns, $R_{open:close}$				
$\beta_{1,Size}$	-0.14	-0.10	-0.08	-1.15	-2.96
	(-0.49)	(-0.37)	(-0.39)	(-0.77)	(-2.67)
	Close-to-end returns, $R_{close:end}$				
$\beta_{T,Size}$	0.52	0.36	0.00	6.85	-5.79
	(0.53)	(0.22)	(0.01)	(0.52)	(-0.64)
Panel B: MLB					
	Open-to-close returns, $R_{open:close}$				
$\beta_{1,Size}$	–	0.03	0.05	–	0.43
	–	(0.89)	(0.59)	–	(3.36)
	Close-to-end returns, $R_{close:end}$				
$\beta_{T,Size}$	–	0.52	0.62	–	2.25
	–	(1.42)	(1.69)	–	(1.34)
Panel C: NHL					
	Open-to-close returns, $R_{open:close}$				
$\beta_{1,Size}$	–	0.27	0.12	–	0.45
	–	(1.16)	(0.74)	–	(3.11)
	Close-to-end returns, $R_{close:end}$				
$\beta_{T,Size}$	–	4.11	-0.52	–	-6.59
	–	(1.71)	(-0.29)	–	(-1.00)
Panel D: All sports					
	Open-to-close returns, $R_{open:close}$				
$\beta_{1,Size}$	0.04	-0.07	0.01	-0.48	-0.01
	(0.23)	(-0.46)	(0.03)	(-0.60)	(-0.17)
	Close-to-end returns, $R_{close:end}$				
$\beta_{T,Size}$	-0.09	1.40	-0.02	-1.22	0.01
	(-0.14)	(1.58)	(-0.17)	(-0.17)	(0.53)

Table XI: **Momentum and Value When Information is Uncertain: Evidence from Sports Betting Markets**

Panel A reports returns from momentum and value strategies in sports betting markets examined separately when uncertainty of the contract values is high versus low. Contract uncertainty is measured by the first 25 percent of games in each sports season, where the quality of teams, and therefore betting contracts, is more uncertain, compared to games during the remaining 75 percent of the season. Momentum and value returns are computed from open-to-close and from close-to-end for Spread, Moneyline, and Over/under contracts averaged across all sports (NBA, NFL, MLB, and NHL), but where Spread contracts are only from the NBA and NFL, over the period 1985 to 2013. The difference in returns between the high and low uncertainty games is also reported, with t-statistics in parentheses. Panel B reports momentum and value returns for betting contracts that have high and low parlay volume. Parlay bets are those that involve a portfolio of bets across multiple games simultaneously, where a payoff only occurs if all the bets involved in the parlay win. Hence, games more likely to be involved in parlay bets should be those where bettors have the most confidence and believe the bet is most likely to win. Ranking betting contracts based on parlay volume (the fraction of outstanding bets involved in a parlay bet), Panel B reports momentum and value returns from the open-to-close and from the close-to-end for betting contracts among the highest third in parlay volume and among the lowest third in parlay volume, as a proxy for bettor uncertainty.

PANEL A: EARLY VERSUS LATE IN THE SEASON

	Spread contracts			Moneyline contracts			O/U contracts		
In-season:	Late	Early	Early-Late	Late	Early	Early-Late	Late	Early	Early-Late
Momentum									
Open-to-close	0.02	0.09	0.11	-0.05	0.08	0.13	0.09	0.20	0.11
	(0.84)	(3.56)	(3.62)	(-1.30)	(2.09)	(2.83)	(1.07)	(2.38)	(1.15)
Close-to-end	-0.04	-0.21	-0.17	-0.10	-0.31	-0.21	-0.21	-0.25	-0.04
	(-0.59)	(-2.89)	(-1.47)	(-1.19)	(-3.78)	(-1.09)	(-0.88)	(-1.05)	(-0.16)
Difference	0.05	0.30	0.27	0.05	0.39	0.34	0.30	0.45	0.15
	(1.39)	(5.04)	(3.35)	(0.71)	(5.34)	(2.48)	(1.56)	(2.65)	(1.10)
Value									
Open-to-close	0.07	-0.10	-0.18	0.24	-0.01	-0.25	0.11	0.08	-0.03
	(2.03)	(-1.46)	(-1.70)	(3.50)	(-0.16)	(-4.01)	(1.62)	(1.16)	(-0.24)
Close-to-end	-0.80	0.07	0.87	-0.99	-0.39	0.60	-0.18	0.07	0.24
	(-3.35)	(0.30)	(3.21)	(-3.52)	(-1.38)	(2.06)	(-1.46)	(0.56)	(0.55)
Difference	0.87	-0.18	-1.05	1.23	0.38	-0.85	0.29	0.01	-0.28
	(3.89)	(-1.26)	(-3.31)	(5.11)	(1.37)	(-3.78)	(2.85)	(0.13)	(-1.68)

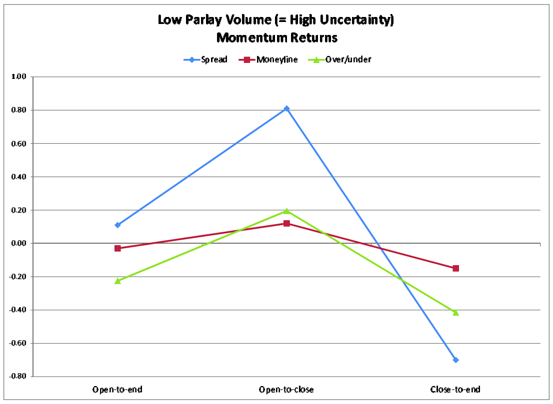
Panel B: Low versus High Parlay Bets

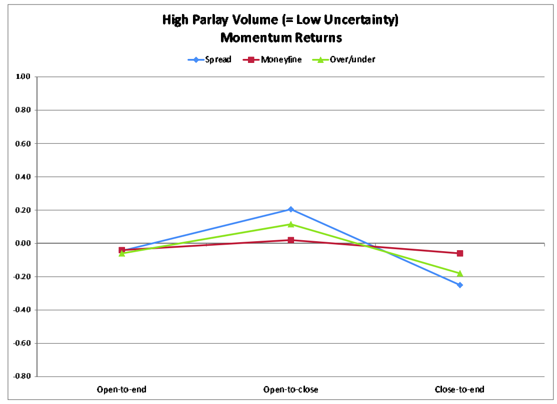
Parlay volume:	Spread contracts High	Spread contracts Low	Spread contracts Low−High	Moneyline contracts High	Moneyline contracts Low	Moneyline contracts Low−High	O/U contracts High	O/U contracts Low	O/U contracts Low−High
Momentum									
Open-to-close	0.01	0.15	0.14	0.02	0.12	0.10	0.02	0.10	0.08
	(1.20)	(2.20)	(2.17)	(1.40)	(2.40)	(2.23)	(1.98)	(2.46)	(2.25)
Close-to-end	-0.05	-0.10	-0.05	-0.06	-0.15	-0.09	-0.02	-0.10	-0.08
	(-0.66)	(-1.43)	(-0.70)	(-0.83)	(-1.50)	(-0.98)	(-1.53)	(-1.72)	(-1.58)
Difference	0.06	0.25	0.19	0.08	0.27	0.19	0.04	0.20	0.16
	(0.93)	(3.63)	(2.86)	(1.38)	(3.47)	(2.83)	(1.45)	(3.05)	(2.71)
Value									
Open-to-close	0.00	0.00	0.01	-0.01	0.00	0.01	-0.02	0.00	0.01
	(-1.05)	(0.54)	(0.99)	(-1.75)	(-0.26)	(1.69)	(-0.71)	(-1.05)	(0.69)
Close-to-end	0.00	0.02	0.02	0.07	0.01	-0.05	0.05	0.03	-0.02
	(0.52)	(0.69)	(0.64)	(2.14)	(1.34)	(-1.89)	(2.86)	(1.68)	(-1.33)
Difference	-0.01	-0.01	-0.01	-0.08	-0.02	0.06	-0.06	-0.03	0.04
	(-1.10)	(-0.60)	(-0.49)	(-2.98)	(-1.48)	(2.66)	(-3.55)	(-1.94)	(2.05)

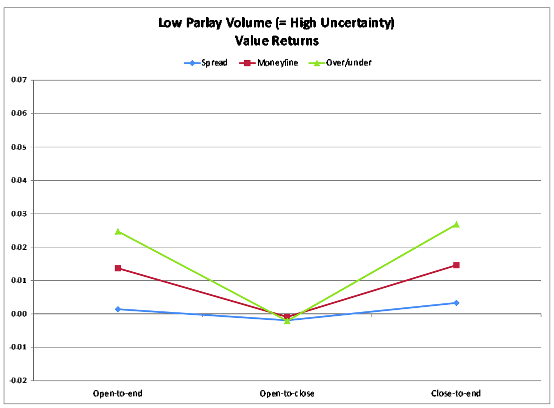
Figure 5: **Momentum and Value Return Patterns When Information is Uncertain – Sports Betting Markets** Plotted are the average returns to momentum and value for open-to-end, open-to-close, and close-to-end returns for each betting contract type: Spread, Moneyline, and Over/under averaged across all sports. The momentum and value indicies are used as the return predictor variable for dollar returns on each of the contracts within each sport and then averaged across sports. The spread contracts only pertain to the NBA and NFL, but the moneyline and over-under contracts are across all four sports leagues: NBA, NFL, NHL, and MLB. Panel A reports results where uncertainty is measured by early games in the season versus later games in the season, with the former containing more uncertainty. Panel B reports results where uncertainty is measured by splitting the sample into the third of games with least parlay volume bets and those with the most parlay volume bets, with the former containing more bettor uncertainty.

Panel A: Early versus Late in Season (All sports)

Panel B: Low versus High Parlay Games (All sports)







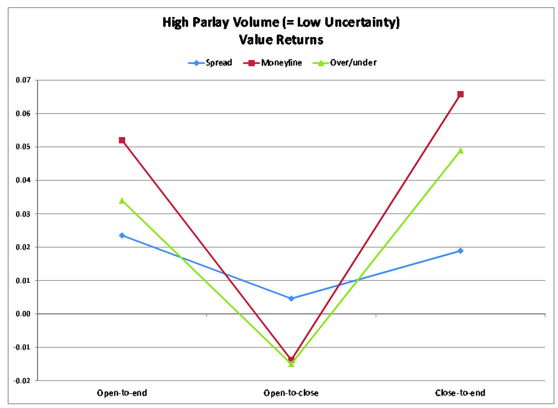


Table XII: **Momentum and Value When Information is Uncertain: Evidence from Financial Markets**

The table reports returns from momentum and value strategies in U.S. equity markets when uncertainty of firm earnings is high versus low. Panel A reports results using the recency of earnings announcements as a proxy for uncertainty, where firms with high earnings uncertainty are those whose most recent earnings number is stale, measured by having their most recent earnings announcement be at least 11 weeks old prior to portfolio formation, while firms with low earnings uncertainty are those with very recent earnings announcements that have occurred within the previous two weeks before portfolio formation. Momentum and value returns are computed for U.S. stocks from the Center for Research in Security Prices over the July 1963 to December 2013 sample period for firms with recent and distant earnings announcements, where firms are first sorted into recent (less than two weeks) or stale (greater than 11 weeks) most recent earnings announcement and then within each group are sorted by their momentum (past 12-month average returns, skipping the most recent month) or value (book-to-price ratio) characteristic into thirds, and a portfolio long the top third of stocks and short the bottom third, where stocks are value weighted by their beginning of month market capitalization within each group, is created within the recent and stale earnings announcement firms. The returns to the long-short momentum (winners minus losers) and value (high BE/ME minus low BE/ME) portfolios within recent and stale earnings announcement groups are reported along with their differences. Panel B reports results using the dispersion in analyst forecasts of one-year earnings per share to capture uncertainty. Following Diether, Malloy, and Scherbina (2002), dispersion in analyst forecasts is measured as the cross-sectional standard deviation across analysts of their fiscal year earnings-per-share forecast, scaled by the mean forecast for a given stock (requiring at least five analysts). Firms are sorted independently into five quintiles based on analyst dispersion and five quintiles based on momentum or value, where the intersection between each group produces 25 analyst dispersion× momentum categories and 25 analyst dispersion× value categories. The value weighted average return of each group is then computed monthly from July 1990 to December 2016. The returns to a strategy that is long the top quintile of momentum and short the bottom quintile of momentum is then computed for each analyst dispersion quintile and reported below are the average annualized returns of the momentum profits within the lowest analyst dispersion stocks (lowest 20%) and the highest analyst dispersion stocks (highest 20%), along with their t-statistics, as well as the differences between mometum profits among low versus high analyst dispersion stocks. The same set of results is reported for value strategy profits that go long the highest quintile of BE/ME stocks and short the lowest within the lowest and highest analyst dispersion firms. For both panels, both the raw average return difference and 4-factor alpha from a regression of the returns on the Fama and French market, size, value, and momentum factors RMRF, SMB, HML, and UMD, are reported.

Panel A: Early versus Late in the Earnings Cycle

	Earnings announcement			Earnings announcement		
	Recent	Stale	Stale-Recent	Recent	Stale	Stale-Recent
	< 2 weeks old	> 11 weeks old		< 2 weeks old	> 11 weeks old	
	Raw returns			4-factor α		
Momentum	1.98	11.81	9.83	-5.39	6.05	11.44
	(0.41)	(2.44)	(1.95)	(-1.39)	(1.44)	(2.18)
Value	12.12	2.96	-9.15	8.80	-1.58	-10.39
	(2.79)	(0.72)	(-1.91)	(2.52)	(-0.44)	(-2.11)

Panel B: Low versus High Analyst Dispersion

	Analyst forecast dispersion			Analyst forecast dispersion		
	Low	High	High - Low	Low	High	High - Low
	Raw returns			4-factor α		
Momentum	1.74	6.36	4.68	3.12	7.92	4.92
	(0.34)	(2.05)	(2.10)	(0.84)	(2.52)	(2.14)
Value	5.64	-0.60	-6.24	5.34	-1.02	-6.30
	(2.42)	(-0.06)	(-1.58)	(2.17)	(-0.01)	(-1.50)

Figure 6: **Price Momentum and Reversals in Financial Markets by Recency of Earnings Announcement.** Plotted are the difference in momentum premia between stale versus recent earnings announcers for various windows of time used to define "stale" and "recent:" > 11 versus < 2 weeks since announcement, > 8 versus < 4 weeks since announcement, and > 7 versus < 6 weeks since announcement. In addition, the figure plots the difference in returns of these momentum portfolios in years two to three *after* portfolio formation, which represent the subsequent reversals that often accompany momentum one-year momentum returns.

Table XIII: **Trading Strategy Profits**

The table reports trading strategy profits from using momentum, value, and size to select sports betting contracts. Using the weighted average index measures for each characteristic (momentum, value, and size) for each betting contract type, a trading strategy is formed using these characteristics to predict upward betting line movements (e.g., continuation) and negative line movements (reversals). The betting strategy invests positively (goes long) when predicted movements are expected to be positive and invests negatively (shorts or takes the other side of the bet) when line movements are predicted to be negative, where the dollars bet are in proportion to how strong the predicted movement is relative to the average. The total bets summed across all games are then rescaled to add up to one dollar. All bets across all contracts from all games in a month are then aggregated and the return recorded. Results are reported separately for each characteristic by itself as well as a combination of all three characteristics ("multi strategy"). Trading strategy profits are computed for both the line movement from open to close (open-to-close returns) and for the return from the closing line to the game outcome (close-to-end return). Average gross returns and net returns that account for transactions costs (the "vig") are reported as well as the gross Sharpe ratio of the strategy. All reutrns and Sharpe ratios are reported as annualized numbers. Panel A reports results for the NBA only and Panel B reports results for all four sports by combining all of the bets across the NBA, NFL, MLB, and NHL into one portfolio, where each sport is given equal weight in the portfolio. Most months contain only two or three sports given that each sport's season lasts less than a year and only has partial overlap with other sports. Finally, Panel C reports a comparison set of returns and Sharpe ratios for momentum, value, and size strategies in financial markets. Using the Fama and French long-short factors for size (SMB), value (HML), and momentum (UMD) from Ken French's website for U.S. stocks and international factors constructed the same way from international equity returns following Asness, Frazzini, Israel, Moskowitz, and Pedersen (2014), the annualized market-adjusted gross return or CAPM alpha of each strategy is reported along with an estimate of the net alpha using the results from Frazzini, Israel, and Moskowitz (2014), who estimate price impact and total trading costs of these three strategies. The annualized information ratio (CAPM alpha divided by residual standard deviation) on these strategies is also reported, gross of trading costs. The sports betting contract returns pertain to the period November 1998 to March 2013 and the financial market returns cover the period January 1972 to December 2013.

	Panel A: NBA (1998 to 2013)			**Panel B: All sports (1998 to 2013)**		
	Gross return	Net return	Sharpe	Gross return	Net return	Sharpe
			Open-to-close returns, $R_{open:close}$			
Multi strategy	0.64	-6.66	0.22	0.47	-4.45	0.31
Momentum	0.60	-6.71	0.11	0.39	-4.54	0.21
Value	0.21	-7.08	0.06	0.11	-4.79	0.08
Size	-0.26	-7.53	-0.04	0.10	-4.78	0.03
			Close-to-end returns, $R_{close:end}$			
Multi strategy	2.56	-1.18	0.13	2.61	-0.58	0.25
Momentum	0.52	-3.15	0.05	0.73	-2.41	0.08
Value	1.56	-2.15	0.08	0.69	-2.45	0.10
Size	-0.58	-4.20	-0.03	0.17	-2.95	0.01

Panel C: Financial market returns

	U.S. Stocks (1972 to 2013)			Global stocks (1972 to 2013)		
	Gross alpha*	Net alpha†	Info ratio	Gross alpha*	Net alpha†	Info ratio
Multi strategy	5.22	3.22	0.81	7.50	5.50	1.38
UMD (momentum)	10.45	6.95	0.67	8.10	4.60	0.80
HML (value)	3.76	2.26	0.33	6.10	4.60	0.54
SMB (size)	1.17	-0.33	0.11	2.36	0.86	0.24

*Alpha is defined relative to the CAPM using the CRSP value-weighted index stock return minus the T-bill rate, $RMRF$.

†Net alpha is estimated as gross alpha minus an estimate of the average trading cost of each strategy from Frazzini, Israel, and Moskowitz (2014), which includes price impact costs at the implied total fund sizes in Frazzini, Israel, and Moskowitz (2014), which are $3.7, $7.0, and $15.1 billion for UMD, HML, and SMB, respectively.

Table XIV: **Covariance Structure Among Momentum and Value Bets**

Reported are regression results of portfolios of games formed on momentum (and value) and regressed on a momentum (value) "factor." The test portfolios are formed by sorting all games in a given month (with at least 40 games) by their momentum (value) characteristic, which is the weighted-average index momentum (value) variable, into five quintiles and then taking the equal-weighted average return of all games within each group. This provides a monthly return to quintile sorted portfolios based on momentum (value), whose returns are then regressed on the monthly returns of the momentum (value) factor, which is the high minus low quintile spread returns, Q5 − Q1. Panel A reports the results for test assets and factors formed from the same games such that there is overlap of the betting contracts that comprise the test portfolios and the factor. Panel B reports results for the same exercise where the test assets/portfolios are formed from one set of games and the factors are formed from a completely different set of games. To compute the test portfolios and the factors independently each month the number of games are split randomly into two groups, with one used to form the test assets and the other used to form the factors, so that there is no overlap in games on the left and right hand side of the regression. Reported are the coefficient estimates (β) on the factor, its t-statistic in parentheses, and the R^2 from each regression. The intercept is not reported for brevity.

	Momentum					Value				
	Low	Q2	Q3	Q4	High	Low	Q2	Q3	Q4	High
Panel A: Overlapping test assets and factors										
Spread contract returns										
β	-0.521	0.109	0.121	-0.118	0.479	-0.427	0.099	0.018	-0.023	0.573
	(-8.98)	(1.29)	(1.18)	(-1.07)	(8.25)	(-5.14)	(1.10)	(0.16)	(-0.29)	(6.89)
R^2	0.52	0.02	0.02	0.02	0.48	0.55	0.06	0.001	0.004	0.68
Moneyline contract returns										
β	-0.356	-0.006	0.143	0.041	0.644	-0.929	-0.232	-0.055	0.206	0.071
	(-3.35)	(-0.04)	(0.98)	(0.13)	(6.06)	(-6.61)	(-0.60)	(-0.11)	(0.58)	(0.51)
R^2	0.26	0.000	0.03	0.001	0.54	0.88	0.06	0.002	0.05	0.04
Over/under contract returns										
β	-0.256	-0.056	-0.106	0.036	0.745	-0.171	-0.251	-0.149	0.088	0.829
	(-4.43)	(-0.56)	(-1.26)	(0.37)	(12.92)	(-1.75)	(-1.96)	(-0.82)	(0.52)	(8.51)
R^2	0.30	0.007	0.03	0.003	0.79	0.18	0.22	0.05	0.02	0.84
Panel B: Non-overlapping test assets and factors										
Spread contract returns										
β	0.016	0.056	0.146	0.008	-0.053	-0.177	-0.123	-0.008	0.031	-0.169
	(0.16)	(0.54)	(1.26)	(0.07)	(-0.62)	(-0.86)	(-0.86)	(-0.08)	(0.30)	(-0.83)
R^2	0.000	0.004	0.02	0.000	0.005	0.03	0.04	0.000	0.005	0.03
Moneyline contract returns										
β	-0.098	-0.131	-0.092	0.086	-0.379	-0.120	-0.059	0.086	0.235	-0.310
	(-0.80)	(-1.02)	(-0.53)	(0.18)	(-2.15)	(-0.30)	(-0.25)	(0.15)	(0.58)	(-1.44)
R^2	0.02	0.03	0.01	0.001	0.13	0.02	0.01	0.004	0.05	0.26
Over/under contract returns										
β	0.102	-0.004	0.067	-0.043	0.076	-0.157	-0.432	-0.366	0.191	-0.125
	(1.35)	(-0.04)	(0.66)	(-0.48)	(0.50)	(-1.25)	(-2.55)	(-1.75)	(0.97)	(-0.52)
R^2	0.04	0.000	0.01	0.005	0.01	0.10	0.32	0.18	0.07	0.02

Appendices

A Appendix

A. Computing Prices and Returns

To compute returns, each contract line is converted into a price by estimating the probability, $\pi(S_t)$, of a payoff occurring and calculating the expected value of the contract based on the probability and value of each payoff state for that contract, where the payoff probability depends on the time t betting line (Spread, Moneyline, or Over/under value). For example, using equation (1) the price of a Spread contract at the terminal date T is,

$$\mathrm{P}_T^S = (210)\,\mathrm{I}\{\text{cover}\} + (110)\,\mathrm{I}\{\text{push}\} + (0)\,\mathrm{I}\{\text{fail}\} \tag{10}$$

where $\mathrm{I}\{\text{cover}\}$ and $\mathrm{I}\{\text{push}\}$ are indicator functions for the payoff outcomes of the contract at time T. For dates $t < T$ before the terminal payoff the price of the Spread contract is,

$$\mathrm{P}_t^S = 210 \times \pi_{cover}\,(S_t) + 110 \times \pi_{push}\,(S_t) + 0 \times \pi_{fail}(S_t) \tag{11}$$

where the probability of cover, push, and fail may vary with the value of the spread, S_t.

For Moneyline contracts, using equation (2) (assuming contracts payoff \$100 if $m < 0$ and cost \$100 for $m > 0$) prices are

$$\begin{aligned} \mathrm{P}_T^m &= (|m_T| + 100)\,\mathrm{I}\{\text{win}\} + [\max(m_T, 100)]\,\mathrm{I}\{\text{tie}\} \\ \mathrm{P}_t^m &= \begin{cases} (100 - m_t)\,\pi_{win}\,(m_t) + (-m_t)\,\pi_{tie}\,(m_t) + 5 & \text{if } m_t < 0 \\ (100 + m_t)\,\pi_{win}\,(m_t) + (100)\,\pi_{tie}\,(m_t) + 5 & \text{if } m_t > 0 \end{cases} \end{aligned} \tag{12}$$

where m_t is the moneyline value at time t and the +5 accounts for the vig charged to both sides of the contract (assumed to be split evenly between both sides). A "tie" for the Moneyline contract means the two teams actually tied, which is very rare, hence ties are effectively ignored.[23]

Finally, the prices of Over/under contracts are very similar to those of Spread contracts. Using equation (3) prices for betting on the "over" are,

$$\begin{aligned} \mathrm{P}_T^{O/U} &= (210)\,\mathrm{I}\{\text{over}\} + (110)\,\mathrm{I}\{\text{push}\} + (0)\,\mathrm{I}\{\text{under}\} \\ \mathrm{P}_t^{O/U} &= 210 \times \pi_{over}\,(O/U_t) + 110 \times \pi_{push}\,(O/U_t) + 0 \times \pi_{under}(O/U_t). \end{aligned} \tag{13}$$

Returns for all contracts $c \in \{S, m, O/U\}$ over the three horizons above are then simply,

$$\mathrm{R}_{open:end}^c = \frac{P_T^c}{P_0^c}, \qquad \mathrm{R}_{close:end}^c = \frac{P_T^c}{P_1^c}, \qquad \mathrm{R}_{open:close}^c = \frac{P_1^c}{P_0^c}. \tag{14}$$

A.1. Estimating payoff probabilities

To calculate prices and returns, the probabilities of payoff states for each contract for both opening and closing lines are calculated. Specifically, cover and push probabilities for opening and closing Spreads, home team win probabilities for opening and closing Moneylines, and over and push probabilities for opening and closing Over/under totals are computed.

[23] Across all of the 59,592 games in the dataset, only 10 games ended in a tie, and only five of those had opening betting lines, all of which are in the NFL. Ties are not allowed in the NBA, MLB, and are not allowed in the NHL since the 2005-2006 season, which is the start of the data, when the NHL instituted shoot outs at the end of tie games to eliminate ties. None of the results are affected by excluding the five tied games in the NFL.

Two approaches are adopted for computing probabilities. The first is a theoretical approach, where contract prices are set such that the odds of winning and losing each bet are equal. The second approach empirically estimates the probabilities of the payoff states from the data, where a logit, probit, and non-parametric kernel density estimator are used to extract empirical probabilities of the payoff states. Levitt (2004) argues and finds that bookmakers try to set their lines/prices such that each side has roughly an equal chance of winning, but also shows that bookmakers will sometimes deviate from this strategy when they can better predict game outcomes or betting interest. Hence, the theoretical equal probability approach likely captures the majority of cases, while the empirical approach should capture any deviations from this, subject to estimation uncertainty. In both cases, the probability assumptions will only affect the results if either model error or estimation error is correlated with the cross-sectional characteristics examined. The returns (and all of the results in the paper) are robust across a variety of probability estimation methods.

The four methods for computing probabilities obtain nearly identical results. Panel A of Figure A1 plots the probability of covering or pushing for Spread contracts at every spread value for NBA contracts. A small amount of random noise is added to each data point to gauge the number of contracts at each spread value. The probabilities associated with full-point spreads and half-point spreads are indicated separately on the graphs, where because half-point spreads cannot result in pushes, the estimated probabilities exhibit a jagged saw-tooth pattern as spreads move from full-point to half-point values. Panels B and C of Figure A1 plot the probability of winning the bet for opening and closing Moneylines and O/U contracts, respectively in the NBA. (Pushing is not an issue for Moneyline contracts [see the discussion in footnote 23], so I do not estimate push probabilities.) The probability estimates are all very similar and match closely the theoretical ones. Given the probabilities of the payoff states for each contract and betting line, I compute prices and returns using equations (10)–(14).

Figure A2 plots the payoff state probabilities for the other sports (NFL, MLB, and NHL), which exhibit similar patterns.

A.2. Probability Estimators

Theoretical probabilities. The first set of probabilities assumes all contract prices are set such that winning and losing is roughly equal. Under this scenario, all contracts face the same probability of covering or failing, so a spread of 20 points is just as likely to pay off as is a spread of two points. This turns out to be very close to what bookmakers actually do, but there are some deviations (see Levitt (2004)), which are taken into account below.

For spread and over/under payoff probabilities, 0.5 is used for cover and fail (or over and under) probabilities for half-point spreads, where pushes are impossible. For full-point spreads, where a push is possible, I first compute the empirical unconditional push probability for that contract type (separately for opening and closing lines) using the full sample of data for all contracts facing the exact same point spread or over/under total, where a minimum of five observations is required to compute this probability. Then, using that empirical push probability I divide the remaining probability equally for winning and losing the bet, $\pi(\text{cover}) = \pi(\text{fail}) = (1 - \hat{\pi}(\text{push}))/2$.

For Moneyline contracts, where the probability of winning and losing is clearly not equal, since payoffs adjust rather than spreads, I solve for π as follows:

$$\begin{aligned} (100)\,\pi\,(\text{win}) + (m)\,[1 - \pi\,(\text{win})] &= -5 \qquad \text{if } m < 0 \\ (m)\,\pi\,(\text{win}) + (-100)\,[1 - \pi\,(\text{win})] &= -5 \qquad \text{if } m > 0 \end{aligned}$$

where m is the actual Moneyline quote. This simplifies to

$$\pi\,(\text{win}) \quad = \quad \frac{-\,(m+5)}{100-m} \qquad \text{if } m < 0$$

$$\pi\left(\text{win}\right) = \frac{95}{100+m} \qquad \text{if } m > 0.$$

The assumption here is that payoffs × probability are set to equal half of the commission or vig on both sides of the contract.

Empirically estimated probabilities. Rather than assume all probabilities of winning or losing the bet are the same across all lines, I also estimate empirically the probability of payoffs for different lines, to take into account that bookmakers may deviate from equal probability bets.

For Spread contracts, for example, the frequency of cover, push, and fail for all contracts facing the same spread for all spreads within each sport is estimated. The assumption here is that probabilities for a given spread are the same across games within a given sport. This is similar in spirit to the literature that tries to recover real probabilities from risk-neutral probabilities from option contracts (e.g., Ross (2013), Andersen, Fusari, and Todorov (2014), Borovicka, Hansen, and Scheinkman (2014), Jensen, Lando, and Pedersen (2015)). However, the difference here is that sports betting contracts are not confounded by risk premia, hence extraction of the true probabilities from prices is simpler. However, since some spreads have a limited number of observations, it is difficult to reliably use sample frequencies to estimate probabilities at all spreads.[24] Therefore, I use several discrete choice models—logit, probit, and a non-parametric Gaussian kernel density estimator—to estimate probabilities. Specifically, I require a minimum of five observations for each spread, and estimate probabilities of payoff states for each opening and closing line for each contract under each of the three discrete choice models. For contracts with many observations, the empirical frequency matches the probability estimates nicely, including the theoretically-motivated probabilities.

The logit model estimates probabilities from the following regression

$$\pi_i(\text{cover, push}) = 1/\left(1+exp\left(-X_i\beta\right)\right),$$

where π_i is the probability of each payoff state at spread (moneyline or over/under total) i and X_i is the spread (moneyline, or over/under line).

The probit model estimates probabilities from,

$$\pi_i(\text{cover, push}) = \Phi\left(X_i\beta\right)$$

where Φ is the CDF of the normal distribution.

The non-parametric kernel density estimator weights observations with a normal weighting function multiplied by the number of observations (for each spread, moneyline, over/under line) in order to better approximate the probabilities for extreme spreads, moneylines, and over/under totals where very few contracts exist. The normal weighting function uses half the standard deviation of spreads, moneylines, or over/under totals when estimating the probabilities for the cover, win, and over probabilities, respectively. Specifically, the probability of covering/winning for the ith spread/moneyline/OU, line_i, is estimated by using the average cover probabilities for all other lines, weighted by the normal density where the mean is line_i and the standard deviation is half the standard deviation of the distribution of unique lines. Each weight is then multiplied by the number of observations. The weights are normalized by dividing by the sum of all the individual weights.

$$\begin{aligned}\pi_i(\text{cover, push}) &= \sum_{j=1}^{N}\left\{\frac{w_{i,j}}{\sum_{k=1}^{N} w_{i,k}} \times \Pr\left(\text{cover, push}|\text{line}_j\right)\right\} \\ w_{i,j} &= \Psi\left(\text{line}_j - \text{line}_i, \frac{\sigma\left(\text{all unique lines}\right)}{2}\right) \times \left(\text{N}_i\right).\end{aligned}$$

[24] For example, because of the few number of observations, using the empirical frequency of payoffs to estimate probabilities implies that a 14.5 point favorite in the NBA pays off 100% of the time, while a 15 point favorite never pays off. Clearly, these are poor estimates of the true probabilities of payoffs for these spreads, which should intuitively be nearly identical.

where Ψ is the normal pdf, N_i are the number of contracts with spread (moneyline or over/under total) equal to i, and j refers to the jth contract with spread i.

For Moneyline contracts, the equations are estimated separately for moneylines > 0 and < 0 since there is a discontinuity between -100 and 100 (because of the vig).

The distribution of returns is nearly identical using other probability estimates such as logit, probit, or the model-implied probabilities. Table A1 reports return correlations across the different probability estimates, where correlations are between 0.993 and 1.000 for open-to-end and close-to-end returns and between 0.90 to 0.99 for open-to-close returns. Consequently, results in the paper are virtually identical under all probability measures. The paper's results use return estimates for the non-parametric probability estimates for the Spread and O/U contracts and the probit estimates for the Moneyline contract. Given the extreme values of some of the Moneyline contracts and the small number of observations at those extremes, the non-parametric estimator is less reliable than the more constraining probit (or logit) estimator.

Table B1: Return Correlations Under Different Probability Models

The table reports return correlations across different estimates of the probability of outcome payoffs for each of the three betting contracts in the NBA: the Spread contract (S), the Moneyline contract (ML), and the Over/under contract (O/U). Three sets of returns are calculated for each contract: the return from the opening line to the outcome (open:end), the return from the closing line to the outcome (close:end), and the return from the opening line to the closing line (open:close). Returns are calculated following equations (10-14), using the probabilities estimated from four models: logit, probit, a model that assumes winning and losing are equally likely, and the non-parametric function. Correlation of the returns estimated across these various probability models are reported below for each set of returns and for each contract type on each game.

	Open-to-end returns				Close-to-end returns				Open-to-close returns			
	Logit	Probit	Model	Non-par	Logit	Probit	Model	Non-par	Logit	Probit	Model	Non-par
						Spread contract						
Logit	1.000	1.000	0.998	0.997	1.000	1.000	0.995	0.996	1.000	1.000	0.906	0.909
Probit	1.000	1.000	0.998	0.997	1.000	1.000	0.995	0.996	1.000	1.000	0.906	0.908
Model	0.998	0.998	1.000	1.000	0.995	0.995	1.000	1.000	0.906	0.906	1.000	0.999
Non-par	0.997	0.997	1.000	1.000	0.996	0.996	1.000	1.000	0.909	0.908	0.999	1.000
						Moneyline contract						
Logit	1.000	1.000	0.992	0.994	1.000	1.000	0.993	0.995	1.000	0.996	0.807	0.951
Probit	1.000	1.000	0.989	0.994	1.000	1.000	0.991	0.995	0.996	1.000	0.764	0.958
Model	0.992	0.989	1.000	0.987	0.993	0.991	1.000	0.989	0.807	0.764	1.000	0.705
Non-par	0.994	0.994	0.987	1.000	0.995	0.995	0.989	1.000	0.951	0.958	0.705	1.000
						Over/under contract						
Logit	1.000	1.000	0.999	0.999	1.000	1.000	0.997	0.997	1.000	1.000	0.967	0.976
Probit	1.000	1.000	0.999	0.999	1.000	1.000	0.997	0.997	1.000	1.000	0.967	0.976
Model	0.999	0.999	1.000	1.000	0.997	0.997	1.000	1.000	0.967	0.967	1.000	0.999
Non-par	0.999	0.999	1.000	1.000	0.997	0.997	1.000	1.000	0.976	0.976	0.999	1.000

Table B2: Correlation of Point Returns

Panel A reports return correlations between dollar-denominated returns and point-denominated returns for the Spread contract (S) and the Over/under contract (O/U) for three sets of returns for each contract: the return from the opening line to the outcome (open:end), the return from the closing line to the outcome (close:end), and the return from the opening line to the closing line (open:close). Panel B reports the correlations among the point-denominated returns across the different contracts and different return horizons. Returns are calculated following equations (10-14), using the probabilities estimated from the non-parametric specification.

Panel A: Correlation Between Dollar and Point Returns

	Spread contract	Over/under contract
Correlation($R^{\$}_{open:end}, R^{pts.}_{open:end}$) =	0.79	0.79
Correlation($R^{\$}_{close:end}, R^{pts.}_{close:end}$) =	0.79	0.79
Correlation($R^{\$}_{open:close}, R^{pts.}_{open:close}$) =	0.28	0.26

Panel B: Point Return Correlations

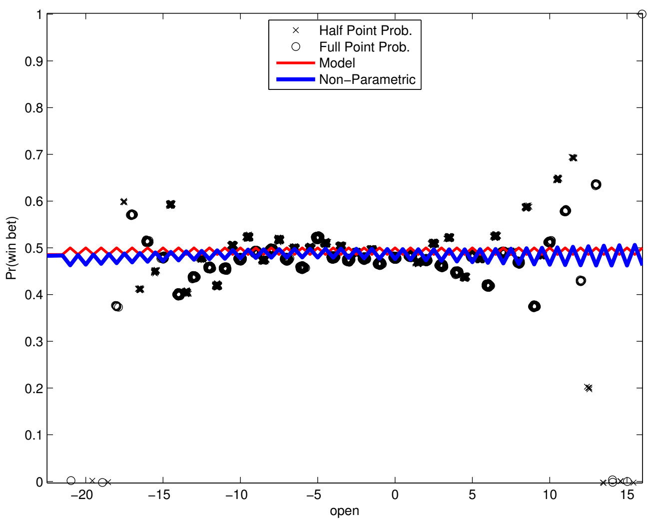
	$R^{\mathrm{S}}_{open:end}$	$R^{\mathrm{S}}_{close:end}$	$R^{\mathrm{S}}_{open:close}$	$R^{\mathrm{O/U}}_{open:end}$	$R^{\mathrm{O/U}}_{close:end}$	$R^{\mathrm{O/U}}_{open:close}$
$R^{\mathrm{S}}_{open:end}$	1.00	0.99	0.10	-0.01	-0.01	0.01
$R^{\mathrm{S}}_{close:end}$		1.00	-0.01	-0.01	-0.01	0.01
$R^{\mathrm{S}}_{open:close}$			1.00	0.00	0.00	0.03
$R^{\mathrm{O/U}}_{open:end}$				1.00	0.99	0.13
$R^{\mathrm{O/U}}_{close:end}$					1.00	0.01
$R^{\mathrm{O/U}}_{open:close}$						1.00

Figure A1: **Estimating Payoff Probabilities for NBA Contracts.**

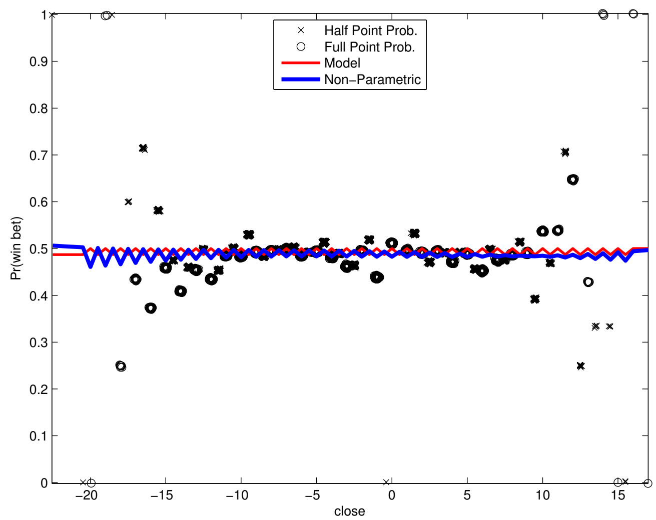
Panel A plots the probability of winning the bet (Pr(win)) or tieing (Pr(push)) for spread contracts at every spread level. The first graph plots the empirical probability (i.e., frequency) of winning at each spread level using the opening spread, as well as two estimates of the probability of winning using 1) a model that assumes the spread is set so that winning and losing are equally likely and 2) a non-parametric estimate of the probability following the estimators in Appendix A. The second graph plots the same probabilities for closing spreads. Full-point spreads and half-point spreads are indicated separately on the graphs. In addition, a small amount of random noise is added to each data point to gauge the number of contracts at each spread level. The third and fourth graphs of Panel A plot the probabilities of "pushing" (tie bet) for opening and closing spreads, respectively, using both the empirical frequency and the non-parametric estimator. Panel B plots the probability of winning the bet for opening and closing Moneylines at each Moneyline value. In addition to the empirical frequency of winning, probability estimates using the 1) model that assumes winning and losing are equally likely, 2) non-parametric estimator, 3) probit model, and 4) logit model are also plotted. Panel C plots the probabilities of winning and pushing for Over/under contracts using the same estimators as Panel A.

Panel A: Spread contracts

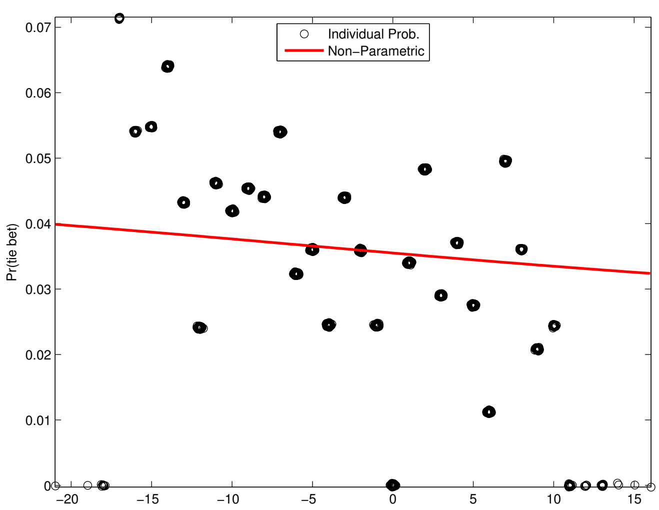
Opening spread Pr(win)



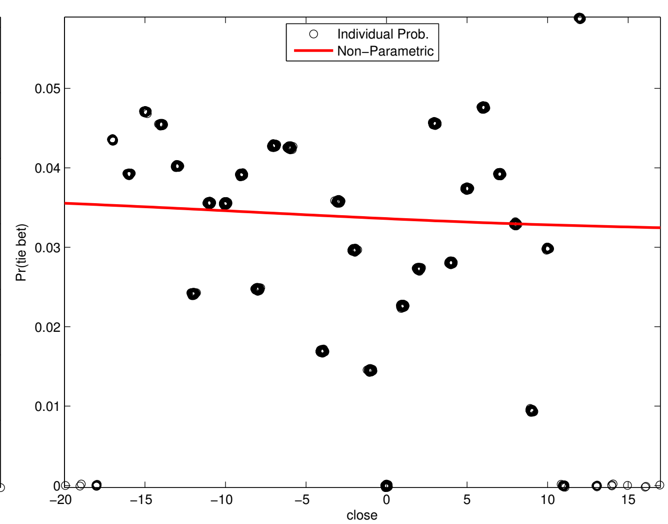
Closing spread Pr(win)



Opening spread Pr(push)

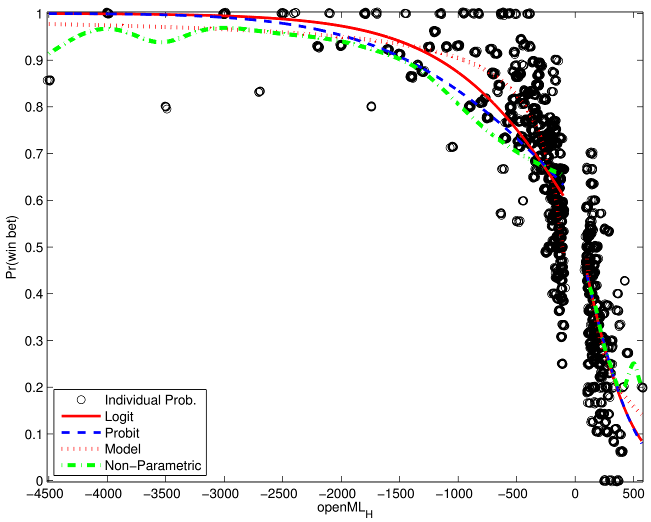


Closing spread Pr(push)

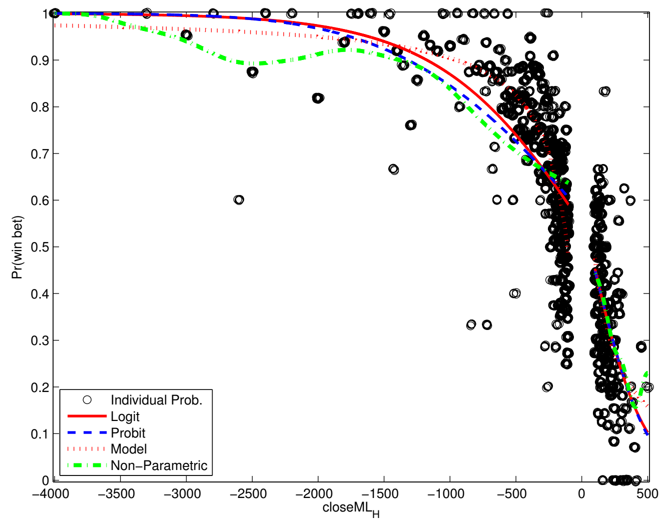


Panel B: Moneyline contracts

Opening spread Pr(win)

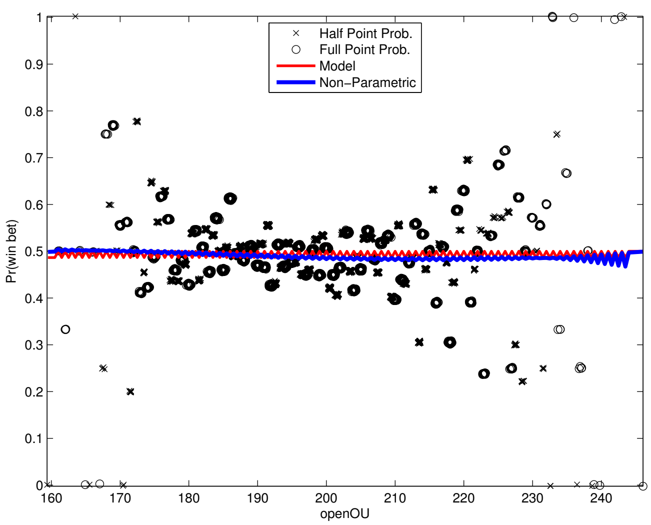


Closing spread Pr(win)

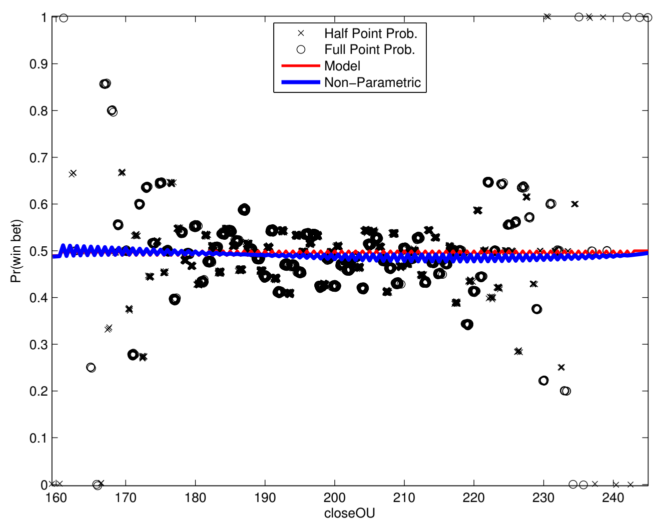


Panel C: Over/under contracts

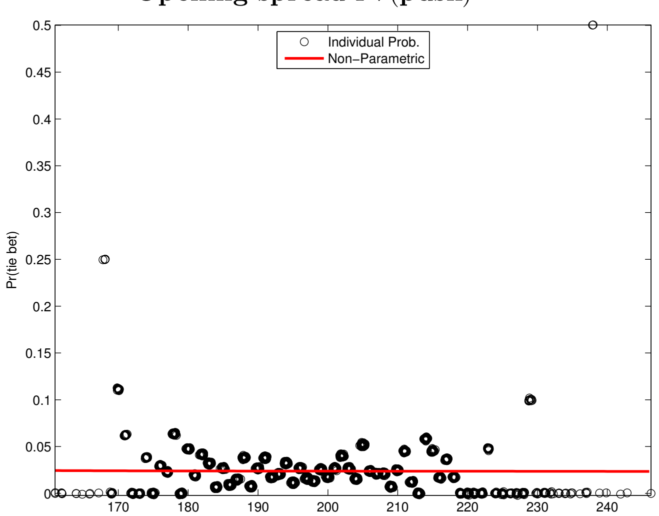
Opening spread Pr(win)



Closing spread Pr(win)



Opening spread Pr(push)



Closing spread Pr(push)

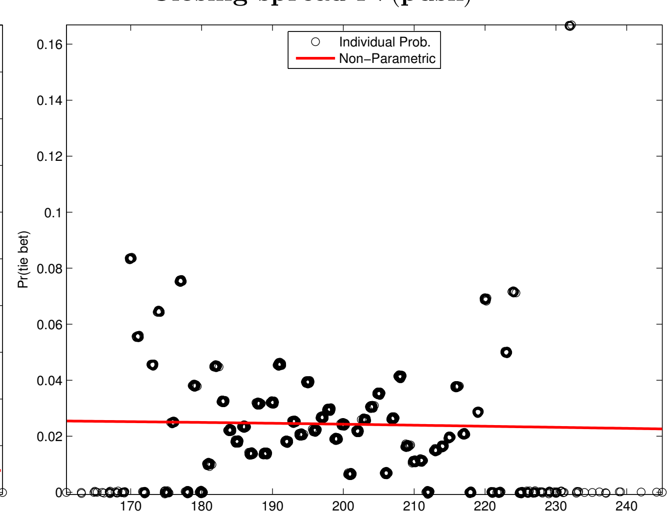
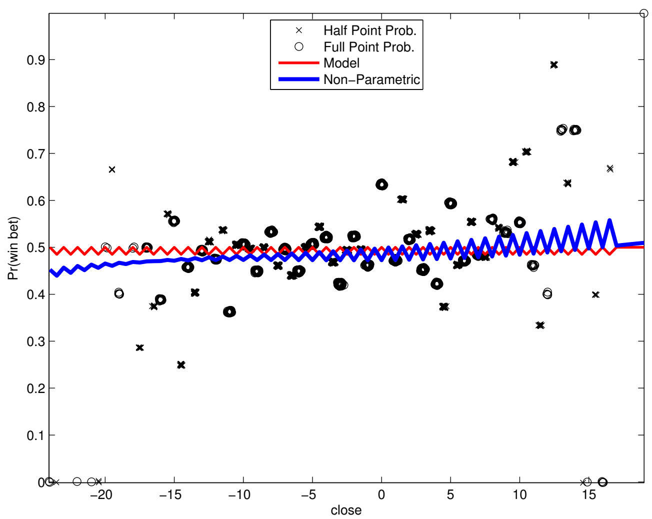


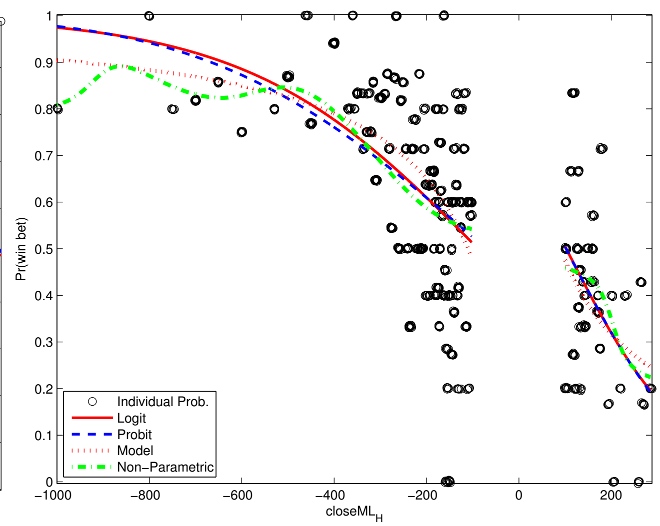
Figure A2: **Estimating Payoff Probabilities for NFL Contracts.**

Panel A plots the probability of winning the bet ($Pr(\text{win})$) or tieing ($Pr(\text{push})$) for spread contracts at every spread level. The first graph plots the empirical probability (i.e., frequency) of winning at each spread level using the closing spread, as well as two estimates of the probability of winning using 1) a model that assumes the spread is set so that winning and losing are equally likely and 2) a non-parametric estimate of the probability. Full-point spreads and half-point spreads are indicated separately on the graphs. In addition, a small amount of random noise is added to each data point to gauge the number of contracts at each spread level. Panel B plots the probability of winning the bet for closing moneylines at each moneyline value. In addition to the empirical frequency of winning, probability estimates using the 1) model that assumes winning and losing are equally likely, 2) non-parametric estimator, 3) probit model, and 4) logit model are also plotted. Panel C plots the probabilities of winning for over/under contracts using the same estimators as Panel A.

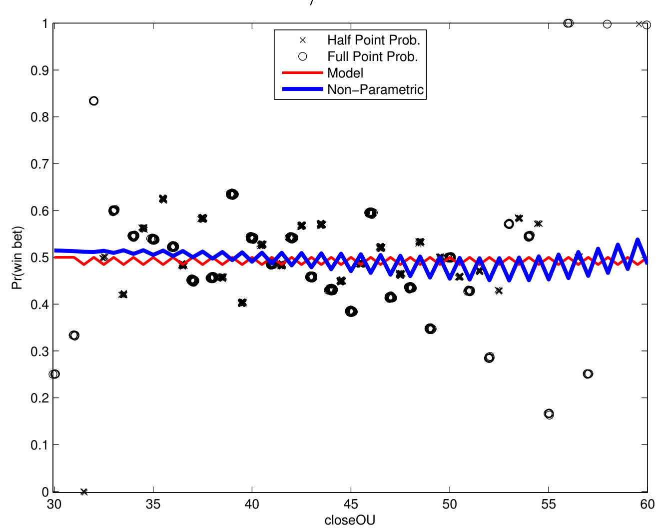
Panel A: Spread contracts



Panel B: Moneyline contracts



Panel C: Over/under contracts

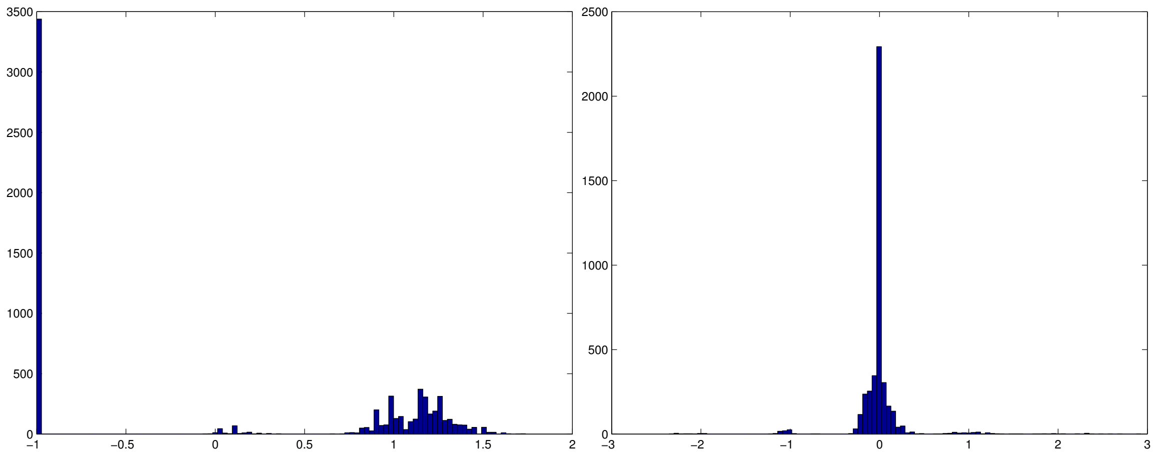


B Supplementary Tables and Figures

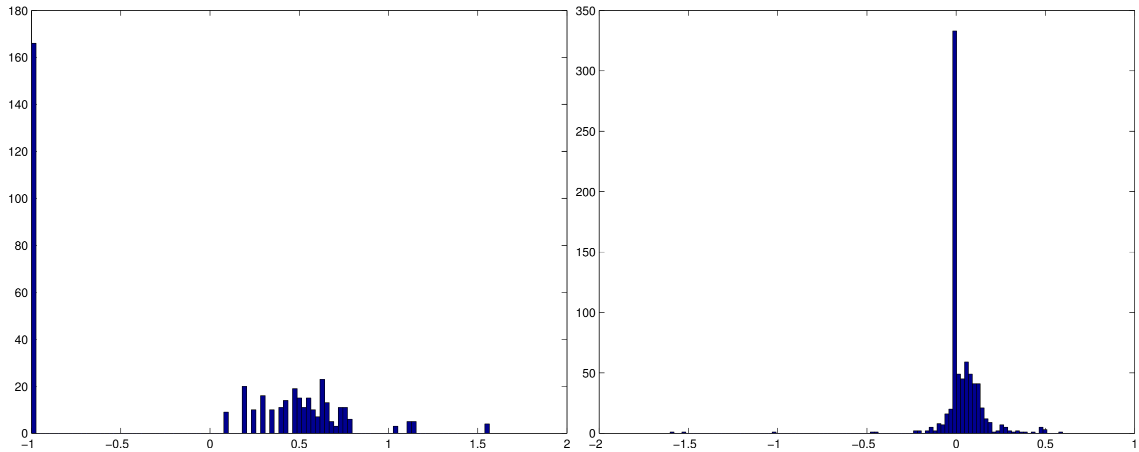
Figure B1: **Return Distributions for NFL, MLB, and NHL Betting Contracts.**

The figure plots the distribution of gross returns to all NFL game betting contracts. Panel A shows returns to spread contracts, Panel B to Moneyline contracts, and Panel C to over/under contracts. In each panel two sets of returns are shown: close-to-end and open-to-close. Returns are calculated following equations (**??**-**??**), using the probabilities estimated from the non-parametric specification. A table is included at the bottom of the three panels that reports the mean, standard deviation, skewness, and excess kurtosis of the net returns to each contract.

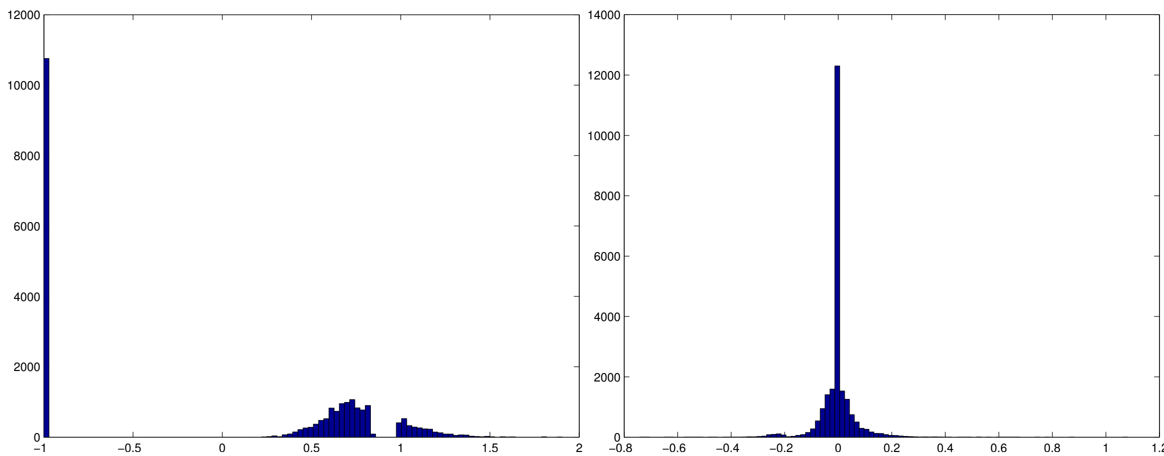
Panel A: NFL Spread contracts



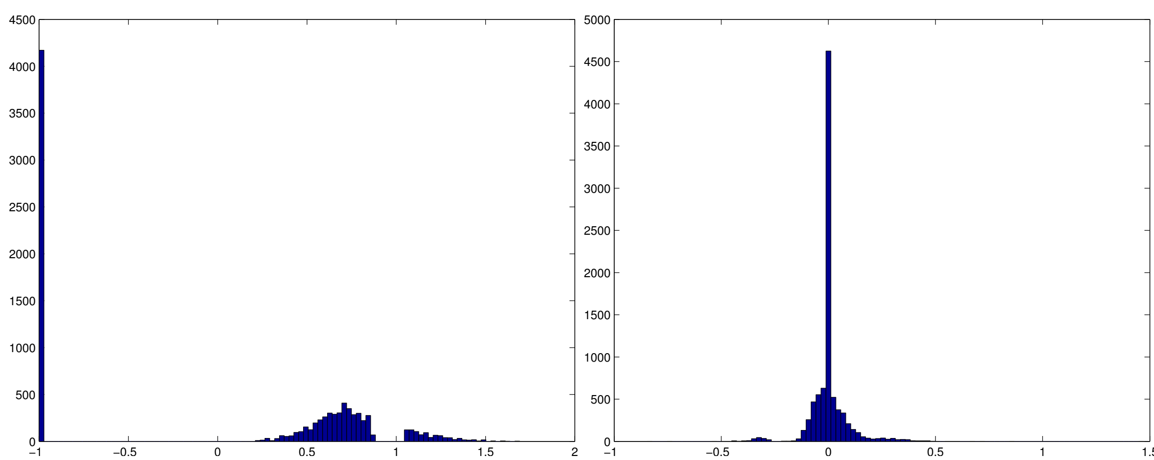
Panel B: NFL Moneyline contracts



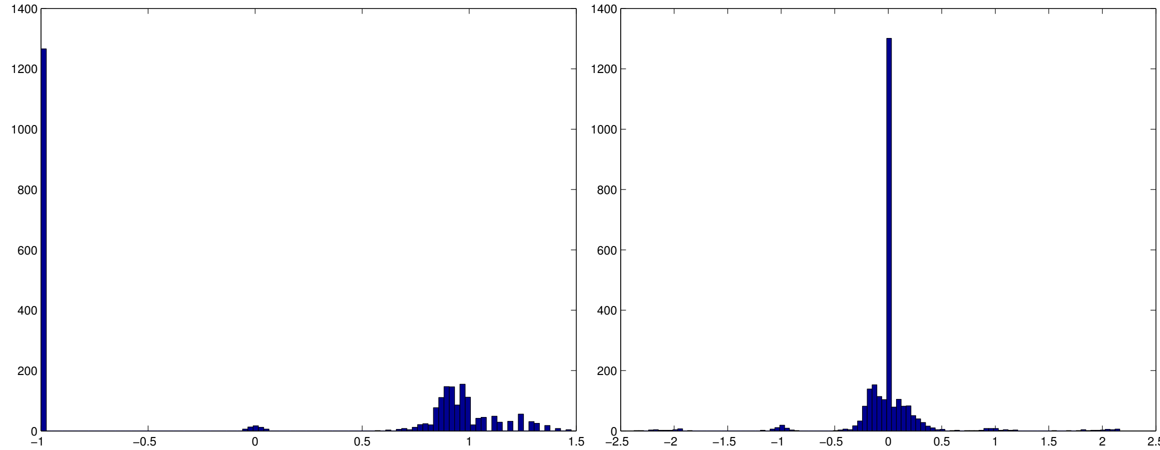
Panel C: MLB Moneyline contracts



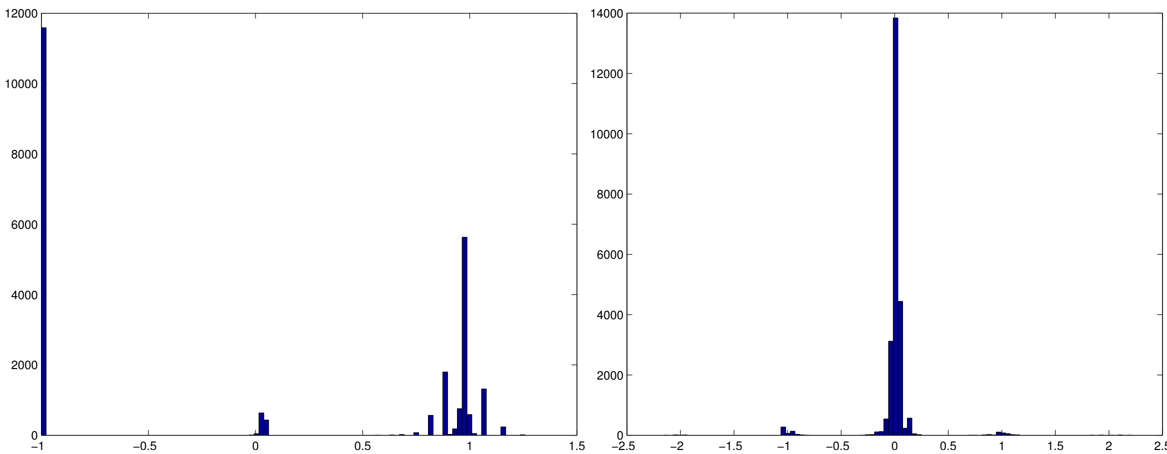
Panel D: NHL Moneyline contracts



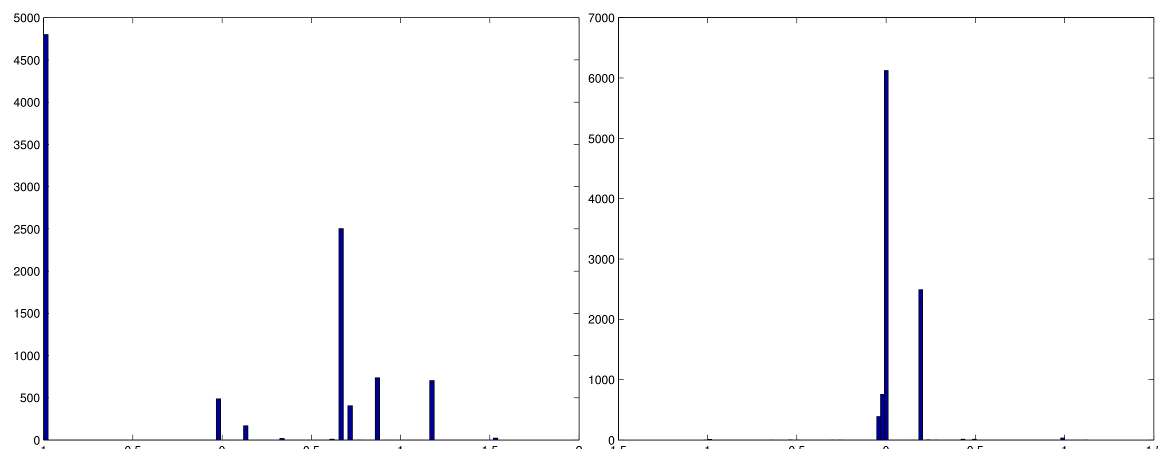
Panel E: NFL Over/under contracts



Panel F: MLB Over/under contracts



Panel G: NHL Over/under contracts



Summary statistics on net returns				
	mean	stdev	skew	ex-kurt
Panel A: NFL contracts (%)				
Spread contract				
$R_{close:end}$	0.09	99.3	0.01	-1.97
$R_{open:close}$	1.58	28.5	0.35	29.44
Moneyline contract				
$R_{close:end}$	-3.23	85.7	0.08	-1.08
$R_{open:close}$	5.01	11.7	-1.95	29.31
Over/under contract				
$R_{close:end}$	-0.04	95.7	-0.06	-1.97
$R_{open:close}$	0.21	33.9	0.29	22.28
Panel B: MLB contracts (%)				
Moneyline contract				
$R_{close:end}$	-3.90	89.84	-0.06	-1.83
$R_{open:close}$	-0.20	6.50	0.24	20.95
Over/under contract				
$R_{close:end}$	-0.13	98.58	0.02	-1.95
$R_{open:close}$	0.07	20.83	-0.90	32.88
Panel C: NHL contracts (%)				
Moneyline contract				
$R_{close:end}$	-3.41	88.90	-0.08	-1.80
$R_{open:close}$	0.34	8.72	0.99	20.13
Over/under contract				
$R_{close:end}$	0.11	100.78	0.09	-1.91
$R_{open:close}$	-0.02	8.31	3.98	53.75

Table B3: Momentum and Saliency

The table reports regression results of open-to-close and close-to-end returns on the momentum index interacted with various saliency measures for Spread contracts, Moneyline contracts, and Over/under contracts in the NBA (Panel A) and for all four sports (NBA, NFL, MLB, and NHL) combined (Panel B). Results are reported for the actual dollar returns of the contracts. The saliency measures for the NBA contracts are an indicator variable for a player scoring at least 50 points in a game over the last eight games for the Spread and Moneyline contracts and an indicator for the total points of both teams being at least 250 points in any game in the last eight games for the Over/under contract. For the NFL, the saliency measure is an indicator for a player scoring at least three touchdowns in a game over the last four games for the Spread and Moneyline contracts and for both teams scoring at least 60 points in any game in the last four games. For MLB, the saliency measure is an indicator for a player hitting more than one home run, or getting at least four hits, or pitching at least six scoreless innings, in any game over the last eight games for the Moneyline contract, and for both teams scoring at least 12 runs combined in any game in the last eight games. For NHL the saliency measure is an indicator for a player scoring at least three goals (e.g., a "hat trick") in any game over the last eight games for the Moneyline contract, and for both teams scoring at least eight goals combined in any game in the last eight games.

	Open-to-close returns, $R_{open:close}$			Close-to-end returns, $R_{close:end}$		
	Spread	Moneyline	O/U	Spread	Moneyline	O/U
Panel A: NBA						
Momentum	0.06	0.02	0.15	-0.04	-0.18	-0.05
	(2.45)	(2.22)	(3.88)	(-0.47)	(-1.50)	(-0.45)
Momentum × salient	0.01	0.02	0.09	-0.03	-0.37	-0.36
	(1.80)	(1.26)	(0.61)	(-0.09)	(-0.70)	-(0.87)
Salient	0.51	0.81	-8.14	-0.06	3.54	-47.55
	(0.55)	(1.28)	(-0.48)	(-0.02)	(0.84)	(-0.98)
Panel B: All sports						
Momentum	0.05	0.03	0.02	-0.11	-0.21	-0.01
	(2.26)	(1.66)	(1.89)	(-1.56)	(-1.95)	(-0.25)
Momentum × salient	0.02	0.03	0.01	-0.24	-0.35	-0.01
	(1.34)	(0.39)	(1.26)	(-1.22)	(-0.72)	(-0.03)
Salient	-0.35	0.43	2.26	0.02	3.02	-3.03
	(-0.49)	(0.62)	(1.27)	(0.01)	(0.70)	(-0.60)

Table B4: Multivariate Regressions

The table reports multivariate regression results of close-to-end returns and open-to-close returns on momentum, value, and size measures simultaneously for Spread contracts, Moneyline contracts, and Over/under contracts in the NBA (Panel A), NFL (Panel B), MLB (Panel C), NHL (Panel D), and all four sports (Panel E). Results are reported for both actual dollar returns from the contracts as well as hypothetical returns based on points scored, except for the Moneyline contract. The momentum, value, and size measures are the index weighted average measures for each set of variables, as described in Section III.

	Dollar returns			Point returns	
	Spread	Moneyline	O/U	Spread	O/U
Panel A: NBA					
Open-to-close returns, $R_{open:close}$					
Momentum	0.05	0.01	0.15	0.89	2.23
	(1.61)	(0.41)	(3.18)	(6.93)	(8.07)
Value	0.12	-0.92	-0.19	-6.81	-14.41
	(0.30)	(-2.90)	(-0.32)	(-3.86)	(-4.19)
Size	0.04	-0.07	-0.22	-1.04	-1.99
	(0.45)	(-1.23)	(-1.74)	(-3.13)	(-2.64)
Close-to-end returns, $R_{close:end}$					
Momentum	0.03	0.16	-0.16	-2.22	-3.42
	(0.30)	(1.05)	(-1.17)	(-1.84)	(-1.45)
Value	0.26	-0.77	-3.99	6.21	-31.85
	(0.18)	(-0.39)	(-2.40)	(0.37)	(-1.08)
Size	-0.14	0.35	-0.32	-2.81	-3.13
	(-0.53)	(0.95)	(-0.87)	(-0.90)	(-0.49)
Panel B: NFL					
Open-to-close returns, $R_{open:close}$					
Momentum	0.07	0.06	0.09	0.63	1.03
	(1.84)	(1.64)	(1.07)	(3.22)	(2.32)
Value	0.32	1.10	-1.27	-1.14	-0.52
	(0.42)	(1.59)	(-1.21)	(-0.29)	(-0.09)
Size	0.03	0.10	-0.17	-0.77	-3.60
	(0.17)	(0.78)	(-0.49)	(-0.94)	(-1.96)
Close-to-end returns, $R_{close:end}$					
Momentum	-0.07	0.10	-0.35	0.87	-5.78
	(-0.54)	(0.37)	(-1.34)	(0.47)	(-1.61)
Value	0.88	6.77	-5.08	24.34	-109.02
	(0.32)	(1.33)	(-1.59)	(0.66)	(-2.46)
Size	-0.04	0.14	-0.04	5.44	-7.69
	(-0.07)	(0.15)	(-0.04)	(0.70)	(-0.52)

Returns =	Panel C: MLB Dollar Moneyline	Dollar O/U	Point O/U	Panel D: NHL Dollar Moneyline	Dollar O/U	Point O/U
			Open-to-close returns, $R_{open:close}$			
Momentum	0.02	0.34	2.03	0.18	0.05	0.06
	(0.81)	(2.87)	(10.74)	(1.02)	(0.14)	(0.22)
Value	-0.52	-0.45	0.11	-1.98	0.68	0.49
	(-2.79)	(-0.90)	(0.13)	(-3.21)	(0.93)	(0.85)
Size	0.07	0.12	0.23	0.26	0.00	0.05
	(2.73)	(1.81)	(2.14)	(2.73)	(-0.02)	(0.49)
			Close-to-end returns, $R_{close:end}$			
Momentum	0.38	0.35	1.46	-0.93	-4.90	-8.35
	(1.29)	(0.61)	(0.56)	(-0.48)	(-1.52)	(-1.14)
Value	-4.29	0.36	6.60	-12.07	12.52	15.94
	(-1.74)	(0.15)	(0.60)	(-1.75)	(1.74)	(0.98)
Size	0.65	0.13	-0.18	1.66	-2.18	-1.91
	(1.98)	(0.41)	(-0.13)	(1.55)	(-1.57)	(-0.61)

Panel E: All sports

	Dollar returns: Spread	Moneyline	O/U	Point returns: Spread	O/U
			Open-to-close returns, $R_{open:close}$		
Momentum	0.09	0.02	0.09	1.51	0.66
	(2.26)	(1.55)	(2.53)	(9.28)	(3.44)
Value	-0.13	-0.46	-0.13	-0.70	-2.94
	(-1.30)	(-2.95)	(-1.53)	(-1.57)	(-6.37)
Size	0.06	0.07	0.02	-0.39	-0.28
	(0.81)	(2.83)	(0.24)	(-1.21)	(-0.74)
			Close-to-end returns, $R_{close:end}$		
Momentum	-0.05	-0.17	-0.13	-2.40	-4.93
	(-0.41)	(-1.31)	(-1.05)	(-1.61)	(-2.99)
Value	0.32	0.31	0.29	6.73	4.19
	(1.37)	(0.94)	(0.86)	(1.86)	(0.91)
Size	-0.15	0.22	-0.10	-1.61	1.12
	(-0.58)	(0.88)	(-0.40)	(-0.54)	(0.35)

C Estimating Expected Contract Price from Fundamentals

One of the measures of value derives a fundamental value of the game itself and divides it by the market price of the contract. The sports analytics community has derived a number of measures of team quality or strength for use in predicting various game outcomes. One of the most popular is known as the Pythagorean win expectation formula, which the sports analytics community has shown is a good predictor of win percentage, across many sports. The formula and the parameter estimates across sports are:

$$E(\text{win}\%) \quad = \quad \frac{P_F^\gamma}{P_F^\gamma + P_A^\gamma} \tag{15}$$

where P_F is the average number of points scored for the team and P_A is the average number of points scored against the team, and γ is the Pythagorean coefficient with $\gamma = 1.83$ for MLB, 13.91 for the NBA, 2.37 for the NFL, and 2.11 for the NHL. These parameters come from the literature and were estimated on historical data prior to and independent of my sample.[25]

Using this formula, I estimate what the expected contract price would be based solely on this formula and the team's fundamentals (e.g., points scored and points allowed) by converting the Pythagorean formula's win percentage estimate into a Spread or Moneyline value. The formula provides an expected win percentage based on the points scored by a team and points scored against a team. I use the most recent scores of each team in their last 40 games (16 for the NFL) including the previous season to estimate win expectation. Taking the difference between the win expectations of both teams provides a measure of relative team strength in units of win probability. Multiplying this probability difference times the Over/under total (which is the market's expectation of the total number of points that will be scored by both teams) converts the probability difference into an expected point difference, which I then divide by the actual betting contract expected point difference or Spread. That is, I take the estimated betting contract price, $E(P)$ from the Pythagorean formula and divide it by the actual market price of the betting contract, P. Intuitively, $E(P)/P$ is a measure of the expected point Spread derived from past scoring information through the Pythagorean model relative to the market's expectation from betting markets. A high value for this ratio implies the Spread contract for a game looks "cheap" or is a value bet and a low ratio looks "expensive" relative to fundamentals.

For the Moneyline contracts I do something similar by matching the Pythagorean-implied Spread to the corresponding Moneyline based on the distributional mapping of actual Spreads to Moneylines in the data.[26] By mapping the predicted Spread from the Pythagorean formula to the Moneyline using the joint distribution of actual Spreads and Moneyline values, I preserve the feature that the Moneyline values are internally consistent with the predicted Spreads, which is appealing since both the Spread and Moneyline contracts are bets on who wins. Alternatively, I could do the opposite and take the expected win probability from the Pythagorean as above and use that to imply a Moneyline, where a rough translation between win probability and Moneyline is as follows: if π = estimated win probability, then if $\pi > 0.5, ML = -(\pi/(1-\pi)) * 100$,

[25] The formula was first used by Bill James to estimate how many games a baseball team "should" have won based on the number of runs they scored and allowed. The name comes from the formula's resemblance to the Pythagorean theorem when the exponent = 2. Empirically, this formula correlates well with how teams actually perform. Miller (2007) shows that if runs for each team follow a Weibull distribution and the runs scored and allowed per game are statistically independent, then the formula gives the probability of winning. The formula makes two assumptions: that teams win in proportion to their "quality", and that their quality is measured by the ratio of their points scored to their points allowed. The different values for the exponent across sports represents the role chance plays in determining the winner across sports. Thus, basketball, in part because so many more points are scored than in other sports, such as baseball, gives the team with higher quality more opportunities to allow that quality to impact the game and diminishes the role of luck, which is why it has a much higher exponent than baseball. For derivations and estimations of the model in each sport see Morey (1994), Miller (2007), Dayartna and Miller (2013), and Football Outsiders (2011).

[26] For example, if the Pythagorean-implied spread is -3.5, I take the Moneyline value associated with a -3.5 point spread from the empirical distribution of actual betting contracts. When there are multiple Moneyline values for a given Spread, I take the average of those Moneylines.

or if $\pi < 0.5, ML = (1-\pi)/(\pi) * 100$. Then, taking the Moneyline estimated from the Pythagorean win probability, I could match the Moneyline to a Spread using the empirical distribution of Moneylines and Spreads to also make the estimates internally consistent. Both methods of computing expected Spreads and Moneylines yield nearly identical results.

For Over/under contracts, which are bets on total points scored by both teams combined, I run a rolling regression model of O/U totals on points scored by the home team, points scored against the home team, points scored by visiting team, and points scored against visiting team over the last 40 games (16 games in the NFL) for each team. Using the regression coefficients for both teams, I then apply them to the average points scored for and against for each team over the last 40 games (16 games in the NFL) and then take an average of the predicted point totals for the two teams, which represents a predicted O/U point total from the number of points scored for and against each team over the last 40 games. This predicted point total is then divided by the actual O/U total from the betting market to obtain a value measure. Alternatively, I could have taken a moving average O/U total from betting markets over the past 40 games involving either team and taken the average across the two teams for my fundamental O/U total. Using this measure instead, I get very similar results.

To estimate the residual value measure, I estimate the expected price of the contract by running a rolling regression each season of the betting line on team $\times$ home and team $\times$ away dummy variables, a variable representing the number of games played in the last three days, a variable representing if the visiting team is playing its second consecutive game away from home, and cumulative points scored for and against both teams. The regression requires a minimum of 20 games per team and hence only examines contracts beginning with the 21st game each season. Using the slope coefficients from this regression, I then forecast the betting line (Spread, Moneyline, or Over/under) on the next game and take the difference between the actual betting line and this forecast as the residual value measure. The predicted regressions based on the above observables have an R^2 of 0.80 on average, indicating that this simple model can predict a significant component of these betting lines. This measure of value changes at a much higher frequency than those above and captures very short-term deviations from expected betting lines.