Does Sentiment Improve Price Efficiency?
The Role of Conditional Market Participation on the Relationship between Sentiment and Asset Prices

James P. Sinclair

School of Business

University of Connecticut

jsinclair@business.uconn.edu

ABSTRACT: Prior research suggests that investor sentiment drives asset prices away from fundamental values. Using a unique dataset of self-reported consumer preferences to proxy for differences in sentiment across active market participants and nonparticipants, I provide evidence that sentiment may actually improve price efficiency. I find evidence that pricing errors decrease as a function of active market participants' sentiment. I also provide evidence that price efficiency is inversely related to the degree of correlation in nonparticipants' shared private beliefs about future asset payoffs. By endogenizing the market participation decision, this paper introduces a rational explanation for the well-documented destabilizing effects of sentiment on asset prices.

Current draft: April 2014

Keywords: sentiment, conditional market participation, asset pricing

JEL Classification Codes: G02, G11, G14

I thank Paul Fischer, Brian Miller, Sarah Rice, Jim Vincent, Luke Watson, and workshop participants at Georgetown University for their helpful comments.

Recent empirical work in behavioral finance suggests that investor sentiment negatively affects asset price efficiency.[1] This relationship is most pronounced during periods of low sentiment and for hard-to-value securities. Such correlation, however, does not necessarily imply causation, as market participation costs are likely to shift with securities' underlying uncertainty. As uncertainty increases, participation costs rise and uncertainty-averse investors rationally choose not to participate in the market. To the extent market nonparticipants withhold value-relevant information from entering the public domain via trading activity, price efficiency is expected to be negatively correlated with investor uncertainty. In this paper, I seek to contribute to the growing debate on the effect of sentiment on asset prices by studying a natural experiment in the form of an online contest whereby contestants forecast the outright winners of 140 college football games and assign a confidence rank-ordering to each pick. I find evidence consistent with contestants' beliefs containing value-relevant information that is not fully impounded into price and that asset-specific investor sentiment is positively associated with price efficiency.

Testing the effect of sentiment on asset prices empirically is challenging for three reasons. First, researchers rarely have access to investors' private expectations. As addressed by Garfinkel (2009), prior literature typically relies on publicly-available proxies for private expectations, such as analyst forecast dispersion, unexplained trading volume, bid-ask spreads, and stock return volatility. However, contractual and implicit incentives, market structure, and investors' risk aversion likely distort these proxies from capturing actual private expectations.[2] Second, parsing sentiment measures between active market participants and nonparticipants requires data from a constant sample of current and potential participants. Publicly-available market prices and trade volume are derived from inconsistent waves of active market participants. Investors choosing to 'stay on the sidelines' are not represented in traditional price and volume data yet may be included in traditional sentiment measures, resulting in incongruent data measurement.

[1] For example, see DeLong et al. (1990), Shleifer and Vishny (1997), Baker and Wurgler (2006), Neal and Wheatley (1998), Shiller (1981), and Dumas, Kurshev, and Uppal (2009).

[2] Several studies have attempted to overcome this problem by collecting Twitter posts, stock picks, or message board text (see Giannini and Irvine (2012), Avery, Chevalier, and Zeckhauser (2012), and Das and Chen (2007)).

Third, conventional sentiment measures, such as those used in Baker and Wurgler (2006), rely on market-wide variables to develop proxies for sentiment. These measures generally fail to consider cross-sectional differences in sentiment across assets, and do not attempt to distinguish sentiment between active market participants and non-participants.

In this study, I overcome these challenges by examining the relationship between sentiment and asset prices in the college football betting market. I collect individuals' private expectations and develop asset-specific sentiment measures from ESPN's 2011 College Pick'em Challenge contest. Contestants were asked to predict winners and assign confidence point values for 140 college football games over the fourteen-week contest period. The sole objective of the contest was to earn points by assigning higher confidence point values to outcomes deemed more likely to occur, and lower confidence point values to outcomes perceived to be less likely to occur. Professional odds-makers and active market participants set market prices (e.g., point spread) for each game. In the spirit of Cao, Wang, and Zhang's (2005) analytical model of market participation constraints, I proxy for market nonparticipants' (participants') beliefs by aggregating the beliefs of all contestants whose confidence ranking falls below (above) a predetermined threshold.[3] By developing an asset-specific sentiment measure for both active market participants and nonparticipants, I am able to further examine the influence of investor sentiment on asset prices.

I present evidence consistent with investor sentiment improving price efficiency. I find that pricing errors decrease as a function of active market participants' sentiment. I also find evidence consistent with reduced price efficiency (i.e., increased mispricing) as the correlation of nonparticipants' common private beliefs about future asset payoffs increases. These findings are consistent with the Diamond and Verrecchia (1987) theoretical framework, and offer an alternative explanation for the results reported in Baker and

[3] Without having knowledge of which contestants actively placed wagers on specific games, I also rely on prior literature for the characteristics of market participants to determine a suitable proxy for active market participation. Specifically, prior research suggests that active market participants are, in general, highly confident and hold relatively few securities (Odean (1998), Odean (1999) and Goetzman and Kumar (2008)).

Wurgler (2006). Specifically, Baker and Wurgler (2006) find evidence of asset mispricing following periods of low sentiment, and interpret this finding as evidence that sentiment (i.e., a seemingly irrational behavioral bias) drives asset prices away from fundamental values. Alternatively, the results presented in this paper suggest that the decision to participate in the market is based on a rational condition (i.e., level of uncertainty), and any residual mispricing stems from the degree of correlation among nonparticipants. By measuring sentiment levels separately for market participants and nonparticipants, the results in this paper further contribute to our understanding of the role of investor sentiment on asset pricing.

I. Sentiment and Asset Prices

Investor sentiment can affect asset prices in two competing ways. Traditional asset pricing models assume full market participation, whereby agents must allocate wealth among existing assets.[4] Under this assumption, market participants with systematically biased private expectations about future asset payoffs adversely affect asset prices if potential arbitrageurs are (at least temporarily) unable to effectively exploit price discrepancies. According to this behavioral asset pricing theory, uninformed demand shocks and limits to arbitrage drive prices away from fundamental values. Consistent with Miller (1977), optimistic investors drive asset values upward, leading to current overvaluation and relatively low future returns. Similarly, pessimistic investors likely demand higher risk premiums, resulting in lower current prices and relatively high future returns. Supporting the view that sentiment drives asset mispricing, Baker and Wurgler (2006) document differences in returns for certain securities following periods of high, relative to low, sentiment.[5]

[4] Allen and Gale (1994) is a notable exception to this generalization. They model the participation decision as endogenous, and show that price volatility is positively related to aggregate uncertainty due to incomplete market participation.

[5] However, mispricing in Baker and Wurgler (2006) appears to be driven by periods following low sentiment. In fact, in almost all sorts of their data, future returns appear to be closer to zero following high sentiment periods, which may be interpreted as evidence consistent with lower mispricing following high sentiment periods.

In contrast with the assumption of unconditional market participation, recent analytical models attempt to explain market nonparticipation by suggesting ambiguity-averse investors may opt to neither buy nor sell a particular asset if their level of uncertainty about the asset's expected value is above a predetermined threshold (see Dow and Werlang (1992) and Cao, Wang, and Zhang (2005)).[6] Price efficiency may actually improve if uncorrelated noise traders withdraw from the market; however systematic nonparticipation of investors sharing common beliefs about future asset payoffs reduces informational efficiency (see Diamond and Verrecchia (1987)). Consequently, the effect of nonparticipation on asset prices is directly related to the degree with which absent traders share correlated beliefs. While factors such as market entry costs, transaction fees, and risk aversion likely impact the decision of whether an investor enters the market at all, cross-sectional differences in uncertainty and confidence levels are also likely to determine which assets investors actively hold and trade. For example, French and Poterba (1991) demonstrate that investors hold more optimistic expectations about domestic, relative to foreign, stocks, resulting in higher nonparticipation rates in foreign asset markets. Likewise, Cao, Wang, and Zhang (2005) analytically show that when uncertainty dispersion is high, investors with high uncertainty choose to not participate in the market. These results suggest that investors selectively choose to hold assets based on their sentiment for one security relative to another.

How does conditional market participation affect the predicted association between investor sentiment and asset prices? During periods of low investor uncertainty (i.e., high sentiment), both overconfident and rational investors actively participate in the market. Although, Grossman (1976) contends that rational investors have no incentive to expend additional resources to acquire new information, Ko and Huang (2007) assert that "overconfident investors believe they can earn extraordinary returns and will consequently invest resources in acquiring information pertaining to financial assets."

[6] Prior literature has examined the general role of nonparticipation on asset prices. Easley and O'Hara (2009) analytically model the disruptive effects of nonparticipation on asset prices and Brav, Constantinidis, and Geczy (2002) and Vissing and Jorgenson (2002) empirically illustrate pricing inefficiencies related to nonparticipation.

Given that price discovery depends on the integration of new information into asset prices, it appears that active participation by overconfident investors improves price efficiency. Thus, under the assumption of conditional market participation, investor sentiment is predicted to reduce asset mispricing.

Turning to periods of high investor uncertainty (i.e., low sentiment), Cao, Wang, and Zhang (2005) suggest that when investors are uncertainty-averse, there is a region of prices over which the investor does not participate in the market, with more uncertain investors exhibiting larger nonparticipation ranges. The remaining distribution of active market participants is truncated with respect to the level of uncertainty aversion in the entire population. Thus, price efficiency is reduced by removing the information content otherwise revealed in the trading activity of highly uncertain investors.[7] Conditional market participation likely affects asset pricing more when correlated, as opposed to uncorrelated, private expectations are systematically removed from the distribution of active market participants. For example, consider an investment mandate that effectively restricts all investors from selling securities short. Beber and Pagano (2011) find that short-selling bans delay price discovery, increase bid-ask spreads, and induce greater pricing errors for all affected stocks. In contrast, Shive (2012) finds evidence of slightly lower return volatility and decreasing bid-ask spreads for certain securities during local power outages[8], suggesting that the removal of local traders, with potentially uncorrelated beliefs, has a minimal effect on overall price efficiency. A key difference between these two settings lies in the degree of correlation among the private expectations of the excluded group. Building on this premise, under the assumption of conditional market participation, I predict pricing errors will increase when investors with highly-correlated beliefs abstain from participating in the public market.

[7] The predicted direction of pricing errors resulting from conditional market participation is not clear. Miller (1977) argues that in the presence of short-selling constraints, optimists set prices, resulting in higher current prices and lower future returns. However, Cao, Wang, and Zhang (2005) argue that because fewer investors bear the risk of holding a stock, the equity risk premium must increase, potentially offsetting a lower uncertainty premium. In this paper, I am less concerned about the direction of the pricing error than I am about detecting the existence of pricing errors in general.

[8] Shive (2012) examines price efficiency during power outages affecting the geographic area containing firms' corporate headquarters.

To illustrate this point, consider the following numerical example related to voting outcomes. Four constituent groups (1-4) are asked to vote for a candidate (A or B). The votes are tabulated and the results are as follows:

	Vote for Candidate A	Vote for Candidate B
Group 1	50%	50%
Group 2	75%	25%
Group 3	30%	70%
Group 4	60%	40%
Total	**53.75%**	46.25%
Total (excluding Group 1)	**55.00%**	45.00%
Total (excluding Group 2)	46.67%	**53.33%**

Assuming the groups are equally-weighted and candidates require a voting majority to win, Candidate A is the clear winner with 53.75% of the vote. Group 1 is the equivalent of uninformed noise traders in many asset pricing models, as their votes are uncorrelated and any bias for one candidate is exactly offset by equal bias for the other candidate. Removing Group 1 from the voting records reduces the noise in calculating the winner without changing the outcome (i.e., Candidate A wins with 55.00% of the vote). Now consider the removal of Group 2, a constituent group with highly correlated beliefs. Candidate A is no longer the winner, collecting only 46.67% of the votes. Furthermore, consider the inferences drawn from a hypothetical regression of candidate choice on overall market-wide sentiment. Overall market sentiment favors Candidate A, yet nonparticipation by Group 2 results in Candidate B winning the election. One could (incorrectly) conclude, based on this example, that market sentiment drives inefficiencies in voting outcomes. However, if the market sentiment measure excludes nonparticipants' beliefs, the same hypothetical regression would yield results consistent with sentiment correctly forecasting the election winner. Thus to assess the role investor sentiment plays on asset pricing, one must first measure sentiment separately for active participants and nonparticipants.

Popular measures of market sentiment, such as the University of Michigan Consumer Sentiment Index and the U.S. Consumer Confidence Index, survey a representative sample of all American households to track consumer attitudes and expectations. The surveys do not, however, attempt to distinguish market participants from nonparticipants. As documented earlier with the voting outcome example, failure to distinguish between these two groups may lead to spurious conclusions regarding the relationship between investor sentiment and asset prices. Other popular sentiment measures also fail to distinguish differences in sentiment between active participants and nonparticipants.[9]

Partially overcoming the conditional participation issue, Giannini and Irvine (2012) and Das and Chen (2007) examine Twitter postings and stock message boards, respectively, to examine market sentiment. By choosing settings with low entry-costs and minimal downside-risk to individuals, these studies maximize the distribution of individual participants in constructing their sentiment measures. Likewise, I examine a minimal cost online contest setting to construct sentiment measures for active market participants and nonparticipants. Motivated by Odean (1998), I condition market participation on self-reported confidence levels. Odean (1998) suggests that active traders are more confident than the general population, and that their overconfidence results in more frequent trade activity than that of rational investors.

Similar to financial markets, sports betting markets are characterized by low entry costs, suggesting that almost anyone can participate in the market. Despite the freedom to participate, individuals demonstrate exceptional selectivity in choosing which asset to hold, consistent with investors holding only those stocks for which they are extremely confident in their expectations of future asset payoffs.[10] Consequently, I assume contestants are more likely to enter the betting market when their overall confidence level for

[9] Examples of proxies for investor sentiment used in the prior literature include the closed-end fund discount, NYSE share turnover, insider trading levels, the percent change in margin borrowing, net new cash flow into mutual funds, retail investor BSI, the VIX index, the number and average first-day returns on IPOs, the put/call ratio, the equity share of new issues, and the dividend premium.
[10] Goetzmann and Kumar (2008) provide evidence that the median investor in their sample holds only three stocks.

choosing a particular game's winner is extremely high. Contestants rank each game they pick on a scale from 1-10, with 1 being the least confident and 10 being the most confident. I define high confidence as confidence rankings of 9 and 10. Likewise, I assume contestants are less likely to participate in the betting market if the overall confidence level assigned to a particular game is low (i.e., confidence ranking of 8 or lower). By parsing contestants' game picks into those more or less likely to actively participate in the market, I am able to construct separate sentiment measures for active participants and nonparticipants.

II. College Football Betting Market

A. Institutional Details

A point spread may be defined as a forecast of the margin of victory by which a stronger team (i.e., favorite) is expected to defeat a weaker team (i.e., underdog). Similar to asset prices in financial markets, point spreads are initially set by market makers and then move in response to new information and market forces. In a frictionless market, point spreads are an unbiased predictor of game outcomes. Biased spreads would create profit opportunities, thereby encouraging arbitrageurs to intervene and return spreads to optimal levels. Wagers are settled by considering whether or not the favorite's margin of victory exceeded the point spread. If the underdog wins the game, or loses by a margin less than the point spread, wagers placed on the underdog are declared winners and wagers placed on the favorite are declared losers. The opposite declaration is made if the favorite's margin of victory exceeds the point spread. This 'settling up' period provides a 'true' value at a predetermined date for each asset. As such, the betting market offers an advantage over traditional financial markets in determining whether sentiment has driven asset prices away from fundamental value.

Bettors participating in this market can drive point spreads up (down) by betting on the favorite (underdog) to win the game. For example, assume the Oregon Ducks are initially favored by seven points to beat the Stanford Cardinal. During the week preceding the game, several positive news stories about the Ducks prompt disproportionate gambling activity favoring Oregon to win the game. Assuming the college football betting market is efficient; the point spread would increase to offset the flow of wagers favoring Oregon. Similar to financial market prices, point spreads could be influenced by investor sentiment, new information entering the market, and discretionary market participation. Since placing a wager, just like buying a stock, is not costless, potential investors may sit on the sidelines when their level of uncertainty about a particular asset is so high as to reduce their perceived expected payout to a level below their perceived transaction cost. In other words, investors are more likely to place bets when they are more confident in future outcomes than when they are less confident.

Similar to financial markets, the college football betting market contains a finite, but extensive, list of available assets.[11] A plethora of public information is available on each team, opponent, coaches, players, weather conditions, historical statistics, etc. Some teams are more widely covered than others, but every team is covered by at least one newspaper, with several other news sources (i.e., websites, blogs, and Twitter feeds) also publicly available. Geographic proximity affects information acquisition costs as local media outlets spend disproportionate resources covering local college teams rather than non-local teams. Limited attention also influences confidence levels in predicting game outcomes as individuals likely spend more time watching and collecting information about their preferred teams than other college teams. Some individuals also exhibit strong biases toward (or against) certain schools due to current or past affiliations, geographic location, conference alignment, economic ties, family histories, etc.

B. Academic Studies

Scholarly interest in sports betting markets stems primarily from the many important features common to betting markets and financial markets, such as large volume, liquidity, information availability, and decision-making under uncertainty, but also from several key differences that distinguish the two markets. First, as highlighted in Avery and Chevalier (1999), assets in the betting market reach perfectly

[11] In the college football betting environment, a wager placed on a particular team to win is considered an asset. The bettor stands to gain or lose from the wager's payoff conditional on the game outcome.

observable terminal values in short time periods. In contrast, assets in financial markets may never reveal their 'true value', thus making identification of pricing errors more tenuous. Second, virtually no limits to short-selling exist in sports betting markets. Bettors can just as easily take one side of a wager as the other side. This freedom to profit from either optimistic or pessimistic expectations eliminates a costly trading friction that exists in financial markets. Finally, asset payoffs in sports betting markets all occur at the same time. While this standardization does not effectively eliminate differences in investment horizons among bettors, it does emphasize cross-sectional differences in expectations of common factors influencing future asset payoffs. For example, in financial markets, a rationally optimistic investor with a long investment horizon may not sell an asset even though she expects a short-term decline in price due to temporary macroeconomic conditions. Conversely, a rational investor would likely sell the asset if the payoff was determined at a pre-determined date in the near term. This distinguishing feature of the betting market is particularly advantageous for examining the relationship between sentiment and asset prices.

Early studies on sports betting markets focus primarily on racetrack bettors.[12] These studies generally examine market efficiency, and document the tendency for bettors to under-bet favorites and over-bet underdogs. The favorite-longshot bias and other potential biases have been examined in several other settings, including the football betting market, with mixed results. Woodland and Woodland (1994) conclude that the baseball betting market is efficient, despite limited evidence of a reverse favorite-longshot bias. Several recent studies find little or no evidence of market inefficiencies in the professional football (NFL) market. However, Golec and Tamarkin (1991), using more powerful statistical tests, find evidence of a small bias against home teams and for favorites in the NFL market. Interestingly, Golec and Tamarkin (1991) do not find evidence that these same biases exist in the college football betting market. [13]

[12] Examples include Weitzman (1965), Rosett (1971), Dowie (1976), Ali (1977), Snyder (1978), Tuckwell, (1983), Crafts (1985), Asch, Malkiel, and Quandt (1982), Gabriel and Marsden (1990), Golec and Tamarkin (1998), and Snowberg and Wolfers (2010)

[13] Gray and Gray (1997) test market efficiency in the NFL betting market and find evidence of a profitable trading strategy by betting on home-team underdogs early in their sample period; however this anomaly disappears over time. Paul, Weinbach, and Weinbach (2003) fail to find evidence of market inefficiencies

Avery and Chevalier (1999) examine the role of sentiment on setting point spreads in the NFL betting market. They document that point spreads move in predictable patterns during the week leading up to a football game; suggesting that at least some investors trade on sentiment. Consistent with sentiment driving bettors' behavior, Golec and Tamarkin (1995) suggest that bettors in the NFL market may be overconfident in their relative information set. Gandar et al. (1998) fail to reject a hypothesis of unbiased point spreads, but offer some evidence that known behavioral biases may still influence point spreads, resulting in profitable trading strategies.

III. Data

A. Data and Sample Selection

I collect data from ESPN's 2011 College Pick'em Challenge, available via http://games.espn.go.com. The data contains all game picks and confidence point assignments for the 2011 contest period. The objective of this free online contest is "to accumulate the most points by correctly selecting the 'winning team' for each U.S. college football game offered during the fourteen (14) weeks" from September 3, 2011 through December 3, 2011. Points are earned by assigning 'confidence points' to each of ten games each week. Contestants are asked to assign each game a unique confidence point value (whole number) between one and ten. For every game that a contestant selects the correct winning team, that contestant earns the points assigned to that game. No points are awarded or lost for incorrect picks. Contestants are not offered economic incentives to choose their 'favorite' teams, or assign low point values to a subset of teams therefore any revealed biases are intrinsic to the individual.[14]

---

in the college football gambling market, except for heavy home-team underdogs. Sauer et al. (1988) argue that the results in Zuber, Gandar, and Bowers (1985) claiming inefficiencies in the NFL gambling market are misleading, after including a proxy for the market's expectation into their regression model. Dare and MacDonald (1996) find little evidence of inefficiency in the NFL or college football betting markets.

[14] In low-risk contests, such as this one, the possibility exists that participants will not necessarily reveal truthful beliefs, but rather 'play to win', make fanciful picks, or otherwise deviate from normal behavior.

Contest officials choose ten games each week to include in the contest. The games are typically among the most highly profiled college football games of the week. Various media outlets offer coverage of the competing teams, and several Las Vegas and offshore sports-books publish point spreads for each game. I use the closing point spreads published by www.covers.com to establish a rational expectations rank ordering of games for each week.[15] In an efficient market, the point spread is an unbiased forecast of the final score differential.

I downloaded the predicted winners and confidence rankings from all 283,839 contestants for the 2011 College Pick'em contest. Out of this total, 67,681 unique contestants completed selections for all 140 games, resulting in 947,534 entry-weeks. I omitted all entry-weeks (9.6% of the sample) that did not change the contest's arbitrarily-assigned, default confidence point ranking.[16] The final sample consists of 8,563,390 game predictions with associated confidence point assignments. See Table 1 for the sample breakdown.

For each entry-week, I collected the unique contestant identification number, the ten teams each contestant selected to win, and the confidence points assigned to each game. I then computed an expected confidence ranking for each game selection based on two factors: (i) the closing point spreads published by www.covers.com and (ii) the complete portfolio of each entry's predicted winners. Ten expected confidence points were assigned to the most heavily favored team selected each entry-week, nine points for

---

To the extent that contestants engage in strategic behavior, the information content contained within my measure of sentiment will be dampened.

[15] www.covers.com compiles betting lines from several different offshore and Vegas sports-books. The point spreads used in this paper do not come from a single sports-book, but rather represent the most recently updated spread as of a preset time.

[16] The default confidence point ranking is based primarily on game start time, as opposed to Vegas point spreads. Including these entry-weeks into the regression analysis would bias against finding a significant association between sentiment and mispricing. In untabulated tests, I included these entry-weeks, and consistent with expectations, the results were qualitatively similar, but statistical significance declined.

the second most heavily-favored team selected to win, and so on. For example, consider the following two games[17]:

Game 1	Wildcats (+5)	vs.	Eagles (-5)
Game 2	Cougars (-2)	vs.	Rebels (+2)

If Contestant 1 picks the Wildcats to win Game 1 and the Cougars to win Game 2, I assign higher expected confidence points to Game 2. Alternatively, if Contestant 2 picks the Eagles to win Game1 and the Cougars to win Game 2, I assign higher expected confidence points to Game 1.[18] I collected additional game information (i.e., home team, final game score) from espn.go.com.

B. Descriptive Statistics

B1. Point Spread and Forecast Errors

Table II Panel A displays point spread and forecast error statistics. On average, for games selected to be in the contest, home teams were favored to win by 2.31 points. The actual average margin of victory for home teams in the sample was 2.98, resulting in a mean forecast error of 0.66 points. The low mean forecast error suggests the absence of a directional bias in setting point spreads, however the distribution of forecast errors is relatively large. Forecast errors range from -48 points to +35 points, with an average absolute forecast error of 12.59 points.

B2. Contestants' Selections

Out of 140 college football games included in the contest period, the team favored to win each game actually won 102 times (.729 winning percentage). On average, contestants' winning percentage was

[17] The numbers in parentheses following the team names represent the point spread for each game. In this example, the Eagles are favored to win Game 1 by five points. The Cougars are favored to win Game 2 by two points.
[18] Ties were split randomly. In unreported tests, I assigned the mean confidence level to ties (i.e., if two games were tied at Confidence Levels 5 and 6, the mean Confidence Level would be 5.5), and the results remained unchanged.

only .660. Out of 65,894 contestants that fully participated each week of the contest period, only 334 contestants (0.51 percent) correctly chose the winning team in more than 102 games; and 1,217 contestants (1.85 percent) outperformed market expectations by choosing the correct winner in more than 72.9 percent of games.[19] See Table II Panel B for details regarding contestants' selections. These results are consistent with findings in Barber and Odean (2000), demonstrating individuals' poor trading performance relative to market benchmarks.

In addition to picking fewer winners than is predicted using a simple strategy of choosing favorites to win every game, contestants exhibit poor judgment in assigning confidence points to game selections. Contestants would have earned 75.2 percent of all possible points by following a market-based strategy of choosing favorites to win every game and assigning confidence points based on the magnitude of the closing point spread. After accounting for contestants' actual predicted winners, a similar strategy of assigning confidence points based on the magnitude of the closing point spread would have earned 74.3 percent of all possible points. However, consistent with Fischhoff, Slovic, and Lichtenstein's (1977) finding that people poorly calibrate confidence levels, contestants earned only 70.6 percent of all possible confidence points. Less than nine percent (8.94%) of contestants earned more points than was predicted by the market-based strategy. Together, these results suggest that individuals are relatively weaker at calibrating confidence levels than in predicting winners.

Table III displays winning percentages by confidence level for three categories. Column 1 represents the winning percentages of teams favored to win each game, with each week's games rank-ordered by the magnitude of the closing point spreads (market strategy). For example, the team favored to win by the largest margin in a particular week is assigned to Confidence Level 10; the team favored to win

[19] Entry-weeks where the default confidence ranking assignments were not changed were excluded from all analyses reported in this paper. For example, if a contestant only predicted game winners, and did not make any changes to the arbitrarily-assigned confidence rankings, that entry-week was omitted. As a result, a contestant may correctly choose the winner of more than 72.9 percent on games, but still have less than 102 wins if the contestant failed to assign confidence points in a given week.

by the second-largest margin that week is assigned to Confidence Level 9; and so on. As discussed earlier, favorites won 72.9 percent of the time, with the team favored to win by the largest margin each week winning 92.9 percent of the time.

Column 2 represents the winning percentages of contestants' predicted winners, with each entry-week's picks rank-ordered by the magnitude of the closing point spread (hybrid strategy). While contestants select the correct winner only 66.13 percent of the time, it appears from Column 2 that their poor performance, relative to the market, is concentrated among the three lowest confidence levels. This finding is consistent with the relatively small drop in possible points that would have been earned (75.2% versus 74.3%) by assigning confidence points based on the magnitude of the closing point spread to contestants' actual picks. The most significant driver of contestants' poor performance, however, appears to arise from poor confidence calibration. Column 3 represents the winning percentages of contestants' predicted winners, with games assigned to confidence levels based on contestants' own self-assessed confidence ranking. While the overall winning percentage is identical in Columns 2 and 3 by design, the breakdown by confidence level reveals a distinct behavioral bias. In particular, contestants are overconfident in their own ability to forecast games. Interestingly, contestants appear to be consistently overconfident when assigning confidence rankings in the top half of the ranking spectrum, and consistently under-confident when assigning confidence rankings in the bottom half of the ranking spectrum.[20] Figure 1 illustrates contestants' poor performance, relative to market expectations, when assigning a confidence level of six and higher. Conversely, contestants outperform market expectations on games assigned to lower confidence levels.

---

[20] One may interpret the lower winning percentages above Level 5 as evidence that contestants are overconfident in these games. However, the reverse may also be true. Contestants may be under-confident in games for which they assigned a confidence ranking at or below Level 5. Regardless of which effect dominates, deviating from the market-based ranking is consistent with overconfidence in one's own abilities to forecast game outcomes.

Table IV presents additional evidence of contestant's deviation from market-based expectations. Over 50 percent (51.68%) of all game picks were coupled with a confidence ranking at least two spots away from the market's expectation. Contestants correctly picked game winners 54.17 percent of the time when assigning a confidence ranking at least two spots above the expected market ranking.

IV. Methodology

A. Variable Construction

To construct a measure of game-specific sentiment, I start by taking the difference between the expected confidence ranking and the actual self-assessed confidence ranking for each game pick to calculate *ConDiff*. Similar to a forecast error in the analyst literature, I calculate *ConDiff* as

$$ConDiff_{i,g,t} = \text{Actual } ConfidencePoints_{i,g,t} - \text{E}(ConfidencePoints_{i,g,t}) \quad (1)$$

If *ConDiff* is positive (negative), the contestant is overconfident (under-confident) in her selection relative to an unbiased confidence ranking based on published point spreads. Next, for each game, I calculate *GameSent* as the weighted-average difference between the average *ConDiff* across all entries that selected the home team to win and the average *ConDiff* across all entries that selected the away team to win.[21]

$$GameSent_g = ((\text{Avg}ConDiff_{i,g,t} * Freq_{i,g,t}) - (\text{Avg}ConDiff_{j,g,t} * Freq_{j,g,t})) / (Freq_{i,g,t} + Freq_{j,g,t}) \quad (2)$$

where $Freq_{i,g,}$ is the number of contestants selecting team i to win game g. Finally, *AbsGameSent* is the absolute value of *GameSent*, and is my measure of game-specific sentiment. Table II Panel A shows that the mean (median) *GameSent* for home teams is -0.779 (-0.903) and the mean (median) *AbsGameSent* is 1.481 (1.355).

[21] Neutral site games were randomly assigned a home team.

When examining the relationship between sentiment and asset prices under the assumption of conditional market participation, I calculate two distinct sentiment measures for every game. Nonparticipants' game sentiment includes only those game picks assigned to confidence levels 1 through 8. Consistent with Cao, Wang, and Zhang (2005), I use this range of confidence levels to proxy for an uncertainty threshold, below which contestants choose to not participate in the betting market. Game picks assigned to Confidence Level 9 or higher are used to calculate active market participants' game sentiment.

B. Regression Analysis

Following Sauer, Brajer, Ferris, and Marr (1998), I run the following OLS regression to test the relationship between investor sentiment and pricing errors:

$$AbsPSE_i = \alpha + \beta_1\, AbsGameSent_i + \beta_2\, Spread_i + \beta_3\, Favorite_i + \varepsilon_i \qquad (3)$$

where the dependent variable, $AbsPSE_i$, is the absolute difference between the closing point spread and the actual margin of victory for game *i*. By including the point spread in the construction of the dependent variable, I effectively control for market expectations. Consistent with Gu and Wu (2003), I assume the absolute value of the point spread error represents the markets' loss function, defined as the unsigned deviation of forecasts from realizations. Larger values of *AbsPSE* indicate greater mispricing. *AbsGameSent* is a game-specific sentiment measure. This variable increases the further contestants deviate from market expectations and rely more on their own abilities to forecast games. $Spread_i$ is the closing point spread listed on www.covers.com for game *i*. $Favorite_i$ is an indicator variable set equal to one if the home team is the favorite, and zero otherwise. The intercept captures bettors' home team bias or other undocumented biases, if any, embedded in the point spread.

V. Results

A. Main Results

Table V tests the role of investor sentiment on market mispricing. Column 1 implicitly assumes unconditional market participation by aggregating all game picks together to measure game sentiment. In Columns 2 and 3, I examine conditional market participation by partitioning the sample based on confidence level. In Column 2, *AbsGameSent* is measured by including only those game picks that were assigned a confidence level below nine. In Column 3, *AbsGameSent* is measured by including only those game picks that were assigned to Confidence Level 9 or 10. Under the assumption of conditional market participation, Column 2 represents market nonparticipants and Column 3 represents active market participants. In estimating equation 3, I exclude games played during the first three weeks of the contest period to reduce the influence of learning how the contest works on contestants' game selections. I also exclude games played during the final week of the contest period to reduce the possibility that contestants may alter their strategy when selecting winners and assigning confidence points to win sub-group competitions.

When all contestants' beliefs are aggregated, the results are consistent with the general conclusion found in prior literature that investor sentiment positively influences pricing errors. The coefficient on *AbsGameSent* is 1.939 and is statistically significant at the 5% level. This result falsely relies on market-wide sentiment measures to capture the influence of investor sentiment on market mispricing. Columns 2 and 3 correct for this misspecification by disaggregating contestants' sentiment on the basis of their likelihood of participating in the betting market. The results highlight the drivers of the earlier test. The positive association between sentiment and mispricing is driven by those contestants that are less likely to actively participate in the market. The coefficient on *AbsGameSent* in Column 2 is 1.886 and is statistically significant at the 1% level. This evidence is consistent with price efficiency declining with increasing correlation in nonparticipants' private expectations. Likewise, pricing errors are less pronounced when investors with uncorrelated beliefs are excluded from market participation. Conversely, among contestants that are most likely to participate in the betting market, the relationship between sentiment and mispricing is reversed. Specifically, extreme levels of asset-specific sentiment appear to enhance price efficiency during periods of full market participation. The coefficient on *AbsGameSent* in Column 3 is -0.952 and is statistically significant at the 5% level. In sum, these results support the predictions made under the assumption of conditional market participation.

Spread is positive but insignificant across all model specifications. *Favorite* is also positive and insignificant in all models, suggesting the well-established favorite-underdog bias is not present in the current sample. The intercept, however, is positive and significant, suggesting the well-established home team bias, and possibly other undocumented biases, are present in the current sample.

B. Betting Strategy

The previous subsection provides evidence consistent with market nonparticipants' sentiment partially driving pricing errors. Next, I examine whether the information contained within market nonparticipants' sentiment can be used to generate a profitable betting strategy. Two distinct effects could be driving the results in this paper. Overall game uncertainty could be correlated with both game-specific sentiment and pricing errors. I attempt to control for uncertainty in the primary research design by including *spread* as a control variable and by constructing *ConDiff* using each contestants' own portfolio of selected winners; however, unobserved uncertainty may still affect the results. If the results are driven by uncertainty, pricing errors are expected to occur in both directions, randomly. On the contrary, if market nonparticipants' beliefs contain valuable information in the price discovery process, a betting strategy could exploit this information to generate profits.

Table VI present the results of five betting strategies using ESPN's 2011 College Pick'em Challenge games. During the 100 games played from Week 4 through Week 13 of the contest period, the team favored to win, beat the spread 56 times, and tied the spread on four occasions. If ties push, meaning that all bets ending in a tie are returned to the bettor, an ‘always-bet-on-the-favorite’ strategy wins 58.3% of the time. Similarly, a strategy of always betting on the home team wins 54.2% of the time. Combining the two strategies (i.e., betting on home favorites to win), improves betting outcomes, as 61.1% of home favorites beat the spread.

To determine whether non-participants’ beliefs could be used to enhance betting outcomes; I first took the absolute difference of market participants’ game sentiment and market nonparticipants’ game sentiment to derive *AbsMP_MNP*. Higher values of *AbsMP_MNP* indicate greater deviation in beliefs between market participants and market nonparticipants. A betting strategy of choosing home favorites with *AbsMP_MNP* above the median (in the top decile) each week wins 63.6% (100.0%) of the time.[22] These results suggest that market non-participants beliefs contain valuable information that is not being captured in market prices.

VI Robustness Tests

A. Cross-Sectional Differences in Betting Volume

Every college football game in the 2011 College Football Pick’em contest included at least one high-profile team. However, some games were more high-profile than others, which may lead to cross-sectional differences in betting volume. If increased betting volume improves point spread accuracy and contestants are more confident picking winners to high-profile games, the results for Confidence Level 9-10 in Table V may be driven by only a few games. This concern is addressed in two ways. First, games are equally-weighted in the reported results, with every game from weeks 4-13 of the contest period represented in each model. This research design choice reduces the probability that cross-sectional betting volume is driving the results. Additionally, in untabulated tests, I add an indicator variable to the regression, set equal

[22] These results should be interpreted with caution for two reasons. First, this trading strategy, as defined in this paper, is not currently implementable (i.e., market nonparticipants’ beliefs are not fully revealed until the start of each game). Second, the betting strategies employing large differences in *AbsMP_MNP* restrict the number of bets to a relatively low number, possibly leading to spurious results.

to one if the game was designated as a 'Game of the Week' by ESPN, and zero otherwise.[23] The indicator variable is insignificant in all models. I also interacted *AbsGameSent* with the indicator variable. The interaction term is statistically insignificant. In all cases, *AbsGameSent* remained statistically significant.

B. Scaling the Point Spread

In this paper, I use point spreads to rank the market's expectation of game outcomes (i.e., relative likelihood of a team winning). An alternative approach may consider the ratio of the point spread to the Over/Under for each game when ranking market expectations.[24] This approach explicitly controls for cross-sectional differences in teams' scoring potential; and consequently, teams' likelihood of winning given a particular point spread. In untabulated tests, I re-estimate Equation 3 using the scaled point spread to rank-order market expectations, and ultimately measure game sentiment. The results are qualitatively unchanged.

C. Alternative Proxy for Market Participation

The confidence ranking represents a relative ranking, rather than an absolute confidence level, leading to possible identification issues in the participant/non-participant tests. This identification issue is not unique to this setting, as many confidence assessments are relative, to some degree. To address this concern, I alternatively considered actual confidence point assignments that deviated from expected confidence points by greater than four units to proxy for market participation. I assigned large deviations to the active market participants group due to the contestants' high self-confidence in their selection, and small deviations to the market nonparticipants group. Consistent with earlier results, the coefficient on *AbsGameSent* remains positive and significant for the nonparticipants group. The coefficient on *AbsGameSent* is positive, but remains insignificant, for the participants group.

---

[23] I identify a game as a 'Game of the Week' if ESPN College Gameday performed a live broadcast from that game's venue on the game day. I also included prime-time games broadcast by ABC as a 'Game of the Week', given the implied large national interest in those games.

[24] In addition to point spreads, bettors may also wager on whether both teams will combine to score under or over a predetermined point total in a game. For example, if the Over/Under is 45.5 and the final score is 31-17, wagers placed on the 'Over' win and wagers placed on the 'Under' lose.

D. Participation Cutoffs

The cutoffs used in this paper to proxy for active market participation and nonparticipation are admittedly ad hoc. I re-estimate Equation 3 for each confidence level, effectively generating a distinct game sentiment measure for each confidence level. In untabulated results, *AbsGameSent* is positive and statistically significant at conventional levels for each level from 1-6. *AbsGameSent* is positive but insignificant, and negative but insignificant for confidence levels 7 and 8, respectively. *AbsGameSent* is negative and statistically significant at the 5% level for both confidence levels 9 and 10. This level of granular detail supports my findings that sentiment, among active market participants, improves price efficiency, while the removal of correlated beliefs from the market increases forecast errors.

E. Team Biases

Prior literature generally considers sentiment as a time-varying behavioral bias. However, individuals are also prone to time-insensitive preferences, such as loyalty to an alma mater or favorite team. These constant preferences may manifest in irrational sentiment in favor of a particular asset. Table VII presents evidence of contestants' bias toward a particular team. [25] To measure team bias for Georgia, for example, I first collected all game picks where a contestant selected Georgia to win. I then calculated the level of overconfidence in the game pick by comparing the contestant's self-assessed confidence level to the expected confidence level for each contestant's game pick. Finally, for all contestants selecting Georgia to win a particular game, I compared the average overconfidence level of contestants that chose Georgia to win *every* game they played to the average overconfidence level of all other contestants that predicted Georgia to lose at least one game. Among contestants that picked Georgia to win every game, the average difference between the self-assessed confidence level and the expected confidence level is 0.5189. The average difference among all other contestants is -0.2550, suggesting that contestants biased toward

[25] In calculating team bias, I only included teams with at least six games represented in the contest and were listed as an underdog in at least one game. Stanford and Wisconsin each had six games in the contest but both schools were favored to win each game, so I excluded them from Table VII.

Georgia are significantly more confident in picking Georgia than is the average unbiased contestant. Contestants biased in favor of a particular team are on average, significantly more confident in choosing that team than unbiased contestants for 16 out of 19 teams. In fact, Clemson is the only team for which biased contestants were less confident than unbiased contestants when selecting that team to win.[26]

VII. Conclusions

While prior empirical work examining the relationship between investor sentiment and asset prices has generally concluded that sentiment drives asset prices away from fundamental value, I present evidence that sentiment may actually improve price efficiency. Under the plausible assumption that investors condition their market participation on their level of uncertainty regarding an asset's future payoffs, investors rationally decide to enter the market if, and only if, their uncertainty level is below a certain threshold. Asset prices become less efficient as the degree of correlation among market nonparticipants' beliefs increases. Conversely, as more overconfident investors bring new information into the market by actively trading, price efficiency improves. By distinguishing sentiment between active market participants and nonparticipants, I show that pricing errors decline when market participants reveal strong preferences, and pricing errors increase when correlated beliefs are removed from the price discovery process. Additionally, I provide some evidence consistent with market nonparticipants' beliefs containing valuable information that is not being captured by market prices, resulting in opportunities for profitable betting strategies.

The results in this paper support recent evidence that 'crowdsourcing' websites may be used to predict future returns (see Avery, Chevalier, and Zeckhauser (2012). Since prices typically are set by active market participation (i.e., active trading), nonparticipants are underrepresented in market-based data. Capturing the private expectations of market nonparticipants is vital to improving price efficiency. Future research may consider the role of conditional market participation on price efficiency when market entry

[26] The results in this paper are unaffected by the removal of biased contestants.

costs are not uniform across all investor classes. For example, a tax on foreign investors or major shift in accounting regimes may impact market participation unequally across investor classes. If the affected classes share private expectations of future asset payoffs, price efficiency is likely to be affected.

References

Ali, Mukhtar M., 1977, Probability and utility estimates for racetrack bettors, *Journal of Political Economy* 85, 803-815.

Allen, Franklin, and Douglas Gale, 1994, Limited market participation and volatility of asset prices, *American Economic Review* 84, 933–955.

Asch, Peter, Burton G. Malkiel, and Richard E. Quandt, 1982, Racetrack betting and informed behavior, *Journal of Financial Economics* 10, 187-194.

Avery, Christopher, and Judith A. Chevalier, 1999, Identifying investor sentiment from price paths: The case of football betting, *Journal of Business* 72, 493–521.

Avery, Christopher, Judith A. Chevalier, and Richard J. Zeckhauser, 2012, The "CAPS" prediction system and stock market returns, Working paper, NBER working paper series.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773–806.

Beber, Alessandro and Marco Pagano, 2011, Short-selling bans around the world: Evidence from the 2007-09 crisis, *Journal of Finance* 68, 343-381.

Brav, Alon, George Constantinides, and Christopher Gezcy, 2002, Asset pricing with heterogeneous consumers and limited participation: Empirical evidence, *Journal of Political Economy* 110,793–824.

Cao, Henry H., Tan Wang, and Harold H. Zhang, 2005, Model uncertainty, limited participation, and asset prices, *Review of Financial Studies* 18, 1219–1251.

Crafts, N. F. R., 1985, Some Evidence of Insider Knowledge in Horse Race Betting in Britain, *Economica* 52, 295-304.

Dare, William H., and S. Scott MacDonald, 1996, A generalized model for testing the home and favorite team advantage in point spread markets, *Journal of Financial Economics* 40, 793–808.

Das, Sanjiv R., and Mike Y. Chen, 2007, Yahoo! for Amazon: Sentiment extraction from small talk on the web, *Management Science* 53, 1375-1388.

DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Diamond, Douglas W., and Robert E. Verrecchia, 1987, Constraints on short-selling and asset price adjustment to private information, *Journal of Financial Economics* 18, 277–311.

Dow, James, and Sergio Ribeiro da Costa Werlang, 1992, Uncertainty aversion, risk aversion, and the optimal choice of portfolio, *Econometrica* 60, 197–204.

Dowie, Jack, 1976, On the efficiency and equity of betting markets, *Econometrica* 43, 139-150.

Dumas, Bernard, Alexander Kurshev, and Raman Uppal, 2009, Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility, *Journal of Finance* 64,579–629.

Easley, David, and Maureen O'Hara, 2009, Ambiguity and nonparticipation: The role of regulation, *Review of Financial Studies* 22, 1817-1843.

Fischhoff, Baruch, Paul Slovic, and Sarah Lichtenstein, 1977, Knowing with certainty: The appropriateness of extreme confidence, *Journal of Experimental Psychology* 3, 552–564.

French, Kenneth, and James Poterba, 1991, Investor diversification and international equity markets, *American Economic Review* 81, 222–226.

Gabriel, Paul E., and James R. Marsden. 1990, An examination of market efficiency in the British racetrack betting, *Journal of Political Economy* 98, 874-885.

Gandar, John, Richard Zuber, Thomas O'Brien, and Ben Russo, 1988, Testing market rationality in the point spread betting market, *Journal of Finance* 43, 995-1007.

Garfinkel, Jon A., 2009, Measuring investors' opinion divergence, *Journal of Accounting Research* 47 (5): 1317–48.

Giannini, Robert C., and Paul J. Irvine, 2012, The impact of divergence of opinion about earnings within a social network, Working paper, University of Georgia.

Goetzmann,William N., and Alok Kumar, 2008, Equity portfolio diversification, *Review of Finance* 12, 433-463.

Golec, Joseph, and Maurry Tamarkin, 1991, The degree of inefficiency in the football betting market: Statistical tests, *Journal of Financial Economics* 30, 311-323.

Golec, Joseph and Maurry Tamarkin, 1995, Do bettors prefer long shots because they are risk-lovers, or are they just overconfident?, *Journal of Risk and Uncertainty* 11, 51–64.

Golec, Joseph, and Maurry Tamarkin, 1998, Bettors love skewness, not risk, at the horse track, *Journal of Political Economy* 106, 205–225.

Gray, Philip, and Stephen Gray, 1997, Testing market efficiency: Evidence from the NFL sports betting market, *Journal of Finance* 52, 1725-1737.

Greenwood, Robin, 2009, Trading restrictions and stock prices, *Review of Financial Studies* 22, 509-539.

Grossman, Sanford, 1976, On the efficiency of competitive stock markets where traders have diverse information, *The Journal of Finance* 31, 573–85.

Gu, Zhaoyang, and Joanna Shuang Wu, 2003, Earnings skewness and analyst forecast bias, *Journal of Accounting and Economics* 35, 5-29.

Ko, K. Jeremy, and Zhijian Huang, 2007, Arrogance can be a virtue: Overconfidence, information acquisition, and market efficiency, *Journal of Financial Economics* 84, 529-560.

Miller, Edward M.,1977, Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32,1151–1168.

Neal, Robert, and Simon M. Wheatley, 1998, Do measures of investor sentiment predict returns?, *Journal of Financial and Quantitative Analysis* 33, 523–547.

Odean, Terrance, 1998, Volume, volatility, price and profit when all traders are above average, *Journal of Finance* 53, 1887–1934.

Odean, Terrance, 1999, Do investors trade too much? *American Economic Review* 89, 1279–1298.

Paul, Rodney J., Andrew P. Weinbach, and Chris J. Weinbach, 2003, Fair bets and profitability in college football gambling, *Journal of Economics and Finance*, 27, 236-242.

Rosett, Richard N., 1971, Weak experimental verification of the expected utility hypothesis, *Review of Economic Studies* 38, 481-492.

Sauer, Raymond, Vic Brajer, Stephen Ferris, and M. Wayne Marr, 1988, Hold your bets: Another look at the efficiency of the gambling market for National Football League games, *Journal of Political Economy* 96, 206-213.

Shiller, Robert J., 1981, Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71, 421–498.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits to arbitrage, *Journal of Finance* 52, 35–55.

Shive, Sophie, 2012, Local investors, price discovery, and market efficiency, *Journal of Financial Economics* 104, 145-161.

Snowberg, Erik and Justin Wolfers, 2010, Explaining the favorite-longshot bias: Is it risk-love or misperceptions?, *Journal of Political Economy* 118, 723-746.

Snyder, Wayne W., 1978, Horse racing: Testing the efficient markets model, *Journal of Finance* 33, 1109-1118.

Tuckwell, R. H., 1983, The thoroughbred gambling market: Efficiency, equity and related issues, *Australian Economic Papers* 22, 106-118.

Vissing-Jørgensen, Annette, 2002, Limited asset market participation and the elasticity of intertemporal substitution, *Journal of Political Economy* 110, 825–853.

Weitzman, Martin, 1965, Utility Analysis and Group Behavior: an Empirical Study, *Journal of Political Economy* 73, 18-26.

Woodland, Linda M., and Bill M. Woodland, 1994, Market efficiency and the favorite-longshot bias: The baseball betting market, *Journal of Finance* 49, 269-279

Zuber, Richard A., John M, Gandar, and Benny D. Bowers, 1985, Beating the spread: Testing the efficiency of the gambling market for National Football League games, *Journal of Political Economy* 93, 800-806.

Figure 1. Winning Percentage by Confidence Level: Contestants' picks. This figure compares winning percentages of contestants' picks when confidence levels are self-assessed versus market expectations. The 'Self-Assessed Confidence' line tracks contestants' winning percentage according to each contestant's self-reported calibrated confidence level. The 'Market Expectations' line tracks contestants' winning percentage under the hypothetical assumption that contestants ranked games according to closing point spreads.

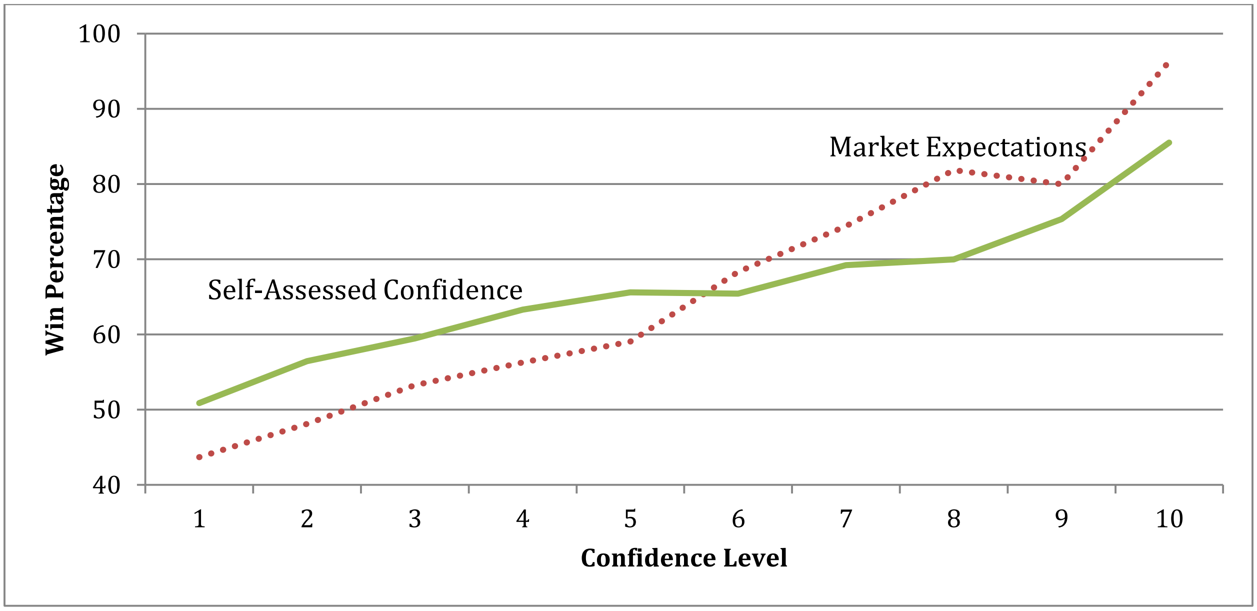

Table I Sample Composition

Total unique entries	283,839
Entries picking less than 140 games	(216,158)
Entries picking all 140 games	67,681
Number of weeks	14
Number of entry-weeks	947,534
Entry-weeks using default confidence point assignment	(91,195)
Total entry-weeks	856,339
Games per week	10
Total game picks	**8,563,390**

Table II **Summary Statistics**

This table shows summary statistics of the key variables. *Panel A* statistics are indexed by the home team. *Panel B* compares contestants' performance to a market-based strategy of selecting favorites to win every game and assigning confidence points based on the magnitude of the point spread. The 'hybrid strategy' assigns confidence points to contestants' actual picks based on the magnitude of the closing point spread. Raw wins and confidence points are downwardly biased due to the exclusion of entry-weeks that did not change the default confidence point assignments.

Panel A: Game variables (indexed by home team)

Variable	Mean	Median	Std. Dev.	Min	Max	Obs.
Point spread	-2.314	-2.500	9.674	-29.5	32	140
Point differential	2.979	4.000	19.496	-60	44	140
Forecast error	0.664	-1.250	15.599	-48	35	140
Absolute forecast error	12.593	11.000	9.167	0	48	140
Game sentiment	-0.779	-0.903	1.613	-3.931	3.993	140
Absolute game sentiment	1.481	1.355	0.999	0.042	3.993	140

*Panel B: Contestant selections**

Variable	Market Strategy	Hybrid Strategy	Contestant Picks
Average Wins	102	86	86
Average Win %	.729	.660	.660
Average Confidence Points	579	531	506
Average Confidence Point %	.752	.743	.706
Percentage of contestants surpassing Market Strategy Win %			1.85%
Percentage of contestants surpassing Market Strategy CP %			8.94%

* based on 65,894 contestants

Table III Win Percentage by Confidence Level

Confidence Level	Market Strategy	Hybrid Strategy	Contestant Picks
10	92.86	96.27	85.53
9	78.57	80.00	75.34
8	85.71	81.84	70.00
7	71.43	74.43	69.24
6	71.43	68.32	65.45
5	57.14	59.04	65.62
4	50.00	56.28	63.29
3	71.43	53.28	59.46
2	71.43	48.13	56.47
1	78.57	43.68	50.87
Total	**72.86**	**66.13**	**66.13**

This table displays winning percentages by confidence level for three different scenarios.

Market Strategy represents the winning percentages of teams favored to win each game. Favorites are rank-ordered each week based on closing point spreads. Teams favored to win by the largest margin each week are assigned to Confidence Level 10; whereas teams favored to win by the lowest margin each week are assigned to Confidence Level 1. Ties (i.e., games with the exact same point spread) are randomly assigned to a higher or lower confidence level.

Hybrid Strategy represents the winning percentages of contestants' predicted winners. Contestants' picks are rank-ordered each week based on closing point spreads. Among the portfolio of teams selected by a contestant in a given week, teams favored to win by the largest margin each week are assigned to Confidence Level 10; whereas teams favored to win by the lowest margin each week are assigned to Confidence Level 1. Since contestants often pick underdogs to win, I assign negative values to the point spread for selected underdogs prior to rank-ordering contestants' weekly picks. Ties (i.e., games with the exact same point spread) are randomly assigned to a higher or lower confidence level.

Contestant Picks represents the winning percentages of contestants' predicted winners. Contestants' picks are assigned to Confidence Levels based on each contestant's self-assessed game confidence ranking. For example, contestants correctly picked the winner 70.00% of the time when assigning a confidence score of 8 to the game.

Winning percentages for *Hybrid Strategy* and *Contestant Picks* are calculated from 856,339 game picks for each confidence level.

Table IV Win Percentage by Confidence Difference

ConDiff	Count	Win Percent – CP	Cumulative Percentage
+9	9,396	27.67	0.11%
+8	22,821	37.44	0.38%
+7	56,796	44.04	1.04%
+6	124,669	45.58	2.50%
+5	237,060	44.87	5.26%
+4	401,113	49.76	9.95%
+3	599,654	55.68	16.95%
+2	819,469	60.70	26.52%
+1	1,194,695	65.29	40.47%
0	1,794,976	71.08	61.43%
-1	1,148,171	71.54	74.84%
-2	739,323	71.96	83.47%
-3	509,622	71.93	89.42%
-4	354,991	72.65	93.57%
-5	249,500	72.20	96.48%
-6	154,899	76.39	98.29%
-7	99,151	64.34	99.45%
-8	35,289	70.98	99.86%
-9	11,795	97.55	100.00%
Total	8,563,390	66.13	

This table displays frequency counts and winning percentages by *ConDiff* levels.

ConDiff is calculated as the difference between the expected confidence ranking and the actual self-assessed confidence level for each game pick. Expected confidence rankings are constructed for each game pick based on each entry-week's portfolio of predicted winners and published closing point spreads. Among the teams selected to win for a particular entry-week, the team favored to win by the smallest (largest) margin is assigned one (ten) expected confidence point(s).

A ConDiff of +5 (-5) suggests a contestant assigned a confidence ranking five spots higher (lower) than would be predicted solely on the basis of point spread magnitudes; thus represents excessive optimism (pessimism) by the contestant.

Table V **Sentiment and Forecast Error**

$$AbsPSE = \alpha + \beta_1 \, AbsGameSent + \beta_2 \, Spread + \beta_3 \, Favorite + \varepsilon \qquad (3)$$

	All	Market Nonparticipants	Market Participants
Intercept	8.92486***	8.44405***	14.14623***
	(1.90324)	(1.90390)	(2.03857)
AbsGameSent	1.93880**	1.88543***	-0.95249**
	(0.85707)	(0.71195)	(0.45785)
Spread	0.17115	0.22514	0.21251
	(0.13858)	(0.13854)	(0.14022)
Favorite	1.84161	1.95607	2.10861
	(2.69711)	(2.66355)	(2.69774)
Adj. R-squared	0.0363	0.0540	0.0581
Observations	100	100	100

Data is compiled from weeks 4-13 of the 2011 College Pick‘em contest hosted by games.espn.go.com. Column 1 includes all game picks. Column 2, Market Nonparticipants, includes only those game picks which are assigned to confidence levels 1 thru 8. Column 3, Market Participants, includes only those game picks which are assigned to confidence levels 9 and 10. For each game, I compute the absolute point spread error (*AbsPSE*) by subtracting the closing Vegas line for that game from the actual final point differential [(Home team score – Away team score) – Point Spread]. *AbsGameSent* is a game-specific sentiment measure computed via a three step procedure: First, I calculate the expected confidence ranking for each game based on each entry-week's ten predicted winners and the corresponding closing point spread for that game. I rank games based on the magnitude of the point spread, with 10 points allocated to the selected team with the most favorable point spread, and 1 point allocated to the selected team with the least favorable point spread (ties are broken randomly). Second, I calculate the average difference between the reported confidence ranking and the expected confidence ranking for each team in the sample. Finally, I compute the weighted average of the difference between the average confidence difference of the home team and the average confidence difference of the away team. *Spread* is the closing Vegas point spread. *Favorite* is an indicator variable equal to one if the home team is favored to win the game; zero otherwise. Standard errors are in parentheses. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

Table VI Betting Strategy

Strategy	Number of bets	Bets Won	Ties	Winning Percentages Ties push*	Ties lose
Home Favorite	55	33	1	61.1%	60.0%
Favorite	100	56	4	58.3%	56.0%
Home Team	100	52	4	54.2%	52.0%
Home Favorite/High Diff	23	14	1	63.6%	60.9%
Home Team/Top Decile	10	6	1	66.7%	60.0%
Home Favorite /Top Decile	6	5	1	100.0%	83.3%

* A tie means that the margin of victory exactly matched the point spread. In some betting circles, all bets are returned in the event of a tie (i.e., ties push). Otherwise, ties are considered a loss. In the *Ties push* column, winning percentages are computed by removing pushes from the number of bets.

This table presents the profitability (winning percentage) of five betting strategies using ESPN's 2011 College Pick'em Challenge. *High Diff* (*Top Decile*) includes all games where the absolute difference between market participants' game sentiment and market nonparticipants' game sentiment was above the median (in the top decile).

Table VII Biased versus All Others

Team	Games	Biased For	All others	Difference
Notre Dame	6	-1.4807	-2.8340	1.3533***
USC	6	0.6149	-0.6879	1.3027***
Ohio State	8	0.9005	-0.3370	1.2375***
Texas A&M	6	-0.7696	-1.6896	0.9200***
Georgia	7	0.5189	-0.2550	0.7740***
Texas	7	0.9118	0.2097	0.7021***
Penn State	6	0.7381	0.0707	0.6674***
Oklahoma	6	-0.9607	-1.5631	0.6025***
Nebraska	6	-0.6290	-1.0698	0.4408***
Auburn	7	1.4408	1.1162	0.3246***
Oregon	6	-1.0702	-1.3295	0.2594***
Oklahoma State	6	0.2543	-0.0044	0.2587***
South Carolina	8	0.4607	0.2470	0.2137***
Michigan State	8	1.5084	1.3392	0.1691***
Michigan	9	0.4962	0.3909	0.1053***
Baylor	6	2.1787	2.0738	0.1049
Miami	6	0.7478	0.6476	0.1002**
Iowa	6	1.4853	1.5248	-0.0394
Clemson	8	0.4193	0.5983	-0.1790***

This table presents mean *ConDiff* scores for two subgroups of contestants. *Games* represents the number of football games listed in the contest for each team The 'Biased For' column represents contestants who predicted a particular team would win every possible football game for which that team played within the 2011 College Pick'em contest lineup. The 'All Others' column represents contestants who predicted a particular team would lose at least one football game for which that team played within the 2011 College Pick'em contest lineup. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.