Skewness Sentiment and Market Anomalies*

Alok Kumar, *University of Miami*

Mehrshad Motahari, *City, University of London*

Richard J. Taffler, *University of Warwick*

January 2023

Abstract

This study demonstrates that the skewness preference of investors is an important driver of various market anomalies. Using a combined measure of mispricing based on 11 prominent anomaly strategies, we show that return predictability associated with the mispricing component of market anomalies is stronger among firms with higher idiosyncratic skewness. The predictability differences are driven by the higher underperformance of high-skewness firms in Short anomaly portfolios. Skewness does not affect the performance of Long anomaly portfolios. Portfolio holdings data from a retail brokerage firm show that investors with stronger skewness preferences assign relatively larger weights to stocks in Short anomaly portfolios.

Keywords: Market anomalies, skewness preference, mispricing, idiosyncratic skewness, investor sentiment.

JEL Classification: G12, G14.

*Please address all correspondence to Mehrshad Motahari, mehrshad.motahari@city.ac.uk. Alok Kumar can be reached at akumar@miami.edu. Richard Taffler can be reached at richard.taffler@wbs.ac.uk. We are grateful for helpful comments and valuable suggestions from two anonymous referees, Constantinos Antoniou, Elias Badri, Turan Bali, Nick Baltas, Söhnke Bartram, Daniele Bianchi, Karl Diether, Jihoon Goh, Amit Goyal, Arie Gozluklu, Xiao Han, Jay Jung, Sonia Konstantinidi, Ingomar Krohn, Bart Lambrecht, Ashley Lester, Michael Moore, Vikas Raman, as well as participants at the 2018 Behavioural Finance Working Group Conference at Queen Mary University of London, the 2018 FMA European Conference, the 2018 Social Finance and Financial Technology Conference at Edinburgh Business School, the 26th Annual Conference of the Multinational Finance Society at Jerusalem School of Business Administration, the CEPR Third Annual Spring Symposium at Imperial College Business School, and the seminars at Bayes Business School (formerly Cass), Cambridge Endowment for Research in Finance (CERF), Cambridge Judge Business School, Schroders Asset Management, and Warwick Business School. We are responsible for all remaining errors and omissions. Previous versions of the paper circulated under the title "Skewness Preference and Market Anomalies."

1 Introduction

Many groups of market participants exhibit a preference for skewness, perhaps to satisfy their desire to gamble and engage in speculation (e.g., Kumar 2009, Kumar et al. 2011, Bali, Gunaydin, Jansson & Karabulut 2020). A preference for skewness may also reflect a broader set of preferences such as dream utility (e.g., Mitton & Vorkink 2007, Brunnermeier et al. 2007, Barberis & Huang 2008). If these preferences are systematic and investors facing a tradeoff between average return and skewness hold under-diversified portfolios, systematic demand shifts induced by a preference for skewness (i.e., skewness sentiment) may influence asset prices. Specifically, stocks with positively skewed or lottery-like return distributions would become overpriced and earn lower average returns in the future (e.g., Mitton & Vorkink 2007, Boyer et al. 2010, Bali et al. 2011, Conrad et al. 2014).

Recent studies have used investor preference for skewness to explain various empirical patterns in asset prices, including the underpricing of initial public offerings (IPOs) (Green & Hwang 2012), the underperformance of distressed stocks (Conrad et al. 2014), and irregularities in out-of-the-money option returns (Boyer & Vorkink 2014). In this study, we extend this literature and examine whether systematic preference for skewness can explain broader mispricing patterns in stock returns. Specifically, we investigate whether the mispricing-related component of market anomalies, as identified in Stambaugh et al. (2012), is associated with systematic investor preference for skewness.

Market anomalies reflect cross-sectional patterns in stock returns that are not fully explained by differences in systematic risk exposures, as measured by existing asset pricing models. In particular, stocks with certain characteristics generate average returns that are not commensurate with their level of risk. While it is often difficult to determine whether anomalies reflect mispricing or imperfect accounting for risk, recent studies suggest that anomalies, at least partly, reflect mispricing.

For example, Nagel (2005) and Stambaugh et al. (2015) demonstrate that anomalies are more prevalent among stocks that are riskier to arbitrage and have higher arbitrage costs.

A significant component of abnormal anomaly returns can be attributed to the underperformance of overpriced stocks. These stocks need to be shorted to correct the potential mispricing, but many investors are reluctant to, or unable to do so (see, e.g., Hirshleifer et al. 2011, Stambaugh et al. 2012, Avramov et al. 2013).

Mispricing also exhibits commonality across stocks. Stambaugh et al. (2012) show that a common, time-varying component across a wide range of anomalies is related to investor sentiment. Specifically, investor sentiment influences the returns of "Short-leg" stocks that are expected to underperform in the near future (Stambaugh et al. 2012).

Motivated by these empirical findings, we examine whether the skewness preference of investors interacts with market anomalies and amplifies mispricing. Our main conjecture is that the mispricing-related component of a wide range of market anomalies at least partly reflects systematic investor preference for skewness. Specifically, we posit that higher levels of systematic demand from skewness-loving investors would generate greater overpricing among stocks with positively-skewed return distributions.[1] And this effect would be more pronounced among firms that are more difficult to arbitrage.[2]

To measure the mispricing-related component of anomalies, we follow Stambaugh et al. (2015) and define a mispricing-based anomaly measure. This measure is constructed by taking the average of each stock's decile ranks for 11 anomaly variables. The set of anomalies we consider includes accruals (Sloan 1996), asset growth (Cooper et al. 2008), composite equity issues (Daniel & Titman 2006), distress (Campbell et al. 2008), gross profitability (Novy-Marx 2013), investment-to-assets (Titman et al. 2004), momentum (Jegadeesh & Titman 1993), net operating assets (Hirshleifer et al. 2004), net stock issues (Ritter 1991, Loughran & Ritter 1995), O-score (Ohlson 1980), and return on assets (Fama & French 2006). This approach diversifies any anomaly-specific effect by taking the average of anomaly decile

[1]Investor preference for skewness is unlikely to be the sole driver of anomalies, as other factors are likely to influence mispricing. In this study, we demonstrate that skewness preference is *one* of the economically important drivers of market anomalies. Specifically, we show that skewness preference exacerbates the mispricing patterns in the cross-section.

[2]This conjecture is consistent with the findings in Stambaugh et al. (2012), who demonstrate that excess noise trader demand only influences the returns of Short-leg stocks as they are harder to arbitrage.

ranks across a range of strategies and provides a measure of the likelihood for every stock to be mispriced (see Stambaugh et al. 2015, Stambaugh & Yuan 2016).

To define skewness, we consider several skewness measures used in the recent empirical asset pricing literature. This set includes jackpot probability (Conrad et al. 2014), lottery index (Kumar et al. 2016), maximum daily return (Bali et al. 2011), expected idiosyncratic skewness (Boyer et al. 2010), and options-based idiosyncratic skewness (Conrad et al. 2013).

In our first set of empirical tests, we examine whether the performance of anomaly-based trading strategies is stronger among stocks with higher skewness. We find that the anomaly-based Long-Short portfolio earns 1.22-1.71% higher risk-adjusted returns among high-skewness firms than firms in the lowest skewness quintile. Similarly, using a regression framework, depending upon the measure of skewness used, we find that a one standard deviation increase in skewness is associated with 30-60% stronger anomaly-based predictability in returns.

Next, we examine whether the effect of skewness sentiment on anomalies is driven by the underperformance of stocks in the Short portfolio. We find that Short-leg stocks with high levels of skewness generate 3 to 9 times larger negative abnormal returns than those with low levels of skewness. In contrast, the returns of Long-leg stocks do not vary significantly with the level of skewness.

We also find that Short-leg stocks with low levels of skewness do not exhibit significant underperformance. This evidence indicates that the presence of short-selling impediments is not sufficient to explain the commonly reported finding in the literature that anomaly spreads are mostly driven by Short-leg stocks (e.g., Hirshleifer et al. 2011, Stambaugh et al. 2012, Avramov et al. 2013). In fact, skewness plays a key role in explaining why overpricing is more prevalent than underpricing in extreme anomaly portfolios.

In the next set of tests, we examine directly whether investors with a preference for skewness invest disproportionately more in Short-leg stocks. Specifically, we consider the portfolio holdings of a sample of retail investors at a large U.S. discount brokerage house.

The sample period is from 1991 to 1996. We find that investors who overweight stocks with high skewness by one standard deviation allocate 11.6-18.4% higher weight (8.7-13.9% higher excess weight, relative to the market) to Short-leg stocks relative to Long-leg stocks.

For robustness, we also use an exogenous geographical proxy for the preference for skewness used in Kumar et al. (2011): the ratio of Catholics to Protestants (i.e., CPRATIO) in the local population. We find that the Catholics to Protestants ratio is associated with a higher portfolio weight on Short-leg stocks in investor portfolios.

We consider a range of alternative explanations for our findings. First, we investigate whether idiosyncratic volatility (IVOL) can explain our findings since IVOL and skewness are highly correlated. Also, Stambaugh et al. (2015) shows that anomalies are significantly stronger for high-IVOL stocks because arbitragers are reluctant to trade them. It is possible that our skewness measures capture the arbitrage deterrent effect of IVOL rather than investor preference for skewness.

Second, we test whether the relation between skewness and anomaly returns is due to a missing systematic coskewness factor in asset pricing models, rather than potential mispricing generated by investor preference for skewness. This test is motivated by the findings in Harvey & Siddique (2000), who show that the return differences between the extreme coskewness-based portfolios are partly explained by the loading on a coskewness factor.

Next, we investigate whether our skewness measures indirectly reflect arbitrage costs instead of stock characteristics that attract investors who like skewness. This test is motivated by previous studies, which document a close association between skewness and limits to arbitrage (e.g., Bris et al. 2007, Chang et al. 2007, Xu 2007). In particular, we investigate whether short-sales constraints can explain our key findings.

Last, we examine whether skewness preference reflects known return anomalies induced by investor overreaction (e.g., Barberis et al. 1998, Daniel et al. 1998).

We find that our main results are robust and do not change in a significant manner when we account for IVOL, coskewness, and a wide range of proxies for limits to arbitrage, short-selling costs, and investor overreaction. In addition, we demonstrate that our results are concentrated among stocks headquartered in high CPRATIO regions. We can only think of one plausible explanation for this finding: Investors in regions with high ratios of Catholics to Protestants have a stronger preference for skewness (Kumar et al. 2011) and generate stronger skewness-induced sentiment. Alternative factors that could potentially explain our findings are unlikely to vary geographically with CPRATIO.

In the last part of the paper, we examine whether asset pricing models with a skewness factor are more successful in explaining anomaly returns. This analysis is motivated by the evidence in the Stambaugh & Yuan (2016) study, which demonstrates that factors associated with a common source of cross-sectional mispricing can help capture abnormal returns associated with a number of anomaly strategies. We use a similar approach and construct a skewness factor by combining four skewness measures: jackpot probability, lottery index, maximum daily return, and expected idiosyncratic skewness.

We find that adding this factor to traditional asset pricing models significantly enhances their performance in explaining market anomalies. Our skewness factor is particularly useful in explaining anomalies that are more likely to be driven by skewness in returns, such as those related to financial distress.

These findings contribute to the growing anomalies literature that focuses on mispricing, including Nagel (2005), Stambaugh et al. (2012, 2015), Avramov et al. (2013), Hanson & Sunderam (2014), Chordia et al. (2014), and McLean & Pontiff (2016). Specifically, we identify skewness preference as a new explanation for commonality in mispricing across a wide set of market anomalies.

Harvey & Siddique (2000) and Bali, Del Viva, Lambertides & Trigeorgis (2020) also examine whether skewness in returns is associated with average returns. In particular, Harvey & Siddique (2000) paper develops a rational asset pricing model where exposure to coskewness reflects undiversifiable downside risk. In contrast, we show that the preference for idiosyncratic skewness is associated with mispricing and predictable patterns in stock returns.

Similarly, our economic interpretation of the empirical findings is different from Bali, Del Viva, Lambertides & Trigeorgis (2020), even though both studies establish that idiosyncratic skewness may be associated with market anomalies. Specifically, Bali, Del Viva, Lambertides & Trigeorgis (2020) posit that idiosyncratic skewness proxies for riskiness arising from firm inflexibility or lack of availability of growth options. In contrast, our study shows that skewness sentiment induced by investor preference for idiosyncratic skewness amplifies mispricing associated with various anomalies.

The remainder of the paper is organized as follows. We summarize the evidence from related papers and develop our main hypothesis in Section 2. We present our data sources and main variables in Section 3. Section 4 reports the main empirical results and Section 5 examines a range of alternative explanations for these findings. Section 6 concludes with a summary.

2 Related Literature and Testable Hypotheses

In this section, we review the related literature and develop three testable hypotheses to examine whether the mispricing-related component of market anomalies is driven by the skewness preference of investors.

2.1 Skewness, Mispricing, and Market Anomalies

The previous finance literature has examined the role of skewness in portfolio decisions and asset prices. One strand of this literature focuses on systematic skewness or coskewness. The key prediction is that risk-averse investors would exhibit a preference for portfolios with positive skewness. Consequently, assets that increase portfolio skewness (i.e., coskewness) are more desirable and earn lower average returns (e.g., Kraus & Litzenberger 1976, Harvey & Siddique 2000, Dittmar 2002). Consistent with these predictions, Harvey et al. (2010)

show that firms with high positive coskewness or systematic skewness earn lower average returns. In this economic setting, idiosyncratic, or firm-level, return skewness does not influence investment decisions or asset prices.

Recent empirical findings, however, indicate that certain types of retail investors, and even institutional investors, exhibit a propensity to gamble in financial markets by overweighting high idiosyncratic volatility or idiosyncratic skewness firms (Kumar 2009, Kumar et al. 2016, Han et al. 2021, Agrawal et al. 2022). The empirical evidence also indicates that idiosyncratic skewness is negatively related to average future returns, even more strongly than coskewness (e.g., Kumar 2009, Boyer et al. 2010, Bali et al. 2011).

These empirical findings are consistent with the theoretical predictions of Barberis & Huang (2008) who demonstrate that investors with cumulative prospect-theoretic preferences (Tversky & Kahneman 1992) would prefer assets with high idiosyncratic skewness. These individuals overweight the tails of return distributions and overvalue securities that generate positively skewed or lottery-like payoffs.[3]

Skewness preference of investors may generate mispricing more broadly as the recent anomalies literature finds that underperforming stocks have large positive skewness. This feature can attract investors with a preference for skewness and has the potential to generate overpricing (underpricing) among more (less) positively skewed stocks[4]. The potential mispricing is likely to persist because it may be too risky or costly for other investors who do not exhibit a preference for skewness to adjust the prices (e.g., Barberis & Huang 2008, Conrad et al. 2014).

Examples of market anomalies that can potentially be attributed to the mispricing induced by skewness preference include IPOs (Green & Hwang 2012), distressed firms (Conrad et al. 2014), and out-of-the-money options (Boyer & Vorkink 2014). More recently, Jiang et al. (2020) construct a lottery factor to examine the impact of lottery preference on various

[3]Previous empirical studies also find direct support for the role of cumulative prospect theory preferences in the pricing of skewness. In particular, Barberis et al. (2016) show that the prospect-theoretic value function assigns a higher value to positively skewed stocks and those stocks are overvalued internationally.

[4]See Barberis (2013) for a review.

anomalies.

The existence of market anomalies does not necessarily imply market mispricing, but evidence from a number of studies shows that they might be connected. For example, anomalies are more pronounced among stocks that are riskier to arbitrage (e.g., Nagel 2005, Stambaugh et al. 2015) and an increase in arbitrage activity is associated with quicker decay in anomaly strategy returns (e.g., Hanson & Sunderam 2014, Chordia et al. 2014, McLean & Pontiff 2016). Further, the profitability of anomaly strategies is largely generated by the short side, which consists of overpriced stocks (e.g., Hirshleifer et al. 2011, Stambaugh et al. 2012, Avramov et al. 2013).[5] More recently, Stambaugh et al. (2012, 2014) further develop a link between market anomalies and mispricing. They identify a common mispricing component across a wide range of anomaly strategies driven by investor sentiment.

2.2 Testable Hypotheses

Our empirical tests focus on the cross-sectional variation in the anomalies-mispricing relation. Our main conjecture is that the effect of mispricing on returns is more pronounced for stocks with higher levels of skewness, i.e., there is an interaction between skewness and mispricing. Specifically, the impact of skewness on returns would be stronger for overpriced stocks than for underpriced stocks.

The greater mispricing among high-skewness firms would be induced by systematic trading by skewness-loving investors that is unrelated to broad macroeconomic variables, i.e., skewness sentiment. Skewness-induced sentiment would vary cross-sectionally with the level of skewness, and it would be distinct from time-varying market-level sentiment. However, similar to the effects of market-level sentiment, the impact of skewness sentiment on returns would depend upon the risk and costs associated with the arbitrage process.

Higher levels of skewness may also attract more noise traders and deter arbitrageurs if they perceive greater noise trader risk among high-skewness firms. As a result, relatively

[5]This observation is consistent with Miller (1977)'s conjecture that mispricing exists because short-selling impediments make it more difficult to adjust overpricing compared with underpricing.

overpriced assets would exhibit a stronger negative correlation with future returns if they have higher levels of skewness.

These conjectures are based on the empirical observation that stocks in the Short (Long) leg of anomaly strategy portfolios that generate the greatest abnormal returns in the future often have the highest (lowest) levels of skewness. We are also motivated by theoretical studies, which predict a connection between skewness and mispricing. In particular, skewness is known to be negatively related to past returns (e.g., Chen et al. 2001, Cao et al. 2002, Xu 2007, Del Viva et al. 2017). Since stocks in the Short (Long) legs generate (higher) lower future returns, they are likely to be associated with relatively higher (lower) levels of skewness in the future.

Further, short-sale constraints increase the skewness of individual stocks (e.g., Chang et al. 2007, Xu 2007) and anomaly strategy returns are mostly generated by stocks in the Short leg, especially those with significant short-sale constraints (Nagel 2005). As a result, short-sale constraints induce most mispriced stocks to also have higher levels of skewness.

These previous results jointly imply that, in the presence of short-sale constraints, stocks with higher levels of overpricing (underpricing) would be associated with higher (lower) levels of skewness. Specifically, Short-leg stocks that are overpriced are more likely to attract skewness-loving investors and would be subject to higher levels of noise-trader demand. In contrast, similar to the evidence in Stambaugh et al. (2012), skewness-driven noise trading is unlikely to affect underpriced, Long-leg stocks.

Combining these theoretical predictions with the findings in Stambaugh et al. (2012, 2014) that market anomalies have a common mispricing component, we propose three testable hypotheses. We first posit that in the presence of arbitrage constraints, systematic demand shifts of skewness-loving investors have the potential to exacerbate market anomalies. More formally, our first hypothesis is:

> **H1:** *The negative relation between the mispricing component of market anomalies and future returns is stronger among stocks with higher idiosyncratic skewness.*

There are two justifications for this interactive effect. First, Stambaugh et al. (2012) show that shifts in noise trader demand only influence the Short leg of market anomalies due to arbitrage asymmetry. Consequently, potential mispricing is concentrated in the Short-leg portfolios where short-sale constraints and limits to arbitrage are likely to be binding.

Based on the same logic, we expect any mispricing generated by the sentiment of skewness-loving investors to mainly affect Short-leg stocks where arbitrage forces are less effective. Since stocks in the Short legs are more positively skewed, we also expect the level of systematic trading by skewness-loving investors to increase with skewness.

In contrast, Long-leg stocks are less likely to be mispriced as they do not face significant limits to arbitrage (Miller 1977). Although the skewness characteristics of these stocks may induce underpricing due to reduced investor demand, arbitrage forces can correct any potential underpricing quicker and more effectively.

Second, An et al. (2020) show that investors exhibit a stronger preference for positively-skewed stocks when they operate in the loss domain since gambles that provide an opportunity to "break-even" are more attractive. The implication of this finding in our setting is that relatively overpriced Short-leg stocks with poor past performance are likely to be more attractive to skewness-loving investors as these stocks would trigger investors' loss-aversion proclivities that induce a stronger preference for skewness. Consequently, lower past returns would amplify the skewness effect (i.e., skewness-induced sentiment would be stronger), and thus the interaction between skewness and anomalies would be stronger.[6]

In contrast, Long-leg stocks often have better return performance and, therefore, they are unlikely to be as attractive as stocks in the Short-leg, even if they have large positive skewness. As a result, prices of stocks in the Long leg anomaly portfolios are unlikely to be affected by the demand shifts of skewness-loving investors.

More formally, our second hypothesis posits that the effect of skewness on anomalies

[6]We expect low past returns to strengthen the interactive effect between skewness and anomalies. Our empirical tests ensure that we are not merely capturing the effects of an omitted variable that captures the interaction between skewness and past returns.

reflects the underperformance of overpriced stocks with higher skewness.

> **H2:** *Skewness preference has a stronger impact on firms in the Short anomaly portfolio. High skewness firms in the Short leg earn lower average future returns, while similar firms in the Long anomaly portfolio are likely to be fairly priced.*

Last, we focus on the mechanism through which investors with skewness proclivities affect market anomalies. Investors with strong preferences for skewness are likely to invest disproportionately more in Short-leg stocks as compared to stocks in the Long-legs. In particular, as predicted by the Barberis & Huang (2008) model, those investors would overweight high skewness stocks. Since stocks in the Short legs of anomalies are more positively skewed than those in the Long legs, Short-leg stocks should be relatively more attractive to investors with skewness preferences. We formulate this prediction as follows:

> **H3:** *Investors with a preference for idiosyncratic skewness assign a higher (lower) weight to stocks in the Short (Long) anomaly portfolio.*

3 Data and Measures

3.1 Main Data Sources

We test our three predictions using data from several sources. Our main tests are based on a sample of all common NYSE, AMEX, and NASDAQ stocks (share code 10 or 11) available in the Center for Research in Security Prices (CRSP) database during the January 1963 to December 2015 period. We exclude firms with negative book equity, those belonging to the financial sector ($6000 \leq SIC \leq 6999$), or those with a share price below $1.[7] When the returns data for a stock are missing, we use delisting returns.

To construct our main skewness and anomaly variables, we use accounting data from Compustat Fundamentals Annual and Quarterly files and option price data from Option-

[7] We consider other share price cutoffs in the robustness tests and show that our results do not depend on the price filter.

Metrics. Our factor returns and risk-free rates are from Professor Kenneth French's data library.[8] We also use the end-of-month portfolio positions of a sample of retail investors from a U.S. discount brokerage house covering the 1991 to 1996 time period. In some of our robustness tests, we use short interest data from Compustat, shorting fee scores from Markit Data Explorers, and quarterly data on institutional stock holdings from Thomson Reuters.

Table A.1 presents the definitions and sources of all variables.

3.2 Skewness and Anomaly-Based Mispricing Measures

We consider four firm-level skewness measures: jackpot probability ($JACKPOT$), lottery index ($LIDX$), maximum daily return ($MAXRET$), and expected idiosyncratic skewness ($ESKEW$). $JACKPOT$ is based on Conrad et al. (2014) and is defined as the out-of-sample probability of a stock generating a log return greater than 100% during the next 12 months. $LIDX$ is a lottery index introduced in Kumar et al. (2016), which ranks securities by how much they share lottery-like features (i.e., low price, high volatility, and high skewness) associated with a preference for skewness. $MAXRET$ is the stock's maximum 1-day return in the past month, as defined in Bali et al. (2011). $ESKEW$ is defined as an out-of-sample measure of expected idiosyncratic skewness, following Boyer et al. (2010).

To define skewness based on option prices, we use the options-based idiosyncratic skewness (OS) measure of Conrad et al. (2013). This measure is defined as the third moment of the (risk-neutral) density function of individual securities formulated in Bakshi et al. (2003). The advantage of OS over the previous measures is that it is based on a nonparametric *ex ante* estimate of future return expectations. Therefore, it should be able to capture investors' expectations of future return skewness.[9] However, OS is only available for a small subset of stocks with traded options, and thus, we do not use it in all of our tests. Finally, we use the

[8] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/.

[9] Conrad et al. (2013) note that with a relatively constant pricing kernel over shorter periods, the cross-sectional variation in risk-neutral and physical moments will capture the same information. Thus, the risk-neutral skewness measure we employ directly captures the cross-sectional information contained in physical skewness, with the advantage of being forward-looking compared to any physical measure of skewness.

coskewness (*COSKEW*) measure of Harvey & Siddique (2000) in our robustness tests.

Beyond the skewness measures, we consider 11 prominent anomaly strategies analyzed in Stambaugh et al. (2012, 2014, 2015). This set of anomalies consists of accruals (Sloan 1996), asset growth (Cooper et al. 2008), composite equity issues (Daniel & Titman 2006), distress (Campbell et al. 2008), gross profitability (Novy-Marx 2013), investment-to-assets (Titman et al. 2004), momentum (Jegadeesh & Titman 1993), net operating assets (Hirshleifer et al. 2004), net stock issues (Ritter 1991, Loughran & Ritter 1995), O-score (Ohlson 1980), and return on assets (Fama & French 2006).

We also use the mispricing (*MIS*) measure defined in Stambaugh et al. (2015). *MIS* is constructed by taking the average of each stock's decile ranks with respect to the 11 anomaly variables. Decile ranks are defined at the end of each month. The 1st and the 10th deciles consist of stocks that each anomaly strategy predicts will outperform and underperform in the following month, respectively. Considering anomalies may not fully capture mispricing, *MIS* is a less noisy measure of mispricing. By taking the average of the anomaly decile ranks, any anomaly-specific effect is diversified and the mispricing component that is common across strategies remains (see Stambaugh et al. 2015, Stambaugh & Yuan 2016).

3.3 Skewness, Mispricing, and Average Returns

Table 1, Panel C presents the performance of *MIS* and the four key skewness measures (i.e., *JACKPOT*, *LIDX*, *MAXRET*, and *ESKEW*) in predicting future returns. We sort stocks into quintiles at the end of every month based on the five variables. Then, we measure the value-weighted return of each quintile group, together with the return of the hedge portfolio (i.e., Long Quintile 5 and Short Quintile 1) in the following month. To compute risk-adjusted performance measures, we regress the monthly returns of each portfolio on the three (Fama & French 1993), the four (Carhart 1997), and the five (Fama & French 2015) factors separately and report the alphas.

The Long-Short strategies associated with all five measures generate statistically significant abnormal returns at the 1% level. The exception is the alpha of the *MAXRET* Long-Short portfolio, which is only significant at the 10% level. The *MIS* Long-Short portfolio yields highly statistically significant alphas with all three models, ranging from 63 to 109 basis points per month. In line with Stambaugh et al. (2015), we find that the majority of *MIS* Long-Short portfolio returns come from the Short leg. With all the factor models, the Short *MIS* portfolios (Quintile 5) generate alphas that are more than twice the alphas of corresponding Long portfolios (Quintile 1).

3.4 Summary Statistics: Characteristics of Mispriced Stocks

To identify the characteristics of stocks with different levels of mispricing, we present the mean cross-sectional characteristics of *MIS* quintiles in Panel A of Table 1. Quintile rankings are determined monthly by sorting stocks using their end-of-month *MIS* value. We measure the characteristics at the end of the month in which we define the quintiles.

We find that the Short-leg (Quintile 5) firms, on average, are smaller (lower market capitalization), are more volatile, and have cheaper shares with poorer past return performance, when compared to firms in the Long-leg (Quintile 1). Short-leg stocks are also relatively less liquid, according to the illiquidity measure of Amihud (2002), and are more heavily shorted. Average holdings indicate that institutional investors target the right stocks by holding more of the shares of Long-leg stocks. In contrast, our brokerage sample suggests that retail (individual) investors place a higher weight on Short-leg stocks.

To examine the relation between skewness and mispricing, we compare the mean skewness measures across *MIS* quintiles. Together with our four main skewness proxies, we look at coskewness (*COSKEW*), options-based skewness (*OS*), idiosyncratic skewness (*ISKEWNESS*), and total skewness (*SKEWNESS*). *ISKEWNESS* and *SKEWNESS* are computed using daily returns for the same month as *MIS*.

Panel B of Table 1 presents the results. The average values of all seven skewness measures monotonically increase across *MIS* quintiles 1 to 5. In all cases, a t-test indicates that the difference between the skewness values of quintiles 1 and 5 is statistically significant at the 5% level. These results are similar to those reported in Harvey & Siddique (2000) and Conrad et al. (2014), who find that skewness increases when moving from the Long to the Short legs of anomaly strategies.

4 Main Empirical Results

To examine the link between skewness and anomalies (hypotheses *H1* and *H2*), we perform a series of double sorts and estimate Fama-Macbeth type regressions. Using the retail brokerage data, we also test hypothesis *H3* and gather additional support for our main conjecture.

4.1 Skewness and Anomalies

4.1.1 Double Sorts

We begin by analyzing the performance of portfolios double sorted on the combined anomaly variable *MIS* and the following four skewness measures: *JACKPOT*, *LIDX*, *MAXRET*, and *ESKEW*. Following previous studies that examine the determinants of anomalies (e.g., Stambaugh et al., 2015), portfolios are formed by independently sorting stocks into quintiles based on each of the two variables at the end of each month. We then compute the value-weighted returns of the 25 portfolios over the following month and regress these on the four Carhart (1997) factors to measure abnormal returns.[10] The sample excludes stocks priced below $1 and covers the period from January 1963 to December 2015, except for sorts based on *ESKEW*, which start in January 1988.

Table 2, Panel A presents the monthly abnormal returns of the double-sorted portfolios. As expected, the degree of mispricing, as captured by the *MIS* spread (most overpriced − most underpriced), monotonically increases for all four skewness measures. The *MIS*

[10]We find similar results if we use the five-factor model of Fama & French (2015).

spread in the high skewness quintile is 1.22-1.71% higher than the spread in the low skewness quintile. The differences in *MIS* spreads across high and low skewness groups are both statistically and economically significant. Among the four skewness measures, sorts based on *JACKPOT* yield the strongest results. High *JACKPOT* stocks generate a *MIS* spread of −2.06%, which is about 6 times larger than the −0.35% spread associated with low *JACKPOT* stocks. These results support our first hypothesis (*H1*), which posits that mispricing is concentrated among stocks with higher levels of skewness.

The results in Panel A also show that the differences in *MIS* spreads across the skewness quintiles mostly come from differences in the returns of the Short portfolio that contain the most overpriced stocks. In fact, within the set of most underpriced stocks, the return difference between the high and the low skewness quintile portfolios is statistically insignificant. This evidence suggests that skewness differences do not affect the returns of underpriced stocks.

Perhaps more interestingly, we find that the negative abnormal returns for the most overpriced stocks portfolio (quintile 5) in the low *MAXRET* or low *JACKPOT* quintiles are statistically insignificant. This evidence suggests that stocks with low levels of skewness are unlikely to become overpriced even if the anomaly variables suggest they will. Therefore, the commonly reported finding in the literature that anomaly spreads are mostly driven by Short-leg stocks crucially depends on the level of skewness. These results support our second hypothesis (*H2*), which posits that the effect of skewness on anomaly returns is concentrated in the Short leg portfolio.

To examine the relative distribution of firms across the most mispriced groups, we compute the average number of observations in each of the double-sorted portfolios. The results presented in Panel B of Table 2 indicate that in the most overpriced stocks portfolio, the average number of stocks increases with each of the four skewness measures. In contrast, there are fewer firms in higher skewness quintiles among most underpriced stocks (i.e., quintile 1).

Taken together, our double-sorting results are consistent with our main conjecture, which posits that the mispricing-related component of anomalies is largely driven by stocks with higher levels of skewness. Further, we find that the impact of skewness on anomaly returns is concentrated among Short-leg stocks, which are relatively more difficult to arbitrage.

4.1.2 Fama-Macbeth Regression: Baseline Estimates

We test our first hypothesis (*H1*) again using a series of Fama & MacBeth (1973) regressions. A natural implication of hypothesis *H1* is that the anomaly premium is higher for stocks with higher levels of skewness. So, the interaction between the skewness measure and the anomaly variable is expected to have a negative sign.

At the end of each month t, we use various stock characteristics and our skewness and mispricing measures to predict stock returns in month $t+1$. The main variable of interest is the interaction between each of the skewness measures and the anomaly variable *MIS*. In all regression specifications, we control for market value, the book-to-market ratio, and past returns for the previous month and for the prior 12 months but skipping the last month. For ease of interpretation, we standardize all variables in the regressions to have a mean of zero and a standard deviation of one. All variables are also Winsorized at the 0.5 and 99.5 percentiles to ensure that extreme values do not affect our results.

Panel A of Table 3 presents the time-series averages of the baseline Fama-Macbeth regression coefficients, along with Newey & West (1987) t-statistics.[11] The first five regression specifications (Columns (1) to (5)) exclude interaction terms and test whether MIS and skewness variables are individually linked to future returns. Each of the five main variables is statistically significant at the 5% level. The *MIS* coefficient is larger and more significant than that of any of the individual skewness measures. A one-standard-deviation increase in *MIS* is associated with a 0.5% decline (t-statistic $= -11.72$) in the following month's return, after controlling for major firm characteristics. Among the skewness measures, $JACKPOT$ is the strongest return predictor with a coefficient of -0.004 (t-statistic $= -4.16$).

[11] Following Newey & West (1994), we choose the optimum number of lags (L^*) according to the following plug-in estimator: $L^* = floor(4 \times (\frac{T}{100})^{(2/9)})$, where T is the panels time dimension.

The set of independent variables in specifications (6) to (9) includes one of the skewness measures, its interaction with *MIS*, and *MIS* itself. In these regression specifications, we examine whether the skewness-anomaly interaction predicts future returns beyond what is captured by each of the two variables individually. We find that all four interaction variants are highly statistically significant, with t-statistics larger than the target threshold figure of 3 suggested by Harvey et al. (2016).[12] A one-standard-deviation increase in skewness adds 0.1-0.3% to the predictive power of *MIS* on a monthly basis. These estimates amount to between 30-60% of the predictive power of *MIS* itself.

An interesting observation among the Fama-MacBeth regression estimates is that the interaction terms fully absorb the statistical significance of *JACKPOT* and *LIDX*. In other words, the return premia of these two variables are fully generated by stocks that are likely to be mispriced, as suggested by the combined anomaly measure.

4.1.3 Fama-Macbeth Regression: Robustness Checks

To ensure that our regression estimates are not sensitive to the chosen data filters or are driven by specific parts of the sample, we perform a series of robustness tests. For brevity, in Panel B of Table 3, we only report the coefficients on our main variables of interest, i.e., the interaction terms.

We find that our regression estimates are robust. Skipping Winsorization and excluding firms with a share price lower than $5 have negligible effects on the interaction coefficients. Further, our results become slightly stronger when we drop micro-cap stocks. Excluding mega-cap stocks, however, has a limited effect on the coefficient estimates. Following Fama & French (2008), we define micro- and mega-cap stocks as those with market capitalizations below the 20th and above the 80th percentiles of NYSE market capitalization, respectively. We also experiment with removing all NASDAQ stocks from our sample. In this case, although the coefficients remain highly significant, their magnitudes shrink slightly in some

[12] Due to possibility of data-mining, Harvey et al. (2016) argue that suggest that a t-statistic of 3 is a more appropriate significance cutoff for Fama-Macbeth regressions than the usual cutoff of 2.

specifications.

We also consider different time periods in the sample. First, we divide the whole sample into recession and expansion periods, based on the NBER Recession Indicator,[13] and estimate the interaction coefficients separately for each subsample. Our goal is to examine whether the effect of skewness on mispricing is particular to recession times when the market is highly volatile.

The results, reported in rows (6) and (7) of Panel B of Table 3, indicate that the interaction coefficients remain significant in both the recession and the expansion sub-periods. The exception, however, is the $ESKEW \times MIS$ coefficient, which is only significant for expansion periods, probably because the $ESKEW$ data start in 1988. Consequently, the estimates fail to capture the recessions of the 1970s and the 1980s. Among the interaction terms based on the other three skewness measures, the coefficient estimates are slightly larger but less significant during the recession periods.

Last, we divide the sample period into two parts: (i) 1962 to 1990 and (ii) 1991 to 2015. Our goal is to examine whether our key results vary over time. We observe that the Fama-MacBeth regression coefficient estimates are much larger for the second subperiod. This evidence suggests that the skewness effect is stronger during more recent time periods.

Overall, the Fama-MacBeth regression results provide additional evidence to support our first hypothesis. The results show that the level of mispricing associated with anomaly strategies is stronger among firms with high skewness.

4.1.4 Evidence Using Option-Based Skewness Measures

Our results using the four skewness measures are in line with our key predictions, but all our measures are potentially noisy proxies for investors' perceptions about future return skewness. To further ensure that our results reflect the role of skewness in predicting returns, we repeat our main tests with a skewness measure constructed using option prices.

[13] The data are available at https://fred.stlouisfed.org/series/USREC.

Specifically, we use the options-based idiosyncratic skewness measure of Bakshi et al. (2003), and Conrad et al. (2013). This options-based measure provides information regarding expected future return skewness and does not suffer from hindsight bias. Another advantage of this measure is that it does not require a parametric model for the estimation of skewness (Conrad et al. 2013).[14]

Table 4 present the results for double sorts (Panel A) and the Fama-Macbeth regressions (Panel B) using the options-based idiosyncratic skewness measure (*OS*). *OS* is constructed following the methodology of Conrad et al. (2013), as explained in Table A.1.[15] The double-sorting results in Panel A of Table 4 are similar to our baseline results. The spread between the most overpriced and the most underpriced stocks is largest among stocks in the high-*OS* quintile. As *OS* increases, *MIS* spreads do not grow with a clear monotonic pattern. However, there is a 2.06% difference between the monthly abnormal returns (t-statistic = -2.10) of the low- and the high-*OS* quintiles. Also, most of the increase in the *MIS* spread in the high skewness group comes from the change in the returns of the Short-leg (most overpriced) stocks. Again, these observations support our first and second hypotheses.

The Fama-Macbeth regression results in Panel B of Table 4 are also in line with our first hypothesis. In specification (1), we find that *OS* by itself cannot significantly predict returns. Conrad et al. (2013) conjecture that, because of the limited number of firms with available option data, the relation between *OS* and returns cannot be reliably estimated. Nevertheless, our tests do not require us to have a reliable estimate for the premium associated with *OS*. We are instead interested to see whether *OS* exacerbates the mispricing captured by *MIS*. In specification (2), we test this conjecture by adding an interaction term between *OS* and *MIS* to the model.

The coefficient of the interaction term is -0.003 (t-statistics = -2.29), indicating that a one-standard-deviation increase in *OS* increases the return predictability of *MIS* by 0.3%.

[14] We are unable to use this option prices based skewness measure in all our tests because option prices are available only for a small subset of firms in our sample.

[15] The sample period for these tests starts in 1996 because option price data for earlier years are not available in the OptionMetrics database.

This estimate is also economically significant. Considering that the *MIS* coefficient is equal to -0.003, the interaction coefficient suggests that a one-standard-deviation increase in *OS* doubles the premium associated with *MIS*.

Collectively, the regression estimates and double-sorting results using the options-based skewness measure support our previous results on the impact of skewness on anomaly-based mispricing.

4.2 Do Skewness-Loving Investors Overweight Overpriced Stocks?

Our results so far suggest that the common mispricing-related component of anomaly strategies is strongly concentrated among stocks with higher levels of skewness. In addition, we show that this relation is mostly driven by the amplifying effect of skewness on the prices of stocks that anomaly strategies suggest are overpriced. In this subsection, we test our third hypothesis (*H3*), which posits that a higher level of skewness among stocks in the Short portfolio induces skewness-loving investors to assign a relatively higher (lower) weight to overpriced (underpriced) stocks.

We use portfolio holdings data of a sample of retail investors obtained from a large U.S. discount brokerage house for the 1991 to 1996 period. We use data for retail investors because previous papers show that such investors are more likely to have a preference for skewness (Kumar 2009). Our main dependent variables are the end-of-month raw and excess weights allocated to overpriced (Short-leg) stocks in each investor portfolio.

Raw and excess relative weights are defined as $[W_{i,t}^{overpriced} - W_{i,t}^{underpriced}]$ and $[(W_{i,t}^{overpriced} - W_{mkt,t}^{overpriced}) - (W_{i,t}^{underpriced} - W_{mkt,t}^{underpriced})]$, respectively. $W_{i,t}^{overpriced}$ is the raw weight allocated to overpriced stocks in portfolio i at the end of month t, and $W_{i,t}^{underpriced}$ is the raw weight allocated to underpriced stocks in portfolio i at the end of month t. Similarly, $W_{mkt,t}^{overpriced}$ is the raw weight allocated to overpriced stocks in the market portfolio at the end of month t, and $W_{mkt,t}^{underpriced}$ is the raw weight allocated to underpriced stocks in the market portfolio at the end of month t. Underpriced and overpriced stocks are defined as those in the 1st and

5th *MIS* quintiles, respectively.

We regress the relative weight measures on a series of variables that capture investor preference for skewness. We also control for various socioeconomic and portfolio characteristics. We estimate the regressions each month and then compute the time-series averages of the coefficient estimates.

Since the preference for skewness is not directly measurable, we adopt an indirect proxy by computing the average portfolio weight each investor allocated to stocks with high levels of positive skewness in the past. We define stocks with high levels of positive skewness as those having skewness measures above the monthly cross-sectional median. At the end of each month t, we take the average of the weight each investor allocates to stocks with high levels of positive skewness over the previous 12 months ending in month $t-1$. The stronger an investor's preference for skewness, the more likely she is to have allocated a higher weight to high-skewness stocks in the past.

As before, we measure skewness using four proxies: *JACKPOT*, *LIDX*, *MAXRET*, and *ESKEW*. In addition, we incorporate the Catholic-to-Protestant ratio (*CPRATIO*) measure used in Kumar et al. (2011) and Kumar et al. (2016) as a measure of the local preference for skewness. Kumar et al. (2011) show that investors living in Catholic regions have stronger gambling tendencies and are more likely to be attracted to investments with positively skewed payoffs than are those in high Protestant regions. *CPRATIO* is defined as the number of Catholic adherents divided by the number of Protestant adherents in the portfolio holder's county. Table A.1 presents details about the construction of all variables, including the socioeconomic and portfolio characteristics controls. We standardize all independent variables to have a mean of zero and a standard deviation of one. We also Winsorize them at the 0.5 and 99.5 percentile levels.

Panel A of Table 5 presents the baseline results. Columns (1) to (4) show that investors who overweight stocks with high levels of skewness in the past year by one standard deviation of the cross-sectional distribution allocate between 11.6% and 18.4% higher raw weight to overpriced stocks, relative to underpriced stocks. Excess weight regression estimates (Columns (5) to (8)) provide a clearer picture of investors' skewness preferences as they are based on weights adjusted for benchmark (market) weights. A one-standard-deviation increase in an investor's past weight on high-skewness stocks is associated with 8.7-13.9% higher relative excess weight on overpriced stocks.

The coefficient estimates of past weights on high-skewness stocks are highly statistically significant for all four skewness measures, even after controlling for a wide range of controls and adjusting standard errors for heteroscedasticity and autocorrelation using the Newey & West (1987) approach. *CPRATIO* coefficients are also statistically significant in all cases, but have relatively small magnitudes. The estimates indicate that a one-standard-deviation increase in regional *CPRATIO* is associated with between 0.4-0.6% higher raw weights (0.3-0.5% higher excess weights) on overpriced stocks relative to underpriced stocks.

Other coefficients in Panel A of Table 5 are also informative, as they further highlight the characteristics of investor clientele who place higher weights on stocks expected to perform poorly. The regression estimates indicate that such investors hold smaller and less diversified portfolios with significantly poorer past performance and higher portfolio variance. These investors are also less likely to concentrate their positions on a specific industry or geographical location. The latter finding is in line with Ivković & Weisbenner (2005), who show that local investors have more knowledge about local stocks and are less likely to buy local stocks that perform poorly.

Investors who assign a higher relative weight to overpriced stocks are also likely to be male, single, old, and living in rental properties. Furthermore, they reside in less-populated regions with greater income inequality and poorer levels of education. Most of these characteristics are similar to those documented in previous studies as features of unsophisticated investors who exhibit stronger behavioral biases (e.g., Goetzmann & Kumar 2008, Korniotis & Kumar 2013) or a stronger preference for skewness (e.g., Mitton & Vorkink 2007, Kumar 2009).

A possible concern with our results in Table 5, Panel A is that we use the same stocks in each portfolio to compute the weights on both mispriced and high-skewness stocks. Even though all our independent variables are lagged by one month, most investors do not change their positions regularly. Therefore, the relation between the weights on skewed and overpriced stocks may just reflect the correlation between the skewness measures and our mispricing indicator *MIS*. In other words, an investor may overweight stocks with high levels of *MIS* (i.e., overpriced) for reasons other than a preference for skewness and still have a relatively high portfolio weight on skewed stocks simply because overpriced stocks have higher skewness levels.

To address this potential concern, we adjust our measures of past weight on high-skewness stocks by excluding all stocks in *MIS* quintiles 1 and 5. With this approach, we compute the average weight an investor allocates to skewed stocks after excluding those that are mispriced according to *MIS*.

Panel B of Table 5 presents the results based on our alternative weight measures. We use the same regression specification used in Panel A. For brevity, control variable coefficients are not reported as they remain very similar. Our main results remain both statistically and economically significant with the new weight measures. A one-standard-deviation increase in an investor's past weight on high-skewness stocks that are not in the extreme *MIS* quintiles predicts 7.3-13.3% higher relative raw weight on overpriced stocks (t-statistics range from 8.51 to 21.44). In relative excess weight regressions, the estimates range between 5.5% and 8.5%.

Taken together, the weight regression estimates support our third hypothesis (*H3*). Investors who have a history of holding stocks with higher levels of skewness are more (less) likely to hold stocks that will underperform (outperform), as suggested by anomaly strategies. Investors who overweight underperforming stocks relative to outperforming stocks are also likely to come from Catholic regions, where the propensity to gamble is stronger. Lastly, we observe that such investors possess other characteristics that have been previously linked to investor sophistication, skewness preference, and other behavioral biases.

4.3 An Idiosyncratic Skewness Factor

Stambaugh & Yuan (2016) show that mispricing has common drivers across stocks. A returns-based factor that captures these commonalities may incrementally explain cross-sectional differences in returns that do not reflect compensation for systematic risk. In this section, we examine whether a firm-specific skewness based *mispricing factor* is able to explain cross-sectional differences in stock returns. Considering that skewness has a significant association with the common mispricing-related component of anomaly strategies, a skewness factor may capture at least part of the commonality in anomaly returns.

We follow the approach in Stambaugh & Yuan (2016) and use our four skewness measures - *JACKPOT*, *LIDX*, *MAXRET*, and *ESKEW* - to construct a skewness factor. We first compute the average decile rank of each stock at the end of each month with respect to the four skewness measures. Next, we independently sort stocks based on their average skewness decile ranks and their market capitalization into three and two portfolios, respectively. We then compute the value-weighted monthly return of each of the six ($= 2 \times 3$) intersecting portfolios. Last, we take the average of the returns of the two size portfolios with the highest skewness tercile rank and deduct it from the average return of the two size portfolios with the lowest tercile rank to derive monthly factor returns. We call this skewness factor *nonskewed minus skewed* (*NMS*).

The results presented in Internet Appendix Table A.2 show that asset pricing models with the *NMS* factor perform better than their peers that do not have the *NMS* factor. The *NMS* factor helps capture part of the commonality in mispricing that is linked to skewness. The *NMS* factor is particularly useful for explaining distress-related anomalies, which are known to be influenced by skewness (e.g., Conrad et al. 2014).

5 Additional Tests and Alternative Explanations

Our empirical findings so far suggest that the effect of skewness on anomalies is driven by the preference of skewness-loving investors. Systematic demand shifts of skewness-loving investors, i.e., skewness sentiment, amplify the negative mispricing-return relation. In this section, we examine a range of alternative explanations for our results.

We begin by looking at the effects of IVOL and coskewness more closely to determine whether one of these variables can explain our findings. We then examine the roles of arbitrage costs and market sentiment in our setting and compare our findings with those in Stambaugh et al. (2012, 2014). Next, we investigate whether growth options and investor overreaction to news can explain our findings. Finally, to further rule out the effects of other potential confounding factors on our results, we test whether the skewness-anomalies relation varies geographically with the Catholics to Protestants ratio ($CPRATIO$).

5.1 Skewness or Idiosyncratic Volatility?

Idiosyncratic volatility (IVOL) has been used as a proxy for arbitrage constraints. Stambaugh et al. (2015) show that high-IVOL stocks are considerably more mispriced in the cross-section because high levels of IVOL make arbitrage more difficult. Measures of volatility and skewness are highly correlated, and in fact, some of the skewness measures depend mechanically on idiosyncratic volatility. In the first set of robustness tests, we ensure that our results do not merely repackage the IVOL effect documented in Stambaugh et al. (2015).

It is difficult to directly control for idiosyncratic volatility in our setting because IVOL itself can capture lottery-like characteristics that attract skewness-loving investors (e.g., Kumar 2009, Boyer et al. 2010). Further, IVOL is either included in the definition of skewness measures or is highly correlated with them, leading to potential multicollinearity issues. In spite of these difficulties, we conduct a range of tests to ensure that IVOL cannot fully explain our results.

In our first test, we control for IVOL and its interaction with *MIS* in our baseline regression model presented in Panel A of Table 3. To address potential multicollinearity concerns in these regression specifications, we replace the continuous IVOL and skewness variables with dummy variables that are equal to one if the measure has values above its cross-sectional median, and zero otherwise.[16]

The results in Panel A of Table 6 show that relative anomaly-based return predictability, as captured by the *MIS* coefficient estimates (i.e., the sum of *MIS* and *MIS* interaction coefficients relative to the *MIS* coefficient), is about 1.6-2 times as large for above-median IVOL and 1.3-2.1 times as large for above-median skewness stocks. Both IVOL and skewness interaction coefficients remain statistically and economically significant, and neither term absorbs the effect of the other.[17] Also, the relative strength of IVOL and skewness interactions varies, depending on the skewness measure used. These results suggest that both IVOL and skewness measures have an economically meaningful impact on anomalies.

We further investigate whether IVOL and skewness interactions with anomalies reflect distinct economic mechanisms. Specifically, we examine whether the coefficients in Panel A vary geographically with the ratio of Catholics to Protestants (CPRATIO) in firm headquarters counties.[18] Kumar et al. (2011, 2016) show that firms headquartered in regions with higher CPRATIO are more likely to be targeted by skewness-loving investors. In light of these findings, we expect the skewness interaction to be stronger for stocks headquartered in high-CPRATIO regions. In contrast, the IVOL interaction should remain unaffected if the IVOL captures arbitrage constraints since there is no obvious reason why arbitrage costs would vary geographically with CPRATIO.

[16]In Panel A of Internet Appendix Table A.3, we use the continuous forms of these variables in the regression specification and verify that there is significant multicollinearity. The variance inflation factor (VIF) is above 10 for several cases. In contrast, there is no indication of multicollinearity when we replace the continuous variables with their dummy versions.

[17]In Panel B of Internet Appendix Table A.3, we triple sort stocks based on *MIS*, *IVOL*, and skewness measures. The results support our finding that the effect of skewness on anomalies is unlikely to be explained by IVOL.

[18]We provide additional results using CPRATIO in Section 5.7. Here, we use it to differentiate between the skewness and idiosyncratic volatility explanations for our findings.

We split our sample into high-CPRATIO and low-CPRATIO subsamples using cross-sectional CPRATIO medians as breakpoints. We then compare the results of our baseline regression model (Panel A of Table 3) for each subsample.

The results in Panel B of Table 6 show that there is no significant difference in the IVOL interaction coefficients for high-CPRATIO and low-CPRATIO stocks. However, there is a 40-60% increase in the magnitude of the interaction coefficients of three of our skewness measures (i.e., *JACKPOT*, *LIDX*, and *MAXRET*), as we move from the low-CPRATIO to the high-CPRATIO subsample.[19] The relative anomaly-based return predictability of above-median skewness stocks (i.e., the sum of *MIS* and *MIS* interaction coefficients relative to the *MIS* coefficient) is between 1.7-2.1 times larger than that of below-median skewness stocks for the high-CPRATIO subsample. In comparison, for the low-CPRATIO subsample, the skewness-induced amplification is 1.4-1.7 times. Thus, unlike the IVOL effect, the skewness effect on mispricing varies geographically according to variation in the skewness preference of investors, as proxied by CPRATIO.

Overall, these robustness test results suggest that IVOL and skewness exacerbate anomalies due to different reasons. IVOL primarily deters arbitrageurs and prolongs mispricing, as demonstrated in Stambaugh et al. (2015), whereas skewness attracts speculative traders whose systematic trades amplify mispricing and strengthen the negative mispricing-return relation.

5.2 Idiosyncratic Skewness or Coskewness?

Harvey & Siddique (2000, 2022) conjecture that only a security's coskewness with the market portfolio should be priced as fully diversified investors should not care about the skewness of individual securities. In this subsection, we use coskewness measures developed in Harvey & Siddique (2000) and test whether the relation between skewness and anomaly returns reflects the effect of systematic coskewness factor rather than mispricing generated

[19] When we include both CPRATIO and *ESKEW* measures in the regression specifications, the sample size drops considerably, and the $DESKEW \times MIS$ coefficient loses its statistical power in both subsamples.

by skewness preference.

Panels A and B of Table 7 report the results of our baseline Fama-Macbeth regressions with the addition of the two coskewness measures of Harvey & Siddique (2000) as additional variables. We also consider an interaction term between coskewness and *MIS* to capture any possible effect coskewness might have on the coefficients of our skewness interaction terms.

We find that controlling for coskewness has almost no impact on the interaction coefficients in Table 3. The interaction between coskewness and *MIS* is not significant in any of our regressions and has negligible coefficients in all cases. The coskewness term does not have a statistically significant coefficient, even in Column (1) of Panel A, where none of the main variables are included in the specification.

These findings indicate that the relation between skewness and anomaly returns is unlikely to reflect the impact of coskewness. It is firm-specific skewness rather than systematic skewness that affects the predictability of anomaly strategies.

5.3 Skewness Sentiment or High Arbitrage Costs?

An important component of our skewness-based explanation for market anomalies relies on the presence of limits to arbitrage. In the absence of arbitrage risks and arbitrage costs, any skewness-related mispricing would disappear quickly as expected utility investors would reverse the pricing effect of skewness-loving investors (Barberis & Huang 2008).

Previous studies document a close link between skewness and limits to arbitrage. In particular, Conrad et al. (2014) find that their *JACKPOT* measure, which has the best performance in our tests, is strongly correlated with measures of arbitrage costs. Given these earlier findings, a potential concern with our results is that our skewness measures might indirectly reflect arbitrage costs instead of return features that trigger investor preference for skewness. In this instance, the amplifying effect of skewness on anomalies would simply indicate that stocks with high idiosyncratic skewness are more difficult to arbitrage.

We address this potential concern by adding several direct measures of arbitrage cost as control variables in our Fama-Macbeth regression specifications. We also interact the limits-to-arbitrage proxies with *MIS* and add them to the regression specifications alongside our main interaction terms. If our results reflect the impact of skewness preference of investors, our main interactions between skewness and *MIS* would not lose their economic and statistical significance when the arbitrage cost-based interaction variables are included in the regression specification.

Following previous studies, we consider five direct indicators of limits to arbitrage. This set includes the illiquidity measure of Amihud (2002), bid-ask spread (motivated by Amihud & Mendelson (1986) and Hasbrouck (2009)), frequency of zero daily returns as suggested by Lesmond et al. (1999), percentage institutional holding as in D'Avolio (2002), and the short-selling fee score (see Porras Prado et al. (2016)). Table A.1 presents the construction details for each of these measures.

Panels A to E of Table 8 report the results from these extended regression specifications. We find that the coefficients of our main interaction terms between skewness and *MIS* remain almost unchanged when the five limits-to-arbitrage proxies and their interactions with *MIS* are added to the regressions. Some of the tests use shorter sample periods leading to *MIS* coefficients having different magnitudes. Nevertheless, *MIS* coefficients are statistically significant in all cases. This finding indicates that although skewness is correlated with arbitrage cost proxies, our results do not merely reflect the known impact of arbitrage costs on anomalies.

5.4 Skewness Sentiment or Time-Varying Market Sentiment?

Next, we examine whether time-varying aggregate market-level investor sentiment during up and down market periods affects our results. These tests are motivated by the findings in Stambaugh et al. (2012, 2014) who show that anomalies are more strongly linked to future returns in periods of high market sentiment. Since we find that skewness has a similar effect in amplifying anomalies, we ensure that our results do not reflect these previous findings.

To investigate the impact of market sentiment on our results, we split our sample into months with positive market returns (up) and negative market returns (down). The sorting is performed using market returns from the same month at the end of which our skewness measures are calculated. We use the CRSP value-weighted index return as our measure of the market return. We estimate the baseline regression model (Panel A of Table 3) for the up and down subsamples separately. The results are presented in Panel A of Table 9. In all but one case, we find that the interaction coefficients are virtually unaffected. The exception is the LIDX interaction coefficient, which is larger in magnitude during market downturns.

For additional robustness, we use another measure to determine whether the effect of skewness on anomalies varies over time with investor sentiment. We obtain the sentiment indices of Baker & Wurgler (2006) from Professor Jeffrey Wurgler's website.[20] We use both the original sentiment index of Baker & Wurgler (2006) based on the first principal component of five sentiment proxies (*SENT*) and their alternative index based on the first principal component of five sentiment proxies that are first orthogonalized with respect to a set of macroeconomic indicators ($SENT^{\perp}$).

We divide the sample period into high and low sentiment periods using the time-series medians of the two sentiment indices separately. We use the lagged sentiment scores for each month so that a month is categorized as a low (high) sentiment month if the sentiment index during the previous month is below (above) the time-series median. For each sub-period, we estimate the baseline regression model (Panel A of Table 3).

Panels B and C of Table 9 present the results based on the two sentiment indices. These results are almost indistinguishable across the two panels, suggesting that *SENT* and $SENT^{\perp}$ are highly correlated. The estimates in the High−Low column indicate that the interaction coefficients are significantly more negative in high sentiment periods. A one-standard-deviation increase in skewness adds between 0.1-0.2% more to anomaly returns during periods of high sentiment compared to periods of low sentiment. These differences

[20] See http://people.stern.nyu.edu/jwurgler/.

are statistically significant, except for the $ESKEW \times MIS$ coefficients. This insignificant estimate is likely because of the shorter time period. $ESKEW$ data start in 1988, unlike other skewness variables which begin in 1963. While the difference between the interaction coefficients of the two subsamples is small, all coefficients are statistically and economically significant even in low sentiment periods.

Overall, market sentiment-based subsample estimates indicate that the skewness-anomalies relation is stronger during periods of high market sentiment but this relation is not affected by market conditions. This evidence is consistent with the findings in Stambaugh et al. (2012) where a high market sentiment period is associated with a higher level of noise trading, which exacerbates mispricing.

5.5 Skewness Sentiment or Growth Options?

Bali, Del Viva, Lambertides & Trigeorgis (2020) conjecture that certain anomalies such as profitability and distress are likely to arise because growth options or firm rigidity risk is not accounted for in asset pricing models. They posit that idiosyncratic skewness can capture this effect and show that an idiosyncratic skewness factor can explain a number of anomalies. In contrast to Bali, Del Viva, Lambertides & Trigeorgis (2020), we adopt a different perspective and show directly that idiosyncratic skewness can explain the link between anomalies and mispricing. Specifically, we use retail investor portfolio holdings and geographic proxies for gambling propensity to show that investor preference for positively skewed payoffs generates mispricing associated with market anomalies.

Further, in our study, idiosyncratic skewness does not explain any particular anomaly completely. Bali, Del Viva, Lambertides & Trigeorgis (2020) suggest that inflexibility risk explains a number of anomalies. In contrast, we demonstrate that investor preference for skewness directly contributes to a wide range of anomalies because overpriced stocks in the cross section have high skewness. While both explanations are possible, our local gambling propensity results reported in Subsection 5.7 suggest that investor preference for skewness is a more likely explanation.

5.6 Skewness Sentiment or Overreaction to News?

Another possible explanation for our findings is investor overreaction to news (e.g., Barberis et al. 1998, Daniel et al. 1998). Stocks with high levels of skewness are likely to be those with some form of recent good news. Therefore, it is likely that good news is associated with right-tail returns. If investors overreact to such news, such overreaction could be an alternative explanation for our results. To address this issue, we disentangle the skewness sentiment explanation from the overreaction effect by controlling for the effect of investor overreaction.

Motivated by Barberis et al. (1998) and Daniel et al. (1998), we conjecture that stocks facing higher investor overreaction are likely to be those with relatively higher levels of turnover (Byun et al. 2016) or those having a recent earnings announcement. To examine these two alternative explanations for our findings, we divide the sample into two parts based on stock turnover and the presence of a recent earnings announcement and re-estimate our baseline regressions for each subsample. The results presented in Internet Appendix Table A.4 show that our main findings are robust even after excluding the set of stocks that experience high investor overreaction.

5.7 Evidence from Local Gambling Propensity

The results from various robustness checks suggest that our main findings are consistent with our skewness preference-based conjectures rather than other potential explanations that may generate similar results. To further rule out other possible explanations for our findings, we use a religion-based proxy for skewness preference. We replace our skewness measures with an exogenous variable that is unlikely to capture any financial effect other than a preference for skewness. Specifically, we investigate whether our results vary geographically with the ratio of Catholics to Protestants (CPRATIO) in firms' headquarter counties.

Kumar et al. (2011, 2016) show that CPRATIO is a proxy for local preference for skewness. There is no reason to believe that other factors such as arbitrage costs, overreaction to news, or other confounds vary with CPRATIO. Therefore, if we find that our results are stronger for firms headquartered in counties with high CPRATIOs, we can conclude with greater confidence that investors' skewness preference is the main driver of our empirical findings.

Similar to the previous robustness checks, each month, we divide all stocks into two subsamples using the lagged monthly CPRATIO. We estimate the baseline regression models (see Table 3, Panel A) for each subsample. The estimation results are presented in Table 10.

We find that, consistent with our main conjecture, the interaction terms are all larger and more statistically signification for high-CPRATIO stocks. The differences in the High−Low column are also all statistically significant. A one-standard-deviation increase in skewness increases the anomaly returns of high-CPRATIO stocks between 0.1-0.3% more than low-CPRATIO stocks. Interestingly, *MIS* coefficient estimates are almost identical in the two subsamples. This is in line with our prediction that the effect of investor preference for skewness on mispricing is not uniform but depends on stock-level skewness.

These CPRATIO-based subsample results show that the skewness amplifies anomalies mainly among stocks headquartered in high-CPRATIO regions. It is hard to find another explanation for this finding other than our conjecture that systematic demand shifts of skewness-loving investors, who are likely to be concentrated in high CPRATIO regions, amplify various market anomalies. This evidence rules out a large number of alternative explanations for our findings and further confirms the role of skewness-induced systematic demand shifts in generating mispricing.

6 Summary and Conclusions

This study examines whether investor preference for skewness is a common driver of cross-sectional mispricing patterns identified by various anomaly strategies. Using a composite mispricing measure based on 11 strategies, we demonstrate that anomalies are significantly stronger among stocks with higher skewness. We find consistent results across a wide range of skewness measures used in the literature. Skewness interacts with anomalies and amplifies them by making stocks in the Short portfolio more overpriced. The returns of stocks in the Long portfolio do not vary significantly with skewness.

We attribute the effect of skewness on anomalies to the proclivity of a group of investors to hold positively skewed positions. Portfolio holdings from a large U.S. retail brokerage house suggest that investors with a history of holding positively skewed positions are considerably more likely to overweight stocks that anomaly strategies predict will underperform relative to those that will outperform. Investors who overweight underperforming stocks relative to outperforming ones also possess characteristics that have been previously linked to investor sophistication and preference for skewness.

We rule out a range of alternative explanations for the relation between skewness and anomalies. In particular, we consider explanations based on IVOL, coskewness, limits to arbitrage, short-selling costs, investor overreaction, and time-varying market sentiment. We find that none of these measures can absorb the explanatory power of skewness.

We also show that skewness exacerbates anomalies more strongly for firms headquartered in geographical regions where investors are known to exhibit a higher propensity to gamble, as measured by the ratio of Catholics to Protestants (CPRATIO). This evidence further rules out alternative explanations for our results as there is no obvious reason why other confounding variables would vary with CPRATIO.

Our results do not fully explain various market anomalies. Numerous underlying mechanisms, though not all related to mispricing, are likely to drive each individual anomaly. We demonstrate mispricing-related commonalities across a range of strategies and show that the preference for skewness plays an important role. In this sense, our work is related to papers that look for common drivers of anomalies. Stambaugh et al. (2012), for example, highlights the role of investor sentiment. While investor sentiment can explain time-series variation in the performance of anomalies, our paper explains performance variation in the cross-section. In particular, we provide one economic mechanism through which some stocks in the Short legs of anomaly portfolios could be more overpriced.

References

Agrawal, V., Jiang, L. & Wen, Q. (2022), 'Why do mutual funds hold lottery stocks?', *Journal of Financial and Quantitative Analysis* **57**(3), 825–856.

Amihud, Y. (2002), 'Illiquidity and stock returns: Cross-section and time-series effects', *Journal of Financial Markets* **5**(1), 31–56.

Amihud, Y. & Mendelson, H. (1986), 'Asset pricing and the bid-ask spread', *Journal of Financial Economics* **17**(2), 223–249.

An, L., Wang, H., Wang, J. & Yu, J. (2020), 'Lottery-related anomalies: The role of reference-dependent preferences', *Management Science* **66**(1), 473–501.

Avramov, D., Chordia, T., Jostova, G. & Philipov, A. (2013), 'Anomalies and financial distress', *Journal of Financial Economics* **108**(1), 139–159.

Baker, M. & Wurgler, J. (2006), 'Investor sentiment and the cross-section of stock returns', *Journal of Finance* **61**(4), 1645–1680.

Bakshi, G., Kapadia, N. & Madan, D. (2003), 'Stock return characteristics, skew laws, and the differential pricing of individual equity options', *Review of Financial Studies* **16**(1), 101–143.

Bali, T. G., Cakici, N. & Whitelaw, R. F. (2011), 'Maxing out: Stocks as lotteries and the cross-section of expected returns', *Journal of Financial Economics* **99**(2), 427–446.

Bali, T. G., Del Viva, L., Lambertides, N. & Trigeorgis, L. (2020), 'Growth options and related stock market anomalies: Profitability, distress, lotteryness, and volatility', *Journal of Financial and Quantitative Analysis* **55**(7), 2150–2180.

Bali, T. G., Gunaydin, A. D., Jansson, T. & Karabulut, Y. (2020), 'Do the rich gamble in the stock market? Low risk anomalies and wealthy households,', **Available at SSRN: https://ssrn.com/abstract=3664501**.

Barberis, N. (2013), 'The psychology of tail events: Progress and challenges', *American Economic Review* **103**(3), 611–616.

Barberis, N. & Huang, M. (2008), 'Stocks as lotteries: The implications of probability weighting for security prices', *American Economic Review* **98**(5), 2066–2100.

Barberis, N., Mukherjee, A. & Wang, B. (2016), 'Prospect theory and stock returns: An empirical test', *Review of Financial Studies* **29**(11), 3068–3107.

Barberis, N., Shleifer, A. & Vishny, R. (1998), 'A model of investor sentiment', *Journal of Financial Economics* **49**(3), 307–343.

Boyer, B. H. & Vorkink, K. (2014), 'Stock options as lotteries', *Journal of Finance* **69**(4), 1485–1527.

Boyer, B., Mitton, T. & Vorkink, K. (2010), 'Expected idiosyncratic skewness', *Review of Financial Studies* **23**(1), 169–202.

Bris, A., Goetzmann, W. N. & Zhu, N. (2007), 'Efficiency and the bear: Short sales and markets around the world', *Journal of Finance* **62**(3), 1029–1079.

Brunnermeier, M. K., Gollier, C. & Parker, J. A. (2007), 'Optimal beliefs, asset prices, and the preference for skewed returns', *American Economic Review* **97**(2), 159–165.

Byun, S. J., Lim, S. S. & Yun, S. H. (2016), 'Continuing overreaction and stock return predictability', *Journal of Financial and Quantitative Analysis* **51**(6), 2015–2046.

Campbell, J. Y., Hilscher, J. & Szilagyi, J. A. N. (2008), 'In search of distress risk', *Journal of Finance* **63**(6), 2899–2939.

Cao, H. H., Coval, J. D. & Hirshleifer, D. (2002), 'Sidelined investors, trading-generated news, and security returns', *Review of Financial Studies* **15**(2), 615–648.

Carhart, M. M. (1997), 'On persistence in mutual fund performance', *Journal of Finance* **52**(1), 57–82.

Chang, E. C., Cheng, J. W. & Yu, Y. (2007), 'Short-sales constraints and price discovery: Evidence from the hong kong market', *Journal of Finance* **62**(5), 2097–2121.

Chen, J., Hong, H. & Stein, J. C. (2001), 'Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices', *Journal of Financial Economics* **61**(3), 345–381.

Chordia, T., Subrahmanyam, A. & Tong, Q. (2014), 'Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?', *Journal of Accounting and Economics* **58**(1), 41–58.

Conrad, J., Dittmar, R. F. & Ghysels, E. (2013), 'Ex ante skewness and expected stock returns', *Journal of Finance* **68**(1), 85–124.

Conrad, J., Kapadia, N. & Xing, Y. (2014), 'Death and jackpot: Why do individual investors hold overpriced stocks?', *Journal of Financial Economics* **113**(3), 455–475.

Cooper, M. J., Gulen, H. & Schill, M. J. (2008), 'Asset growth and the cross-section of stock returns', *Journal of Finance* **63**(4), 1609–1651.

Daniel, K., Hirshleifer, D. & Subrahmanyam, A. (1998), 'Investor psychology and security market under-and overreactions', *the Journal of Finance* **53**(6), 1839–1885.

Daniel, K. & Titman, S. (2006), 'Market reactions to tangible and intangible information', *Journal of Finance* **61**(4), 1605–1643.

Del Viva, L., Kasanen, E. & Trigeorgis, L. (2017), 'Real options, idiosyncratic skewness, and diversification', *Journal of Financial and Quantitative Analysis* **52**(1), 215–241.

Dittmar, R. F. (2002), 'Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns', *Journal of Finance* **57**(1), 369–403.

D'Avolio, G. (2002), 'The market for borrowing stock', *Journal of Financial Economics* **66**(2–3), 271–306.

Fama, E. F. & French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**(1), 3–56.

Fama, E. F. & French, K. R. (2006), 'Profitability, investment and average returns', *Journal of Financial Economics* **82**(3), 491–518.

Fama, E. F. & French, K. R. (2008), 'Dissecting anomalies', *Journal of Finance* **63**(4), 1653–1678.

Fama, E. F. & French, K. R. (2015), 'A five-factor asset pricing model', *Journal of Financial Economics* **116**(1), 1–22.

Fama, E. F. & MacBeth, J. D. (1973), 'Risk, return, and equilibrium: Empirical tests', *Journal of Political Economy* pp. 607–636.

Gibbons, M. R., Ross, S. A. & Shanken, J. (1989), 'A test of the efficiency of a given portfolio', *Econometrica: Journal of the Econometric Society* pp. 1121–1152.

Goetzmann, W. N. & Kumar, A. (2008), 'Equity portfolio diversification', *Review of Finance* **12**(3), 433–463.

Green, T. C. & Hwang, B.-H. (2012), 'Initial public offerings as lotteries: Skewness preference and first-day returns', *Management Science* **58**(2), 432–444.

Han, X., Roussanov, N. L. & Ruan, H. (2021), 'Mutual fund risk shifting and risk anomalies', *Available at SSRN 3931449* .

Hanson, S. G. & Sunderam, A. (2014), 'The growth and limits of arbitrage: Evidence from short interest', *Review of Financial Studies* **27**(4), 1238–1286.

Harvey, C. R., Liechty, J. C., Liechty, M. W. & Müller, P. (2010), 'Portfolio selection with higher moments', *Quantitative Finance* **10**(5), 469–485.

Harvey, C. R., Liu, Y. & Zhu, H. (2016), '... and the cross-section of expected returns', *Review of Financial Studies* **29**(1), 5–68.

Harvey, C. R. & Siddique, A. (2000), 'Conditional skewness in asset pricing tests', *Journal of Finance* **55**(3), 1263–1295.

Harvey, C. R. & Siddique, A. (2022), 'Conditional skewness in asset pricing: 25 years of out-of-sample evidence', *Available at SSRN: https://ssrn.com/abstract=4085027* .

Hasbrouck, J. (2009), 'Trading costs and returns for u.s. equities: Estimating effective costs from daily data', *Journal of Finance* **64**(3), 1445–1477.

Hirshleifer, D., Hou, K., Teoh, S. H. & Zhang, Y. (2004), 'Do investors overvalue firms with bloated balance sheets?', *Journal of Accounting and Economics* **38**, 297–331.

Hirshleifer, D., Teoh, S. H. & Yu, J. J. (2011), 'Short arbitrage, return asymmetry, and the accrual anomaly', *Review of Financial Studies* **24**(7), 2429–2461.

Ivković, Z. & Weisbenner, S. (2005), 'Local does as local is: Information content of the geography of individual investors' common stock investments', *Journal of Finance* **60**(1), 267–306.

Jegadeesh, N. & Titman, S. (1993), 'Returns to buying winners and selling losers: Implications for stock market efficiency', *Journal of Finance* **48**(1), 65–91.

Jiang, L., Wen, Q., Zhou, G. & Zhu, Y. (2020), 'Lottery preference and anomalies', *Available at SSRN 3595419* .

Korniotis, G. M. & Kumar, A. (2013), 'Do portfolio distortions reflect superior information or psychological biases?', *Journal of Financial and Quantitative Analysis* **48**(01), 1–45.

Kraus, A. & Litzenberger, R. H. (1976), 'Skewness preference and the valuation of risk assets', *Journal of Finance* **31**(4), 1085–1100.

Kumar, A. (2009), 'Who gambles in the stock market?', *Journal of Finance* **64**(4), 1889–1933.

Kumar, A., Page, J. K. & Spalt, O. G. (2011), 'Religious beliefs, gambling attitudes, and financial market outcomes', *Journal of Financial Economics* **102**(3), 671–708.

Kumar, A., Page, J. K. & Spalt, O. G. (2016), 'Gambling and comovement', *Journal of Financial and Quantitative Analysis* **51**(01), 85–111.

Lesmond, D. A., Ogden, J. P. & Trzcinka, C. A. (1999), 'A new estimate of transaction costs', *Review of Financial Studies* **12**(5), 1113–1141.

Loughran, T. & Ritter, J. R. (1995), 'The new issues puzzle', *Journal of Finance* **50**(1), 23–51.

McLean, R. D. & Pontiff, J. (2016), 'Does academic research destroy stock return predictability?', *Journal of Finance* **71**(1), 5–32.

Miller, E. M. (1977), 'Risk, uncertainty, and divergence of opinion', *Journal of Finance* **32**(4), 1151–1168.

Mitton, T. & Vorkink, K. (2007), 'Equilibrium underdiversification and the preference for skewness', *Review of Financial Studies* **20**(4), 1255–1288.

Nagel, S. (2005), 'Short sales, institutional investors and the cross-section of stock returns', *Journal of Financial Economics* **78**(2), 277–309.

Newey, W. K. & West, K. D. (1987), 'A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica* **55**(3), 703–708.
URL: *http://www.jstor.org/stable/1913610*

Newey, W. K. & West, K. D. (1994), 'Automatic lag selection in covariance matrix estimation', *The Review of Economic Studies* **61**(4), 631–653.

Novy-Marx, R. (2013), 'The other side of value: The gross profitability premium', *Journal of Financial Economics* **108**(1), 1–28.

Ohlson, J. A. (1980), 'Financial ratios and the probabilistic prediction of bankruptcy', *Journal of Accounting Research* **18**(1), 109–131.

Porras Prado, M., Saffi, P. A. & Sturgess, J. (2016), 'Ownership structure, limits to arbitrage, and stock returns: Evidence from equity lending markets', *The Review of Financial Studies* **29**(12), 3211–3244.

Ritter, J. R. (1991), 'The long-run performance of initial public offerings', *Journal of Finance* **46**(1), 3–27.

Sloan, R. G. (1996), 'Do stock prices fully reflect information in accruals and cash flows about future earnings?', *Accounting Review* **71**(3), 289–315.

Stambaugh, R. F., Yu, J. & Yuan, Y. (2012), 'The short of it: Investor sentiment and anomalies', *Journal of Financial Economics* **104**(2), 288–302.

Stambaugh, R. F., Yu, J. & Yuan, Y. (2014), 'The long of it: Odds that investor sentiment spuriously predicts anomaly returns', *Journal of Financial Economics* **114**(3), 613–619.

Stambaugh, R. F., Yu, J. & Yuan, Y. (2015), 'Arbitrage asymmetry and the idiosyncratic volatility puzzle', *Journal of Finance* **70**(5), 1903–1948.

Stambaugh, R. F. & Yuan, Y. (2016), 'Mispricing factors', *Review of Financial Studies* **30**(4), 1270–1315.

Titman, S., Wei, K. C. J. & Xie, F. (2004), 'Capital investments and stock returns', *Journal of Financial and Quantitative Analysis* **39**(04), 677–700.

Tversky, A. & Kahneman, D. (1992), 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and Uncertainty* **5**(4), 297–323.

Xu, J. (2007), 'Price convexity and skewness', *Journal of Finance* **62**(5), 2521–2552.

Table 1: Summary Statistics

This table reports the average characteristics of MIS quintiles in Panels A and B and the monthly value-weighted abnormal returns of quintiles based on MIS and the four skewness measures of JACKPOT, LIDX, MAXRET, and ESKEW in Panel C. MIS is a combined measure of mispricing based on 11 prominent anomaly strategies, following Stambaugh et al. (2015). Higher (lower) values of MIS indicate a higher likelihood for the stock to be overpriced (underpriced). Table A.1 defines MIS and all other variables. Quintile portfolios are formed by sorting stocks into five groups at the end of every month. The t-statistics for the difference between the values of quintiles 1 and 5 (5 - 1) in Panels A and B are based on Newey-West heteroscedasticity- and autocorrelation-consistent standard errors using a lag of 6. The three-, four-, and five-factor models used to adjust returns in Panel C correspond to the models of Fama & French (1993), Carhart (1997), and Fama & French (2015), respectively. The sample excludes penny stocks and covers January 1963 to December 2015, except for the sorts based on ESKEW, which start in January 1988. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Full Sample	MIS Portfolios					
		1	2	3	4	5	5 - 1
Panel A: Key Statistics of MIS Portfolios							
ME ($ Billion)	1.56	3.50	1.88	1.14	0.78	0.51	-2.99 (-3.61)
PRICE ($)	21.16	29.17	24.4	20.43	17.81	13.99	-15.18 (-8.04)
VOLATILITY (%)	3.05	2.50	2.78	3.03	3.27	3.68	1.19 (7.83)
IVOL (%)	2.95	2.37	2.67	2.95	3.19	3.60	1.23 (6.98)
RET[-12,-2] (%)	3.13	18.41	10.67	4.08	-3.07	-14.45	-32.86 (-11.13)
TURNOVER (%)	9.39	8.87	8.83	8.98	9.67	10.63	1.76 (3.54)
SHORTRATIO	2.14	1.82	1.96	2.12	2.41	2.80	0.98 (5.38)
ILLIQ (10^{-6})	4.28	2.91	4.22	5.10	4.92	4.26	1.35 (7.59)
LEVERAGE	0.24	0.13	0.19	0.25	0.29	0.33	0.20 (15.52)
B/M	0.84	0.67	0.81	0.90	0.92	0.90	0.23 (5.66)
RHOLDING (%)	0.10	0.08	0.09	0.10	0.11	0.11	0.03 (4.37)
IHOLDING (%)	30.71	36.89	34.00	32.04	28.02	22.61	-14.28 (-7.93)
Panel B: Skewness Characteristics of MIS Portfolios							
ESKEW	0.78	0.62	0.70	0.77	0.84	0.96	0.33 (5.48)
JACKPOT (%)	2.00	1.30	1.63	1.98	2.24	2.82	1.52 (4.36)
LIDX	0.49	0.42	0.45	0.49	0.51	0.57	0.15 (8.17)
MAXRET (%)	6.87	5.49	6.18	6.83	7.42	8.45	2.97 (7.23)
OS	-0.28	-0.32	-0.28	-0.25	-0.24	-0.21	0.11 (6.86)
ISKEWNESS	0.18	0.17	0.18	0.18	0.19	0.20	0.03 (2.24)
SKEWNESS	0.25	0.23	0.24	0.25	0.26	0.28	0.05 (2.55)

Table 1 (Continued)

Panel C: Abnormal Returns of MIS and Skewness Measures							
Variable	Model	1	2	3	4	5	5 - 1
MIS	3-Factor	0.29***	0.07*	-0.07	-0.22***	-0.80***	-1.09***
		(6.66)	(1.77)	(-1.24)	(-3.57)	(-8.30)	(-8.90)
	4-Factor	0.20***	0.08*	-0.05	-0.11*	-0.56***	-0.76***
		(4.87)	(1.76)	(-0.82)	(-1.78)	(-6.31)	(-6.92)
	5-Factor	0.18***	0.05	0.00	-0.08	-0.45***	-0.63***
		(4.46)	(1.20)	(0.08)	(-1.26)	(-5.32)	(-5.99)
JACKPOT	3-Factor	0.08***	0.02	-0.09	-0.51***	-0.97***	-1.05***
		(3.47)	(0.29)	(-1.24)	(-4.23)	(-5.39)	(-5.50)
	4-Factor	0.05**	0.07	-0.03	-0.34***	-0.65***	-0.70***
		(2.16)	(1.31)	(-0.35)	(-2.82)	(-3.73)	(-3.80)
	5-Factor	0.01	0.17***	0.15**	-0.08	-0.35**	-0.37**
		(0.62)	(3.17)	(2.36)	(-0.73)	(-2.29)	(-2.27)
LIDX	3-Factor	0.11***	-0.02	-0.09	-0.32***	-0.99***	-1.09***
		(4.13)	(-0.32)	(-1.17)	(-3.13)	(-6.45)	(-6.77)
	4-Factor	0.08***	0.00	0.01	-0.14	-0.66***	-0.74***
		(2.97)	(0.08)	(0.07)	(-1.36)	(-4.55)	(-4.84)
	5-Factor	0.07***	0.07	0.14**	0.00	-0.57***	-0.64***
		(3.07)	(1.34)	(1.99)	(-0.02)	(-4.09)	(-4.46)
MAXRET	3-Factor	0.10**	0.03	0.08	-0.10	-0.61***	-0.70***
		(2.03)	(0.60)	(1.38)	(-1.07)	(-4.92)	(-4.68)
	4-Factor	0.08	0.06	0.12**	-0.04	-0.47***	-0.55***
		(1.56)	(1.09)	(1.97)	(-0.44)	(-3.81)	(-3.62)
	5-Factor	0.01	0.06	0.18***	0.12	-0.24**	-0.25*
		(0.20)	(1.29)	(2.84)	(1.35)	(-2.17)	(-1.86)
ESKEW	3-Factor	0.10**	0.11*	-0.09	-0.25**	-0.62***	-0.72***
		(2.55)	(1.73)	(-0.83)	(-2.10)	(-4.21)	(-4.46)
	4-Factor	0.08**	0.10	-0.01	-0.13	-0.44***	-0.52***
		(2.06)	(1.52)	(-0.09)	(-1.07)	(-2.96)	(-3.22)
	5-Factor	0.09**	0.14**	0.10	0.00	-0.25*	-0.34**
		(2.30)	(2.29)	(0.88)	(-0.02)	(-1.87)	(-2.32)

Table 2: Sorting Results

Panel A reports benchmark adjusted returns for double-sorted portfolios based on MIS and one of the four skewness measures of JACKPOT, LIDX, MAXRET, and ESKEW. Table A.1 defines all the variables. The portfolios are formed by independently sorting stocks into five portfolios at the end of every month with respect to each variable. We then compute the value-weighted returns of the 25 intersecting portfolios for the following month and regress the time series of returns on the four factors of Carhart (1997). The regression intercept is the abnormal return estimate reported in the table. Panel B presents the average number of stocks in each portfolio. Standard errors are adjusted following the Newey & West (1987) approach using a lag of 6. The sample excludes penny stocks and covers January 1963 to December 2015, except for sorts based on ESKEW, which start in January 1988. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

		Most Underpriced	2	3	4	Most Overpriced	Most Overpriced - Most Underpriced
JACKPOT	Low	0.19***	0.02	-0.11*	-0.11	-0.16	-0.35***
		(4.14)	(0.54)	(-1.82)	(-1.41)	(-1.14)	(-2.61)
	2	0.37***	0.37***	0.17*	-0.19	-0.68***	-1.05***
		(4.22)	(4.08)	(1.87)	(-1.51)	(-4.50)	(-5.75)
	3	0.42***	0.48***	0.08	-0.09	-0.92***	-1.34***
		(3.51)	(3.91)	(0.63)	(-0.77)	(-6.28)	(-7.40)
	4	0.72***	0.35**	0.07	-0.10	-1.11***	-1.83***
		(4.46)	(2.09)	(0.53)	(-0.69)	(-6.36)	(-8.04)
	High	0.50**	0.45**	-0.21	-0.29	-1.55***	-2.06***
		(2.23)	(2.22)	(-1.07)	(-1.63)	(-7.12)	(-7.62)
	High -	0.31	0.43**	-0.09	-0.18	-1.40***	-1.71***
	Low	(1.29)	(2.06)	(-0.56)	(-0.98)	(-5.45)	(-6.26)
LIDX	Low	0.22***	0.06	-0.04	-0.06	-0.35***	-0.56***
		(4.74)	(1.32)	(-0.44)	(-0.44)	(-2.60)	(-4.14)
	2	0.21**	0.05	-0.01	-0.17*	-0.34***	-0.55***
		(2.18)	(0.43)	(0.18)	(-1.94)	(-2.72)	(-3.46)
	3	0.51***	0.35***	0.13	-0.2	-0.75***	-1.27***
		(4.34)	(3.38)	(1.01)	(-1.34)	(-5.15)	(-6.51)
	4	0.43***	0.52***	0.05	-0.04	-1.06***	-1.5***
		(3.04)	(3.33)	(0.34)	(-0.42)	(-5.59)	(-6.30)
	High	0.47**	0.15	-0.23	-0.40	-1.43***	-1.90***
		(2.05)	(0.61)	(-1.10)	(-1.63)	(-6.19)	(-7.41)
	High -	0.25	0.09	-0.18	-0.34	-1.08***	-1.37***
	Low	(1.04)	(0.31)	(-0.93)	(-1.30)	(-4.36)	(-5.21)
MAXRET	Low	0.20**	0.10	0.00	0.04	-0.15	-0.34**
		(2.57)	(1.40)	(0.05)	(0.63)	(-1.03)	(-2.38)
	2	0.21***	0.14*	0.06	-0.11	-0.34***	-0.54***
		(2.95)	(1.80)	(0.82)	(-0.94)	(-2.68)	(-3.79)
	3	0.49***	0.14	-0.05	-0.04	-0.62***	-1.11***
		(4.85)	(1.25)	(-0.25)	(-0.13)	(-4.21)	(-5.98)
	4	0.54***	0.14	0.03	-0.34**	-0.80***	-1.34***
		(3.27)	(0.92)	(0.19)	(-2.41)	(-5.03)	(-5.91)
	High	0.20	-0.08	-0.37**	-0.39**	-1.51***	-1.70***
		(0.99)	(-0.40)	(-2.46)	(-2.22)	(-7.16)	(-6.80)
	High -	0.00	-0.17	-0.37**	-0.43**	-1.37***	-1.37***
	Low	(0.03)	(-0.80)	(-2.12)	(-2.23)	(-5.00)	(-4.91)

Table 2 (Continued)

		Most Underpriced	2	3	4	Most Overpriced	Most Overpriced - Most Underpriced
ESKEW	Low	0.31***	0.03	-0.05	-0.09	-0.32**	-0.62***
		(4.65)	(0.60)	(-0.46)	(-0.85)	(-2.32)	(-4.31)
	2	0.22**	0.15*	0.11	-0.01	-0.51***	-0.73***
		(2.45)	(1.64)	(1.17)	(-0.94)	(-3.72)	(-4.77)
	3	0.32**	0.18	0.09	-0.11	-0.86***	-1.18***
		(2.34)	(1.27)	(0.75)	(-0.73)	(-4.51)	(-5.10)
	4	0.57***	0.35*	-0.10	-0.06	-1.00***	-1.58***
		(3.55)	(1.75)	(-0.51)	(-0.28)	(-6.01)	(-8.92)
	High	0.58***	0.35*	0.22	-0.24	-1.30***	-1.88***
		(2.86)	(1.85)	(1.12)	(-1.31)	(-5.02)	(-6.62)
	High -	0.28	0.32	0.27	-0.16	-0.98***	-1.22***
	Low	(1.19)	(1.55)	(1.22)	(-0.80)	(-3.55)	(-4.39)

Panel B: Number of Stocks

		Most Underpriced	2	3	4	Most Overpriced
JACKPOT	Low	209	172	135	106	68
	2	154	144	132	118	86
	3	112	116	122	124	119
	4	92	106	120	134	157
	High	65	95	124	150	202
LIDX	Low	196	163	131	106	66
	2	154	141	130	120	94
	3	121	123	123	125	123
	4	89	105	120	133	158
	High	63	93	120	140	182
MAXRET	Low	176	146	124	104	78
	2	156	143	130	117	96
	3	126	130	129	128	120
	4	97	112	124	135	150
	High	69	94	117	139	181
ESKEW	Low	190	160	133	115	83
	2	168	151	135	118	90
	3	121	126	129	132	132
	4	102	114	124	135	153
	High	74	104	128	146	181

Table 3: Baseline Fama-Macbeth Regression Estimates

This table presents estimates from the monthly Fama-MacBeth cross-sectional regressions. At the end of each month t, we use a set of independent variables including stock characteristics and our skewness and mispricing measures to predict the stock returns for month $t+1$. Our primary independent variable is the interaction between each of the four skewness measures of JACKPOT, LIDX, MAXRET, and ESKEW and the combined anomaly variable, MIS. Table A.1 defines all the variables. All independent variables in our regressions are standardized to have a mean of zero and a standard deviation of 1 and are winsorized at the 0.5 and 99.5 percentiles. Standard errors are adjusted following the Newey & West (1987) approach using a lag of 6. Panel A reports the baseline regression results, and Panel B presents the results based on alternative samples or data filters. For brevity, we only report the interaction coefficients in Panel B. The sample excludes penny stocks and covers January 1963 to December 2015, except for the regression that includes ESKEW, which starts in January 1988. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Baseline Estimates									
Intercept	0.010***	0.010***	0.010***	0.010***	0.010***	0.009***	0.010***	0.010***	0.009***
	(4.33)	(3.98)	(4.35)	(4.11)	(3.89)	(3.83)	(4.29)	(4.11)	(3.60)
MIS	-0.005***					-0.005***	-0.004***	-0.004***	-0.005***
	(-11.72)					(-12.37)	(-13.69)	(-11.74)	(-12.25)
JACKPOT		-0.004***				-0.001			
		(-4.16)				(-1.56)			
LIDX			-0.002**				0.000		
			(-2.06)				(-0.44)		
MAXRET				-0.003***				-0.002***	
				(-5.14)				(-3.39)	
ESKEW					-0.002***				-0.002**
					(-2.60)				(-2.03)
MIS × JACKPOT						-0.003***			
						(-6.12)			
MIS × LIDX							-0.002***		
							(-7.53)		
MIS × MAXRET								-0.002***	
								(-7.60)	
MIS × ESKEW									-0.001***
									(-3.95)
log(ME)	-0.002**	-0.002***	-0.002***	-0.002**	-0.001*	-0.002***	-0.002***	-0.002***	-0.002**
	‘(-2.26)	(-2.67)	(-3.50)	(-2.55)	(-1.72)	(-3.18)	(-3.12)	(-3.18)	(-2.44)
log(B/M)	0.003***	0.004***	0.004***	0.003***	0.004***	0.003***	0.003***	0.003***	0.003***
	(5.29)	(6.01)	(5.98)	(5.83)	(6.14)	(5.06)	(4.95)	(4.96)	(5.14)
RET[-12,-2]	0.004***	0.006***	0.006***	0.006***	0.005***	0.004***	0.004***	0.004***	0.003***
	(4.56)	(5.89)	(6.56)	(6.25)	(5.35)	(4.07)	(4.54)	(4.50)	(3.47)
RET[-1,0]	-0.007***	-0.007***	-0.007***	-0.006***	-0.007***	-0.007***	-0.007***	-0.006***	-0.007***
	(-11.47)	(-11.01)	(-11.01)	(-8.06)	(-10.06)	(-11.86)	(-12.02)	(-9.39)	(-11.09)
Average Number of Observations	3,033	3,146	3,082	3,083	3,200	3,072	3,032	3,033	3,153
Average Adjusted R^2	0.042	0.042	0.046	0.045	0.039	0.047	0.048	0.047	0.044

Table 3 (Continued)

Panel B: Robustness Tests								
Test	MIS × JACKPOT	Avg N	MIS × LIDX	Avg N	MIS × MAXRET	Avg N	MIS × ESKEW	Avg N
Baseline	-0.003***	3,072	-0.002***	3,032	-0.002***	3,033	-0.001***	3,153
	(-6.12)		(-7.53)		(-7.60)		(-3.95)	
Basic Robustness Checks								
(1) No Winsorization	-0.002***	3,072	-0.002***	3,032	-0.002***	3,033	-0.001***	3,153
	(-4.37)		(-7.66)		(-6.85)		(-3.91)	
(2) Excludes Price ≤ 5	-0.003***	2,385	-0.002***	2,355	-0.002***	2,356	-0.002***	2,433
	(-8.14)		(-8.33)		(-7.37)		(-5.82)	
(3) Excludes Micro-Cap Stocks	-0.009***	1,351	-0.002***	1,335	-0.002***	1,336	-0.002***	1,385
	(-5.49)		(-6.92)		(-6.04)		(-4.55)	
(4) Excludes Mega-Cap Stocks	-0.003***	2,811	-0.002***	2,773	-0.002***	2,775	-0.001***	2,889
	(-6.00)		(-8.03)		(-7.41)		(-4.12)	
(5) Excludes NASDAQ Stocks	-0.002***	1,549	-0.002***	1,531	-0.001***	1,532	-0.001***	1,582
	(-4.81)		(-6.71)		(-4.95)		(-4.22)	
Subperiods								
(6) Recession Periods	-0.005***	1,731	-0.002**	1,730	-0.003***	1,731	-0.001	1,647
	(-2.67)		(-2.56)		(-3.83)		(-1.50)	
(7) Expansion Periods	-0.003***	2,782	-0.001***	2,745	-0.002***	2,746	-0.001***	2,858
	(-5.39)		(-6.70)		(-6.50)		(-2.96)	
(8) 1962–1990	-0.001***	2,658	-0.001***	2594	-0.001***	2,596	—	
	(-3.18)		(-3.01)		(-3.35)			
(10) 1991–2015	-0.004***	3,526	-0.003***	3,526	-0.003***	3,526	-0.004***	3,456
	(-5.79)		(-7.53)		(-8.38)		(-4.55)	

Table 4: Estimates Using Options-Based Skewness Measure

This table presents double sorting and Fama-Macbeth regression results based on the options-based idiosyncratic skewness measure (OS) of Conrad et al. (2013). The double sorting and regression methodologies are the same as those described in Tables 2 and 3, respectively. Table A.1 defines all the variables. Standard errors are adjusted following the Newey & West (1987) approach using a lag of 6. The sample period covers January 1996 to December 2015, as the option price data for older periods are not available in the OptionMetrics database. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Double Sorts

OS Quintile	Most Underpriced	Most Overpriced	Most Overpriced - Most Underpriced
Low	0.21*	0.28	0.10
	(1.81)	(-0.11)	(-0.46)
2	0.00	-1.86***	-1.70***
	(1.37)	(-2.93)	(-3.15)
3	0.44*	-1.72***	-2.04***
	(1.27)	(-3.12)	(-3.20)
4	0.08*	-0.82	-0.99**
	(1.94)	(-1.54)	(-2.40)
High	0.52	-1.66***	-2.13***
	(1.37)	(-4.82)	(-4.00)
High - Low	0.31	-1.55**	-2.06**
	(0.74)	(-2.03)	(-2.10)

Panel B: Fama-Macbeth Estimates

	(1)	(2)
Intercept	-0.001	0.005
	(-0.16)	(0.61)
MIS		-0.003**
		(-2.31)
OS	0.000	-0.001
	(-0.59)	(-1.43)
MIS × OS		-0.003**
		(-2.29)
log(ME)	0.003	0.000
	(1.22)	(0.08)
log(B/M)	0.001	0.001
	(0.39)	(0.71)
RET[-12,-2]	0.003	0.002
	(1.30)	(0.79)
RET[-1,0]	0.002	0.002
	(1.27)	(1.27)
Average Number of Observations	279	278
Average Adjusted R^2	0.10	0.11

Table 5: Individual Investor Portfolio Weight Regression Estimates

This table presents estimates from the Fama-Macbeth regressions, where the dependent variables are the raw weight (columns (1) to (4)) and the excess weight (columns (5) to (8)) allocated to overpriced stocks relative to underpriced ones in each investor portfolio at the end of every month. Overpriced (underpriced) stocks are defined as those in the fifth (first) quintile of MIS. The raw and the excess relative weights are defined as $W_{i,t}^{overpriced} - W_{i,t}^{underpriced}$ and $EW^{overpriced} - EW^{underpriced} = [(W_{i,t}^{overpriced} - W_{mkt,t}^{overpriced}) - (W_{i,t}^{underpriced} - W_{mkt,t}^{underpriced})]$, respectively. $W_{i,t}^{overpriced}$ is the raw weight allocated to overpriced stocks in portfolio i at the end of month t; $W_{i,t}^{underpriced}$ is the raw weight allocated to underpriced stocks in portfolio i at the end of month t; $W_{mkt,t}^{overpriced}$ is the raw weight allocated to overpriced stocks in the market portfolio at the end of month t; and $W_{mkt,t}^{underpriced}$ is the raw weight allocated to underpriced stocks in the market portfolio at the end of month t. In Panel A, our main independent variables are the average portfolio weight an investor allocated to stocks with skewness levels above the sample median over the past 12 months. We use four different skewness measures of JACKPOT, LIDX, MAXRET, and ESKEW to compute this weight. In Panel B, we estimate the same models but modify our measures of past weight on skewed stocks to exclude all stocks allocated to MIS quintile 1 or 5. We include a wide range of socioeconomic and portfolio characteristics control variables in both panels. For brevity, we do not report the control variable coefficients in Panel B. Table A.1 defines all the variables. We standardize all independent variables in our regressions to have a mean of zero and a standard deviation of 1 and winsorize them at the 0.5 and 99.5 percentiles. Standard errors are adjusted for heteroscedasticity and autocorrelation following the Newey & West (1987) approach using a lag of 6. The sample period is January 1991 to December 1996. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Baseline Estimates								
	$W^{overpriced} - W^{underpriced}$				$EW^{overpriced} - EW^{underpriced}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	-0.317***	-0.319***	-0.317***	-0.313***	-0.075***	-0.076***	-0.075***	-0.073***
	(-9.57)	(-9.97)	(-10.34)	(-10.96)	(-4.63)	(-4.74)	(-4.59)	(-4.90)
$W_{JACKPOT}$	0.142***				0.107***			
	(20.45)				(16.61)			
W_{LIDX}		0.170***				0.129***		
		(20.19)				(23.05)		
W_{MAXRET}			0.184***				0.139***	
			(21.09)				(20.67)	
W_{ESKEW}				0.116***				0.087***
				(17.45)				(20.63)
Portfolio α	-0.004	-0.006	-0.008	0.000	-0.005	-0.006	-0.008	-0.003
	(-0.13)	(-0.24)	(-0.27)	(-0.01)	(-0.20)	(-0.31)	(-0.34)	(-0.09)
Portfolio Return	-0.153***	-0.161***	-0.167***	-0.158***	-0.121***	-0.127***	-0.131***	-0.124***
	(-4.65)	(-5.83)	(-5.30)	(-4.59)	(-4.74)	(-5.90)	(-5.41)	(-4.70)
Portfolio Variance	0.183***	0.154***	0.132***	0.217***	0.148***	0.125***	0.109***	0.173***
	(8.22)	(9.94)	(7.42)	(9.43)	(8.73)	(10.55)	(8.03)	(10.07)
Local Weight	-0.012***	-0.014***	-0.022***	-0.008**	-0.01***	-0.011***	-0.018***	-0.007**
	(-2.84)	(-4.24)	(-4.79)	(-2.23)	(-2.74)	(-3.85)	(-4.35)	(-2.24)
Industry	-0.118***	-0.118***	-0.119***	-0.121***	-0.122***	-0.122***	-0.122***	-0.124***
Concentration	(-17.95)	(-21.09)	(-21.05)	(-19.13)	(-20.45)	(-23.35)	(-23.27)	(-21.61)
Diversification	0.013***	0.009***	0.007***	0.017***	0.013***	0.011***	0.009***	0.016***
	(5.31)	(4.38)	(3.56)	(5.89)	(6.77)	(5.72)	(5.37)	(7.01)
ln(Portolio Size)	-0.013**	-0.001	-0.005	-0.014**	-0.001	0.008*	0.005	-0.002
	(-2.02)	(-0.14)	(-0.98)	(-2.33)	(-0.25)	(1.70)	(1.06)	(-0.35)
Age (Years)	0.015***	0.015***	0.021***	0.012***	0.012***	0.011***	0.016***	0.009***
	(7.26)	(7.43)	(10.59)	(4.76)	(7.07)	(6.87)	(10.09)	(4.65)
Male Dummy	0.009***	0.008***	0.006***	0.01***	0.007***	0.006***	0.005***	0.007***
	(8.83)	(6.25)	(5.75)	(8.19)	(9.52)	(7.12)	(5.92)	(8.99)

Table 5 (Continued)

Panel A (Continued):

	$W^{overpriced} - W^{underpriced}$				$EW^{overpriced} - EW^{underpriced}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Married Dummy	-0.010***	-0.010***	-0.009***	-0.011***	-0.007***	-0.007***	-0.006***	-0.008***
	(-4.98)	(-4.71)	(-4.44)	(-6.00)	(-4.22)	(-3.85)	(-3.72)	(-5.06)
Tenant Dummy	0.005**	0.006***	0.005***	0.006***	0.003**	0.004***	0.003***	0.004***
	(2.64)	(3.20)	(3.10)	(3.33)	(2.35)	(2.90)	(2.70)	(3.07)
CPRATIO	0.005***	0.004***	0.004**	0.006***	0.004**	0.003**	0.003*	0.005***
	(3.20)	(2.78)	(2.12)	(4.34)	(2.64)	(2.31)	(1.82)	(3.60)
ln(Population)	-0.012***	-0.011***	-0.012***	-0.013***	-0.010***	-0.009***	-0.010***	-0.011***
	(-3.65)	(-3.02)	(-2.95)	(-3.75)	(-4.25)	(-3.49)	(-3.43)	(-4.26)
Income Equality (%)	-0.024***	-0.027***	-0.035***	-0.023***	-0.020***	-0.022***	-0.029***	-0.02***
	(-4.48)	(-6.26)	(-6.70)	(-4.21)	(-4.86)	(-6.64)	(-6.61)	(-4.46)
ln(Household Income)	-0.034***	-0.038***	-0.044***	-0.034***	-0.028***	-0.031***	-0.036***	-0.028***
	(-12.22)	(-13.57)	(-24.58)	(-14.28)	(-14.1)	(-14.62)	(-28.5)	(-15.99)
Minority (%)	0.004	0.001	0.001	0.003	0.003	0.001	0.001	0.003
	(1.05)	(0.19)	(0.18)	(0.86)	(0.94)	(0.22)	(0.22)	(0.79)
Rural (%)	-0.003	-0.004*	-0.003	-0.003*	-0.003	-0.004*	-0.003	-0.003*
	(-1.46)	(-1.93)	(-1.21)	(-1.94)	(-1.66)	(-2.00)	(-1.38)	(-1.96)
Education (%)	-0.006*	-0.005**	-0.011***	-0.005	-0.004*	-0.004*	-0.008***	-0.003
	(-1.91)	(-2.04)	(-3.99)	(-1.64)	(-1.86)	(-1.87)	(-3.78)	(-1.42)
Average Number of Observations	6477	6477	6,477	6,477	6,477	6,477	6,477	6,477
Average Adjusted R^2	0.248	0.267	0.272	0.234	0.245	0.261	0.266	0.233

Panel B: Skewness Weights Excluding Overpriced and Underpriced Stocks

	$W^{overpriced} - W^{underpriced}$				$EW^{overpriced} - EW^{underpriced}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$W_{JACKPOT}$	0.076***				0.056***			
	(15.88)				(13.6)			
W_{LIDX}		0.113***				0.085***		
		(21.44)				(21.31)		
W_{MAXRET}			0.104***				0.077***	
			(8.77)				(8.33)	
W_{ESKEW}				0.073***				0.055***
				(8.51)				(8.14)

Table 6: Fama-Macbeth Regression Estimates with Idiosyncratic Volatility

Panel A of this table presents the Fama-Macbeth regression estimates after controlling for the effect of IVOL. We take the regression specifications in Table 3 and add a dummy for IVOL values above the cross-sectional median (DIVOL) and the dummy's interaction with MIS to all regressions. Similarly, we use the dummy forms of our four skewness measures (DJACKPOT, DLIDX, DMAXRET, and DESKEW). Panel B, provides the Fama-Macbeth regression estimates for subsamples of stocks with high and low CPRATIOs. We sort stocks every month based on their CPRATIO values measured at the end of the previous month and allocate them into two groups using the cross-sectional median. The regression specifications estimated for each group are similar to those in Panel A. Table A.1 explains the construction details for all variables. For brevity, we only report the coefficients on MIS and the interaction terms in Panel B. All independent variables in the regressions are standardized to have a mean of zero and a standard deviation of 1 and are winsorized at the 0.5 and 99.5 percentiles. Standard errors are adjusted for heteroscedasticity and autocorrelation following the Newey & West (1987) approach using a lag of 6. All coefficients are multiplied by 100 for better comparison. The sample excludes penny stocks and covers January 1973 to December 2015, except for the regression that includes ESKEW, which starts in January 1988. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Panel A: Controlling for IVOL				
Intercept	4.544***	4.262***	4.402***	4.400***
	(8.36)	(8.21)	(8.14)	(7.88)
MIS	-0.243***	-0.263***	-0.288***	-0.293***
	(-8.09)	(-8.24)	(-9.50)	(-7.97)
DIVOL	-0.110	-0.189**	-0.125*	-0.170
	(-1.43)	(-2.25)	(-1.70)	(-1.62)
DJACKPOT	-0.149*			
	(-1.73)			
DLIDX		0.058		
		(0.76)		
DMAXRET			-0.063	
			(-1.03)	
DESKEW				-0.125*
				(-1.88)
MIS × DIVOL	-0.152***	-0.216***	-0.213***	-0.294***
	(-4.88)	(-6.92)	(-5.74)	(-8.11)
MIS × DJACKPOT	-0.264***			
	(-7.36)			
MIS × DLIDX		-0.171***		
		(-5.01)		
MIS × DMAXRET			-0.123***	
			(-3.82)	
MIS × DESKEW				-0.100***
				(-2.79)
Log(ME)	-0.123***	-0.100***	-0.106***	-0.091***
	(-3.91)	(-3.27)	(-3.13)	(-2.81)
Log(B/M)	0.261***	0.269***	0.275***	0.300***
	(4.54)	(4.58)	(4.68)	(4.79)
RET[-12,-2]	0.491***	0.503***	0.516***	0.434***
	(4.12)	(4.18)	(4.32)	(3.36)
RET[-1,0]	-4.360***	-4.349***	-4.235***	-4.373***
	(-11.94)	(-11.80)	(-11.09)	(-11.31)
Average Number of Observations	3054	3054	3054	3133
Average Adjusted R2	0.047	0.047	0.046	0.046

Table 6 (Continued)

Panel B: CPRATIO Subsamples												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Low CPRATIO				High CPRATIO				High - Low			
MIS	-0.236***	-0.258***	-0.255***	-0.285***	-0.233***	-0.253***	-0.270***	-0.281***	0.000	0.000	0.000	0.000
	(-6.90)	(-7.11)	(-7.32)	(-6.90)	(-7.37)	(-7.71)	(-8.71)	(-7.48)	(-0.46)	(-0.87)	(-0.61)	(-0.27)
MIS × DIVOL	-0.149***	-0.186***	-0.156***	-0.243***	-0.093**	-0.172***	-0.171***	-0.241***	0.046	0.014	-0.015	0.002
	(-3.26)	(-4.18)	(-3.10)	(-5.57)	(-1.97)	(-3.71)	(-3.18)	(-4.82)	(0.72)	(0.22)	(-0.20)	(0.03)
MIS × DJACKPOT	-0.168***				-0.254***				-0.086***			
	(-3.44)				(-5.78)				(-2.76)			
MIS × DLIDX		-0.112**				-0.180***				-0.068**		
		(-2.30)				(-5.26)				(-2.12)		
MIS × DMAXRET			-0.129**				-0.181***				-0.052**	
			(-2.50)				(-3.10)				(-2.08)	
MIS × DESKEW				-0.050				-0.055				-0.005
				(-1.04)				(-1.14)				(-0.07)
Average Number of Observations	1191	1191	1191	1242	1188	1188	1188	1234				
Average Adjusted R2	0.051	0.051	0.050	0.049	0.052	0.052	0.052	0.051				

Table 7: Fama-Macbeth Regression Estimates with Coskewness

This table presents the Fama-Macbeth regression estimates after controlling for the effect of coskewness. We take the regression specifications in Table 3 and add a measure of coskewness and its interaction with MIS to all regressions. In Panel A, we define coskewness (COSKEW) following the original Harvey and Siddique (2000) definition. In Panel B, we adopt Harvey and Siddique's (2000) alternative measure of coskewness, which is defined as the regression coefficient on a squared market factor. Table A.1 defines all the variables. We standardize all independent variables in our regressions to have a mean of zero and a standard deviation of 1 and winsorize them at the 0.5 and 99.5 percentiles. Standard errors are adjusted for heteroscedasticity and autocorrelation following the Newey & West (1987) approach using a lag of 6. The sample excludes penny stocks and covers January 1963 to December 2015, except for the regression that includes ESKEW, which starts in January 1988. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
Panel A: Coskewness Based on the Original Harvey and Siddique (2000) Definition (COSKEW)					
Intercept	0.011***	0.010***	0.010***	0.010***	0.009***
	(4.70)	(3.90)	(4.32)	(4.13)	(3.69)
MIS		-0.004***	-0.004***	-0.004***	-0.004***
		(-10.53)	(-12.96)	(-11.18)	(-11.17)
JACKPOT		-0.001			
		(-1.12)			
LIDX			0.000		
			(-0.16)		
MAXRET				-0.002**	
				(-2.55)	
ESKEW					-0.001
					(-1.40)
COSKEW	0.000	0.000	0.000*	0.000	0.000
	(-1.2)	(-1.21)	(-1.77)	(-1.28)	(-1.52)
MIS × JACKPOT		-0.002***			
		(-3.6)			
MIS × LIDX			-0.002***		
			(-6.62)		
MIS × MAXRET				-0.002***	
				(-6.35)	
MIS × ESKEW					-0.001***
					(-3.85)
MIS × COSKEW		0.000	0.000	0.000	0.000
		(-0.24)	(-0.68)	(-0.29)	(-0.81)
log(ME)	-0.001	-0.002***	-0.002***	-0.002***	-0.001**
	(-1.47)	(-2.98)	(-2.71)	(-2.98)	(-1.98)
log(B/M)	0.003***	0.003***	0.003***	0.003***	0.003***
	(5.21)	(4.84)	(4.62)	(4.66)	(4.92)
RET[-12,-2]	0.005***	0.003***	0.003***	0.003***	0.003***
	(5.22)	(3.48)	(3.71)	(3.69)	(3.11)
RET[-1,0]	-0.007***	-0.008***	-0.008***	-0.007***	-0.007***
	(-11.34)	(-12.24)	(-12.42)	(-10.04)	(-11.69)
Average Number of Observations	2,173	2,154	2,154	2,154	2,198
Average Adjusted R^2	0.043	0.052	0.053	0.052	0.048

Table 7 (Continued)

Panel B: Coskewness Defined as the Coefficient on the Squared Market Factor (β_{m^2})					
β_{m^2}	-0.001	0.000	-0.001	0.000	-0.001
	(-0.53)	(-0.55)	(-0.72)	(-0.40)	(-0.60)
MIS × JACKPOT		-0.003***			
		(-6.05)			
MIS × LIDX			-0.002***		
			(-7.65)		
MIS × MAXRET				-0.002***	
				(-7.20)	
MIS × ESKEW					-0.001***
					(-3.96)
MIS × β_{m^2}		0.000	0.000	0.000	0.000
		(0.15)	(-0.34)	(-0.16)	(-0.19)
Average Number of Observations	3,084	3,072	3,032	3,033	3,153
Average Adjusted R^2	0.041	0.048	0.05	0.048	0.045

Table 8: Fama-Macbeth Regression Estimates with Arbitrage Cost Measures

This table presents the Fama-Macbeth regression estimates after controlling for limits to arbitrage. We take the regression specifications in Table 3 and add five proxies for limits to arbitrage and their interactions with MIS to each specification, separately. Panels A to E report the results based on each of the five proxies. ILLIQ is the illiquidity measure of Amihud (2002); BIDASK is the bid-ask spread; %ZEROS is the frequency of zero daily returns devised by Lesmond et al. (1999); IHOLDING is the percentage institutional holding; and CBS is the short-selling fee score as in Porras Prado et al. (2016). Table A.1 explains the construction details for all variables. For brevity, we only report the coefficients on MIS, the interaction terms, and the proxies for limits to arbitrage. All independent variables in the regressions are standardized to have a mean of zero and a standard deviation of 1 and are winsorized at the 0.5 and 99.5 percentiles. Standard errors are adjusted for heteroscedasticity and autocorrelation following the Newey & West (1987) approach using a lag of 6. The sample excludes penny stocks and covers January 1963 to December 2015, except for the regressions that include ESKEW, IHOLDING, and CBS, which because of data availability start in January 1988, January 1980, and July 2006, respectively. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Panel A: ILLIQ				
MIS	-0.005***	-0.005***	-0.005***	-0.005***
	(-11.03)	(-12.4)	(-11.29)	(-10.92)
ILLIQ	0.003***	0.003***	0.003***	0.002***
	(3.37)	(2.94)	(3.12)	(2.71)
MIS × JACKPOT	-0.004***			
	(-5.82)			
MIS × LIDX		-0.002***		
		(-8.96)		
MIS × MAXRET			-0.003***	
			(-10.5)	
MIS × ESKEW				-0.002***
				(-5.01)
MIS × ILLIQ	0.002***	0.002***	0.002***	0.002***
	(3.99)	(3.51)	(2.94)	(2.9)
Average Number of Observations	3097	3096	3097	3047
Average Adjusted R^2	0.049	0.051	0.049	0.048
Panel B: BIDASK				
MIS	-0.007***	-0.006***	-0.006***	-0.006***
	(-7.33)	(-8.22)	(-8.38)	(-8.2)
BIDASK	0.003	0.003*	0.004**	0.004*
	(1.63)	(1.67)	(2.11)	(1.85)
MIS × JACKPOT	-0.004***			
	(-3.19)			
MIS × LIDX		-0.003***		
		(-6.98)		
MIS × MAXRET			-0.003***	
			(-5.62)	
MIS × ESKEW				-0.003***
				(-5.68)
MIS × BIDASK	0.000	0.002*	0.000	0.002
	(0.43)	(1.67)	(-0.16)	(1.34)
Average Number of Observations	2860	2859	2860	2810
Average Adjusted R^2	0.043	0.044	0.042	0.041

Table 8 (Continued)

Panel C: %ZEROS				
MIS	-0.006***	-0.004***	-0.006***	-0.005***
	(-7.19)	(-5)	(-7.1)	(-6.2)
%ZEROS	-0.001	-0.001	-0.001	-0.002
	(-0.54)	(-0.36)	(-0.65)	(-0.74)
MIS × JACKPOT	-0.004***			
	(-6.96)			
MIS × LIDX		-0.002***		
		(-9.44)		
MIS × MAXRET			-0.002***	
			(-10.88)	
MIS × ESKEW				-0.002***
				(-6.18)
MIS × %ZEROS	0.000	0.002**	0.000	0.001
	(0.55)	(2.18)	(-0.28)	(1.55)
Average Number of Observations	3370	3368	3370	3252
Average Adjusted R^2	0.047	0.049	0.047	0.046

Panel D: IHOLDING				
MIS	-0.006***	-0.006***	-0.005***	-0.006***
	(-10.44)	(-11.78)	(-9.99)	(-10.34)
IHOLDING	0.001**	0.001*	0.001*	0.001*
	(2.11)	(1.68)	(1.91)	(1.91)
MIS × JACKPOT	-0.003***			
	(-5.38)			
MIS × LIDX		-0.002***		
		(-7.24)		
MIS × MAXRET			-0.002***	
			(-9.04)	
MIS × ESKEW				-0.001***
				(-3.28)
MIS × IHOLDING	0.001***	0.001**	0.001***	0.001***
	(4.33)	(2.41)	(4.87)	(4.76)
Average Number of Observations	3493	3493	3493	3407
Average Adjusted R^2	0.039	0.041	0.039	0.038

Table 8 (Continued)

Panel E: CBS				
MIS	-0.002**	-0.002**	-0.002**	-0.002**
	(-2.22)	(-2.38)	(-2.19)	(-2.31)
CBS	-0.004***	-0.004***	-0.004***	-0.004***
	(-3.38)	(-3.91)	(-3.61)	(-4.21)
MIS × JACKPOT	-0.002**			
	(-2.22)			
MIS × LIDX		-0.002***		
		(-2.86)		
MIS × MAXRET			-0.002**	
			(-2.27)	
MIS × ESKEW				-0.002**
				(-2.33)
MIS × CBS	-0.001	-0.001	-0.002	-0.001
	(-1.25)	(-1.33)	(-1.52)	(-0.98)
Average Number of Observations	2306	2306	2306	2173
Average Adjusted R^2	0.034	0.032	0.033	0.033

Table 9: **Fama-Macbeth Regression Estimates with Market Sentiment Indicators**

This table presents the Fama-Macbeth regression estimates within subsamples of market ups and downs and high and low sentiment. The market subsamples are created by dividing the sample into months with positive (up) and negative (down) CRSP value-weighted index returns. The sentiment subsamples are created by dividing the sample using the time series median level of the Baker & Wurgler (2006) index. We then take the regression specifications in Table 3 and estimate them separately for each group. In Panel A, we look at market ups and downs using the CRSP index. In Panel B, we use the original sentiment index of Baker & Wurgler (2006) based on the first principal component of five sentiment proxies (SENT). In Panel C, the index is based on the first principal component of five sentiment proxies that are first orthogonalized with respect to a set of macroeconomic indicators ($SENT^{\perp}$). For brevity, we only report the coefficients on MIS and the interaction terms. All independent variables in the regressions are standardized to have a mean of zero and a standard deviation of 1 and are winsorized at the 0.5 and 99.5 percentiles. Standard errors are adjusted for heteroscedasticity and autocorrelation following the Newey & West (1987) approach using a lag of 6. The sample excludes penny stocks and covers January 1963 to December 2015, except for the regression that includes ESKEW, which starts in January 1988. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Market Ups and Downs				
		Market Down	Market Up	Up - Down
JACKPOT	MIS	-0.007***	-0.004***	0.003***
		(-10.47)	(-8.61)	(3.69)
	MIS × JACKPOT	-0.003***	-0.003***	0.000
		(-3.6)	(-5)	(0.03)
	Avg N	3062	3078	
LIDX	MIS	-0.006***	-0.004***	0.002***
		(-10.74)	(-10.18)	(2.93)
	MIS × LIDX	-0.002***	-0.001***	0.001***
		(-6.35)	(-4.53)	(2.60)
	Avg N	3036	3029	
MAXRET	MIS	-0.006***	-0.003***	0.003***
		(-10.48)	(-7.92)	(4.37)
	MIS × MAXRET	-0.002***	-0.002***	0.000
		(-5.5)	(-5.6)	(0.00)
	Avg N	3038	3030	
ESKEW	MIS	-0.006***	-0.004***	0.002***
		(-10.06)	(-8.68)	(2.65)
	MIS × ESKEW	-0.002***	-0.001**	0.001
		(-3.2)	(-2.54)	(1.35)
	Avg N	3109	3182	

Table 9 (Continued)

Panel B: Sentiment Index (SENT)				
		Low SENT	High SENT	High - Low
JACKPOT	MIS	-0.004***	-0.007***	-0.003***
		(-7.15)	(-10.89)	(-3.52)
	MIS × JACKPOT	-0.002***	-0.004***	-0.002**
		(-3.94)	(-5.24)	(-2.18)
	Avg N	2841	3481	
LIDX	MIS	-0.003***	-0.006***	-0.003***
		(-8.56)	(-11.43)	(-4.75)
	MIS × LIDX	-0.001***	-0.002***	-0.001***
		(-4.19)	(-7.16)	(-2.72)
	Avg N	2838	3480	
MAXRET	MIS	-0.003***	-0.006***	-0.003***
		(-7.55)	(-9.94)	(-4.15)
	MIS × MAXRET	-0.002***	-0.003***	-0.001*
		(-5.25)	(-8.08)	(-1.90)
	Avg N	2840	3481	
ESKEW	MIS	-0.003***	-0.007***	-0.004***
		(-6.72)	(-10.86)	(-5.10)
	MIS × ESKEW	-0.001**	-0.002***	-0.001
		(-2.04)	(-3.39)	(-1.30)
	Avg N	2866	3418	
Panel C: Orthogonalized Sentiment Index ($SENT^{\perp}$)				
		Low $SENT^{\perp}$	High $SENT^{\perp}$	High - Low
JACKPOT	MIS	-0.003***	-0.007***	-0.004***
		(-6.97)	(-11.2)	(-5.27)
	MIS × JACKPOT	-0.002***	-0.004***	-0.002**
		(-3.81)	(-5.46)	(-2.20)
	Avg N	2881	3441	
LIDX	MIS	-0.003***	-0.006***	-0.003***
		(-8.42)	(-11.49)	(-4.75)
	MIS × LIDX	-0.001***	-0.002***	-0.001***
		(-3.97)	(-7.47)	(-2.70)
	Avg N	2879	3440	
MAXRET	MIS	-0.003***	-0.006***	-0.003***
		(-7.13)	(-10.31)	(-4.18)
	MIS × MAXRET	-0.002***	-0.003***	-0.001*
		(-5.01)	(-8.24)	(-1.90)
	Avg N	2881	3441	
ESKEW	MIS	-0.003***	-0.007***	-0.004***
		(-6.52)	(-10.99)	(-5.09)
	MIS × ESKEW	-0.001*	-0.002***	-0.001
		(-1.81)	(-3.71)	(-1.30)
	Avg N	2919	3367	

Table 10: Fama-Macbeth Regression Estimates For Geographic Gambling Propensity Subsamples

This table presents the Fama-Macbeth regression estimates for subsamples of stocks with high and low CPRATIOs. We sort stocks every month based on their CPRATIO values measured at the end of the previous month and allocate them into two groups using the cross-sectional median as the breakpoint. Then, we take the regression specifications in Table 3 and estimate them separately for each group. CPRATIO is defined as the ratio of the Catholic population to the Protestant population in the stock headquarters's county. Table A.1 explains the construction details for all variables. For brevity, we only report the coefficients on MIS and the interaction terms. All independent variables in the regressions are standardized to have a mean of zero and a standard deviation of 1 and are winsorized at the 0.5 and 99.5 percentiles. Standard errors are adjusted for heteroscedasticity and autocorrelation following the Newey & West (1987) approach using a lag of 6. The sample excludes penny stocks and covers January 1973 to December 2015, except for the regression that includes ESKEW, which starts in January 1988. The sample starts in January 1973, unlike the previous tests because of short interest data availability. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

		Low CPRATIO	High CPRATIO	High - Low
JACKPOT	MIS	-0.004***	-0.004***	0.000
		(-9.7)	(-10.2)	(0.00)
	MIS × JACKPOT	-0.001***	-0.004***	-0.003***
		(-3.42)	(-3.84)	(-2.80)
	Avg N	1194	1198	
LIDX	MIS	-0.004***	-0.004***	0.000
		(-10.47)	(-11.29)	(0.00)
	MIS × LIDX	-0.001***	-0.002***	-0.001**
		(-4.64)	(-4.67)	(-2.10)
	Avg N	1179	1182	
MAXRET	MIS	-0.004***	-0.004***	0.000
		(-9.23)	(-9.22)	(0.00)
	MIS × MAXRET	-0.001***	-0.002***	-0.001*
		(-3.22)	(-4.75)	(-1.90)
	Avg N	1179	1182	
ESKEW	MIS	-0.005***	-0.004***	0.001
		(-9.92)	(-10.26)	(1.57)
	MIS × ESKEW	-0.001***	-0.002***	-0.001*
		(-2.6)	(-4.3)	(-1.70)
	Avg N	1242	1250	

Appendix Tables

Table A.1: Variable Descriptions

This table defines the main variables used in the empirical analysis.

Variable Name	Source	Description
		Panel A: Skewness and Anomaly Variables
β_{m^2}	CRSP	This is computed following Harvey and Siddique (2000) by estimating the following model: $$R_{i,t} - R_{f,t} = \alpha_i + \beta_{m,i}(R_{m,t} - R_{f,t}) + \beta_{m^2,i}(R_{m,t} - R_{f,t})^2 + \epsilon_{i,t},$$ where $R_{i,t}$ is the return on stock i on day t, $R_{m,t}$ is the market return on day t, and $R_{f,t}$ is the risk-free rate on day t. We estimate the above regression using daily returns for the most recent month.
COSKEW	CRSP	Harvey and Siddique (2000) use this as their main measure of coskewness computed as follows: $$COSKEW_{i,t} = \frac{E[\epsilon_{i,t}\epsilon_{m,t}^2]}{\sqrt{E[\epsilon_{i,t}^2]}E[\epsilon_{m,t}^2]},$$ where $\epsilon_{i,t} = R_{i,t} - R_{f,t} - \alpha_i - \beta_i(R_{m,t} - R_{f,t})$, $R_{i,t}$ is the return on stock i on month t, $R_{m,t}$ is the market return on month t, and $R_{f,t}$ is the risk-free rate on month t. We estimate the above regression using monthly returns for the past 60 months.
DESKEW	CRSP	A dummy that is set to one if ESKEW is above its cross-sectional median
DJACKPOT	CRSP and Compustat	A dummy that is set to one if JACKPOT is above its cross-sectional median
DLIDX	CRSP	A dummy that is set to one if LIDX is above its cross-sectional median
DMAXRET	CRSP	A dummy that is set to one if MAXRET is above its cross-sectional median
ESKEW	CRSP	Following Boyer et al. (2010), this is defined by running a cross-sectional regression at the end of every month using the most recent 5 years of data to predict the daily idiosyncratic skewness of stocks estimated over the following 5 years. Variables used in the regression include the historical estimates of daily idiosyncratic volatility and skewness relative to the Fama-French three-factor model over the past 60 months, momentum as the cumulative returns over months $t-12$ through $t-1$, turnover as the average daily turnover

Table A.1 (Continued)

Variable Name	Source	Description
		Panel A (Continued): Skewness and Anomaly Variables
ESKEW (Continued)	CRSP	in month $t-1$, small- and medium-sized market capitalization dummies (based on sorts of firms by market capitalization into three groups of small, medium, and large), an industry dummy based on the Fama-French 17 industries, and a NASDAQ dummy. After estimating the model at the end of every month t, we use the parameters together with the most recent data to get out-of-sample expected idiosyncratic skewness estimates for months $t+61$ through $t+120$. Our estimates start in 1988 because detailed data on the trading volume of NASDAQ stocks become available in 1983.
JACKPOT	CRSP and Compustat	Conrad et al. (2014) compute this by running a logit model at the end of June for every year to predict the out-of-sample probability of a stock generating a log return greater than 100% in the next 12 months. Variables used in the logit regression are the stock's (log) return over the last 12 months, volatility and skewness of daily log returns over the past 3 months, detrended stock turnover ([6-month volume/shares outstanding] − [18-month volume/shares outstanding]), and log market capitalization. The model is estimated following a rolling-window approach using data from the past 10 years. Unlike Conrad et al. (2014), who use data from the past 20 years, we only require 10 years of historical data for each rolling window estimation. Considering that the Compustat Fundamentals database started in 1950, a shorter estimation window enables us to start our parameter estimates from 1963. After estimating the logit model at the end of June of year t, the estimated parameters are used together with the most recently available data to estimate a jackpot score for every stock from July of year t to the end of June of year $t+1$.
LIDX	CRSP	Following Kumar et al. (2016), this is defined as the sum of the vigintile allocation of stocks with respect to price, idiosyncratic volatility, and idiosyncratic skewness divided by 60. Vigintiles are defined such that stocks with the lowest price, the highest idiosyncratic skewness, and the highest idiosyncratic volatility are allocated to the highest corresponding vigintile groups. All stocks in the sample are sorted at the end of each month based on the three characteristics to compute the lottery index for the following month. Price is the monthly closing price. Idiosyncratic volatility is defined as the standard deviation of the residuals from fitting the four-factor model of Carhart (1997) to the daily return data for the past 6 months. Idiosyncratic skewness refers to the skewness of residuals obtained from a two-factor model estimated using daily return data for the past 6 months, with the two factors being the market factor and its square.

Table A.1 (Continued)

Variable Name	Source	Description
		Panel A (Continued): Skewness and Anomaly Variables
MAXRET	CRSP	Bali et al. (2011) define this as the maximum daily return in the previous month.
MIS	CRSP and Compustat	Following Stambaugh et al. (2015), MIS is the average of decile ranks of a stock with respect to 11 prominent anomalies. Sorting for each anomaly is performed at the end of every month. Deciles 1 and 10 include stocks that each anomaly strategy predicts will outperform and underperform the most in the following month, respectively. Unlike Stambaugh et al. (2015), we determine our decile cutoffs using our whole sample, not just NYSE stocks. We require at least five non-missing anomaly decile ranks to compute MIS for a stock. The 11 anomaly strategies considered are accruals (Sloan, 1996), asset growth (Cooper et al., 2008), composite equity issues (Daniel and Titman, 2006), distress (Campbell et al., 2008), gross profitability (Novy-Marx, 2013), investment-to-assets (Titman et al., 2004), momentum (Jegadeesh and Titman, 1993), net operating assets (Hirshleifer et al., 2004), net stock issues (Ritter, 1991; Loughran and Ritter, 1995), O-score (Ohlson, 1980), and return on assets (Fama and French, 2006). We follow the detailed description of Stambaugh et al. (2012, 2015), together with the corresponding anomaly literature, to replicate each strategy.
ISKEWNESS	CRSP	Skewness of residuals obtained from running the three-factor model of Fama and French (1993) on daily returns for the most recent month.
OS	OptionMetrics	This follows Conrad et al. (2013) and Bakshi et al. (2003) and is defined as the third moment of the risk-neutral density function of a security constructed using a set of out-of-the-money option prices with different strike prices on that security. Our sample of out-of-the-money calls and puts include securities that have expiration dates close to 0.250 years (3 months). We choose this time to maturity because the measure based on options with 3 months to maturity has the strongest return predictability in Conrad et al. (2013). Our estimation technique and option data filters closely follow those used in Conrad et al. (2013).
SKEWNESS	CRSP	Skewness of daily returns for the most recent month.

Table A.1 (Continued)

Variable Name	Source	Description
		Panel B: Control Variables
%ZEROS	CRSP	This was devised by Lesmond et al. (1999) as the percentage of daily returns of each stock equal to zero. We measure this using the past 12 months of daily returns for each firm.
B/M	CRSP and Compustat	This is the ratio of the book value to the market capitalization of the firm.
BIDASK	CRSP	This is the average daily bid-ask spread over the past 12 months.
CBS	Markit Data Explorers	This is the cost of borrowing score devised by Markit Data Explorers. The variable ranges from 1 to 10 indicating how expensive it is to borrow each stock based on the loan-weighted fees charged by lenders. We compute a monthly score by taking the average of the daily shorting fee scores of each stock over the past month.
DIVOL	CRSP	A dummy that is set to one if IVOL is above its cross-sectional median
IHOLDING	Thomson Reuters	The fraction of a stock's outstanding shares held by institutional investors. We obtain the stock's institutional holdings by aggregating the positions of its institutional investors. If the Thomson Reuters database does not have data on a particular stock, we set the stock's institutional holdings to zero.
ILLIQ	CRSP	This is the annual average of the daily ratio of absolute stock return to daily dollar trading volume, following Amihud (2002).
IVOL	CRSP	Volatility of residuals obtained from running the three-factor model of Fama and French (1993) on daily returns for the most recent month (Stambaugh et al., 2015).
LEVERAGE	CRSP and Compustat	This is the sum of total debt from current liabilities plus total long-term debt, all divided by total assets.
ME	CRSP	Price times shares outstanding.
PRICE	CRSP	Monthly closing price.
RET[-1,0]	CRSP	Buy-and-hold return over the previous month.
RET[-12,-2]	CRSP	The prior years monthly compounded buy-and-hold return skipping the last month.
RHOLDING	Brokerage	Percentage of total shares outstanding owned by individuals in the brokerage sample.

Table A.1 (Continued)

Variable Name	Source	Description
		Panel B (Continued): Control Variables
SHORTRATIO	Compustat	Average ratio of short interest to shares outstanding over the past 12 months.
		Panel C: Variables Used in the Individual Holdings Regressions
TURNOVER	CRSP	Total trading volume over the last month divided by shares outstanding.
VOLATILITY	CRSP	Volatility of daily returns for the most recent month.
Age (Years)	Brokerage	The portfolio holder's age.
CPRATIO	ARDA	This is the ratio of the Catholic population to the Protestant population in the portfolio holder's county.
Diversification	Brokerage and CRSP	Portfolio variance divided by the average variance of all stocks in the portfolio.
Education	1990 Census	This is the proportion of residents in the portfolio holder's county with a Bachelor's degree or higher.
Income Equality	1990 Census	This is the ratio of the number of households in the lowest annual income group (less than $10,000) to those in the highest annual income group ($150,000 or more) in the portfolio holder's county.
Industry Concentration	Brokerage and CRSP	Largest weight allocated to one of the 48 Fama-French industries.
ln(Household Income)	1990 Census	This is the natural log of annual household income in the portfolio holder's county.
ln(Population)	1990 Census	This is the natural log of the portfolio holder's home county population.
ln(Portolio Size)	Brokerage	This is the natural log of the size of the portfolio.
Local Weight	Brokerage	Portfolio weight allocated to stocks located in the portfolio holder's home state.
Male Dummy	Brokerage	This is equal to 1 if the portfolio holder is a male.
Married Dummy	Brokerage	This is equal to 1 if the portfolio holder is married.

Table A.1 (Continued)

Variable Name	Source	Description
Panel C (Continued): Variables Used in the Individual Holdings Regressions		
Minority	1990 Census	This is the proportion of the population that is not white in the portfolio holder's county.
Portfolio Return	Brokerage and CRSP	Monthly compounded portfolio returns over the past 12 months.
Portfolio Variance	Brokerage and CRSP	Variance of the portfolio estimated using the past 12 months of returns.
Portfolio α	Brokerage and CRSP	This is the intercept of the regression of monthly portfolio returns for Carharts (1997) four factors estimated using the past 12 months of data.
Rural	1990 Census	This is the proportion of the population that lives in rural areas in the portfolio holder's county.
Tenant Dummy	Brokerage	This is equal to 1 if the portfolio holder lives in a rental property.
W_{ESKEW}	Brokerage	Average monthly weight allocated to stocks with ESKEW values above the cross-sectional median over the past 12 months.
$W_{JACKPOT}$	Brokerage	Average monthly weight allocated to stocks with JACKPOT values above the cross-sectional median over the past 12 months.
W_{LIDX}	Brokerage	Average monthly weight allocated to stocks with LIDX values above the cross-sectional median over the past 12 months.
W_{MAXRET}	Brokerage	Average monthly weight allocated to stocks with MAXRET values above the cross-sectional median over the past 12 months.

Table A.2: Ability of a Skewness Factor to Capture Anomaly Alphas

Panels A and B present the alphas and the t-statistics of 11 anomaly strategies based on various asset pricing models. In Panel C, we present a set of summarizing performance measures for each model, including the average absolute alpha, the average absolute t-statistic of alpha, the Gibbons et al. (1989) statistics (GRS), and the p-value corresponding to the GRS statistics. We devise a skewness factor, denoted *NMS*, and add it to the following four prominent models: Fama & French (1993) three-factor (FF3), Carhart (1997) four-factor (CAR), Fama & French (2015) five-factor (FF5), and Fama & French (2015) with the addition of momentum (FF6). We construct *NMS* in four steps. First, we compute the average decile rank of each stock at the end of every month with respect to the four skewness measures of JACKPOT, LIDX, MAXRET, and ESKEW. Next, we independently sort stocks based on their average skewness decile ranks and their market capitalization into three and two portfolios, respectively. We then compute the value-weighted monthly return of each of the six intersecting portfolios. Lastly, we take the average of the returns of the two size portfolios with the highest skewness tercile rank and deduct it from the average return of the two size portfolios with the lowest tercile rank to obtain the monthly factor returns. The sample excludes penny stocks and covers January 1963 to December 2015, except for distress and return-on-assets anomalies, which because of data availability start in January 1973. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Anomaly	FF3	FF3 + NMS	CAR	CAR + NMS	FF5	FF5 + NMS	FF6	FF6 + NMS
				Panel A: Alphas				
Accruals	0.009	0.011	0.007	0.010	0.009	0.011	0.008	0.010
Asset Growth	0.004	0.003	0.004	0.003	0.001	0.001	0.002	0.001
Composite Equity Issues	0.004	0.005	0.004	0.005	0.005	0.006	0.005	0.005
Distress	0.019	0.012	0.010	0.005	0.012	0.008	0.005	0.003
Gross Profitability	0.008	0.005	0.007	0.004	0.001	0.001	0.001	0.001
Investment-to-Assets	0.006	0.006	0.005	0.006	0.005	0.005	0.004	0.005
Momentum	0.022	0.020	0.006	0.007	0.019	0.018	0.006	0.007
Net Operating Assets	0.009	0.009	0.008	0.009	0.009	0.009	0.008	0.008
Net Stock Issues	0.008	0.006	0.008	0.006	0.004	0.003	0.004	0.004
O-Score	0.009	0.006	0.008	0.005	0.006	0.005	0.005	0.004
Return on Assets	0.012	0.008	0.010	0.007	0.006	0.005	0.005	0.005

Table A.2 (Continued)

Anomaly	FF3	FF3 + NMS	CAR	CAR + NMS	FF5	FF5 + NMS	FF6	FF6 + NMS
Panel B: t-Statistics								
Accruals	4.467	5.221	3.802	4.688	4.676	5.195	4.150	4.760
Asset Growth	2.632	1.941	2.665	2.002	0.720	0.496	1.090	0.807
Composite Equity Issues	2.726	3.357	2.603	3.234	3.769	3.898	3.579	3.740
Distress	6.241	3.637	3.986	2.044	4.170	2.536	2.320	1.210
Gross Profitability	4.285	2.587	3.575	2.104	0.596	0.706	0.363	0.512
Investment-to-Assets	4.633	4.537	3.891	3.974	4.001	3.941	3.528	3.559
Momentum	6.886	5.710	4.253	4.144	5.773	5.044	4.214	4.155
Net Operating Assets	6.090	5.612	5.420	5.120	6.000	5.515	5.498	5.124
Net Stock Issues	5.915	4.114	5.884	4.184	3.217	2.622	3.620	2.975
O-Score	5.865	3.664	5.146	3.196	4.510	3.414	4.037	3.057
Return on Assets	5.618	3.504	4.621	2.933	3.439	2.748	2.773	2.320
Panel C: Model Performance								
Average $\|\alpha\|$	0.010	0.008	0.007	0.006	0.007	0.006	0.005	0.004
Average $\|t\|$	5.033	3.989	4.168	3.420	3.715	3.283	3.197	2.929
GRS	10.436	7.269	8.293	6.701	7.852	6.524	6.870	6.471
p(GRS)	0.00×10^{-14}	1.44×10^{-11}	1.91×10^{-13}	1.60×10^{-10}	1.24×10^{-12}	3.40×10^{-10}	7.86×10^{-11}	4.25×10^{-10}

Table A.3: Additional Idiosyncratic Volatility (IVOL) Tests

Panel A of this table presents the Fama-Macbeth regression estimates after controlling for the effect of IVOL. We take the regression specifications in Table 3 and add IVOL and its interaction with MIS to all regressions. N(VIF>10) presents the number of cross-sectional regressions with variance inflation factor (VIF) values above 10. All coefficients are multiplied by 100 for better comparison. Panel B, reports benchmark adjusted returns for triple-sorted portfolios based on MIS, one of the four skewness measures (JACKPOT, LIDX, MAXRET, and ESKEW), and IVOL. Table A.1 defines all the variables. The portfolios are formed by independently sorting stocks into three MIS and three skewness portfolios at the end of every month. We then further sort the stocks in each skewness portfolio into three IVOL portfolios and compute the value-weighted four-factor Carhart (1997) alphas of the 27 intersecting portfolios. For brevity, we only report the results for terciles 1 and 3. Standard errors are adjusted following the Newey & West (1987) approach using a lag of 6. The sample excludes penny stocks and covers January 1963 to December 2015, except for sorts based on ESKEW, which start in January 1988. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Panel A: Continuous Variables				
MIS	-0.460***	-0.452***	-0.447***	-0.464***
	(-12.30)	(-13.49)	(-12.39)	(-12.36)
IVOL	-0.251***	-0.269***	-0.198**	-0.273***
	(-3.93)	(-5.16)	(-2.26)	(-3.99)
JACKPOT	-0.004			
	(-0.05)			
LIDX		0.053		
		(0.83)		
MAXRET			-0.037	
			(-0.63)	
ESKEW				-0.115*
				(-1.65)
MIS × IVOL	-0.185***	-0.171***	-0.172***	-0.211***
	(-6.67)	(-6.10)	(-3.79)	(-8.58)
MIS × JACKPOT	-0.094**			
	(-2.05)			
MIS × LIDX		-0.065***		
		(-2.81)		
MIS × MAXRET			-0.040	
			(-0.97)	
MIS × ESKEW				0.004
				(0.13)
Log(ME)	-0.121***	-0.113***	-0.125***	-0.118***
	(-3.98)	(-3.88)	(-3.94)	(-3.61)
Log(B/M)	0.269***	0.259***	0.267***	0.284***
	(4.51)	(4.43)	(4.49)	(4.56)
RET[-12,-2]	0.485***	0.480***	0.487***	0.404***
	(3.94)	(4.00)	(3.99)	(3.13)
RET[-1,0]	-4.254***	-4.260***	-4.206***	-4.261***
	(-11.27)	(-11.23)	(-10.58)	(-10.62)
N(VIF>10)	4	0	37	21
Average Number of Observations	3054	3054	3054	3133
Average Adjusted R2	0.050	0.050	0.048	0.048

Table A.3 (Continued)

Panel B: Triple Sorts										
		Low IVOL			High IVOL			High IVOL - Low IVOL		
		Most Underpriced	Most Overpriced	Most Overpriced - Most Underpriced	Most Underpriced	Most Overpriced	Most Overpriced - Most Underpriced	Most Underpriced	Most Overpriced	Most Overpriced - Most Underpriced
JACKPOT	Low	0.11*	-0.07	-0.18	0.21***	-0.43***	-0.64***	0.11	-0.36**	-0.46***
		(1.77)	(-0.69)	(-1.55)	(2.60)	(-3.66)	(-4.63)	(0.91)	(-2.31)	(-2.64)
	High	0.89***	-0.38**	-1.27***	-0.14	-1.7***	-1.56***	-1.02***	-1.31***	-0.29
		(5.38)	(-2.29)	(-7.18)	(-0.58)	(-7.78)	(-6.50)	(-4.18)	(-6.59)	(-1.05)
	High - Low	0.78***	-0.31	-1.09***	-0.35	-1.27***	-0.92***	-1.13***	-0.96***	0.17
		(4.41)	(-1.55)	(-5.36)	(-1.44)	(-5.34)	(-3.50)	(-4.32)	(-4.02)	(0.53)
LIDX	Low	0.12**	-0.05	-0.16	0.30***	-0.38***	-0.68***	0.18	-0.33**	-0.51***
		(2.03)	(-0.45)	(-1.46)	(3.21)	(-3.27)	(-4.87)	(1.44)	(-2.12)	(-2.99)
	High	0.76***	-0.34**	-1.11***	-0.42*	-1.7***	-1.28***	-1.19***	-1.36***	-0.17
		(5.52)	(-2.13)	(-5.77)	(-1.9)	(-6.96)	(-4.46)	(-4.72)	(-5.97)	(-0.52)
	High - Low	0.65***	-0.30	-0.94***	-0.72***	-1.32***	-0.60*	-1.37***	-1.02***	0.35
		(4.12)	(-1.44)	(-3.99)	(-3.14)	(-5.08)	(-1.88)	(-4.96)	(-3.72)	(0.90)
MAXRET	Low	0.12*	-0.09	-0.22*	0.28***	-0.28**	-0.56***	0.15	-0.19	-0.35*
		(1.74)	(-0.86)	(-1.83)	(2.80)	(-2.17)	(-3.77)	(1.23)	(-1.35)	(-1.93)
	High	0.45***	-0.53***	-0.98***	-0.34	-1.65***	-1.31***	-0.79***	-1.12***	-0.33
		(3.38)	(-4.02)	(-5.84)	(-1.50)	(-7.40)	(-4.80)	(-3.25)	(-4.88)	(-1.09)
	High - Low	0.32*	-0.44**	-0.76***	-0.62**	-1.37***	-0.75**	-0.94***	-0.93***	0.02
		(1.95)	(-2.42)	(-3.64)	(-2.44)	(-5.12)	(-2.48)	(-3.48)	(-3.40)	(0.05)
ESKEW	Low	0.12*	-0.03	-0.16	0.30***	-0.70***	-1.01***	0.18	-0.67***	-0.85***
		(1.95)	(-0.31)	(-1.33)	(2.79)	(-5.73)	(-6.65)	(1.30)	(-3.95)	(-4.39)
	High	0.60***	-0.44**	-1.04***	0.09	-1.67***	-1.76***	-0.51*	-1.23***	-0.72**
		(3.71)	(-2.37)	(-5.01)	(0.32)	(-6.49)	(-5.33)	(-1.74)	(-5.27)	(-2.06)
	High - Low	0.47**	-0.41*	-0.88***	-0.22	-0.97***	-0.75**	-0.69**	-0.56**	0.13
		(2.54)	(-1.73)	(-3.56)	(-0.77)	(-3.58)	(-2.09)	(-2.17)	(-2.05)	(0.32)

Table A.4: Effect of Investor Overreaction on Fama-Macbeth Regression Estimates

This table presents the Fama-Macbeth regression estimates within subsamples of stocks facing high and low levels of investor overreaction. We take the regression specifications in Table 3 and estimate them separately for each group. In Panel A, we sort stocks each month based on their TURNOVER values measured at the end of the previous month and allocate them into two groups using the cross-sectional median as the breakpoint. TURNOVER is defined as total trading volume divided by shares outstanding. In Panel B, the subsamples include stocks with an earnings announcement during the previous month and those without. For brevity, we only report the coefficients on MIS and the interaction terms. All independent variables in the regressions are standardized to a standard deviation of 1 and are winsorized at the 0.5 and 99.5 percentiles. Standard errors are adjusted for heteroscedasticity and autocorrelation following the Newey & West (1987) approach using a lag of 6. The sample excludes penny stocks and covers January 1963 to December 2015, except for the regression that includes ESKEW, which starts in January 1988. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Turnover				
		Low TURNOVER	High TURNOVER	High - Low
JACKPOT	MIS	-0.005***	-0.006***	-0.001
		(-12.81)	(-11.32)	(-1.52)
	MIS × JACKPOT	-0.002***	-0.004***	-0.002*
		(-2.77)	(-4.86)	(-1.80)
	Avg N	1416	1424	
LIDX	MIS	-0.004***	-0.005***	-0.001**
		(-14.74)	(-12.28)	(-2.04)
	MIS × LIDX	-0.001***	-0.002***	-0.001***
		(-5.33)	(-7.33)	(-3.00)
	Avg N	1399	1407	
MAXRET	MIS	-0.004***	-0.005***	-0.001*
		(-11.33)	(-10.19)	(-1.65)
	MIS × MAXRET	-0.001***	-0.002***	-0.001**
		(-2.99)	(-6.07)	(-2.10)
	Avg N	1399	1407	
ESKEW	MIS	-0.004***	-0.006***	-0.002***
		(-12.79)	(-12.59)	(-3.51)
	MIS × ESKEW	-0.001***	-0.002***	-0.001*
		(-2.94)	(-4.56)	(-1.80)
	Avg N	1472	1491	

Table A.4 (Continued)

Panel B: Earnings Announcement				
		Recent Announcement	No Recent Announcement	Announcement - No Announcement
JACKPOT	MIS	-0.004***	-0.006***	-0.002***
		(-7.78)	(-11.66)	(-2.75)
	MIS × JACKPOT	-0.003***	-0.003***	0.000
		(-3.38)	(-3.87)	(0.00)
	Avg N	983	2042	
LIDX	MIS	-0.004***	-0.005***	-0.001*
		(-9.34)	(-13.39)	(1.76)
	MIS × LIDX	-0.001***	-0.002***	-0.001***
		(-4.39)	(-6.86)	(-2.70)
	Avg N	982	2041	
MAXRET	MIS	-0.004***	-0.005***	-0.001
		(-7.74)	(-11.75)	(-1.49)
	MIS × MAXRET	-0.003***	-0.002***	0.001*
		(-6.32)	(-6.21)	(1.70)
	Avg N	983	2042	
ESKEW	MIS	-0.004***	-0.006***	-0.002***
		(-7.66)	(-12.06)	(-2.77)
	MIS × ESKEW	-0.001*	-0.001***	0.000
		(-1.72)	(-3.46)	(0.00)
	Avg N	961	1999	