Essays of Asset Pricing

BEI CHEN

presented to

THE DISCIPLINE OF FINANCE, BUSINESS SCHOOL

THE UNIVERSITY OF SYDNEY, AUSTRALIA

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF

Doctor of Philosophy

2021

Statement of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Signature

Bei Chen

Publications

Some of the work in this thesis has been published in refereed journals or is currently under the peer-review process in refereed journals.

The work that forms the basis of Chapter 1 – "Does the options market underreact to firms' left-tail risk?", is currently under revise & resubmit as follows:
Chen, B., and Gan, Q., 2021, Does the options market underreact to firms' left-tail risk?, *Journal of Financial and Quantitative Analysis* (revise-and-resubmit).

The work that forms the basis of Chapter 2 – "Anticipating Jumps: Decomposition of Straddle Returns Around Earnings Announcements", is currently under revise & resubmit as follows:
Chen, B., and Gan, Q., 2021, Anticipating Jumps: Decomposition of Straddle Returns Around Earnings Announcements, *Journal of Banking and Finance* (revise-and-resubmit).

The work that forms the basis of Chapter 3 – "Measuring gambling activity in options market", is currently published as follows:

Chen, B. and Gan, Q., 2021, "Measuring gambling activity in options market", *Review of Behavioral Finance*, *forthcoming.*

Authorship Attribution Statement

This is to certify that I have made a substantial original contribution to the above three co-authored papers, including research idea generation, empirical data analysis, and drafting.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Student Name, Signature, Date

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Supervisor Name, Signature, Date

Acknowledgements

I would like to sincerely thank Dr. Quan Gan at the University of Sydney Business School during my PhD study. His insightful feedback pushed me to sharpen my thinking and brought my work to a higher level. I would also like to express my deepest gratitude to my auxiliary supervisor, A/Prof. Eliza Wu, for her help and guidance.

I gratefully acknowledge the financial support provided by the University of Sydney Business School and Rich Data Corporation. I am also indebted to all the course coordinators who have provided me the great opportunity to teach under your guidance and with those wonderful colleagues.

I would like to thank my friends, research centre mates, and colleagues – Yiyuan Cao, Maiying Sui, Zhou Zhou, and Ke Peng for a cherished time spent together in the research centre, and in social settings. I could not have completed this dissertation without the support of them who provided stimulating discussions as well as happy distractions to rest my mind outside of my research.

Most importantly, I like to sincerely thank my parents and sister for their support over the past four years. To my husband, thank you for your love, support, encouragement, and visits. To my puppy Hachi, I cannot get through the lonely and tough times in Sydney without you. I dedicate this thesis to them.

Abstract

This dissertation explores issues related to asset pricing anomalies by focusing on options market. It consists of 3 chapters. In Chapter 1, we find that firms' left-tail risk is a strong positive predictor of future bear spread returns, suggesting that the options market underreacts to firms' left-tail risk and the downside protection provided by bear spreads is not adequately priced. We provide a behavioral explanation for this phenomenon. We find that the underreaction to firms' left-tail risk is stronger when the underlying stocks experience larger recent losses, are closer to their 52-week lows, and have higher information uncertainty. The options market's underreaction to firms' left-tail risk mainly happens in high investor sentiment periods.

In Chapter 2, we decompose and analyze the straddle returns around firms' earnings announcements. Previous study shows that delta-neutral straddles earn positive returns around earnings announcements, indicating an underpricing of earnings-induced risk. This study uses a volatility-jump decomposition to analyze the driving components of the delta-neutral straddle returns. We find that the volatility component consistently generates positive returns. The jump component's return is positive over the pre-announcement period and becomes negative after announcement. Our findings suggest that options market anticipates earnings-induced jumps. The

overall pattern of delta-neutral straddle's cumulative return is mainly driven by its jump component but the positive cumulative return after announcement is mainly driven by its volatility component.

In Chapter 3, we propose a gambling activity measure by jointly considering open interest and moneyness of out-of-the-money individual equity call options. The new measure, $CallMoney$, captures excessive optimism during the dot-com bubble, the oil price bubble, and the pre-GFC stock market bubble. $CallMoney$ robustly and negatively predicts both out-of-the-money and at-the-money call option returns cross-sectionally. The option return predictability of $CallMoney$ is stronger when stock price is further from its 52-week high, capital gains overhang is lower, and when information uncertainty of the underlying stock is higher. $CallMoney$ also robustly and negatively predicts cross-sectional stock returns. Comparing to lottery-like-payoffs based (indirect) gambling measures, $CallMoney$ performs better at predicting both option and stock returns.

Contents

List of Figures

List of Tables

Chapter 1

Does the options market underreact to firms' left-tail risk?

1.1 Introduction

Loss aversion plays an important role in economic decisions. The utility of a loss-averse investor is steeper for losses than for gains (Kahneman and Tversky, 1979; Tversky and Kahneman, 1991). Berkelaar et al. (2004) and Jarrow and Zhao (2006) show that the optimal portfolios for loss-averse investors include hedge and insurance against left-tail risk. The hedging demand of left-tail risk has its profound impact on asset prices. A growing body of literature investigates the impact of left-tail risk on the cross section of stock returns. Lu and Murray (2019) use the returns of bear spread constructed on Standard & Poor's (S&P) 500 index options to capture bear market risk and find that bear market risk is priced in the cross section of stock returns. Kelly and Jiang (2014) construct a tail risk facor by identifying the common fluctuation of crash events for individual firms and find that stocks with higher tail-risk factor loadings earn higher future returns. By contrast, Atilgan et al. (2020*a*) show that the risk-return tradeoff between firms' left-tail risk and stock returns breaks down as firms' left-tail risk and future stock returns have a negative relation, generating the "left-tail momentum". The left-tail momentum indicates that equity investors underreact to firms' left-tail risk.

In this paper, we present evidence of underreaction to firms' left-tail risk in options market by analyzing the returns of bear spreads constructed on equity options. We document a positive relation between firms' left-tail risk and future bear spread returns, suggesting that the protection against firms' left-tail risk provided by bear spreads is underpriced. The bear spreads on high left-tail risk stocks earn a monthly five-factor alpha 0.84% to 1.31% higher than the bear spreads on low left-tail risk stocks.

We explore several behavioral explanations to our main finding. First, investors may underestimate the persistence of losses and underprice the risk protection provided by bear spreads. Atilgan et al. (2020*a*) show that the left-tail momentum is stronger for the stocks that experience recent large losses. They argue that investors anticipate short-term stock price mean-reversion and underestimate the persistence in left-tail risk. Empirical evidence (George and Hwang, 2004; Driessen et al., 2013) also show that stock losses persist and option implied volatilities decrease, when stock prices are close to the 52-week low. Investors with anchoring bias underestimate the probability of downward breakthroughs toward new 52-week lows. We find that the positive relation between firms' left-tail risk and future bear spread returns is stronger when the underlying stocks experience large recent losses (left-tail momentum effect) and when stock prices are close to the 52-week low (anchoring effect). The anchoring effect alone may not offer an explanation to our main finding. But together with the left-tail momentum effect, we have strong evidence that investors' underestimation of loss persistence is one of the main sources to the underpricing of bear spreads on high left-tail risk stocks.

Second, information uncertainty may contribute to the underreaction to firms' left-tail risk. Hong et al. (2000) show that firm-specific, negative information difusses slowly and generates stock return momentum. Zhang (2006*a*,*b*) finds that information uncertainty amplifies investors' and analysts' underreaction to new information and causes greater price drift. We find that the underreaction to firms' left-tail risk is more pronounced when information uncertainty of a firm is higher. When information uncertainty is high, rational investors should demand more downside protection rather than less. Our finding of stronger bear spread underpricing when information uncertainty is high suggests that such underpricing is likely to be driven by behavioral biases.

Last, investor sentiment can be another behavioral explanation to the underpricing of bear spreads. High investor sentiment may reduce the perceived value of downside protection, causing underpricing of bear spreads. The literature shows that high investor sentiment may cause unwarranted investor optimism, overvalued stocks, and underestimated risk. Yu and Yuan (2011) find that the positive mean-variance tradeoff breaks down during high market sentiment periods. Stambaugh et al. (2012) show that a large set of anomalies in cross-sectional stock returns are amplified by high investor sentiment. Han (2008) shows that risk hedging demand, reflected by index options' implied volatility smile and risk-neutral skewness, decreases when market sentiment is high. Using Baker and Wurgler (2006) market-based sentiment index to measure investor sentiment, we find that the underreaction to firms' left-tail risk in options market mainly happens in high investor sentiment periods. While prior literature (Baker and Wurgler, 2006; Stambaugh et al., 2012; Byun and Kim, 2016) shows that high sentiment leads to stronger overpricing of risky assets; we show, as two sides of the same coin, that high sentiment also leads to stronger underpricing of bear spread, an option strategy providing downside protection.

Our rationale of using bear spreads instead of out-of-the-money (OTM) or deep-out-of-the-money (DOTM) put options to examine left-tail risk is similar to Lu and Murray (2019). Bear spread, as a popular option trading strategy, is frequently used to hedge against left-tail risk. A bear spread is comprised of opposite positions in two put options. The long-short positions help mitigate the options market forces that simultaneously impact options at different strike prices. For example, Garleanu et al. (2008) show that demand pressure could cause positively correlated option price deviations from the fundamental values across strike prices. Using OTM or DOTM options alone is more prone to such options market forces and thus provides less desirable identification of left-tail risk. Moreover, from the option valuation point of view, OTM or DOTM option prices are determined by the relevant discounted conditional expectation times the tail probability, using the risk-neutral probability measure (Bates, 1991). Therefore, OTM or DOTM option prices not only capture the tail probability, but also capture the expected tail distribution of the underlying asset price. By contrast, the scaled bear spread price can be interpreted as the discounted tail probability, representing the Arrow-Debreu state price of left-tail events that is comparable in cross section.

Our study contributes to the growing literature of tail risk and asset pricing (Kelly and Jiang, 2014; Van Oordt and Zhou, 2016; Chabi-Yo et al., 2018; Lu and Murray, 2019; Atilgan et al., 2020*a*) by documenting a positive relation between firms' left-tail risk and future bear spread returns. Contingent claims traded in options market offer unique opportunities to isolate, hedge and analyze tail risk. Using bear spread option strategy, we show that the risk-return tradeoff breaks down as the options market underreacts to firms' left-tail risk. Option traders' behavioral biases help explain such underreaction. Our study highlights that although investors frequently emphasize the importance of tail risk management, the protection against downsides is likely underpriced in the options market. The underpricing of bear spreads shows that adequately pricing tail risk in financial markets is more challenging than merely recognizing its importance.

The remainder of the paper proceeds as follows. Section 1.2 describes the data, the construction of bear spreads and variables. Section 1.3 presents the main empirical results. Section 1.4 provides behavioral explanations for the main findings. Section 1.5 presents further discussion and Section 1.6 concludes.

1.2 Data

1.2.1 Sample construction

Our sample period is from January 1996 to December 2017. We obtain stock price and accounting data from the Center for Research in Security Prices (CRSP) and Compustat. Option data are from OptionMetrics, which include daily closing bid and ask prices, open interest, volume, implied volatility, and option Greeks. To avoid the bid–ask bounce, the mid points of bid and ask prices are used to compute option returns.

Following previous literature (e.g. Kelly and Jiang (2014), Gao et al. (2018), and Ruan (2020*a*)), five filters are applied to the option data: (1) The option prices are at least \$0.125. (2) the underlying stock prices are at least \$5. (3) options must have nonmissing bid and ask price quotes and positive open interests. (4) bid and ask prices must satisfy basic arbitrage bounds to filter out erroneous observations. Arbitrage boundaries include: bid > 0, bid < ask, bid ≤ strike and ask ≥ max(0, strike price – stock price); (5) options' embedded leverage calculated following Frazzini and Pedersen (2012) is not in the top or bottom 1% of the distribution.

The risk factors, including Fama and French (1993) three factors (MKT, SMB, HML) and Carhart (1997) momentum factor (UMD) are from the Kenneth R. French data library. Coval and Shumway (2001) systematic volatility risk factor (Zb-strad-rf) data are from OptionMetrics. Zb-strad-rf factor is the excess return of a zero-beta at-the-money (ATM) S&P 500 index straddle.

1.2.2 Bear spread construction

A bear spread is constructed by taking a long position in one OTM put option, denoted as PUT_1, with price P_1, strike price K_1 and delta Δ_1 and a short position in a further OTM put option, denoted as PUT_2, with price P_2, strike price K_2 and delta Δ_2 ($K_1 > K_2$ and $\Delta_1 < \Delta_2$). The bear spread generates a payoff of $K_1 - K_2$ when the stock price at expiration is below K_2 and zero when the stock price at expiration is above K_1. The bear spread payoff linearly decreases from $K_1 - K_2$ to zero for stock price between K_2 and K_1.

Choosing K_1 and K_2 in empirical studies deserves careful consideration. As discussed in Lu and Murray (2019), if the bear region boundary K_2 is set to be at a constant percentage below the forward price, the bear region would correspond to left-tail events with different probability when the underlying assets possess different volatility levels. To address this issue, for index option bear spreads, Lu and Murray (2019) set K_2 and K_1 to be 1.5 and 1 standard deviation below the index forward price.

Since equity options have more sparse strike prices compared to index options, we use option deltas instead of strike prices to select put options in bear spreads. Extensive literature uses Black-Scholes deltas to identify options with the same moneyness across assets as the absolute delta approximates the probability that an option will be in the money at expiration (Bollen and Whaley, 2004; Driessen et al., 2009; Bali and Murray, 2013; Jin et al., 2012; Kelly, Pástor and Veronesi, 2016; Kelly, Lustig and Van Nieuwerburgh, 2016). The typical ranges of OTM and DOTM put option deltas are $[-0.40, -0.20)$ and $[-0.20, 0]$ (e.g. Kelly and Jiang (2014); Muravyev (2016)). We construct each of bear spreads by a long (short) position in PUT_1 (PUT_2) as the OTM (DOTM) put option with Δ_1 (Δ_2) closest to -0.30 (-0.10), the midpoint of the OTM (DOTM) delta range.

A bear spread has a negative delta ($\Delta_1 - \Delta_2$), embedding an equivalent short position in the underlying stock. Therefore, unhedged bear spread returns also capture movements of the underlying stock. Atilgan et al. (2020*a*) document a negative relation between stocks' left-tail risk and their future returns, the left-tail momentum. To remove the contribution of the underlying stocks' left-tail momentum, we use delta-hedged bear spreads in our empirical tests.[1] We use static delta-hedging which is also used by other equity option studies (Goyal and Saretto, 2009; Bali and Murray, 2013; Byun and Kim, 2016). Following Goyal and Saretto (2009) and Kelly and Jiang (2014), we use one-month options to construct bear spreads, and form delta-hedged bear spreads on the first trading day immediately following the third Saturday in month t and close all positions at the option maturity on the third Friday in month $t+1$. Our sample consists of $155,003$ cross-sectional monthly returns of delta-hedged bear spreads.

1.2.3 Left-tail risk measures

We estimate left-tail risk using two standard measures following Atilgan et al. (2020a): value-at-risk (VaR) and expected shortfall (ES). VaRx (ESx) is calculated as (the average of the observations that are less than or equal to) the x percentile of the daily returns over the past 250 trading days. As the left-tail loss measures are typically negative, we multiple these measures by -1 so that higher value of VaR or ES corresponds to higher left-tail risk. At the portfolio formation date in month t, we compute VaR1, VaR5, ES1, and ES5 with the restriction that there are at least 200 non-missing past trading day returns.

[1] Lu and Murray (2019) show in a theoretical model that delta-hedged bear spread returns expose only to the left-tail risk.

1.2.4 Control variables

We construct three groups of control variables that are commonly used to examine the cross section of equity option returns in the literature (Goyal and Saretto, 2009; Bali and Murray, 2013; Cao and Han, 2013; Byun and Kim, 2016; Ruan, 2020*a*).

First, we construct three variables related to firm characteristics. Firm size ($SIZE$) is the natural logarithm of firm market capitalization observed at the end of month $t-1$. Book-to-market ratio (BTM) is the ratio of a firm's net asset's book value at the previous fiscal year-end to the market capitalization of the stock at the end of month $t-1$. Firm leverage (DTA) is the ratio of a firm's total liability to the book value of total asset at the previous fiscal year-end.

Second, we construct six variables related to stock returns and stock trading activities. Momentum (MOM) is the cumulative stock return from month $t-6$ to month $t-2$. Short-term reversal (REV) is the stock return in month $t-1$. Illiquidity ratio ($ILLIQ$) is the natural logarithm of the average ratio of the absolute daily stock return to its daily dollar trading volume multiplied by 10^8 in month $t-1$. Idiosyncratic volatility of stock return ($IVOL$) is the standard deviation of the residuals of the daily stock excess return regressed on daily market excess return in month $t-1$. Stock return skewness ($SKEW$) and kurtosis ($KURT$) are calculated using last year's daily stock return data.

Finally, we construct three variables related to options. Variance risk premium (VRP) is the difference between the average implied volatility of ATM short-term options (with moneyness between 0.95 to 1.05 and 10 to 60 day-to-maturity) and the annualized last-quarter's daily stock return standard deviation observed at the end of month $t-1$. Volatility of volatility (VOV) is calculated following Baltussen et al. (2018) by scaling the standard deviation of ATM short-term option implied volatility by the average ATM short-term option implied volatility over month $t-1$. Risk-neutral skewness (RNS) at the end of month $t-1$ is calculated using OTM call and put options prices following Bakshi et al. (2003).

1.2.5 Summary statistics

Table 1.1 presents the summary statistics for delta-hedged bear spread returns, option characteristics for PUT_1 and PUT_2 in bear spreads, the left-tail risk measures and control variables. The reported mean, standard deviation, 25th percentile, median, and 75th percentile are computed as the time-series average of their cross-sectional values.

Table 1.1: Summary Statistics

This table presents the descriptive statistics (mean, standard deviation, 25th percentile, median, 75th percentile) for delta-hedged bear spread returns, characteristics of put options in the bear spreads, the left-tail risk measures, and control variables. Statistics are computed as the time-series averages of the monthly cross-sectional means, standard deviations and percentiles. The sample period is from January 1996 to December 2017.

	Mean	Std Dev	25th	Median	75th
Panel A: Delta-hedged bear spread returns					
	-0.17%	13.22%	-8.56%	-1.74%	6.77%
Panel B: Option characteristics					
PUT_1					
Delta	-0.30	0.05	-0.34	-0.30	-0.27
Implied Volatility	0.48	0.20	0.33	0.44	0.57
PUT_2					
Delta	-0.12	0.03	-0.14	-0.11	-0.09
Implied Volatility	0.53	0.22	0.37	0.49	0.63
Panel C: Left-tail risk measures					
VaR5	4.24%	1.57%	3.02%	3.99%	5.19%
VaR1	6.96%	2.82%	4.86%	6.44%	8.48%
ES5	6.15%	2.38%	4.34%	5.75%	7.51%
ES1	9.28%	4.41%	6.14%	8.33%	11.28%
Panel D: Control variables					
SIZE	22.23	1.58	21.09	22.18	23.30
BTM	4.51	49.70	0.35	0.69	1.43
DTA	0.18	0.19	0.02	0.14	0.28
MOM	0.15	0.34	-0.04	0.13	0.31
REV	0.03	0.13	-0.04	0.02	0.09
ILLIQ	-7.96	1.56	-9.03	-8.02	-6.89
IVOL	0.02	0.01	0.01	0.02	0.03
SKEW	0.24	1.22	-0.18	0.18	0.59
KURT	8.61	11.12	4.16	5.46	8.63
VRP	0.09	0.27	-0.06	0.09	0.25
VOV	0.10	0.06	0.06	0.09	0.12
RNS	-0.63	0.58	-0.95	-0.56	-0.22

In Panel A, both mean and median of delta-hedged bear spread returns are negative, consistent with the negative risk premium carried by the bear spreads as they provide insurance against left-tail risk. The 75th percentile is 6.77%, indicating that at least 25% of return observations are positive. The summary statistics reveal that although on average the delta-hedged bear spread returns are negative, there are still substantial occurrences of returns are positive.

In Panel B, PUT_1 (OTM put option) and PUT_2 (DOTM put option) have the median deltas of -0.30 and -0.11, close to our selection targets of -0.30 and -0.10 and the standard deviations deltas are moderately small, suggesting a satisfactory selection result of option pairs in bear spread construction. PUT_1's mean implied volatility is 48%, which is smaller than PUT_2's mean implied volatility of 53%. This is consistent with the typical shape of volatility smirk for put options (Xing et al., 2010).

In Panel C, $VaR5$ ($VaR1$) has a mean of 4.24% (6.96%), implying that on average there is a 5% (1%) probability that the daily loss that a firm experiences in the prior year exceeds 4.24% (6.96%). $ES5$ ($ES1$) has a mean of 6.15% (9.28%), which is larger than the mean of $VaR5$ ($VaR1$) as expected.[2]

In Panel D, the median BTM is 0.69 and the median DTA is 0.14. The median MOM is 0.13, and the median REV is 0.02. These are similar to those reported in Byun and Kim (2016). The median $ILLIQ$ is -8.02, similar to that reported in Cao and Han (2013). The median $IVOL$ is 0.02, same as reported in Atilgan et al. (2020*a*). The median $SKEW$ is 0.18 and the median $KURT$ is 5.46, similar to those reported in Ruan (2020*a*). The median VRP is 0.09, similar to that reported in Cao and Han (2013). The median VOV is 0.09, similar to that reported in Baltussen et al. (2018). The median RNS is -0.56, similar to that reported in Stilger et al. (2016).

[2] We also calculate correlations between the left-tail risk measures. The correlation coefficients are in the range of 0.72 to 0.96, similar to those in Atilgan et al. (2020*a*).

1.3 Empirical Analysis

1.3.1 Univariate portfolio analysis

We conduct univariate portfolio analysis to examine the relation between the left-tail risk measures and future delta-hedged bear spread returns. In month t, we form decile portfolios of delta-hedged bear spreads by sorting the underlying stocks based on one of the left-tail risk measures ($VaR1$, $VaR5$, $ES1$, $ES5$). The decile 10 (decile 1) portfolio contains delta-hedged bear spreads on stocks with the highest (lowest) left-tail risk.

We compute both equal-weighted and dollar-open-interest-weighted (DOI-weighted) hold-to-maturity monthly portfolio returns. The DOI weighting puts more weights on options or option strategies with higher liquidity (open interests). Our construction of DOI weighting is similar to Cao and Han (2013) and Gao et al. (2018). The dollar-open-interest weight is calculated on the bear spread formation date as the cost of each bear spread, multiplied by the minimum of the open interests of two put options in that bear spread:

$$\begin{aligned} DOI =&(PUT_1\, PRICE - PUT_2\, PRICE) \\ &\times min(PUT_1\, OPEN\, INTEREST, PUT_2\, OPEN\, INTEREST). \end{aligned}$$

DOI captures the maximum possible dollar open interest to form a bear spread.

Table 1.2 reports the time-series average monthly returns of each decile portfolio, together with the return spreads and their associated alphas between the highest and lowest ("10-1") left-tail risk decile portfolios. To adjust for serial correlation, robust Newey-West (1987) t-statistics are reported in brackets.

Panel A reports the results for equal-weighted portfolios. In the first column of Panel A, the decile 1 portfolio (with the lowest $VaR5$) has an average monthly return of -0.62%, while decile 10 portfolio (with the highest $VaR5$) has an average monthly return of 0.40%. The decile portfolio returns in general increase from decile 1 to decile 10. The "10-1" monthly return spread is 1.03% (t-statistic=3.60). The corresponding "10-1" CAPM alpha is 1.23% (t-statistic=4.59). The three-factor alpha after adjusting for the Fama-French three factors (MKT, SMB, and HML) is 1.16% (t-statistic=4.34). The four-factor alpha after adjusting the Fama-French three factors and the momentum factor (MKT, SMB, HML, and UMD) is 1.18% (t-statistic=4.62). The five-factor alpha after adjusting for the four factors and the systematic volatility factor (MKT, SMB, HML, UMD, and Zb-strad-rf) is 1.20% (t-statistic=4.48). The other three columns in Panel A exhibit similar patterns when portfolios are sorted on $VaR1$, $ES5$, and $ES1$.

In Panel B, the DOI-weighted portfolio returns have patterns similar to Panel A. In the first column, for example, the "10-1" monthly return spread is 1.04% (t-statistic=2.44). The corresponding "10-1" CAPM alpha is 1.28% (t-statistic=3.02). The three-factor alpha is 1.25% (t-statistic=2.92). The four-factor alpha is 1.35% (t-statistic=3.16) and the five-factor alpha is 1.31% (t-statistic=3.15). In Panels A and B, the five-factor alphas range from 0.84% to 1.31% per month.

The bear spreads with bear regions concentrated in left tails provide protection against rare events such as stock price crashes. In theory, option traders should pay

Table 1.2: Univariate Portfolio Analysis

This table reports the time-series average monthly returns for the delta-hedged bear spread decile portfolios sorted on the left-tail risk measures ($VaR5$, $VaR1$, $ES5$, $ES1$), along with the return spreads ("10-1") and the associated alpha spreads between decile 10 portfolio (with the highest left-tail risk metric) and decile 1 portfolio (with the lowest left-tail risk metric). Panels A and B present the results for equal-weighted and dollar-open-interest-weighted portfolio returns. CAPM alphas are calculated after adjusting for CAPM market risk factor. Three-factor alphas are calculated after adjusting for Fama-French three factors. Four-factor alphas are calculated after adjusting for Fama-French three factors and Carhart (1997) momentum factor. Five-factor alphas are calculated after adjusting for Fama-French three factors, Carhart (1997) momentum factor, and Coval and Shumway (2001) systematic volatility factor. Newey-West (1987) adjusted t-statistics are presented in parentheses. The sample period is from January 1996 to December 2017.

	Panel A: Equal-weighted				Panel B: Dollar-open-interest-weighted			
	VaR5	VaR1	ES5	ES1	VaR5	VaR1	ES5	ES1
1	-0.62%	-0.61%	-0.61%	-0.60%	-0.59%	-0.55%	-0.53%	-0.56%
2	-0.46%	-0.43%	-0.44%	-0.44%	-0.38%	-0.53%	-0.41%	-0.38%
3	-0.37%	-0.36%	-0.40%	-0.33%	-0.23%	-0.25%	-0.40%	-0.15%
4	-0.19%	-0.17%	-0.14%	-0.23%	-0.18%	-0.05%	0.10%	-0.04%
5	-0.13%	-0.13%	-0.19%	-0.02%	0.16%	0.03%	0.10%	0.07%
6	-0.11%	-0.07%	-0.06%	-0.05%	0.06%	0.13%	0.21%	-0.25%
7	0.18%	0.09%	0.13%	0.13%	0.12%	0.14%	0.10%	0.01%
8	0.17%	0.16%	0.20%	0.11%	0.14%	0.15%	0.09%	0.27%
9	0.09%	0.16%	0.13%	0.24%	0.01%	0.52%	0.12%	0.50%
10	0.40%	0.30%	0.30%	0.14%	0.45%	0.27%	0.46%	0.29%
10-1	1.03%***	0.91%***	0.91%***	0.74%***	1.04%**	0.83%**	0.99%**	0.85%***
(t-stat)	(3.60)	(3.33)	(3.41)	(3.48)	(2.44)	(2.26)	(2.35)	(2.61)
CAPM alpha	1.23%***	1.08%***	1.10%***	0.87%***	1.28%***	1.01%***	1.19%***	1.00%***
(t-stat)	(4.59)	(4.35)	(4.45)	(4.26)	(3.02)	(2.71)	(2.86)	(3.03)
Three-factor alpha	1.16%***	1.03%***	1.04%***	0.82%***	1.25%***	1.01%***	1.24%***	0.99%***
(t-stat)	(4.34)	(4.12)	(4.19)	(3.92)	(2.92)	(2.75)	(3.01)	(3.01)
Four-factor alpha	1.18%***	1.06%***	1.07%***	0.82%***	1.35%***	1.12%***	1.34%***	1.05%***
(t-stat)	(4.62)	(4.42)	(4.47)	(4.29)	(3.16)	(3.12)	(3.27)	(3.26)
Five-factor alpha	1.20%***	1.09%***	1.09%***	0.84%***	1.31%***	1.09%***	1.30%***	0.98%***
(t-stat)	(4.48)	(4.31)	(4.39)	(4.24)	(3.15)	(3.02)	(3.19)	(3.06)

adequately for such protection, accept negative risk premium, and expect negative future returns. However, empirically we find that high decile portfolios generate higher returns than low decile portfolios and the "10-1" return alphas are all statistically and economically significant in both panels, indicating significant underpricing of bear spreads when left-tail risk is high.

The positive relation between firms' left-tail risk and future delta-hedged bear spread returns deserves more analyses. Since there exist strong positive correlations between the four left-tail risk measures, and the empirical distribution of $VaR5$ is more well-behaved in terms of being closer to normality compared to the other three measures (Atilgan et al., 2020a), we present our subsequent empirical results using $VaR5$ as the key left-tail risk measure except for Section 1.5.1.

1.3.2 Bivariate portfolio analysis

We investigate whether the positive relation between the underlying stocks' left-tail risk measures and future bear spread returns can be explained by other (firm, stock or option related) variables using the dependent (conditional) bivariate portfolio sorting method.

In month t, we first form decile portfolios of delta-hedged bear spreads by sorting the underlying stocks based on one of the control variables in Section 1.2.4. Then, within each decile, we form decile portfolios based on the left-tail risk measure $VaR5$. Each left-tail risk decile portfolio is then averaged across the control variable deciles.

Table 1.3 reports equal-weighted and DOI-weighted returns of decile portfolios in

Panels A and B, together with the raw and risk-adjusted return spreads ("10-1") between the highest and lowest $VaR5$ decile portfolios. Newey-West (1987) t-statistics are reported in brackets.

For equal-weighted portfolio results reported in Panel A, the "10-1" return spreads and their corresponding alphas are positive and statistically significant. The return spreads range from 0.64% (t-statistic = 3.05) to 1.09% (t-statistic = 4.99) per month. The "10-1" CAPM alphas range from 0.78% to 1.25%. The three-factor alphas range from 0.72% to 1.18%. The four-factor alphas range from 0.67% to 1.19%, and the five-factor alphas range from 0.69% to 1.18%. For DOI-weighted portfolio returns reported in Panel B, the "10-1" return spreads range from 0.50% (t-statistic = 1.49) to 1.16% (t-statistic = 3.44) per month. The "10-1" CAPM alphas range from 0.77% to 1.32%. The three-factor alphas range from 0.71% to 1.28%. The four-factor alphas range from 0.67% to 1.23% and the five-factor alphas range from 0.65% to 1.26%. The "10-1" return spreads and their corresponding alphas are all positive, and statistically significant in most cases.

The results indicate that after controlling for various characteristic variables, there is still a strong positive relation between firms' left-tail risk measure $VaR5$ and future returns of delta-hedged bear spreads. The return predictability of $VaR5$ cannot be explained by the characteristic variables commonly used by the literature.

1.3.3 Fama-MacBeth regressions

We perform Fama-MacBeth (1973) regressions to formally test the positive cross-sectional relation between $VaR5$ and future bear spread returns. The dependent

Table 1.3: Bivariate Portfolio Analysis

This table presents results for the equal-weighted delta-hedged bear spread portfolios (in Panel A) and the dollar-open-interest-weighted delta-hedged bear spread portfolios (in Panel B) based on bivariate dependent sorts of one (firm-specific, stock-related or option related) characteristic variable and $VaR5$. In month t, decile portfolios of delta-hedged bear spreads are formed by sorting underlying stocks based on one of the characteristic variables. Then within each decile, additional decile portfolios of delta-hedged bear spreads are formed by sorting underlying stocks based on $VaR5$ observed in the previous year. Each $VaR5$ decile portfolio is then averaged over the control characteristic deciles. This table reports the raw returns for equal-weighted decile portfolios thus obtained, and the return spreads and associated alpha spreads for decile 10 portfolio (highest $VaR5$) and decile 1 portfolio (lowest $VaR5$). CAPM alphas are adjusted for CAPM market risk factor. Three-factor alphas are adjusted for Fama-French three factors. Four-factor alphas are adjusted for Fama-French three factors and Carhart (1997) momentum factor. Five-factor alphas are calculated after adjusting for Fama-French three factors, Carhart (1997) momentum factor, and Coval and Shumway (2001) systematic volatility factor. Newey-West (1987) adjusted t-statistics are presented in parentheses. The sample period is from January 1996 to December 2017.

Panel A: Equal-weighted portfolios												
	SIZE	BTM	DTA	MOM	REV	ILLIQ	IVOL	SKEW	KURT	VRP	VOV	RNS
1	-0.56%	-0.54%	-0.51%	-0.44%	-0.58%	-0.60%	-0.61%	-0.61%	-0.35%	-0.59%	-0.63%	-0.59%
2	-0.47%	-0.38%	-0.28%	-0.29%	-0.44%	-0.36%	-0.37%	-0.35%	-0.27%	-0.39%	-0.45%	-0.36%
3	-0.16%	-0.31%	-0.36%	-0.24%	-0.34%	-0.34%	-0.33%	-0.38%	-0.40%	-0.37%	-0.29%	-0.13%
4	-0.20%	-0.19%	-0.23%	-0.23%	-0.26%	-0.23%	-0.23%	-0.29%	-0.16%	-0.28%	-0.17%	-0.12%
5	-0.19%	-0.08%	-0.02%	-0.20%	-0.10%	-0.18%	-0.05%	-0.12%	0.05%	-0.20%	-0.18%	-0.01%
6	-0.09%	-0.01%	0.03%	0.05%	-0.06%	-0.07%	-0.01%	0.04%	0.04%	0.09%	0.02%	0.18%
7	0.11%	-0.13%	0.01%	0.03%	0.02%	0.01%	0.01%	0.10%	0.14%	-0.04%	0.12%	-0.01%
8	-0.02%	0.09%	0.03%	0.01%	0.07%	0.17%	-0.01%	0.23%	0.03%	0.02%	0.10%	-0.07%
9	0.24%	0.13%	0.14%	0.01%	0.11%	0.15%	0.11%	0.05%	0.01%	0.17%	0.05%	0.26%
10	0.30%	0.10%	0.10%	0.21%	0.51%	0.41%	0.40%	0.33%	0.47%	0.34%	0.17%	0.46%
10-1	0.86%***	0.67%***	0.65%***	0.64%***	1.09%***	1.01%***	1.02%***	0.94%***	0.82%**	0.93%***	0.79%***	0.85%***
(t-stat)	(3.89)	(3.65)	(2.97)	(3.05)	(4.99)	(4.12)	(5.57)	(3.73)	(1.94)	(3.56)	(2.85)	(3.23)
CAPM alpha	1.00%***	0.79%***	0.78%***	0.79%***	1.25%***	1.15%***	1.10%***	1.11%***	1.01%**	1.09%***	0.97%***	0.96%***
(t-stat)	(4.66)	(4.48)	(3.54)	(3.88)	(5.88)	(4.84)	(6.14)	(4.77)	(2.43)	(4.28)	(3.83)	(3.62)
Three-factor alpha	0.94%***	0.75%***	0.72%***	0.73%***	1.18%***	1.09%***	1.07%***	1.06%***	1.02%**	1.02%***	0.92%***	0.92%***
(t-stat)	(4.43)	(4.11)	(3.34)	(3.64)	(5.73)	(4.63)	(5.96)	(4.63)	(2.44)	(4.02)	(3.57)	(3.39)
Four-factor alpha	0.90%***	0.68%***	0.69%***	0.67%***	1.14%***	1.08%***	1.10%***	1.09%***	1.19%***	0.99%***	0.93%***	0.84%***
(t-stat)	(4.49)	(3.99)	(3.48)	(3.64)	(6.00)	(4.91)	(6.22)	(4.93)	(2.87)	(4.22)	(3.80)	(3.40)
Five-factor alpha	0.94%***	0.69%***	0.71%***	0.72%***	1.18%***	1.13%***	1.18%***	1.11%***	1.09%***	1.04%***	0.97%***	0.91%***
(t-stat)	(4.55)	(3.93)	(3.48)	(3.72)	(5.86)	(4.95)	(6.57)	(4.82)	(2.83)	(4.31)	(3.78)	(3.59)

Table 3 continued from previous page

Panel B: Dollar-open-interest-weighted portfolios												
	SIZE	BTM	DTA	MOM	REV	ILLIQ	IVOL	SKEW	KURT	VRP	VOV	RNS
1	-0.76%	-0.42%	-0.54%	-0.29%	-0.45%	-0.90%	-0.46%	-0.35%	-0.42%	-0.61%	-0.46%	-0.42%
2	-0.41%	-0.38%	-0.20%	-0.32%	-0.40%	-0.14%	0.06%	-0.27%	-0.29%	-0.01%	-0.25%	-0.58%
3	-0.28%	0.01%	0.06%	-0.13%	-0.17%	-0.42%	-0.32%	-0.40%	-0.32%	-0.37%	-0.28%	-0.20%
4	-0.30%	0.04%	-0.28%	-0.14%	-0.22%	-0.37%	-0.14%	-0.16%	-0.25%	-0.35%	-0.21%	-0.01%
5	-0.39%	-0.09%	0.00%	-0.40%	0.27%	-0.28%	-0.07%	0.05%	0.01%	0.04%	0.17%	0.16%
6	-0.23%	0.03%	0.42%	0.14%	-0.21%	0.06%	-0.04%	0.04%	0.35%	0.19%	0.20%	0.43%
7	-0.11%	-0.12%	0.07%	-0.08%	0.12%	-0.11%	-0.27%	0.14%	0.14%	0.07%	0.01%	-0.31%
8	0.08%	0.12%	-0.01%	-0.02%	0.01%	0.13%	-0.29%	0.03%	0.13%	0.19%	-0.07%	-0.18%
9	0.17%	0.08%	0.08%	-0.26%	0.07%	0.15%	0.12%	0.01%	0.06%	0.10%	0.05%	0.18%
10	0.21%	0.28%	0.18%	0.56%	0.60%	0.26%	0.57%	0.47%	0.21%	0.11%	0.26%	0.39%
10-1	0.97%***	0.50%	0.62%**	0.85%**	1.06%***	1.16%***	1.03%***	0.82%*	0.63%	0.72%*	0.72%*	0.84%**
(t-stat)	(3.26)	(1.49)	(2.27)	(2.45)	(2.91)	(3.44)	(3.41)	(1.94)	(1.60)	(1.77)	(1.82)	(2.30)
CAPM alpha	1.13%***	0.77%**	0.80%***	1.06%***	1.18%***	1.32%***	1.07%***	1.01%**	0.87%**	0.87%**	0.85%**	0.87%**
(t-stat)	(3.78)	(2.51)	(2.86)	(3.09)	(3.34)	(3.94)	(3.51)	(2.43)	(2.26)	(2.11)	(2.15)	(2.36)
Three-factor alpha	1.11%***	0.71%**	0.74%***	1.05%***	1.16%***	1.28%***	1.08%***	1.02%**	0.86%**	0.82%**	0.86%**	0.79%**
(t-stat)	(3.74)	(2.43)	(2.82)	(3.05)	(3.26)	(3.82)	(3.56)	(2.44)	(2.25)	(2.00)	(2.16)	(2.13)
Four-factor alpha	1.06%***	0.67%**	0.70%***	1.05%***	1.16%***	1.23%***	1.16%***	1.19%***	0.99%***	0.92%**	0.94%**	0.78%**
(t-stat)	(3.56)	(2.30)	(2.71)	(2.98)	(3.23)	(3.78)	(4.04)	(2.87)	(2.62)	(2.26)	(2.37)	(2.11)
Five-factor alpha	1.07%***	0.65%**	0.68%**	1.04%***	1.17%***	1.26%***	1.12%***	1.09%***	0.96%***	0.88%**	0.93%**	0.77%**
(t-stat)	(3.62)	(2.18)	(2.54)	(3.02)	(3.31)	(3.92)	(4.01)	(2.83)	(2.60)	(2.20)	(2.32)	(2.08)

variable is the hold-to-maturity monthly return of delta-hedged bear spread formed in month t, and the variable of interest is $VaR5$. The control variables are from Section 1.2.4.

Table 1.4 presents the time-series averages of the regression coefficient estimates, along with their Newey-West (1987) adjusted t-statistics in the brackets.

There are 14 regression models with their results reported in the corresponding columns. In model 1, we perform a univariate regression on $VaR5$. The coefficient is 0.201 (t-statistic=4.58), indicating a significantly positive relation between $VaR5$ and future delta-hedged bear spread return. In models 2-13, we perform regressions on $VaR5$ together with each of the control variables. The coefficients on $VaR5$ are all positive and significant at the 1% level. In model 14, we perform regressions on $VaR5$ together with all the control variables. The coefficient on $VaR5$ is 0.204 (t-statistic=3.82), still positive and significant.

In model 14, the coefficient on $SIZE$ is -0.004 (t-statistic=-4.37). Larger firms tend to attract more attention from investors, and the protection provided by bear spreads is likely more valuable to investors. Thus the bear spreads on larger firms tend to generate lower future returns. The economic intuition is consistent to Goyal and Saretto (2009), Cao and Han (2013), and Byun and Kim (2016).

The coefficients on $ILLIQ$ and $IVOL$ are -0.003 (t-statistic=-4.59) and -0.284 (t-statistic=-2.56) respectively. The protection of left-tail risk provided by bear spreads is more valuable when such risk is harder to hedge for low liquidity or high idiosyncratic volatility stocks. Thus for highly illiquid and high idiosyncratic volatility stocks, future bear spread returns are lower. The economic intuition is

Table 1.4: Fama-MacBeth Regressions

The table presents the results of Fama and MacBeth (1973) regressions of monthly delta-hedged bear spread returns on $VaR5$ and the control variables. Coefficients are time-series averages and the associated Newey and West (1987) t-statistics are reported in parentheses. The last row reports the average adjusted R-squared. Model 1 is the univariate regression of delta-hedged bear spread returns on $VaR5$. Models 2-13 are regressions of delta-hedged bear spread returns on $VaR5$ and each of the control variables. Model 14 is the regression of delta-hedged bear spread returns on $VaR5$ and all the control variables. The sample period is from January 1996 to December 2017.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
VaR5	0.201***	0.210***	0.118***	0.128***	0.172***	0.186***	0.239***	0.235***	0.182***	0.183***	0.143***	0.154***	0.184***	0.204***
	(4.58)	(4.64)	(2.63)	(2.74)	(3.98)	(4.26)	(5.13)	(5.11)	(4.18)	(4.15)	(3.21)	(3.46)	(3.69)	(3.82)
SIZE		0.000												-0.004***
		(0.82)												(-4.37)
BTM			0.000											0.000
			(1.04)											(0.97)
DTA				-0.002										-0.003
				(-1.16)										(-1.17)
MOM					-0.002									-0.001
					(-1.43)									(-0.68)
REV						-0.004								0.000
						(-1.03)								(0.05)
ILLIQ							-0.001***							-0.003***
							(-3.55)							(-4.59)
IVOL								-0.113***						-0.284**
								(-2.85)						(-2.56)
SKEW									-0.001*					0.000
									(-1.80)					(0.15)
KURT										0.000**				0.000
										(-2.53)				(-1.37)
VRP											-0.002			-0.014*
											(-0.90)			(-1.81)
VOV												-0.008		0.000
												(-0.98)		(0.04)
RNS													-0.001	-0.001
													(-1.10)	(-0.67)
Intercept	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R^2	1.32%	1.47%	1.33%	1.41%	1.90%	1.83%	1.54%	1.81%	1.50%	1.53%	1.68%	1.47%	1.58%	4.68%

consistent to Cao and Han (2013).

The coefficient on VRP is -0.014 (t-statistic=-1.81), suggesting a negative variance premium of bear spreads. This is consistent to the positive vega exposure that bear spreads carry.

The results of Table 1.4 indicate a strong and robust positive relation between firms' left-tail risk and future delta-hedged bear spread returns after controlling for various combinations of control variables. Our findings suggest that options market underreacts to firms' left-tail risk and underprices bear spreads on high left-tail risk stocks, resulting in the high future returns of the bear spreads when firm' left-tail risk is high.

1.4 Why does the options market underreact to firms' left-tail risk?

The positive relation between firms' left-tail risk and future bear spread returns indicates that options market underreacts to firms' left-tail risk. In previous sections, we show that such underreaction cannot be explained by the well-known characteristic variables or risk factors. In this section, we analyse the potential behavioral explanations to the options market's underreaction to firms' left-tail risk.

1.4.1 Persistence of losses

A behavioral explanation to the positive relation between firms' left-tail risk and future bear spread returns is that option traders underestimate persistence of losses. From this perspective, we analyze two effects: the left-tail momentum effect and the anchoring effect. Atilgan et al. (2020a) show that investors underestimate loss persistence and such underestimation contributes to the left-tail return momentum, a negative relation between firms' left-tail risk and expected equity returns. George and Hwang (2004) show that anchoring behaviour helps explain loss momentum around the 52-week low, albeit the anchoring effect is weaker than 52-week high. Driessen et al. (2013) show that option implied volatilities decrease when stock prices approach their 52-week low, suggesting that investors underestimate persistence in risk due to anchoring bias. The anchoring effect by itself cannot directly explain our main finding. By analyzing its impact on bear spread returns and combining with the results from the left-tail momentum effect help us to pin down the underestimation of loss persistence as one of the main behavioral sources to the positive relation between firms' left-tail risk and future bear spread returns.

To analyse the impact of these two effects, we construct two corresponding measures: 1) $\Delta VaR5$, change of $VaR5$, which is the difference between the month t and the month $t-1$ $VaR5$. A positive $\Delta VaR5$ implies that the 5th percentile loss computed in month t is higher than the loss computed in month $t-1$, indicating larger recent losses. 2) NL, nearness to the 52-week low, which is the current stock price divided by the lowest stock price in the previous year. A lower NL indicates that the stock price is closer to its 52-week low. We expect that the positive relation between firms' left-tail risk and future bear spread returns is stronger when $\Delta VaR5$ is positive or when NL is lower.

To test our conjecture, we first sort bear spreads into two subsamples according to the signs of $\Delta VaR5$ of the underlying stocks: $\Delta VaR5 > 0$ and $\Delta VaR5 <= 0$. Alternatively, we sort bear spreads into tercile subsamples according to the value of NL of the underlying stocks: low NL, mid NL, and high NL. Then, within each subsample, we further sort bear spreads into deciles based on the left-tail risk measure $VaR5$.

Table 1.5 reports the time-series average monthly returns for the highest and lowest $VaR5$ decile portfolios in each subsample, together with the return spreads ("10-1") and their corresponding five-factor alphas. Newey-West (1987) adjusted t-statistics are reported in brackets.

Table 1.5: Persistence of Left-Tail Risk

This table presents return comparisons between bear spread decile portfolios sorted based on the monthly change in $VaR5$ (ΔVaR) in Panel A and the nearness to 52-week low (NL) in Panel B. ΔVaR is defined as the difference between $VaR5$ in month t and month $t-1$. NL is calculated as the the previous month-end stock price divided by the minimum price in the previous year. Delta-hedged bear spread portfolios are sorted into $\Delta VaR > 0$ and $\Delta VaR \leq 0$ groups or NL terciles. Then decile portfolios are formed based on $VaR5$ within each group or tercile. The returns for decile 10 and decile 1 portfolios, the return spread ("10-1") and its associated five-factor alpha are reported. Five-factor alphas are calculated after adjusting for Fama-French three factors, Carhart (1997) momentum factor, and Coval and Shumway (2001) systematic volatility factor. Newey-West (1987) adjusted t-statistics are presented in parentheses. The sample period is from January 1996 to December 2017.

Panel A: Sorts based on ΔVaR				
	Equal-weighted		Dollar-open-interest-weighted	
	ΔVaR>0	ΔVaR<=0	ΔVaR>0	ΔVaR<=0
1	-0.53%	-0.64%	-0.56%	-0.44%
10	0.48%	0.15%	1.08%	-0.28%
10-1	1.02%***	0.79%***	1.64%***	0.16%
(t-stat)	(2.76)	(2.70)	(3.01)	(0.43)
Five-factor alpha	1.12%***	0.87%***	1.77%***	0.25%
(t-stat)	(2.85)	(3.34)	(3.11)	(0.59)

Panel B: Sorts based on NL						
	Equal-weighted			Dollar-open-interest-weighted		
	Low NL	Mid NL	High NL	Low NL	Mid NL	High NL
1	-0.65%	-0.68%	-0.21%	-0.55%	-0.67%	-0.62%
10	0.47%	0.04%	0.31%	1.48%	0.52%	-0.17%
10-1	1.12%***	0.72%***	0.52%	2.04%***	1.19%***	0.46%
(t-stat)	(3.28)	(2.74)	(1.45)	(3.74)	(2.63)	(0.73)
Five-factor alpha	1.23%***	0.83%***	0.76%**	2.21%***	1.34%***	0.55%
(t-stat)	(3.26)	(3.58)	(2.05)	(4.02)	(3.03)	(0.78)

Panel A presents results for the sorts based on $\Delta VaR5$. For equal-weighted portfolios, when $\Delta VaR5 > 0$, the "10-1" return spread is 1.02% (t-statistic=2.76). The corresponding five-factor alpha is 1.12% (t-statistic=2.85). When $\Delta VaR5 \leq 0$, the "10-1" return spread is 0.79% (t-statistic=2.70). The corresponding five-factor alpha is 0.87% (t-statistic=3.34). Although the positive relation between $VaR5$ and future bear spread returns are statistically significant for both subsamples, the magnitudes of the "10-1" return spread and the corresponding alpha are larger for $\Delta VaR5 > 0$

subsample than for $\Delta VaR5 <= 0$ subsample.

For DOI-weighted portfolios, when $\Delta VaR5 > 0$, the "10-1" return spread is 1.64% (t-statistic=3.01). The corresponding five-factor alpha is 1.77% (t-statistic=3.11). When $\Delta VaR5 \leq 0$, both the "10-1" return spread and the alpha are statistically insignificant. The positive relation between $VaR5$ and future bear spread returns is only significant for $\Delta VaR5 > 0$ subsample.

Combining the results for equal-weighted and DOI-weighted portfolios, Panel A shows a strong and consistent positive relation between $VaR5$ and future delta-hedged bear spread returns when $\Delta VaR5$ is positive. Since stocks that have experienced recent large loss are more likely to experience similar large losses in the near future (stock price's left-tail risk momentum), the protection provided by the bear spreads on these stocks should be more valuable. However, option traders underestimate the left-tail risk persistence and underprice the bear spreads on the stocks with high recent extreme losses, showing a left-tail risk momentum effect.

Panel B presents results for the sorts based on NL. For equal-weighted portfolios, in the low NL subsample, the "10-1" return spread is 1.12% (t-statistic=3.28). The corresponding five-factor alpha is 1.23% (t-statistic=3.26). In the mid NL subsample, the "10-1" return spread and the alpha are slightly lower but remain significant, being 0.72% (t-statistic=2.74) and 0.83% (t-statistic=3.58) respectively. In the high NL subsample, the "10-1" return spread is 0.52% and insignificant. The alpha is 0.76% (t-statistic=2.05) and is lower than those of the low and mid NL subsamples.

For DOI-weighted portfolios, in the low NL subsample, the "10-1" return spread is 2.04% (t-statistic=3.74). The five-factor alpha is 2.21% (t-statistic=4.02). In the mid NL subsample, the "10-1" return spread is 1.19% (t-statistic=2.63), and the five-factor alpha is 1.34% (t-statistic=3.03). The return spread and the alpha approximately halve those of the low NL subsample. In the high NL subsample, the "10-1" return spread and the alpha are both insignificant. Furthermore, the five-factor alpha difference between the low and high NL subsamples' "10-1" spread is 1.66% (t-statistic=2.14).

The results for equal-weighted and DOI-weighted portfolios in Panel B show that the underestimation of the left-tail risk in the options market is stronger when the stock price is nearer to it's 52-week low price. Option traders anchor their loss expectation around the 52-week low and underestimate the persistence of stock price decline, leading to a stronger positive relation between firms' left-tail risk and the future bear spread returns. Our results are consistent to Driessen et al. (2013) as option traders' underestimation to the chance of downward breakthroughs leads to stronger underpricing of bear spreads when stock prices approach their 52-week low, showing an anchoring effect.

Overall, the results in Table 1.5 suggest that both the left-tail momentum effect and the anchoring effect have strong impact on the underpricing of bear spreads. The existence of both effects indicate that one of the driving forces of the underreaction to firms' left-tail risk in the options market is option traders' underestimation of loss persistence.

1.4.2 Information uncertainty

Prior literature (Hong et al., 2000; Jiang et al., 2005; Zhang, 2006*a*,*b*; Kumar, 2009*a*) shows that information uncertainty amplifies investor behavioural biases. In particular, high information uncertainty may lead to investors' slow reaction to news (especially bad news), causing predictable price drift or momentum.

Following the literature, we construct four information uncertainty proxies: 1) $SIZE$, as defined in Section 1.2.4. 2) AGE, firm age, which denotes the number of years that the firms are listed on Compustat at the previous year-end. 3) AC, residual analyst coverage, which is the residual from the cross-sectional regression of the logarithm of the analyst coverage on the logarithm of firm's market capitalization in the previous quarter. Since analyst coverage is very strongly correlated with firm size, we control for the influence of firms' market capitalization on analyst coverage using the residual analyst coverage following Hong et al. (2000). 4) $DISP$, analysts' forecast dispersion, which is the standard deviation of the analysts' forecasts scaled by the stock price in the previous quarter.

Zhang (2006*b*) use all the four proxies above to measure information uncertainty. Hirshleifer and Teoh (2003) use firm size and analyst coverage as proxies for investor inattention. Kumar (2009*a*) use firm age to measure valuation uncertainty. Taking into account conceptual overlap and mixed interpretation of proxies between information uncertainty, investor inattention, and valuation uncertainty, we use small firm size, young firm age, low analyst coverage, and high dispersion in analyst forecasts as proxies of high information uncertainty. We expect that high information uncertainty amplifies the positive relation between the left-tail risk and the future bear spread returns.

To test our conjecture, we construct subsamples based on the median of each of the four proxies above. Within these subsamples, we sort bear spread returns into deciles based on the underlying stock's $VaR5$, and we calculate the time-series average monthly returns for the decile portfolios in each subsample.

Table 1.6 reports the highest and lowest $VaR5$ decile portfolio returns, together with the return spreads ("10-1") and its five-factor alphas. Newey-West (1987) adjusted t-statistics are reported in brackets.

Panel A presents the results for equal-weighted portfolios. When portfolios are sorted based on $SIZE$, the "10-1" return spread is 1.11% (t-statistic=3.78) for the low $SIZE$ group and 0.80% (t-statistic=2.85) for the high $SIZE$ group. The corresponding five-factor alphas are 1.24% (t-statistic=4.44) and 0.83% (t-statistic=3.27).

When portfolios are sorted based on Age, the "10-1" return spread is 1.32% (t-statistic=4.22) for the low Age group and 0.88% (t-statistic=3.19) for the high Age group. The five-factor alphas are 1.44% (t-statistic=5.11) and 1.02% (t-statistic=3.79) respectively.

When portfolios are sorted based on AC, the "10-1" return spread is 1.13% (t-statistic=3.94) for the low AC group and 0.58% (t-statistic=2.15) for the high AC group. The five-factor alphas are 1.28% (t-statistic=4.72) and 0.81% (t-statistic=2.77) respectively. The difference in alpha spreads between the low and high AC group is 0.47% (t-statistic=1.74).

When portfolios are sorted based on $DISP$, the "10-1" return spread is 1.10%

Table 1.6: Information Uncertainty

This table presents return comparisons between bear spread decile portfolios in subsamples sorted based on the median of one of the four information uncertainty proxies: firm size ($SIZE$), firm age (AGE), analyst coverage (AC), and analysts' forecast dispersion ($DISP$). Decile portfolios are formed based on the underlying stocks' $VaR5$ within each subsample. The returns for decile 10 and decile 1 portfolios, the return spread ("10-1") and its associated five-factor alpha are reported. Five-factor alphas are calculated after adjusting for Fama-French three factors, Carhart (1997) momentum factor, and Coval and Shumway (2001) systematic volatility factor. Newey-West (1987) adjusted t-statistics are presented in parentheses. The sample period is from January 1996 to December 2017.

Panel A: Equal-weighted portfolios								
	Low SIZE	High SIZE	Low AGE	High AGE	Low AC	High AC	High DISP	Low DISP
1	-0.44%	-0.64%	-0.57%	-0.69%	-0.61%	-0.51%	-0.53%	-0.69%
10	0.67%	0.16%	0.75%	0.19%	0.52%	0.07%	0.57%	-0.20%
10-1	1.11%***	0.80%***	1.32%***	0.88%***	1.13%***	0.58%**	1.10%***	0.49%
(t-stat)	(3.78)	(2.85)	(4.22)	(3.19)	(3.94)	(2.15)	(3.71)	(1.54)
Five-factor alpha	1.24%***	0.83%***	1.44%***	1.02%***	1.28%***	0.81%***	1.18%***	0.74%**
(t-stat)	(4.44)	(3.27)	(5.11)	(3.79)	(4.72)	(2.77)	(4.09)	(2.33)
Panel B: Dollar-open-interest-weighted portfolios								
	Low SIZE	High SIZE	Low AGE	High AGE	Low AC	High AC	High DISP	Low DISP
1	-0.65%	-0.57%	-0.43%	-0.73%	-0.62%	-0.37%	-0.61%	-0.53%
10	1.06%	0.03%	1.50%	-0.12%	0.74%	0.49%	1.09%	-0.37%
10-1	1.71%***	0.60%	1.93%***	0.61%	1.36%**	0.86%**	1.69%***	0.16%
(t-stat)	(3.47)	(1.62)	(4.40)	(1.26)	(2.22)	(2.11)	(3.68)	(0.35)
Five-factor alpha	1.75%***	0.66%*	2.00%***	0.95%**	1.52%***	1.14%***	1.74%***	0.62%
(t-stat)	(3.33)	(1.83)	(4.66)	(2.03)	(2.67)	(2.69)	(3.99)	(1.16)

(t-statistic=3.71) for the high $DISP$ group and insignificant for the low $DISP$ group. The corresponding five-factor alphas are 1.18% (t-statistic=4.09) and 0.74% (t-statistic=2.33). The difference in alpha spreads between the high and low $DISP$ group is 0.44% (t-statistic=1.68).

Panel B presents the results for DOI-weighted portfolios. When portfolios are sorted based on $SIZE$, the "10-1" return spread is 1.71% (t-statistic=3.47) for the low $SIZE$ group and insignificant for the high $SIZE$ group. The corresponding five-factor alphas are 1.75% (t-statistic=3.33)and 0.66% (t-statistic=1.83). The difference in alpha spreads between the low and high $SIZE$ group is 1.09% (t-statistic=2.10).

When portfolios are sorted based on Age, the "10-1" return spread is 1.93% (t-statistic=4.40) for the low Age group and insignificant for the high Age group. The five-factor alphas are 2.00% (t-statistic=4.66) and 0.95% (t-statistic=2.03) respectively. The difference in alpha spreads between the low and high Age group is 1.06% (t-statistic=1.76).

When portfolios are sorted based on AC, the "10-1" return spread is 1.36% (t-statistic=2.22) for the low AC group and 0.86% (t-statistic=2.11) for the high AC group. The corresponding five-factor alphas are 1.52% (t-statistic=2.67) and 1.14% (t-statistic=2.69).

When portfolios are sorted based on $DISP$, the "10-1" return spread is 1.69% (t-statistic=3.68) for the high $DISP$ group and insignificant for the low $DISP$ group. The corresponding five-factor alphas are 1.74% (t-statistic=3.99) for the high $DISP$ group and insignificant for the low $DISP$ group. The difference in alpha spreads between the high and low $DISP$ group is 1.12% (t-statistic=1.74).

Overall, the positive relation between firms' left-tail risk and the future bear spread returns is stronger in the high information uncertainty subsamples. Information uncertainty usually amplifies investors' behavioral biases such as investor underreaction to bad news, our finding suggests that the options market's underreaction to firms' left-tail risk is amplified by information uncertainty.

1.4.3 Investor sentiment

Sentiment is a biased investor belief conditional on available information (Barberis et al., 1998). Asset mispricing and risk underestimation are more likely to happen during the high investor sentiment periods (Baker and Wurgler, 2006; Yu and Yuan, 2011; Stambaugh et al., 2012; Lemmon and Ni, 2014; Byun and Kim, 2016). Stambaugh et al. (2012) show that high investor sentiment contributes to the significant profits from the short legs of long-short strategies building upon a large set of anomalies. Byun and Kim (2016) document that the overvaluation of lottery-like options is attributable to high investor sentiment. While prior literature focuses more on the overpricing of risky assets, the underpricing of safer, protective assets may also happen when investor sentiment is high.

Complementary to prior literature, we analyse potential, more pronounced underpricing for bear spread, a protective option strategy, in high sentiment periods. We use the monthly market-based sentiment index (BW sentiment index) constructed by Baker and Wurgler (2006) to classify high and low investor sentiment months. A high (low) sentiment month is one in which the value of the BW sentiment index in the previous month is above (below) the median value for the sample period.

Within the subsample with high (low) sentiment months, we form decile portfolios of delta-hedged bear spreads based on $VaR5$ and calculate the time-series average monthly returns for the decile portfolios.

Table 1.7 reports the returns for the highest and lowest $VaR5$ decile portfolios, together with the return spreads ("10-1") and the corresponding five-factor alphas. Newey-West (1987) t-statistics are reported in brackets.

Table 1.7: Investor Sentiment

This table presents return comparisons between bear spread decile portfolios in subsamples sorted based on the median of Baker and Wurgler (2006) sentiment index. A high (low) sentiment month is one in which the value of the BW sentiment index in the previous month is above (below) the median value for the sample period. Decile portfolios are formed based on the underlying stocks' $VaR5$ within each subsample. The returns for decile 10 and decile 1 portfolios, the return spread ("10-1") and its associated five-factor alpha are reported. Five-factor alphas are calculated after adjusting for Fama-French three factors, Carhart (1997) momentum factor, and Coval and Shumway (2001) systematic volatility factor. Newey-West (1987) adjusted t-statistics are presented in parentheses. The sample period is from January 1996 to December 2017.

	Equal-weighted portfolios		DOI-weighted portfolios	
	High Sentiment	Low Sentiment	High Sentiment	Low Sentiment
1	-0.57%	-0.61%	-0.35%	-0.77%
10	0.85%	-0.08%	1.13%	-0.29%
10-1	1.42%***	0.52%*	1.49%**	0.48%
(t-stat)	(3.25)	(1.67)	(2.25)	(1.09)
Five-factor alpha	1.46%***	0.74%**	1.87%***	0.39%
(t-stat)	(3.13)	(2.46)	(2.75)	(0.84)

Panel A presents the results for equal-weighted portfolios. In the high-sentiment subsample, the decile 1 portfolio has a negative average return of -0.57%, while decile 10 portfolio has a positive average return of 0.85%. The "10-1" monthly return spread is 1.42% (t-statistic=3.25). The corresponding five-factor alpha is 1.46% (t-statistic=3.13).

In the low-sentiment subsample, the decile 1 portfolio has a negative average return of -0.61%, while decile 10 portfolio has an average return of -0.08%. The "10-1" monthly return spread is 0.52% (t-statistic=1.67). The corresponding five-factor alpha is 0.74% (t-statistic=2.46). The return difference between decile 1 portfolios in subsamples is lower than decile 10 portfolios, so the difference in "10-1" return spreads between two subsamples is mainly driven by the decile 10 (high left-tail risk) portfolio returns.

Panel B presents the results for DOI-weighted portfolios. In the high-sentiment periods, the decile 1 portfolio has a negative average return of -0.35%, while decile 10 portfolio has a positive average return of 1.13%. The "10-1" monthly return spread is 1.49% (t-statistic=2.25). The corresponding five-factor alpha is 1.87% (t-statistic=2.75). In the low-sentiment periods, the decile 1 portfolio has a negative average return of -0.77%, while decile 10 portfolio has a negative average return of -0.29%. The "10-1" monthly return spread and the corresponding five-factor alpha are both insignificant.

The results in both panels show that the underreaction to firms' left-tail risk is stronger during the high-sentiment periods. When market sentiment is high, option traders tend to overlook downside risk and and underprice the downside protection provided by bear spreads. Our finding, together with the prior research on risky asset overpricing in high sentiment periods, supports the economic intuition that overvaluation in risky assets and undervaluation in safer, protective assets may happen at the same time (Acharya and Naqvi, 2019). Han (2008) shows that when market sentiment is high, the index option volatility smile is flatter and the risk-neutral skewness of index return extracted from the index option prices is less negative, suggesting decreased risk hedging demand. Our results are consistent to Han (2008) as low hedging demand leads to stronger underpricing of bear spreads during the high market sentiment periods.

1.5 Further Discussion

1.5.1 Arrow-Debreu price

From the option prices in a bear spread strategy, we can compute Arrow-Debreu state price of the left-tail events (Lu and Murray, 2019). If we scale the option positions in the bear spread by $K_1 - K_2$, we would have a price of $(P_1 - P_2)/(K_1 - K_2)$, and a payoff of \$1 when $S_T < K_2$. Therefore, the price of the scaled bear spread should be equal to $e^{-rT}\hat{\mathbb{E}}[\mathbf{1}_{\{S_T<K_2\}}]$, where $\mathbf{1}$ is the indicator function and $\hat{\mathbb{E}}$ represents the expected value under risk-neutral probability. The scaled bear spread price can be interpreted as the discounted risk-neutral state probability of left-tail events.

In our prior results, we show that options market underreact to firms' left-tail risk by presenting evidence of a positive risk-return relation between firms' left-tail risk and future bear spread returns. Now, we examine the Arrow-Debreu state price of left tail events to shed more light on such underreaction. If options market totally ignores and misprices firms' left-tail risk, then Arrow-Debreu state price will exhibit significant misalignment to the left-tail risk measures. Otherwise, if options market prices firms' left-tail risk but fails to fully price such risk, then Arrow-Debreu state price should more or less align to the left-tail risk measures.

We sort firms into deciles based on one of the four left-tail risk measures ($VaR1$, $VaR5$, $ES1$, $ES5$) and calculate the average Arrow-Debreu state price, from scaled bear spread prices, in each decile.

Table 1.8 reports the time series average of Arrow-Debreu state price in each decile portfolio, together with the difference in state prices between the highest and the lowest left-tail risk deciles. Newey-West (1987) adjusted t-statistics are reported in brackets.

Table 1.8: Arrow-Debreu Price

This table presents the time-series average Arrow-Debreu prices (scaled bear spread prices) formed on deciles based on firms' left-tail risk metrics ($VaR5$, $VaR1$, $ES5$, $ES1$), along with the price spread between decile 10 and decile 1 portfolios. The scaled bear spread price is calculated by scaling the bear spread price by the difference of the strike prices of the two puts in the bear spread. Newey-West (1987) adjusted t-statistics are reported in parentheses. The sample period is from January 1996 to December 2017.

	VaR5	VaR1	ES5	ES1
1	0.176	0.176	0.176	0.177
2	0.183	0.183	0.183	0.184
3	0.189	0.190	0.189	0.191
4	0.196	0.197	0.197	0.198
5	0.204	0.204	0.203	0.205
6	0.211	0.211	0.211	0.212
7	0.218	0.218	0.218	0.219
8	0.227	0.227	0.227	0.227
9	0.239	0.238	0.239	0.237
10	0.261	0.258	0.260	0.253
10-1	0.085***	0.082***	0.085***	0.076***
(t-stat)	(61.58)	(64.35)	(63.42)	(62.89)

Arrow-Debreu state price increases monotonically with all the left-tail risk measures. For example, when portfolios are sorted based on $VaR5$, the average Arrow-Debreu state price of decile 1 (10) is 0.176 (0.261). The difference is 0.085 and significant at 1% level. Arrow-Debreu price extracted from bear spread strategy exhibits the correct price order of hedging demand. Together with our prior results, we show that although the options market underreacts to firms' left-tail risk, the bear spread prices at least partially reflect such risk. In other words, firms' left-tail risk is priced in, but is not adequately priced in.

1.5.2 Transaction costs

Transactions on OTM and DOTM equity options usually incur high transaction costs. Our prior finding might be driven by illiquid options as option traders underreact to risk due to these options' high transaction costs. In this section, we examine whether transaction costs in options market severely impact our findings.

We conduct three tests following Bali and Murray (2013). We begin our analysis by restricting our sample to the following subsamples where liquidity or transaction costs are of less concern: 1) Large Open Interest subsample, where we require the minimum open interest of the two puts in the bear spread to be among the upper quartile on portfolio formation date. The size of this subsample is 38,751; 2) Small $Spread subsample, where we require both put options in the bear spread to have bid-ask spreads of less than $0.15. The size of subsample is 63,526; and 3) Small %Spread subsample, where we require both put options in the bear spread to have percentage bid-ask spread (bid-ask spread divided by the midpoint of the bid and ask price) to be below the lower tercile of each corresponding put option group on each portfolio formation date. The size of subsample is 28,222.

Panel A in Table 1.9 reports the five-factor alpha on return spread between the highest and lowest $VaR5$ decile portfolios for each subsample.

Table 1.9: Transaction Costs

This table presents the results when transaction costs are taken into consideration. Panel A reports the the five-factor alphas for return spreads between decile 10 and decile 1 delta-hedged bear spread portfolios formed on $VaR5$ in three subsamples. The *Large Open Interest* subsample is constructed by requiring the minimum open interest of the two puts in the bear spread to be above the upper quartile on each portfolio formation date. The *Small $Spread* subsample is formed by requiring both options in the bear spread to have bid-ask spread of less than $0.15. The *Small %Spread* subsample is formed by requiring both options in the bear spread to have percentage bid-ask spread (bid-ask spread divided by the midpoint of the bid and ask price) to be below the lower quartile on each portfolio formation date. Panel B presents the results of Fama and MacBeth (1973) regressions by adding additional controls for option liquidity and transaction costs. The dependent variable is the delta-hedged bear spread (formed in month t) hold-to-maturity monthly return. The additional controls are for the options (PUT_1 and PUT_2) used to create the bear spread strategy. These controls are: open interest($OpenInt$), bid-ask spread ($\$Spread$), percentage bid-ask spreads ($\%Spread$), volatility spread ($SpreadVol$), and implied volatility spread ($SpreadImvol$). Panel C presents the five-factor alphas of return spreads between decile 10 and decile 1 delta-hedged bear spread portfolios formed on $VaR5$ after paying 10%, 20%, 30% of the quoted spread on the option positions. Five-factor alphas are calculated after adjusting for Fama-French three factors, Carhart (1997) momentum factor, and Coval and Shumway (2001) systematic volatility factor. Newey-West (1987) adjusted t-statistics are presented in parentheses. The sample period is from 1996 to 2017.

Panel A: "10-1" alpha in subsamples	Equal-weighted portfolios	DOI-weighted portfolios
Large Open Interest subsample		
"10-1" five-factor alpha	1.22%***	1.41%***
(t-stat)	(3.57)	(3.35)
Small $Spread subsample		
"10-1" five-factor alpha	1.24%***	0.91%**
(t-stat)	(3.54)	(2.15)
Small %Spread subsample		
"10-1" five-factor alpha	1.54%***	1.43%**
(t-stat)	(3.73)	(2.25)

Table 1.9 continued from previous page

Panel B: Fama-MacBeth regressions	1	2	3	4

VaR5	0.198***	0.213***	0.199***	0.197***
	(3.28)	(3.62)	(3.29)	(3.25)
$OpenInt\_PUT_1$	0.000	0.000	0.000	0.000
	(1.33)	(1.19)	(1.32)	(1.16)
$OpenInt\_PUT_2$	0.000*	0.000**	0.000**	0.000**
	(1.88)	(2.09)	(1.99)	(2.02)
$\$Spread\_PUT_1$	0.000			
	(-0.01)			
$\$Spread\_PUT_2$	-0.009*			
	(-1.73)			
$\%Spread\_PUT_1$		-0.005		
		(-1.59)		
$\%Spread\_PUT_2$		0.016*		
		(1.73)		
$SpreadVol\_PUT_1$			-0.022	
			(-1.02)	
$SpreadVol\_PUT_2$			-0.010	
			(-0.34)	
$SpreadImvol\_PUT_1$				-0.001
				(-0.41)
$SpreadImvol\_PUT_2$				-0.001
				(-0.49)
Controls	YES	YES	YES	YES
Adj. R^2	4.75%	4.78%	5.35%	4.42%
Panel C: Effective bid-ask spreads as percentage of quoted spreads				
	0%	10%	20%	30%
Equal-weighted portfolios				
"10-1" five-factor alpha	1.20%***	0.83%***	0.46%*	0.10%
(t-stat)	(4.48)	(3.09)	(1.72)	(0.38)
DOI-weighted portfolios				
"10-1" five-factor alpha	1.31%***	1.11%***	0.91%**	0.72%*
(t-stat)	(3.15)	(2.66)	(2.19)	(1.73)

Panel A shows that for all the low transaction costs subsamples, the "10-1" alpha spreads remain positive and significant, confirming that our results are not only driven by options with high transaction costs.

Next, we perform Fama-MacBeth regressions by further controlling for option liquidity and transaction costs proxies. We use the open interest of an option ($OpenInt$) as the option liquidity proxy. We use the following four option transaction costs proxies: 1) $\$Spread$, the dollar bid-ask spread of an option; 2) $\%Spread$, the percentage bid-ask spread of an option; 3) $SpreadVol$, the volatility spread of an option, which is calculated as the dollar spread divided by the option vega; and 4) $SpreadImvol$, the implied volatility spread of an option, which is calculated as the dollar spread divided by the option implied volatility.

Panel B in Table 1.9 reports the results from Fama-MacBeth regressions. Columns 1 to 4 in Panel B use different combinations of option liquidity and transaction costs proxies. The results consistently show that the coefficients on $VaR5$ remain positive and significant at 1% level, confirming that our prior results are not driven by illiquid options with transaction costs.

Last, we directly add transaction costs to our analysis by assuming the effective spreads for options are equal to 10%, 20%, and 30% of the quoted spreads. Muravyev and Pearson (2020) provide evidence that the effective spreads of option traders are less than 30% of the quoted spreads when traders time execution.

Panel C in Table 1.9 reports the five-factor alpha of return spread between the highest and lowest $VaR5$ decile portfolios. For comparison, we also include the results with no transaction costs (effective spread = 0% of quoted spread). For both equal-weighted and DOI-weighted portfolios, increasing transaction costs reduces "10-1" alphas. When the effective option spread is 20% of the quoted spread, the five-factor alphas are 0.46% (t-statistic=1.72) and 0.91% (t-statistic=2.19) for equal-weighted and DOI-weighted portfolios respectively. Even at 30% of the quoted spread, the five-factor alpha is 0.72% (t-statistic=1.73) for DOI-weighted portfolios.

Overall, results in Table 1.9 show that our prior results largely hold when transaction costs are taken into consideration.

1.6 Conclusion

Asset pricing research puts a lot of emphasis on tail risk. Practitioners in financial markets also emphasize their loss aversion against investment downsides. Adequately estimating and pricing firms' left-tail risk is important for equity investors, option traders, and well-functioned financial markets in general.

Using bear spreads with bear regions concentrated on firms' left-tail, we show that firms' left-tail risk is a strong positive predictor of future bear spread returns. Our finding suggests that the options market underreacts to firms' left-tail risk and does not adequately price in such risk.

Behavioral biases help to explain the underreaction to firms' left-tail risk. We show that underreaction is stronger for stocks with larger recent losses and closer to their 52-week lows, suggesting that option traders do not adequately factor in the persistence of losses. Higher information uncertainty amplifies investor underreaction to bad news, leading to stronger bear spread underpricing. Investor sentiment also has significant impact on left-tail risk underreaction as we show that the underreaction mainly happens during high market sentiment periods.

Bear spreads provide protection against downsides and such protection should be priced adequately in the options market. Our finding suggests that although the loss aversion against the left-tail risk plays an important role in financial markets, option traders fail to demand adequate price premium for bear spreads to compensate firms' left-tail risk.

Our study contributes to the literature by using an option trading strategy, bear spread, to isolate and analyze firms' left-tail risk and showing that merely recognizing the importance of left-tail risk is not enough, investors need overcome behavioral biases to adequately price left-tail risk.

Chapter 2

Anticipating Jumps: Decomposition of Straddle Returns Around Earnings Announcements

2.1 Introduction

Quarterly earnings announcements contain important information about company fundamentals and are closely followed by not only stock investors but also option traders. Gao et al. (2018) find that trading delta-neutral straddles on individual firms generates significantly positive returns around earnings announcements, indicating that option traders underestimate the magnitude of uncertainty surrounding earnings announcements. One of the most interesting findings of Gao et al. (2018) is that the cumulative returns of delta-neutral straddles remain positive even after the announcements, suggesting a persistent underpricing of straddles constructed before the announcements.

A straddle is formed by simultaneously buying both a put option and a call option on the underlying stock with the same strike price and the same expiration date. Although trading straddles is deemed a popular volatility trading strategy, a straddle buyer actually pays a risk premium for protection against not only volatility risk but also jump risk. The underpricing of risks in straddles could be driven by the underpricing of either volatility risk or jump risk, or both. Coval and Shumway (2001) motivate the use of delta-neutral straddles for examining the asset pricing effect of stochastic volatility because of their large vegas. Cremers et al. (2015) point out that delta-neutral straddles have large gammas as well as vegas thus they have large exposure to jump risk as well as volatility risk. They use a combination of short-term and long-term delta-neutral straddles to construct two tradable portfolios: a volatility factor-mimicking portfolio with exposure only through vegas and a jump factor-mimicking portfolio with exposure only through gammas. Distinguishing the jump risk from the diffusive volatility risk is important in finance because these two types of risk should be treated and priced separately. Jumps represent the discontinuity in asset prices whereas volatility represents variation along the continuous asset price paths ((Bates, 1996), (Pan, 2002)). Lee and Mykland (2008) show that individual stock jumps are driven by earnings announcements and other company-specific news events whereas the S&P 500 index jumps are associated with general market news announcements. Dubinsky et al. (2019) find that options market prices in earnings-induced jump risk prior to the earnings announcements.

In this paper, we extend the volatility-jump factor-mimicking methodology developed in Cremers et al. (2015) to a unique decomposition of delta-neutral straddle into its volatility component (delta and gamma neutral but vega positive) and jump component (delta and vega neutral but gamma positive). Subsequently, the delta-neutral straddle returns can be decomposed into the returns of the volatility-jump components. We use our decomposition method to examine the delta-neutral straddle returns around the earnings announcements and answer the following research question: if options market underprices the uncertainty around earnings announcements and such underpricing manifests itself in positive straddle returns, then what is the constitution of such underpricing with respect to volatility risk and jump risk? Uncovering the answer to this question could significantly enhance our understanding of the uncertainty anticipated by options market and priced in straddles before the earnings announcements.

We analyze the return patterns of straddle trading strategies with different holding periods around the earnings announcement date (EAD) and compare the patterns of straddle returns to the patterns of its volatility-jump component returns. To capture the difference between pre- and post-announcement return patterns, we focus on two buy-and-hold strategies. Strategy [-3,-1] (Strategy [-3,1]) is the trading strategy that delta-neutral straddles are purchased on the three trading days before the EAD, and are sold on one trading day before (after) the EAD. We find that straddle cumulative returns are both positive on [-3,-1] and [-3,1], with the cumulative return on [-3,1] (post-announcement) being lower than the cumulative return on [-3,-1] (pre-announcement).

Using our decomposition method, we find a pattern of run-up of uncertainty over the pre-announcement period [-3,-1], with both volatility and jump component returns being significantly positive. Compared to the volatility component's return, the return of the jump component on [-3,-1] is much larger, suggesting a substantial run-up of priced-in jump risk in straddles before the EAD. When post-announcement period is included in the straddle holding period, we find that the volatility component's cumulative return remains significantly positive on [-3,1] and is larger than its cumulative return on [-3,-1], whereas the jump component's cumulative return on [-3,1] turns to be significantly negative.

Our findings show that the overall pattern of delta-neutral straddle returns around earnings announcements is mainly driven by its jump component. The pattern consists of a run-up of uncertainty priced in straddles, dominated by an accumulation of priced-in jump risk, over the pre-announcement period, and a drop of straddle price after the EAD, dominated by a drop of priced-in jump risk, over the post-announcement period.

Our decomposition results provide a new angle to explain the phenomenon of straddle underpricing. We show that the main driver of underpricing is the volatility component of straddles. After the EAD, the volatility component of a straddle still has positive return, countering the significant return drop of the jump component and keeping the straddle's cumulative return positive. The persistent underpricing of individual company's straddles around earnings announcements mainly comes from the persistent underpricing of its volatility component.

We answer our research question by showing that options market anticipates the stock price jumps induced by earnings announcements. The discontinuity of the stock price around earnings announcements is the focal point of options traders. The return patterns of both the straddle and its jump component demonstrate such anticipation. By contrast, variation of smoothly changing part of stock price is not fully priced in straddles. Options market appears to underestimate the volatility risk and underprice the volatility component of straddles, leading to the positive cumulative returns on both the straddle and its volatility component after the earnings announcements.

Gao et al. (2018) suggest that the underestimation of uncertainty around EAD is due to the noisiness of information received by investors and the transaction costs. We find marked difference between volatility and jump components' responses to information noisiness and transaction costs. On [-3,-1], information noisiness and transaction costs strongly positively predict the straddle and its jump component returns. Instead, they have little or negative impact on the volatility component returns. Our results lend further support that jump component returns dominate the risk run-up pattern of straddle returns before the earnings announcements. On [-3,1], the positive return predictability of these measures on straddle and jump component largely disappears or even reverses, consistent to the substantial decline of information uncertainty after the earnings announcements.

We conduct further analysis on returns over [-3,0] by separating the announcements during the market hours and the after-hours announcements. For firms with earnings announced during the market hours, the volatility component returns are significantly positive. Whereas for firms with after-hours earnings announcements, the straddle and the jump component returns are significantly positive. Our findings show that the pre-announcement period positive straddle returns are mainly driven by its jump component whereas the post-announcement period positive straddle returns are mainly driven by its volatility component. These findings are consistent to our findings on returns over [-3,-1] and [-3,1].

Finally, We find that in recent years (2014-2017) positive straddle returns around the EAD largely disappear. The jump component return is negative, suggesting that options market pays more attention to the straddle mispricing issue by pricing in more jump risk. However, the underpricing of volatility risk is persistent and the volatility component continues to generate significantly positive returns around the earnings announcements.

Our paper contributes to the empirical asset pricing literature in two ways. First, we extend the volatility-jump risk factor mimicking methodology in Cremers et al. (2015) to a new volatility-jump decomposition method. Cremers et al. (2015) examine the pricing of systematic volatility and jump risks in the cross-section of stock returns by constructing factor-mimicking trading strategies using index options. Their focal point is the roles of systematic jump risk and volatility risk in stock prices. Our paper's focal point is individual stock jumps anticipated by options market around earnings announcements. Our decomposition method is used to disentangle volatility-jump components in individual option straddles. We contribute to the literature on disentangling jump and volatility[1] by showing that the methodology in Cremers et al. (2015) can be extended to disentangle volatility-jump components in straddles and straddle returns can be decomposed into volatility-jump component returns.

Second, we investigate the puzzling positive straddle returns around earnings

[1]See Bates (1996), Liu and Pan (2003), Maheu and McCurdy (2004), Santa-Clara and Yan (2010), Todorov (2010), Cremers et al. (2015).

announcements in Gao et al. (2018) and find that the positive straddle cumulative returns up to the post-announcement period are mainly driven by the positive volatility component cumulative returns. We also show that the jump component return is the main driving force for shaping the straddle's pre- and post-announcement return patterns. The literature on option prices around earnings announcement can be traced back to Patell and Wolfson (1979,1981). Patell and Wolfson (1979, 1981) document increases in option prices preceding the EADs and show that option traders anticipate an increase in equity volatility around EADs. Jin et al. (2012) find that option volatility skews and volatility spreads have stronger stock return predictability around EADs and other company-specific information events. We contribute to the literature on option prices around earnings announcements[2] by showing that volatility and jump components have markedly different return patterns around EADs. The return patterns suggest that option traders anticipate jumps but underestimate the volatility risk before the earnings announcements.

The remainder of the paper proceeds as follows. Section 2.2 presents the new decomposition method and describes the data. Section 2.3 presents the empirical results. Section 2.4 concludes.

[2]See also Patell and Wolfson (1979, 1981), Whaley and Cheung (1982), Billings and Jennings (2011), Jin et al. (2012), Dubinsky et al. (2019).

2.2 Methodology and Data

2.2.1 Decomposition method

We develop a new method to decompose a straddle's return into its volatility-jump component returns. Our method is an extension to the volatility-jump factor-mimicking methodology in Cremers et al. (2015). The construction of Cremers et al. (2015) focuses on the disentangling of volatility-jump factors using short-term and long-term straddles. Our method focuses on a decomposition of straddle price and return into the volatility-jump components and their corresponding component returns.

We consider two delta-neutral straddles $S1$ and $S2$ with short and long maturities $T1$ and $T2$. Straddle $S1$ ($S2$) is constructed by combining one unit of call option $c1$ ($c2$) and $-\Delta_{c1}/\Delta_{p1}$ ($-\Delta_{c2}/\Delta_{p2}$) unit of put option $p1$ ($p2$), where Δ_{c1} and Δ_{p1} (Δ_{c2} and Δ_{p2}) are deltas for the call and put options.

We use $S1 = (c1, -\Delta_{c1}/\Delta_{p1}p1)$ and $S2 = (c2, -\Delta_{c2}/\Delta_{p2}p2)$ to denote the composition of straddles. The same notations will be used in the following part to explain the decomposition of a straddle.

Vega and gamma of straddle $S1$ ($S2$) are $Vega_{S1}$ and $Gamma_{S1}$ ($Vega_{S2}$ and $Gamma_{S2}$). Comparing to long-maturity options, short-maturity options have smaller vegas and larger gammas, so typically we have $Vega_{S1} < Vega_{S2}$ and $Gamma_{S1} > Gamma_{S2}$.

We construct a volatility risk factor-mimicking portfolio V and a jump risk factor mimicking-portfolio J by combining straddles $S1$ and $S2$. Portfolio V is a delta-neutral, gamma-neutral, and vega-positive strategy consisting of one unit of $S2$ and $-Gamma_{S2}/Gamma_{S1}$ unit of $S1$. Porfolio J is a delta-neutral, vega-neutral, and gamma-positive strategy consisting of one unit of $S1$ and and $-Vega_{S1}/Vega_{S2}$ unit of $S2$. The composition of these two portfolios are:

$$V = (-(Gamma_{S2}/Gamma_{S1})S1, S2);$$
$$J = (S1, -(Vega_{S1}/Vega_{S2})S2).$$

The construction of V and J is the same as Cremers et al. (2015). We begin our extension by combining portfolio V and J to replicate short-maturity straddle $S1$. In this way, we decompose both the price and the returns of $S1$ into its volatility and jump components. Suppose $S1 = (aV, bJ)$, the decomposition question is to solve the unknown factor-mimicking portfolios' weights a and b. Utilising the straddle notation above, the decomposition can be expressed recursively and then simplified as follows,

$$\begin{aligned} S1 &= (aV, bJ) \\ &= \big(a(-(Gamma_{S2}/Gamma_{S1})S1, S2), b(S1, -(Vega_{S1}/Vega_{S2})S2)\big) \\ &= \big((b - a(Gamma_{S2}/Gamma_{S1}))S1, (a - b(Vega_{S1}/Vega_{S2}))S2\big). \end{aligned}$$

The conditions for the decomposition to succeed are the following two equations:

$$b - a(Gamma_{S2}/Gamma_{S1}) = 1;$$
$$a - b(Vega_{S1}/Vega_{S2}) = 0.$$

Solving the equations, we obtain the *unique* decomposition of straddle $S1$ by portfolios V and J as[3]:

$$S1 = \left(\frac{\frac{Vega_{S1}}{Vega_{S2}}}{1 - \frac{Vega_{S1}}{Vega_{S2}} \frac{Gamma_{S2}}{Gamma_{S1}}} V, \frac{1}{1 - \frac{Vega_{S1}}{Vega_{S2}} \frac{Gamma_{S2}}{Gamma_{S1}}} J \right). \quad (2.1)$$

Eq.(2.1) shows that a straddle is a portfolio of the volatility risk factor-mimicking portfolio V and the jump risk factor-mimicking portfolio J, with the corresponding portfolio weights (a and b) being $\frac{\frac{Vega_{S1}}{Vega_{S2}}}{1 - \frac{Vega_{S1}}{Vega_{S2}} \frac{Gamma_{S2}}{Gamma_{S1}}}$ and $\frac{1}{1 - \frac{Vega_{S1}}{Vega_{S2}} \frac{Gamma_{S2}}{Gamma_{S1}}}$. Therefore, a straddle's price can be decomposed into aV, its volatility component, and bJ , its jump component.

Since V and J are investable portfolios, we can decompose a straddle's return into its volatility-jump component returns as follows:

$$\frac{S_2 - S_1}{S_1} = \frac{a(V_2 - V_1)}{aV_1 + bJ_1} + \frac{b(J_2 - J_1)}{aV_1 + bJ_1}, \quad (2.2)$$

where S_1 (S_2), V_1 (J_2) and J_1 (J_2) are straddle price, portfolio V's price and portfolio J's price at time 1 (2). $\frac{a(V_2 - V_1)}{aV_1 + bJ_1}$ $\left(\frac{b(J_2 - J_1)}{aV_1 + bJ_1}\right)$ is the volatlity (jump) component's return contribution in total straddle return $\frac{S_2 - S_1}{S_1}$.

Our study aims to understand the uncertainty anticipated by options market before the EAD. As the delta-neutral straddles constructed in Gao et al. (2018), we do not conduct daily rebalance of the volatility and jump component portfolios. During the announcement period, the Greeks (delta, vega, and gamma) of options may change,

[3]The linear equation system to determine two unknowns a and b involves four equations matching the units of $c1$, $p1$, $c2$, and $p2$ as 1, $-\Delta_{c1}/\Delta_{p1}$, 0, and 0. This equation system can be reduced to two equations with two unknowns and the solution is unique. See details in Appendix.

leading to changed exposures to the underlying stock's price movement, volatility and jump. In Section 3, we devote additional discussion on how changes of Greeks may impact our results and how to conduct adjustment in face of such changes.

2.2.2 Data

Our sample period is from January 1996 to December 2013. We choose this period for a direct comparison to the results in Gao et al. (2018). In Section 2.3.5, we present the results on the extended period during January 2014 to December 2017. We obtain stock price, firm information, and earnings announcement data from the Center for Research in Security Prices (CRSP), Compustat, and Institutional Brokers' Estimate System (I/B/E/S). We require that the announcement date is available in both Compustat and I/B/E/S to be included in the sample.

We obtain option data from OptionMetrics, which provides end-of-day bid and ask quotes, open interest, volume, implied volatility, and option Greeks for all listed options. To avoid the bid–ask bounce from daily closing prices, we use the mid point of closing bid and ask prices to compute option returns. Following previous literature (Gao et al. (2018), Ruan (2020*b*)), we apply the following filters to the option data: (1) The option prices are at least \$0.125. (2) the underlying stock prices are at least \$5. (3) options must have nonmissing bid and ask price quotes and positive open interests. (4) bid and ask prices must satisfy basic arbitrage bounds to filter out erroneous observations. Arbitrage boundaries include: bid $>$ 0, bid $<$ offer; for put options we require strike $\geq$ bid and offer $\geq$ max(0, strike price – stock price); for call options, we require stock price $\geq$ bid and offer $\geq$ max(0, stock price – strike price). (5) the moneyness of the option is defined as the strike price over the previous day's stock price. To be considered as at-the-money, options must have moneyness between 0.9 and 1.1 and absolute delta between 0.375 and 0.625.

To form straddles in our study, we apply two additional rules: (a1) only paired calls and puts with matched time-to-maturity and strike price are included; (a2) options in short-term straddles have 10 to 60 days to maturity; options in long-term straddles expire in the calendar month that follows the short-term straddles' expiration month, and have within 90 days to maturity.[4]

When there are multiple qualified at-the-money short-term straddles for a stock, we use two weighting schemes[5] to aggregate their returns at firm level: equal weighting and dollar open-interest weighting. The dollar open interest (DOI) for each straddle is computed based on the option information in the previous day as follows:

$$\begin{aligned} DOI =& (CALL\,PRICE + PUT\,PRICE) \\ & \times min(CALL\_OPEN\_INTEREST, PUT\_OPEN\_INTEREST), \end{aligned}$$

which is the maximum possible dollar open-interest for this straddle. For long-term straddles, we pick the call and put option pair that is closest to being at-the-money among all qualified options.

Table 2.1 reports the summary statistics for our data from the pooled sample.

[4]To conduct the decomposition, we need two straddle pairs with short and long maturities. Our choice of short-term straddles is consistent to Gao et al. (2018). Our choice of long-term straddles ensures the options included in straddles are of relatively high liquidity.

[5]Volume weighting produces qualitatively similar results to the dollar open-interest weighting scheme. Results can be obtained from the authors upon request.

Table 2.1: Summary Statistics

This table reports summary statistics on stock characteristics (Panel A), straddle characteristics (Panel B), and daily delta-neutral straddle returns together with its volatility-jump component returns (Panels C and D). [-3,1] represents the period over three pre-anouncement days, the announcement day, and one post-announcement day. The sample period is from Jan. 1996 to Dec. 2013. N is the number of observations, and P25 (P75) represents lower quartile (upper quartile).

	N	P25	Median	P75	Mean	Std.Dev
Panel A: Stock Characteristics						
Market capitalization (in $millions)	26,180	1,045	2,911	9,131	11,782	31,998
Book-to-market ratio	26,179	0.211	0.364	0.589	0.454	0.375
Yearly stock return	26,180	-0.053	0.201	0.474	0.252	0.556
Annualized stock return volatility	26,180	0.289	0.403	0.567	0.463	0.264
Stock return skewness	26,180	-0.220	0.171	0.588	0.182	1.099
Stock return kurtosis	26,180	3.008	3.787	5.576	5.515	5.093
Panel B: Straddle Characteristics						
Moneyness	40,757	0.988	1.007	1.029	1.009	0.033
Days to maturity	40,757	17	29	44	31	16
Open interest (in 100s)	40,757	191	772	2,730	3,116	8,660
Volume (in 100s)	40,757	4	62	332	595	2,388
Implied volatility	40,757	0.334	0.447	0.597	0.491	0.221
Panel C: Daily Returns over the Whole Sample Period						
Straddle	1,592,357	-3.591%	-1.248%	1.527%	-0.575%	5.151%
Volatility	1,592,357	-2.289%	-0.048%	2.267%	0.036%	4.346%
Jump	1,592,357	-4.534%	-1.057%	2.905%	-0.611%	6.877%
Panel D: Daily Returns over [-3,1]						
Straddle	22,897	-3.501%	-1.080%	2.458%	0.208%	6.228%
Volatility	22,897	-0.980%	0.160%	1.368%	0.256%	2.368%
Jump	22,897	-3.969%	-1.066%	2.804%	-0.048%	6.453%

Panel A presents the stock characteristics including market capitalization, book-to-market ratio, yearly stock returns and higher moments of stock returns. For each quarter, we compute past one-year stock return up to last quarter end. Stock return volatility (annualized), skewness and kurtosis are computed using daily stock returns within a quarter. In total, we have more than 26,000 firm-quarter observations. The sample size is smaller compared to Gao et al. (2018) because our analysis requires long-term as well as short-term straddles.

The median market capitalization is about $2.911 billion, and the median book-to-market ratio of our sample is 0.211. The median yearly stock return is 20.1%. The median annualized stock return volatility is 40.3%. The medians for skewness and kurtosis are 0.171 and 3.787. Our summary statistics in Panel A are generally in line with those in Gao et al. (2018).

Panel B presents summary statistics for option straddles. The median moneyness level is 1.007, confirming our options are at-the-money. The median days to maturity are 31 days. We compute open interest (volume) for a straddle as the sum of open interests (volumes) of call and put in the straddle on the straddle construction day. The median open interest is 772 round lots and the median daily volume is 62 round lots, which indicate reasonable level of option liquidity in our sample. The median implied volatility is 44.7%, suggesting high earnings announcement uncertainty priced in options.

Panel C reports the daily delta-neutral straddle returns and the decomposed volatility-jump component returns during the whole sample period. Panel C shows that the median daily returns for straddles (-1.248%), volatility component (-0.048%), and jump component (-1.057%) are all negative. All median returns are more negative than the mean returns, indicating that the returns are positively skewed. These summary statistics suggest that straddle buyers pay premiums for protection against both volatility and jump risks, leading to negative returns for the straddle and its volatility-jump components. Trading straddle sometimes generates large positive returns, leading to positive skewness in returns. Panel D reports the daily delta-neutral straddle returns and the decomposed volatility-jump component returns on [-3,1]. Panel D shows that both the median and the mean daily returns for jump components around EADs are negative, and both the median and the mean volatility component returns are positive. These summary statistics give us some initial evidence that around EAD the volatility component is persistently underpriced, leading to the positive daily returns. The median (mean) of daily delta-neutral straddle returns is -1.080% (0.208%), both are larger than those of the whole sample in Panel C. In particular, the positive mean return suggests potentially underpriced straddles around EADs.

2.3 Empirical Analysis

2.3.1 Straddle return decomposition

We analyze straddle returns and its volatility-jump component returns around the EAD. We focus on a group of buy-and-hold strategies in which the delta-neutral straddles are purchased before the EAD and then sold on or after the EAD. For a strategy $[t_1, t_2]$, t_1 (< 0) is the straddle purchase day and t_2 (≥ 0) is the day when the straddle is sold. Table 2.2 reports the delta-neutral straddle returns and the volatility-jump component returns for each of these strategies. Panel A (B) shows the pooled sample (time-series sample) results. Within a firm, either equal weights or dollar open interest weights are used when there are multiple at-the-money straddles.

Table 2.2: Straddle Returns and Volatility-Jump Component Returns around Earnings Announcements

This table reports the mean values of delta-neutral at-the-money straddle returns and its volatility-jump component returns around earnings announcements, together with their t-statistics. Day 0 is the earnings announcement day. For a stock with more than one pair of at-the-money straddles, average returns are calculated using either equal weights or dollar open interest weights. Panel A reports the pooled sample returns. Panel B reports the time-series sample returns where we first compute the quarterly equal-weighted returns across firms, and then we average returns over all quarters. For Panel B, the t-statistics are computed using Newey–West (1987) standard errors with 3 lags. The sample period is from Jan. 1996 to Dec. 2013. ***, **, and * indicate significance at 1%, 5% and 10% levels.

Holding Period	Straddle	t-Stat.	Volatility	t-Stat.	Jump	t-Stat.
Panel A: Pooled Sample						
Equal Weight within Firms						
[-3,-1]	1.72%***	25.57	0.10%**	2.19	1.62%***	20.05
[-3,0]	1.89%***	15.90	0.41%***	7.31	1.48%***	11.63
[-3,1]	0.63%***	3.82	1.31%***	21.32	-0.67%***	-3.99
[-1,0]	1.03%***	10.39	0.49%***	10.55	0.54%***	4.92
[-1,1]	0.53%***	3.40	1.35%***	24.96	-0.82%***	-5.07
Dollar Open-Interest Weight within Firms						
[-3,-1]	1.79%***	25.86	0.10%**	2.05	1.69%***	20.29
[-3,0]	1.93%***	15.62	0.41%***	7.19	1.51%***	11.46
[-3,1]	0.58%***	3.40	1.33%***	21.15	-0.75%***	-4.27
[-1,0]	1.01%***	9.87	0.50%***	10.66	0.51%***	4.46
[-1,1]	0.45%***	2.84	1.37%***	24.89	-0.92%***	-5.51
Panel B: Time-Series Sample						
Equal Weight within Firms						
[-3,-1]	1.58%***	7.17	0.15%	1.30	1.43%***	5.88
[-3,0]	1.99%***	5.43	0.36%***	2.59	1.63%***	4.05
[-3,1]	0.89%*	1.65	1.16%***	6.70	-0.27%	-0.49
[-1,0]	1.35%***	4.66	0.96%***	2.87	0.39%***	4.90
[-1,1]	1.06%**	2.04	1.18%***	9.85	-0.12%	-0.22
Dollar Open-Interest Weight within Firms						
[-3,-1]	1.65%***	7.25	0.14%	1.21	1.51%***	6.06
[-3,0]	2.03%***	5.32	0.36%***	2.58	1.68%***	4.02
[-3,1]	0.85%	1.52	1.18%***	6.70	-0.33%	-0.58
[-1,0]	1.35%***	4.47	0.41%***	5.07	0.94%***	2.71
[-1,1]	1.01%*	1.88	1.21%***	9.97	-0.20%	-0.34

In Panel A, straddle returns are positive across all strategies with different within-firm weighting schemes, with 1% significance level. Straddle returns on [-3,-1] (1.72% and 1.79%) are lower than the returns on [-3,0] (1.89% and 1.93%) and higher than the returns on [-3,1] (0.63% and 0.58%). Straddle returns on [-1,0] (1.03% and 1.01%) are higher than the returns on [-1,1] (0.53% and 0.45%). There is a pattern that straddle return goes up before the EAD and drops after the EAD. The significantly positive returns on [-3,1] and [-1,1] indicate that the option traders who sell the straddles before the EAD underestimate the earnings announcement uncertainty.

The volatility component and the jump component returns show distinct patterns. The volatility component returns are significantly positive across the five strategies with 1% significance level, except for strategy [-3,-1] (positive return with 5% significance level). The volatility component returns on [-3,0] (0.41% for both weighting schemes) are higher than the returns on [-3,-1] (0.10% for both weighting schemes) and lower than the return on [-3,1] (1.31% and 1.33%). The volatility component returns on [-1,0] (0.49% and 0.50%) are lower than the returns on [-1,1] (1.35% and 1.37%). The volatility component return is always positive and keeps increasing over the announcement period, indicating a persistent underpricing of the volatility component.

The jump component returns are significantly positive for strategies [-3,-1] (1.62% and 1.69%), [-3,0] (1.48% and 1.51%) and [-1,0] (0.54% and 0.51%), with 1% significance level. In contrast, its returns are significantly negative for strategies [-3,1] (-0.67% and -0.75%) and [-1,1] (-0.82% and -0.92%), also with 1% significance level. The results indicate that the jump component returns capture the uncertainty run-up in the pre-announcement period and the uncertainty decline in the post-announcement period. Option traders anticipate earnings-induced jumps. The straddle buyers before the announcements pay jump risk premium and such premium manifests itself in negative post-announcement jump component returns.

The pooled sample results may lose some information on the time variation of straddle returns. The returns calculated using the pooled sample may give more weights to observations in recent years because option trading is more active in recent years.

In Panel B, we examine the time series sample in which the returns are first averaged over all firms in each quarter and then averaged across all quarters. The results in Panel B confirm the general patterns we find in Panel A, although some returns are not significantly different from zero. The straddle and the jump component returns are both significantly positive during the pre-announcement period and decrease after the announcement date, whereas the volatility component return increases over the announcement periods. To provide a more intuitive understanding of the return patterns, we visualize these patterns in Figure 2.1.

Figure 2.1 illustrates the straddle return decomposition on the time period [-3,1] with the time series sample. Within firms, dollar open-interest weights are used when there are multiple at-the-money straddles.

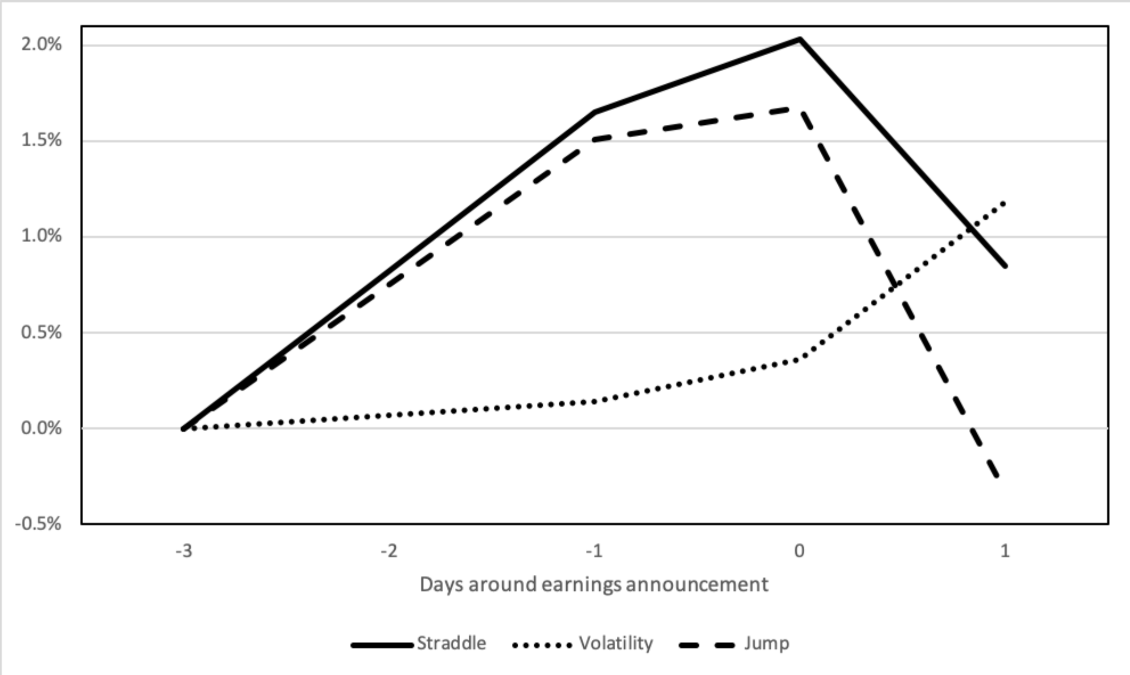

Figure 2.1: Straddle and Volatility-Jump Component Returns around Earnings Announcements

This figure displays the average straddle and volatility-jump component returns over [-3,1], where day 0 is the earnings announcement day. When multiple straddles are available for a firm, the dollar open-interest weighted returns are used. Returns are aggregated across firms in each quarter and the average quarterly returns are displayed. The data are from Jan. 1996 to Dec. 2013. The solid line represents the delta-neutral straddle returns, the dotted line represents the volatility component returns, and the dashed line represents the jump component returns.

Figure 2.1 shows that the straddle return and the volatility-jump component returns are all positive during the pre-announcement period and the jump component return is much higher than the volatility component return. After the EAD, the jump component's cumulative return goes down to negative, while the volatility component's cumulative return continues its increase. The straddle's cumulative return also goes down but is still positive. Figure 2.1 shows that the patterns of straddle return are mainly driven by its jump component and the positive cumulative straddle return after the EAD is mainly driven by its volatility component.

There are some important points we want to discuss before we move on to further tests. The first point is about the magnitude of returns. The cumulative straddle returns in Figure 2.1 are smaller than those in Gao et al. (2018). For example, our cumulative return on [-3,0] is 2.03%, while Gao et al. (2018) report 3.34% on [-3,0] . The main reason is that Gao et al. (2018, p2593) calculate the straddle delta as the dollar-weighted portfolio delta, whereas our study uses the quantity-weighted portfolio delta.[6]

For a portfolio consisting of a quantity w_i of option i, its quantity-weighted portfolio delta is given by:

$$\Delta = \sum_{i=1}^{n} w_i \Delta_i$$

where Δ_i is the delta of the ith option. The quantity-weighted portfolio delta is the common way to compute portfolio delta and determines the position in the options to construct a delta-neutral portfolio (e.g. Hull 2018, Chap 19. p 430). In our paper, we use the quantity-weighted portfolio delta throughout. We note that although the dollar-weighted portfolio delta inflates straddle returns, it does not change the general patterns of straddle returns.

The second point is about the static nature of our decomposition and the dynamic nature of option Greeks. A delta-neutral straddle constructed on day -3 may not be delta-neutral on day 1 due to the change of option deltas. A vega-neutral or gamma-neutral portfolio's vega or gamma may also change over the time. Figure 2.2 illustrates the boxplots of volatility component aV's and jump component bJ's delta, vega, and gamma around the EAD. The straddles are constructed and decomposed

[6]There is one more subtle reason. Our sample leans towards more liquid stocks as our decomposition method requires at least two straddles with short and long maturities. And more liquid stocks may have lower straddle returns around the EAD. But the return difference will not be large.

on day -3. On day t, we use day $t-1$'s closing time option Greeks to calculate portfolio's Greeks.[7]

Figure 2.2 shows that the median deltas of the volatility component aV are around 0. The dispersion increases from day -2 to day 1, but there is no systematic bias toward either positive or negative delta. The distribution of aV's vega is relatively stable with all vega values being positive. The median gammas of aV are approximately 0. Though the dispersion increases over time, they are generally symmetric around 0.

The median deltas of the jump component bJ are approximately 0, with increasing dispersion over time; but there is no systematic bias toward either positive or negative delta. The distribution of bJ's gamma is relatively stable with all gammas being positive. The median, together with upper and lower quartiles of vega, turn negative from day -1, suggesting that bJ's returns also reflect some exposure to negative vega.

For the absolute magnitude of the Greeks, volatility component aV's median vega is much larger than jump component bJ's median vega whereas jump component bJ's median gamma is much larger than volatility component aV's median gamma. aV (bJ) captures more exposure on volatility (jump) risk than bJ (aV).

On balance, the volatility component aV should serve as a reasonable vega-positive proxy of volatility risk. The return patterns of aV is not likely to reflect any systematic bias toward either positive or negative deltas and gammas. The return patterns of bJ need adjustment to remove the exposure to negative vega. This adjustment can be done by examining vega-positive aV portfolio's returns. The cumulative returns of aV on [-3,1] show a monotonically increasing pattern. So the adjustment to bJ's

[7]Using the average of t and $t-1$'s closing time option Greeks does not qualitatively change our findings.

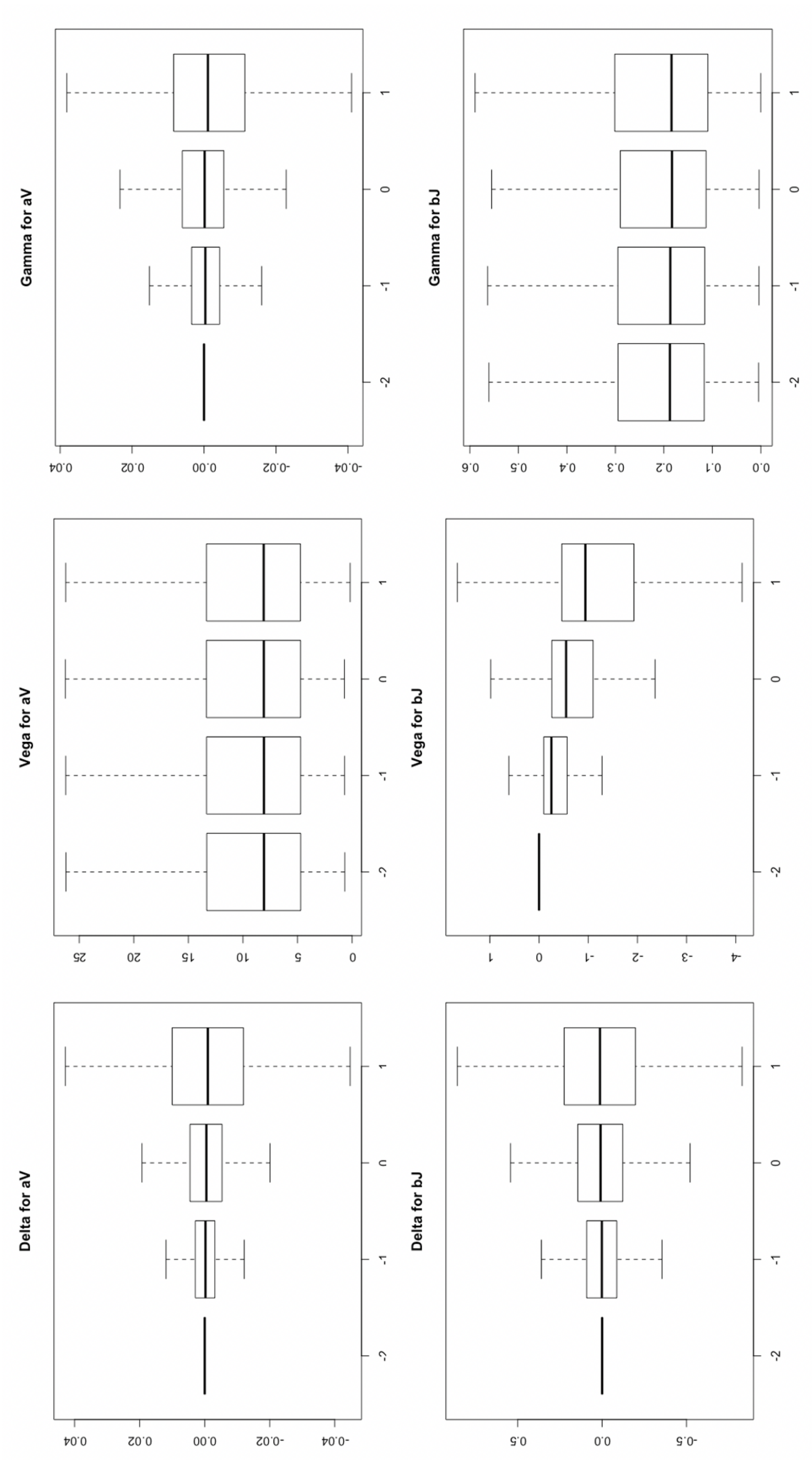

Figure 2.2: Greeks for aV and bJ around Earnings Announcements

This figure displays the boxplots of the Greeks (delta, vega, and gamma) for the volatility component portfolio aV and jump component portfolio bJ on [-2,1], where day 0 is the earnings announcement day and day -3 is the straddle construction and decomposition day. Day $t-1$'s option Greeks at market closing time are used to compute day t's portfolio Greeks.

cumulative returns is a monotonically decreasing return component; with absolute return magnitude being much lower than aV's returns. After such adjustment, the jump component's returns will shift upward, reinforcing our previous finding that straddle return's patterns are mainly driven by its jump component. Since the vega exposure embedded in the jump component is much smaller comparing to the volatility component, the adjustment is not likely to change our previous finding that the positive cumulative straddle return on day 1 (post-announcement) is mainly driven by its volatility component. If there is an underpricing of straddles before the EAD, the underpricing still much likely comes from the volatility component.

The last point is about the data on day 0 (the EAD). Strategies with ending period on day 0 (e.g. [-3,0]) capture a mixture of pre-announcement and post-announcement effects because some earnings announcements are released after the trading hours. Gao et al. (2018) do not separate the earnings announcements inside and outside of the market hours and this may bias some of their results. In our subsequent analysis, we focus on strategies [-3,-1] and [-3,1] to capture the pre-announcement and post-announcement effects separately. In Section 2.3.4, we re-examine the returns on [-3,0] by separating the firms that made announcements before or after the market close.

2.3.2 Portfolio sorting analysis

We use the portfolio sorting approach to investigate the factors that may impact straddle returns and the volatility-jump component returns. The factors fall into four categories: 1) high moments ; 2) stock price jumps; 3) cumulative abnormal returns around EADs; and 4) transaction costs. The first category includes annualized stock return volatility VOL and kurtosis $KURTOSIS$ using daily return data over the previous quarter. The second category includes jump frequency $JUMP\_FREQ$ and jump size $JUMP\_SIZE$ using past one year's daily stock return data. $JUMP\_FREQ$ and $JUMP\_SIZE$ are obtained by using the non-parametric method in Lee and Mykland (2008). The third category includes the variance of the cumulative abnormal return around EAD ($var(CAR)$) over the past eight quarters and the absolute value of the cumulative abnormal return around EAD ($|CAR|$) in the last quarter. The cumulative abnormal return is computed on [-1,1]. The fourth category includes the daily closing time bid-ask spreads scaled by closing price for options ($OPTION\_SPREAD$) and stocks ($STOCK\_SPREAD$), averaged over the previous quarter.

Gao et al. (2018) also examined these four categories of factors. They find that high noisiness of company information (proxied by higher moments, jumps, and the cumulative abnormal returns) and high transaction costs help explain the positive straddle returns around earnings announcements. Their study focuses on straddle returns over [-3,0]. To separate the pre- and post-announcement effects, our study focuses on straddle returns over [-3,-1] and [-3,1]. Information noisiness and transaction costs may impact straddle's volatility and jump component returns in different ways over different periods, and portfolio sorting based on these factors should uncover such difference.

For each quarter, we sort firms into four equal-sized groups based on one factor and calculate the average straddle and component returns in each group. When there is more than one qualified straddle for a firm, we calculate straddle returns using the dollar open interest weighting scheme. Finally, we compute the mean returns over 72 quarters for each group.

Table 2.3 reports the portfolio sorting results for returns on [-3,-1]. These results capture the pre-announcement impacts by the factors.

Panel A shows the portfolio sorting results based on high moments. For straddle returns, the spread between the highest and lowest volatility (kurtosis) groups "High-Low" is 1.52% (1.07%), with 1% significance level. In the pre-announcement period, higher stock return volatility and kurtosis predict higher straddle returns. For volatility component returns, the spread between the highest and lowest volatility groups is -0.83%, with 1% significance level. And the spread based on kurtosis sorting is not significantly different from 0. Stock return volatility negatively predicts the volatility component returns in the pre-announcement period. For jump component returns, the spread between the highest and lowest factor value groups sorted on volatility (kurtosis) is 2.35% (1.11%), with 1% significance level. Both volatility and kurtosis positively predict the jump component returns. In contrast, in untabulated results stock return skewness has insignificant impact on volatility-jump component returns as the vega and gamma exposures to stock movement are undirectional.

Panel B shows the portfolio sorting results based on jumps in previous stock returns. The spread of straddle returns between highest and lowest jump frequency (jump size) groups is 1.35% (0.97%), with 1% (5%) significance level. Jump frequency and jump size positively predict straddle returns. For the volatility component returns, the spreads based on jump frequency and jump size sorting are not significantly different from 0. For the jump component returns, the spread between the highest and lowest

Table 2.3: Portfolio Sorting: Returns on [-3,-1]

This table reports the portfolio sorting results on straddle and its volatility-jump component returns over [-3,-1]. Day 0 is the earnings announcement day. In each quarter, firms are sorted into 4 equal-sized portfolios based on one of the previous-period characteristics, and the average time-series portfolio return together with the volatility-jump component returns are reported. For a stock with more than one pair of at-the-money straddles, average returns within firms are calculated using dollar open interest weights. Returns are sorted by the following characteristics. Historical second moment VOL and fourth moment $KURTOSIS$ are computed using the last quarter daily returns. $JUMP\_FREQ$ and $JUMP\_SIZE$ are jump frequency and jump size measures observed at the last quarter-end, following Lee and Mykland (2008) procedure. The cumulative abnormal return, CAR, is computed over [-1,1] around earnings announcements and adjusted for the market return, and the absolute value of CAR ($|CAR|$) in the last quarter is taken. The variance of CAR ($var(CAR)$) is computed from the previous 8-quarter data. $OPTION\_SPREAD$ is calculated as the average closing time bid-ask spread scaled by the price of the option in the last quarter. $STOCK\_SPREAD$ is calculated as the average closing time bid-ask spread scaled by the price of the stock in the last quarter. The sample period is from Jan. 1996 to Dec. 2013. The t-statistics are computed using Newey-West (1987) standard errors with 3 lags. ***, **, and * indicate significance at 1%, 5% and 10% levels.

	Straddle	*t*-Stat.	Volatility	*t*-Stat.	Jump	*t*-Stat.		Straddle	*t*-Stat.	Volatility	*t*-Stat.	Jump	*t*-Stat.
Panel A: Sort on Past High Moments													
VOL							KURTOSIS						
Low	1.35%***	3.28	0.73%***	4.69	0.62%	1.41	Low	1.48%***	3.23	0.30%*	1.70	1.19%**	2.47
2	1.63%***	3.55	0.64%***	4.33	1.00%**	2.11	2	2.02%***	4.40	0.57%***	2.73	1.45%***	2.64
3	2.49%***	4.62	0.13%	0.59	2.36%***	4.03	3	2.14%***	4.19	0.26%	1.62	1.88%***	3.58
High	2.87%***	6.00	-0.10%	-0.44	2.97%***	5.14	High	2.56%***	5.27	0.26%	1.42	2.30%***	4.20
High-Low	1.52%***	3.55	-0.83%***	-3.94	2.35%***	5.20	High-Low	1.07%***	3.12	-0.04%	-0.26	1.11%***	2.87
Panel B: Sort on Past Jumps													
JUMP_FREQ							JUMP_SIZE						
Low	1.44%***	3.80	0.32%**	2.21	1.12%***	2.92	Low	1.36%***	3.39	0.24%	1.02	1.12%**	2.30
2	1.64%***	3.60	0.36%**	2.15	1.28%***	2.61	2	2.16%***	4.09	0.25%	1.49	1.91%***	3.27
3	2.42%***	4.73	0.37%**	2.25	2.05%***	3.49	3	2.77%***	4.65	0.46%*	1.87	2.31%***	3.32
High	2.79%***	5.82	0.38%**	2.12	2.41%***	4.34	High	2.33%***	5.00	0.44%***	3.03	1.89%***	3.97
High-Low	1.35%***	3.66	0.06%	0.53	1.29%***	3.09	High-Low	0.97%**	2.49	0.20%	0.93	0.77%*	1.80
Panel C: Sort on Past Earning Surprise													
var(CAR)							$\|CAR\|$						
Low	1.27%***	5.24	0.32%**	2.51	0.95%***	3.98	Low	1.54%***	6.90	0.24%*	1.89	1.30%***	5.52
2	1.55%***	5.09	0.19%	1.18	1.35%***	3.76	2	1.52%***	5.58	0.27%*	1.92	1.25%***	4.35
3	1.92%***	6.92	0.12%	0.88	1.81%***	6.60	3	1.60%***	5.17	0.11%	0.66	1.49%***	3.88
High	2.67%***	9.82	0.09%	0.60	2.58%***	7.41	High	2.64%***	8.77	0.02%	0.11	2.62%***	7.18
High-Low	1.40%***	5.70	-0.23%**	-2.49	1.63%***	6.26	High-Low	1.10%***	4.89	-0.22%	-1.22	1.32%***	4.11
Panel D: Sort on Transaction Costs													
OPTION_SPREAD							STOCK_SPREAD						
Low	1.19%***	2.72	0.26%	1.36	0.93%**	2.06	Low	1.11%***	2.66	0.87%***	5.11	0.25%	0.57
2	1.88%***	4.20	0.43%***	2.61	1.45%***	2.76	2	1.83%***	3.93	0.64%***	4.57	1.20%***	2.61
3	2.60%***	5.36	0.40%**	2.02	2.21%***	4.45	3	2.35%***	4.42	0.03%	0.16	2.32%***	3.93
High	2.48%***	4.20	0.32%	1.44	2.16%***	3.27	High	2.91%***	5.61	-0.13%	-0.56	3.04%***	5.05
High-Low	1.29%**	2.10	0.06%	0.25	1.23%*	1.94	High-Low	1.79%***	3.39	-0.99%***	-4.58	2.79%***	5.53

factor value groups sorted on jump frequency (jump size) is 1.29% (0.77%), with 1% (10%) significance level. Jump frequency and jump size positively predict the jump component returns.

Panel C shows the portfolio sorting results based on $var(CAR)$ and $|CAR|$. The spread of straddle returns between the highest and lowest $var(CAR)$ ($|CAR|$) groups is 1.40% (1.10%), with 1% significance level. Both $var(CAR)$ and $|CAR|$ positively predict straddle returns, suggesting that higher past earnings surprise leads to higher straddle returns. For the volatility component returns, the spread between the highest and lowest $var(CAR)$ groups is -0.23%, with 5% significance level. The spread based on $|CAR|$ sorting is not significantly different from 0. The results suggest that $var(CAR)$ negatively predicts the volatility component returns. The spread of the jump component returns between the highest and lowest $var(CAR)$ ($|CAR|$) groups is 1.63% (1.32%), with 1% significance level. Both $var(CAR)$ and $|CAR|$ positively predict the jump component returns.

Panel D shows the portfolio sorting results based on transaction costs. The spread of straddle returns between the highest and lowest $OPTION\_SPREAD$ ($STOCK\_SPREAD$) groups is 1.29% (1.79%), with 5% (1%) significance level. The spread of the volatility component returns between the highest and lowest $OPTION\_SPREAD$ groups is not significantly different from 0 whereas the spread between the highest and lowest $STOCK\_SPREAD$ groups is -0.99% with 1% significance level. The spread of the jump component returns between the highest and lowest $OPTION\_SPREAD$ ($STOCK\_SPREAD$) groups is 1.23% (2.79%), with 10% (1%) significance level.

On balance, the results in Tables 2.3 suggest that in the pre-announcement period, high information noisiness and transaction costs increase both straddle and the jump component returns but decrease the volatility component returns. The increase of both straddle and jump component returns shows that the run-up of uncertainty in the pre-announcement period is mainly driven by the run-up of the uncertainty on earnings-induced jumps. When information noisiness (transaction cost) is higher, option traders focus more on jump risk induced by earnings announcements, driving up the jump component's return. When stock price discontinuity is the focal point, option traders focus less on volatility risk and the volatility component's return declines.

Table 2.4 reports the portfolio sorting results for returns on [-3,1]. These results capture the post-announcement impacts by the factors.

Panel A shows the portfolio sorting results based on high moments. The spread of straddle returns between the highest and lowest volatility groups is not significantly different from 0. The spread based on kurtosis sorting is -0.74%, with 10% significance level. In the post-announcement period, higher kurtosis predicts lower straddle returns. The relation is opposite to what we find in the pre-announcement period.

The spread of the volatility component returns between the highest and lowest volatility groups is -0.61%, with 1% significance level. The spread based on kurtosis sorting is not significantly different from 0. Stock return volatility negatively predicts the volatility component returns in the post-announcement period as well as the pre-announcement period. When historical stock return volatility is higher, options market prices in higher volatility risk in straddle price, leading to lower pre- and post-announcement volatility component returns.

Table 2.4: Portfolio Sorting: Returns on [-3,1]

This table reports the portfolio sorting results on straddle and its volatility-jump component returns over [-3,1]. Day 0 is the earnings announcement day. In each quarter, firms are sorted into 4 equal-sized portfolios based on one of the previous-period characteristics, and the average time-series portfolio return together with the volatility-jump component returns are reported. For a stock with more than one pair of at-the-money straddles, average returns within firms are calculated using dollar open interest weights. Returns are sorted by the following characteristics. Historical second moment VOL and fourth moment $KURTOSIS$ are computed using the last quarter daily returns. $JUMP\_FREQ$ and $JUMP\_SIZE$ are jump frequency and jump size measures observed at the last quarter-end, following Lee and Mykland (2008) procedure. The cumulative abnormal return, CAR, is computed over [-1,1] around earnings announcements and adjusted for the market return, and the absolute value of CAR ($|CAR|$) in the last quarter is taken. The variance of CAR ($var(CAR)$) is computed from the previous 8-quarter data. $OPTION\_SPREAD$ is calculated as the average closing time bid-ask spread scaled by the price of the option in the last quarter. $STOCK\_SPREAD$ is calculated as the average closing time bid-ask spread scaled by the price of the stock in the last quarter. The sample period is from Jan. 1996 to Dec. 2013. The t-statistics are computed using Newey-West (1987) standard errors with 3 lags. ***, **, and * indicate significance at 1%, 5% and 10% levels.

	Straddle	*t*-Stat.	Volatility	*t*-Stat.	Jump	*t*-Stat.		Straddle	*t*-Stat.	Volatility	*t*-Stat.	Jump	*t*-Stat.
Panel A: Sort on Past High Moments													
VOL							KURTOSIS						
Low	0.81%	1.59	1.39%***	6.14	-0.59%	-1.18	Low	1.07%	1.52	1.13%***	5.34	-0.06%	-0.09
2	0.90%	1.29	1.52%***	7.76	-0.62%	-0.89	2	1.41%***	2.59	1.25%***	4.82	0.16%	0.24
3	0.95%	1.02	1.02%***	4.55	-0.07%	-0.07	3	0.58%	1.00	1.19%***	6.47	-0.61%	-1.01
High	0.74%	1.06	0.78%***	3.22	-0.05%	-0.07	High	0.33%	0.39	1.30%***	6.56	-0.97%	-1.12
High-Low	-0.07%	-0.10	-0.61%***	-2.59	0.54%	0.89	High-Low	-0.74%*	-1.82	0.16%	1.03	-0.91%**	-2.08
Panel B: Sort on Past Jumps													
JUMP_FREQ							JUMP_SIZE						
Low	0.64%	1.03	1.12%***	5.81	-0.48%	-0.78	Low	0.63%	1.22	1.01%***	4.46	-0.38%	-0.65
2	0.59%	0.90	1.16%***	5.45	-0.57%	-0.82	2	1.38%*	1.85	1.14%***	5.75	0.24%	0.32
3	1.03%	1.55	1.26%***	6.51	-0.23%	-0.31	3	0.87%	1.17	1.43%***	5.69	-0.56%	-0.73
High	1.33%*	1.95	1.27%***	5.95	0.06%	0.08	High	0.38%	0.56	1.26%***	6.26	-0.88%	-1.31
High-Low	0.69%	1.28	0.15%	1.24	0.54%	0.92	High-Low	-0.25%	-0.59	0.25%	1.25	-0.50%	-1.12
Panel C: Sort on Past Earning Surprise													
var(CAR)							$\|CAR\|$						
Low	0.42%	0.61	1.08%***	4.38	-0.67%	-1.04	Low	0.73%	1.14	1.15%***	4.95	-0.42%	-0.68
2	0.66%	0.99	1.27%***	5.29	-0.61%	-0.85	2	0.22%	0.41	1.24%***	6.97	-1.02%*	-1.81
3	1.03%	1.64	1.05%***	4.95	-0.02%	-0.02	3	1.05%	1.45	1.03%***	5.30	0.02%	0.02
High	1.49%*	1.65	1.25%***	5.59	0.24%	0.26	High	1.63%**	1.97	1.32%***	5.41	0.31%	0.33
High-Low	1.08%	1.44	0.17%	0.63	0.91%	1.25	High-Low	0.90%*	1.67	0.17%	0.91	0.73%	1.38
Panel D: Sort on Transaction Costs													
OPTION_SPREAD							STOCK_SPREAD						
Low	0.93%	1.27	1.14%***	4.25	-0.22%	-0.31	Low	1.21%**	2.10	1.46%***	6.74	-0.25%	-0.44
2	1.05%	1.39	1.32%***	7.05	-0.28%	-0.34	2	0.33%	0.52	1.38%***	6.52	-1.05%	-1.57
3	0.62%	0.95	1.26%***	5.10	-0.63%	-0.91	3	0.53%	0.66	0.99%***	4.66	-0.46%	-0.53
High	-0.01%	-0.01	1.60%***	8.23	-1.61%**	-2.14	High	1.30%*	1.83	0.91%***	3.74	0.40%	0.55
High-Low	-0.93%	-1.05	0.46%*	1.71	-1.39%**	-1.99	High-Low	0.10%	0.16	-0.55%**	-2.37	0.65%	1.17

The spread of the jump component returns between the highest and lowest volatility groups is not significantly different from 0. The spread based on kurtosis sorting is -0.91%, with 5% significance level. In the post-announcement period, kurtosis negatively predicts the jump component returns. The relation is opposite to the findings in the pre-announcement period.

Panel B shows the portfolio sorting results based on jumps in previous stock returns. All spreads are insignificant, suggesting that past jumps have no return predictability on straddles and the volatility-jump components return in the post-announcement period.

Panel C shows the portfolio sorting results based on $var(CAR)$ and $|CAR|$. $var(CAR)$ losses its predicting power on all spreads. The spread of straddle returns between the highest and lowest $|CAR|$ groups is 0.9%, with 10% significance level.

Panel D shows the portfolio sorting results based on transaction costs. The spreads of straddle returns are insignificant. The impact of $OPTION\_SPREAD$ ($STOCK\_SPREAD$) on volatility component return spread is positive (negative), yielding inconclusive results. The spread of the jump component returns between the highest and lowest $OPTION\_SPREAD$ groups is -1.39%, with 5% significance level. The spread based on $STOCK\_SPREAD$ sorting is insignificant.

On balance, the results in Table 2.4 suggest that in the post-announcement period, information noisiness and transaction costs have little or mixed impacts on the volatility component returns. The generally insignificant or negative impacts on the straddle and the jump component returns are in opposite to the observations in the pre-announcement period, suggesting that high information noisiness and transaction costs loss their positive return predictability after the earnings information is released on day 1.

Combining the results in Tables 2.3 and 2.4, we find that in both pre- and post-announcement periods, the jump component return and straddle return respond to the sorting factors in similar ways. Options market puts its emphasis on earnings-induced jumps before the EAD, and such emphasis also dominates the patterns of straddle returns.

2.3.3 Fama-Macbeth regressions

We further examine the impact factors of straddle and its component returns around earnings announcement in the cross section using Fama-Macbeth (1973) regressions. We run quarterly cross-sectional regressions for straddle returns and its components' returns during the pre-announcement period [-3,-1] and the post-announcement period [-3,1]. For each regression, the variable of interest is one of the factors we use in the portfolio sorting analysis. Control variables include option maturity, option moneyness, firm size, book-to-market ratio, and the last yearly stock return. To avoid multicollinearity, we remove firm size in regressions involving $OPTION\_SPREAD$ and $STOCK\_SPREAD$. We add VOL and stock return skewness into control variables in regressions involving $KURTOSIS$. We report the estimated coefficients on the variables of interest, together with Newey-West adjusted t-statistics and the adjusted R-squared. The estimated coefficients on control variables and the intercept are omitted for brevity, but full results are available upon request. Table 2.5 reports the regression results of returns on [-3,-1].

Table 2.5: Fama-Macbeth Regressions: Returns on [-3,-1]

This table reports the results of the Fama-MacBeth (1973) regressions. The dependent variable is the straddle return, the volatility component return or the jump component return over [-3,-1]. Day 0 is the earnings announcement day. The variable of interest is one of the following characteristics. Historical second moment VOL and fourth moment $KURTOSIS$ are computed using the last quarter daily returns. $JUMP\_FREQ$ and $JUMP\_SIZE$ are jump frequency and jump size measures observed at the last quarter-end, following Lee and Mykland (2008) procedure. The cumulative abnormal return, CAR, is computed over [-1,1] around earnings announcements and adjusted for the market return, and the absolute value of CAR ($|CAR|$) in the last quarter is taken. The variance of CAR ($var(CAR)$) is computed from the previous 8-quarter data. $OPTION\_SPREAD$ is calculated as the average closing time bid-ask spread scaled by the price of the option in the last quarter. $STOCK\_SPREAD$ is calculated as the average closing time bid-ask spread scaled by the price of the stock in the last quarter. Control variables include option maturity, option moneyness, firm size, book-to-market ratio, and the last year's stock return. In regressions of $OPTION\_SPREAD$ and $STOCK\_SPREAD$, the control variable firm size is removed. In regression of $KURTOSIS$, additional control variables VOL and stock return skewness are added. The sample period is from Jan. 1996 to Dec. 2013. The average coefficient for the variable of interest and the adjusted R^2 are reported. The t-statistics are computed using Newey-West (1987) standard errors with 3 lags. ***, **, and * indicate significance at 1%, 5% and 10% levels.

	Straddle		Volatility		Jump	
	Coef.	Adj. R^2	Coef.	Adj. R^2	Coef.	Adj. R^2
ln(VOL)	0.222	0.65%	-0.189**	0.48%	0.411**	0.59%
	[1.24]		[-2.31]		[2.47]	
ln(KURTOSIS)	0.001***	0.87%	0.000	0.59%	0.001***	0.58%
	[4.28]		[0.42]		[3.09]	
JUMP_FREQ	0.002**	0.53%	0.000	0.28%	0.001	0.43%
	[2.29]		[1.37]		[1.51]	
JUMP_SIZE	0.000	0.46%	0.000	-0.03%	0.000	0.47%
	[1.36]		[-1.03]		[1.56]	
var(CAR)	0.196***	0.93%	-0.002	0.41%	0.199***	0.80%
	[6.15]		[-0.11]		[5.62]	
$\|CAR\|$	0.123***	0.75%	-0.007	0.40%	0.130***	0.66%
	[5.95]		[-0.44]		[4.68]	
OPTION_SPREAD	0.035	0.27%	-0.001	0.34%	0.037	0.18%
	[1.61]		[-0.12]		[1.53]	
STOCK_SPREAD	0.530***	0.40%	-0.240***	0.52%	0.771***	0.31%
	[3.76]		[-3.42]		[5.68]	

For regressions on straddle returns, the estimated coefficients on $ln(KURTOSIS)$, $var(CAR)$, $|CAR|$, $STOCK\_SPREAD$ are 0.001, 0.196, 0.123, and 0.530 respectively, all being significant at 1% level. The estimated coefficient on $JUMP\_FREQ$ is 0.002, being significant at 5% level. The estimated coefficients on $ln(VOL)$, $JUMP\_SIZE$, and $OPTION\_SPREAD$ are insignificant. Overall, information noisiness and transaction costs positively predict pre-announcement straddle returns.

For regressions on the volatility component returns, the estimated coefficients on $ln(VOL)$ and $STOCK\_SPREAD$ are -0.189 and -0.240, significant at 5% and 1% levels respectively. The estimated coefficients are insignificant for other factors. Overall, a few regression results show that information noisiness and transaction costs negatively predict pre-announcement volatility component returns.

For regressions on the jump component returns, the estimated coefficients on $ln(KURTOSIS)$, $var(CAR)$, $|CAR|$, $STOCK\_SPREAD$ are 0.001, 0.199, 0.130, and 0.771 respectively, significant at 1% level. The estimated coefficient on $ln(VOL)$ is 0.411, significant at 5% level. The estimated coefficients on $JUMP\_FREQ$, $JUMP\_SIZE$, and $OPTION\_SPREAD$ are positive and insignificant. Overall, information noisiness and transaction costs positively predict pre-announcement jump component returns.

Table 2.5 confirms our previous finding that in pre-announcement period, high information noisiness and transaction costs lead to high straddle and jump component returns but low volatility component returns. The run-up of uncertainty is mainly the run-up of jump uncertainty induced by earnings announcements.

Table 2.6 reports the regression results for returns on [-3,1].

For regressions on straddle returns, the estimated coefficient on $JUMP\_SIZE$ is -0.001, significant at 1% level. The estimated coefficients on $var(CAR)$ and $|CAR|$ are 0.125 and 0.077, significant at 10% level. The estimated coefficients on other factors are insignificant.

For regressions on the volatility component returns, the estimated coefficients on

Table 2.6: Fama-Macbeth Regressions: Returns on [-3,1]

This table reports the results of the Fama-MacBeth (1973) regressions. The dependent variable is the straddle return, the volatility component return or the jump component return over [-3,1]. Day 0 is the earnings announcement day. The variable of interest is one of the following characteristics. Historical second moment VOL and fourth moment $KURTOSIS$ are computed using the last quarter daily returns. $JUMP\_FREQ$ and $JUMP\_SIZE$ are jump frequency and jump size measures observed at the last quarter-end, following Lee and Mykland (2008) procedure. The cumulative abnormal return, CAR, is computed over [-1,1] around earnings announcements and adjusted for the market return, and the absolute value of CAR ($|CAR|$) in the last quarter is taken. The variance of CAR ($var(CAR)$) is computed from the previous 8-quarter data. $OPTION\_SPREAD$ is calculated as the average closing time bid-ask spread scaled by the price of the option in the last quarter. $STOCK\_SPREAD$ is calculated as the average closing time bid-ask spread scaled by the price of the stock in the last quarter. Control variables include option maturity, option moneyness, firm size, book-to-market ratio, and the last year's stock return. In regressions of $OPTION\_SPREAD$ and $STOCK\_SPREAD$, the control variable firm size is removed. In regression of $KURTOSIS$, additional control variables VOL and stock return skewness are added. The sample period is from Jan. 1996 to Dec. 2013. The average coefficient for the variable of interest and the adjusted R^2 are reported. The t-statistics are computed using Newey-West (1987) standard errors with 3 lags. ***, **, and * indicate significance at 1%, 5% and 10% levels.

	Straddle		Volatility		Jump	
	Coef.	Adj. R^2	Coef.	Adj. R^2	Coef.	Adj. R^2
ln(VOL)	0.001	1.10%	-0.004	0.86%	0.005	0.66%
	[0.18]		[-1.23]		[0.74]	
ln(KURTOSIS)	-0.003	1.32%	0.004***	0.91%	-0.007*	0.84%
	[-0.61]		[3.53]		[-1.71]	
JUMP_FREQ	0.001	0.97%	0.000	0.66%	0.001	0.53%
	[1.31]		[0.30]		[1.12]	
JUMP_SIZE	-0.001***	0.92%	0.000	0.50%	-0.001***	0.43%
	[-4.05]		[-0.97]		[-3.64]	
var(CAR)	0.125*	1.35%	0.057**	1.00%	0.068	0.92%
	[1.89]		[2.09]		[1.10]	
$\|CAR\|$	0.077*	1.21%	0.027*	0.75%	0.050	0.71%
	[1.89]		[1.84]		[1.16]	
OPTION_SPREAD	-0.011	0.97%	-0.002	0.84%	-0.010	0.71%
	[-0.43]		[-0.16]		[-0.40]	
STOCK_SPREAD	0.061	1.04%	-0.172**	0.86%	0.230	0.65%
	[0.32]		[-2.26]		[1.31]	

$ln(KURTOSIS)$, $var(CAR)$ and $|CAR|$ are 0.004, 0.057 and 0.027, significant at 1%, 5% and 10% levels. The estimated coefficient on $STOCK\_SPREAD$ is -0.172, significant at 10% level.

For regressions on the jump component returns, the estimated coefficients on $ln(KURTOSIS)$ and $JUMP\_SIZE$ are -0.007 and -0.001, significant at 10% and 1% levels. The estimated coefficients on other factors are insignificant.

Overall, the impacts of information noisiness and transaction costs in the post-announcement period are not as strong as in the pre-announcement period. Past jump size significantly negatively predicts both straddle and the jump component returns in the post-announcement period, suggesting that options market anticipated and priced in jump risk before the EAD. High information noisiness in general predicts high volatility component returns, suggesting underpricing of volatility risk is more persistent when information is more noisy.

Results in Tables 2.5 and 2.6 suggest that straddle returns' main patterns in the pre- and post-announcement periods are driven by its jump component returns as they response mostly in the same way to the same factors. Noisiness and transaction costs positively predicts pre-announcement straddle and the jump component returns but such predicting power disappears or reverses in the post-announcement period, suggesting options market anticipate earnings-induced jumps.

2.3.4 After-hours earnings announcements

Our previous empirical analysis focuses on returns over [-3,-1] and [-3,1], because they provide clear identification of pre- and post-announcement effects. Returns over [-3,0] capture a mixture of the pre-announcement effect and the post-announcement effect as the earnings announcements may occur outside of the normal trading hours. Berkman and Truong (2009) show the importance of separately considering after-hours earnings announcements as the stock price response to the announcements will be delayed by one day. If the earnings announcement is made after day 0's market close time, then the earnings information should be mainly reflected in the stock price on day 1, rather than on day 0.

The option markets close at 4:02 p.m. Eastern Standard Time (EST), while the stock markets close at 4:00 p.m. EST. We choose 4:00 p.m. as the cut-off time for after-hours earnings announcements. We examine the straddle returns and its volatility-jump component returns on [-3,0] by dividing the sample into two sub-samples: a) firms making announcements before 4:00 p.m. market close on day 0; and b) firms making announcements after 4:00 p.m. market close on day 0. The cumulative returns over [-3,0] for sub-sample a) mainly capture post-announcement effect, and returns over [-3,0] for sub-sample b) mainly capture pre-announcement effect.

Table 2.7 reports the straddle returns, volatility component returns, and jump component returns on [-3,0].

Table 2.7: After-hours Earnings Announcements

This table separately reports the average time-series straddle returns and its volatility-jump component returns over [-3,0] for the firms that make earnings announcements before (after) the market close time 4:00 p.m. in Panel A (B). Day 0 is the earnings announcement day. The sample period is from Jan. 1996 to Dec. 2013. For a stock with more than one pair of at-the-money straddles, average returns are calculated using equal weight or dollar open interest weight. The quarterly equal-weighted returns across firms are computed and then aggregated over all quarters to obtain the time-series average return. The t-statistics are computed using Newey–West (1987) standard errors with 3 lags. ***, **, and * indicate significance at 1%, 5% and 10% levels.

	Straddle	t-Stat.	Volatility	t-Stat.	Jump	t-Stat.
Panel A: Returns for Firms that Announce Earnings before Market Close						
Equal Weight within Firms	0.25%	0.49	0.83%***	4.75	-0.58%	-0.99
Dollar Open Interest Weight within Firms	0.22%	0.40	0.83%***	4.67	-0.61%	-1.03
Panel B: Returns for Firms that Announce Earnings after Market Close						
Equal Weight within Firms	3.70%***	6.54	-0.17%	-0.89	3.87%***	5.87
Dollar Open Interest Weight within Firms	3.82%***	6.57	-0.13%	-0.71	3.95%***	5.87

Panel A shows that for sub-sample a), both straddle and the jump component returns are insignificant, whereas the volatility component returns are significantly positive. The results are consistent to our findings on post-announcement returns.

Panel B shows that for sub-sample b), both straddle and the jump component returns are significantly positive, whereas the volatility component returns are insignificant. The results are consistent to our findings on pre-announcement returns.

Table 2.7 shows that the patterns of straddle and the jump component returns are similar. When we adjust for the after-hours earnings announcements, the high straddle return around the EAD is mainly a pre-announcement phenomenon.

2.3.5 Returns in recent years

In the previous sections, we examine the straddle and the components' returns during year 1996-2013 for a direct comparison with the results in Gao et al. (2018). Now we analyze the returns using the sample of 2014-2017, which includes the most recent available option data. We calculate the time-series average of straddle returns and the corresponding components' returns around earnings announcements, where we use equal weight across firms in each quarter and two different weighting schemes within firms. We report the results in Table 2.8.

Table 2.8: Straddle Returns and Volatility-Jump Component Returns: 2014-2017

This table reports time-series average straddle returns and its volatility-jump component around earnings announcements. Day 0 is the earnings announcement day. The sample period is from Jan. 2014 to Dec. 2017. At the firm level, for a stock with more than one pair of at-the-money straddles, average returns are calculated using equal weight or dollar open interest weight. The quarterly equal-weighted returns across firms are computed and then aggregated over all quarters to obtain the time-series average return. The *t*-statistics are computed using Newey–West (1987) standard errors with 3 lags. ***, **, and * indicate significance at 1%, 5% and 10% levels.

Holding Period	Straddle	*t*-Stat.	Volatility	*t*-Stat.	Jump	*t*-Stat.
Equal Weight within Firms						
[-3,-1]	0.66%***	2.74	0.91%	1.40	-0.25%	-0.34
[-3,0]	0.06%	0.26	0.76%***	2.73	-0.70%	-1.57
[-3,1]	-2.71%***	-4.54	1.45%***	3.83	-4.17%***	-11.19
[-1,0]	-0.03%	-0.08	0.93%***	3.43	-0.96%***	-2.88
[-1,1]	-2.26%***	-3.83	1.61%***	6.54	-3.87%***	-7.50
Dollar Open-Interest Weight within Firms						
[-3,-1]	0.69%***	3.46	0.84%	1.34	-0.15%	-0.23
[-3,0]	0.00%	0.00	0.77%***	2.70	-0.77%*	-1.87
[-3,1]	-2.88%***	-4.95	1.50%***	3.91	-4.38%***	-12.00
[-1,0]	-0.16%	-0.49	0.94%***	5.24	-1.10%***	-3.79
[-1,1]	-2.49%***	-4.21	1.57%***	8.25	-4.06%***	-9.45

Panel A (B) presents the returns using equal weights (dollar open interest weights) within firms. The return patterns in both panels are very consistent. As can be seen from the table, first, the straddle returns are lower for all strategies comparing to the returns in Table 2. Straddle returns on [-3,1] and [-1,1] even turn to be significantly negative. Second, the volatility component returns are positive across all strategies, and are being significant over [-3,0], [-3,1], [-1,0], and [-1,1]. Third, the jump component returns are negative across all strategies, and are being significant over [-3,1], [-1,0], and [-1,1].

Overall, the disappearance of the positive cumulative straddle returns after the EAD is evident in recent years. This result suggests that financial markets pay more attention to the straddle (option) underpricing issue in recent years. Options market puts even more emphasis on jumps, which is manifested by the significantly negative jump component returns around EAD in recent years. The volatility risk is still underestimated by options market around earnings announcement, leading to the significantly positive volatility component returns. Although the positive straddle returns around EADs disappear in recent years, the marked difference between volatility-jump component return patterns persists.

2.4 Conclusion

Option straddles have positive exposures to both volatility risk and jump risk. In this study, we decompose the straddles into one component with only volatility risk exposure, and the other component with only jump risk exposure. We examine the underpricing issue of the straddle around earnings announcements by analyzing the return patterns of its volatility component and jump component around the announcements. We find that the volatility component is consistently underpriced, generating significantly positive returns. Jump component's return is significantly positive over the pre-announcement period and becomes either insignificant or significantly negative after the announcements. We find that option traders anticipate the earnings-induced jump risk, while they underestimate the uncertainty of diffusive volatility surrounding the announcements.

Chapter 3

Measuring gambling activity in options market

3.1 Introduction

Recent research shows that investors have gambling preference for assets with lottery-like payoffs, causing overvaluation and low returns on these assets (Brunnermeier et al., 2007; Mitton and Vorkink, 2007; Barberis and Huang, 2008; Kumar, 2009*b*; Bali et al., 2011; Conrad et al., 2013; Boyer and Vorkink, 2014). Options, especially out-of-the-money (OTM) ones, exhibit dramatic lottery features, exemplified by their high ex-ante return skewness. Boyer and Vorkink (2014) find that OTM options have several times higher ex-ante return skewness compared to stocks. Blau et al. (2016) and Doran et al. (2011) document that calls, especially OTM calls, are actively traded when investors have gambling demand. These studies provide the foundation of using OTM call options to measure investors' gambling activity.

In this paper, we propose a gambling activity measure, $CallMoney$, using open interest and moneyness of OTM individual equity call options. Specifically, we calculate weighted average moneyness with the weight being the associated open interest as a percentage of shares outstanding. This measure is monotonically increasing with both OTM call open interest and moneyness, capturing gambling demand and optimistic expectation among option traders. Our work is closely related to Bergsma et al. (2020). Bergsma et al. (2020) propose a dollar volume-weighted average moneyness measure, $aveMoney$, to capture informed option trading. They show that $aveMoney$ positively predict future stock returns. Inspired by their work, we use OTM option moneyness and its associated open interest to capture gambling activity in options market.

To provide initial evidence that $CallMoney$ is a gambling measure, we conduct four analyses. First, we show that market and industry-wide $CallMoney$ time series successfully captures excessive investor optimism during the dot-com bubble, the oil price bubble, and the pre-GFC stock market bubble. $CallMoney$ time series exhibits conspicuous spikes during the bubbles. Second, we use a variety of univariate and multivariate tests to show that $CallMoney$ is higher for stocks with lottery-like characteristics identified by previous literature. Third, single sort portfolio analysis shows a negative relation between $CallMoney$ and next month's delta-hedged returns for OTM and at-the-money (ATM) call options. The return spread between the highest and lowest $CallMoney$ decile portfolios of OTM (ATM) calls is -1.321% (-0.831%) per month. Fama-French-Carhart four-factor alpha is -1.285% (-0.778%) per month. Fourth, we use a non-parametric approach based on stochastic dominance to compare option returns in high and low $CallMoney$ groups. We show that delta-hedged option returns in low $CallMoney$ group dominate those in high $CallMoney$ group by the second-order stochastic dominance. The result implies that risk-averse investors will prefer option returns in low $CallMoney$ group, which lends support for using $CallMoney$ as a gambling measure.

Previous empirical asset pricing literature proposes several indirect gambling measures based on lottery-like asset payoffs. Assets with lottery-like payoffs tend to attract investors with gambling preference and gambling activity will lead to high contemporaneous asset prices and subsequently earn low returns. Bali et al. (2011) find that stocks with high maximum daily return (MAX) over the past month earn low subsequent returns. Kumar (2009b) considers high idiosyncratic volatility ($IVOL$) and high idiosyncratic skewness ($ISKEW$) as the lottery-like characteristics of stocks, and show that investors have a greater propensity to gamble in stocks with these characteristics. Byun and Kim (2016) use both MAX and $ISKEW$ as the features of lottery-like stocks and find individual call options on lottery-like stocks tend to be overvalued and earn low returns. Boyer and Vorkink (2014) show that ex-ante skewness for options ($OSKEW$) negatively predicts option returns.

Compared to the lottery-like-payoffs based (indirect) gambling measures, our measure *directly* gauges option traders' gambling activities. As investors with excessive optimism overweight the probabilities of extreme positive returns, they will participate aggressively in OTM calls with high moneyness, driving up the value of $CallMoney$. For financial market regulators and policy makers, a direct gambling measure such as $CallMoney$ can be much more useful than indirect ones. Widespread gambling activities on financial instruments like OTM individual equity call options provide instant warning signals to financial market regulators. By contrast, assets with lottery-like-payoffs always exist but may not always lead to excessive investor optimism, asset overpricing, and subsequent low returns. Regulators and policy makers are not likely to use indirect gambling measures to help determine or prevent speculative excess that leads to financial bubbles and crashes.

We perform double-sort tests to compare our new gambling measure $CallMoney$ against the indirect gambling measures (MAX, $IVOL$, $ISKEW$, and $OSKEW$). After controlling one of the indirect gambling measures, the spread and alpha between the highest and lowest $CallMoney$ option quintile portfolios are significantly negative. While after controlling $CallMoney$, the spread and alpha between highest and lowest quintile portfolios sorted based on the indirect gambling measures are small and not statistically significant.

Moreover, Fama-MacBeth (1973) regression analyses show that $CallMoney$ robustly and negatively predicts future OTM and ATM call option returns after controlling indirect gambling measures as well as other stock and option characteristics. By contrast, in OTM call return regressions, the coefficients of $ISKEW$ and $OSKEW$ are insignificant, and the coefficient of $IVOL$ is only significant at 10% significance level. In ATM call return regressions, the coefficients of $ISKEW$ and $OSKEW$ are insignificant. $CallMoney$, as a direct gambling measure, performs more robustly comparing to indirect gambling measures.

We investigate how gambling behaviour, measured by $CallMoney$, interacts with investors' reference-dependent preferences and informational environment. Recent studies relate investors behaviour biases to lottery (skewness) preference. An et al. (2020) find that the lottery-like-payoffs related anomalies are state dependent. Choi et al. (2019) show that disposition effect and skewness preference are correlated. We find that the option return predictability of $CallMoney$ is stronger when the underlying stock's price is further from its 52-week high, or when shareholders experience larger capital losses. This finding is consistent to the explanation that investors are more likely to gamble when the stock price is far from its 52-week high or when they experience capital losses, causing overpricing on call option prices. We also find that the $CallMoney$'s option return predictability is accentuated by low analysts' coverage and high analysts forecast dispersion on the underlying stock. This finding is consistent to the explanation that investors exhibit stronger behavioural bias and are more likely to gamble when there is greater information uncertainty. The findings give more support on using $CallMoney$ as a gambling measure and distinguish it from Bergsma et al. (2020)'s informed option trading measure $aveMoney$.

Investor optimism of the underlying stock, captured by gambling activity in options market and measured by $CallMoney$, should also negatively predicts underlying stock returns. Empirical results from Fama-MacBeth regressions confirm this intuition. $CallMoney$ robustly and negatively predicts future stock returns, after controlling stock characteristics and indirect gambling measures. By contrast, when $CallMoney$ is included in regressions, MAX has no stock return predictive power while the coefficient of $ISKEW$ is only significant at 5% level. Again, $CallMoney$ performs more robustly than indirect gambling measures.

In further analyses, we show that option (stock) return predictability is more pronounced when the market volatility is high, consistent with the argument that gamblers are more active following high market volatility periods (Clark et al., 2018). We also show that option (stock) return predictability is more pronounced when the market sentiment is high, indicating more pronounced overpricing of the assets with high $CallMoney$ when investors are more optimism. Furthermore, we show that option (stock) return predictability is robust to alternative measure constructions using the point estimate or the change of $CallMoney$. $CallMoney$'s option and stock return predictability cannot be explained overpricing caused by option (calls plus puts) demand.

Our study has three important contributions to the literature. First, we put forth $CallMoney$ as a gambling activity measure. $CallMoney$ has the advantages of being economically intuitive, model-free, easy to measure. And its estimation is not necessarily dependent on past time series. Second, we distinguish between direct and indirect gambling measures. $CallMoney$ is a direct gambling measure based on investors' activity, whereas MAX, $IVOL$, $ISKEW$, and $OSKEW$ are indirect measures based on lottery-like payoffs. We show that $CallMoney$ performs more robustly than indirect measures with respect to option and stock return predictability, and more reliably capturing the overpricing of options and stocks. Our work helps understanding the gambling related anomalies in equity option returns and stock returns.[1] Third, $CallMoney$ provides a new way to construct and analyse weighted-average of moneyness comparing to Bergsma et al. (2020)'s $aveMoney$. While $aveMoney$ captures informed trading in options market, $CallMoney$ captures optimism and gambling activity in options market.

The rest of the paper is organized as follows. Section 3.2 defines the measure and Section 3.3 describes the data used in this study. Section 3.4 discusses our empirical

[1] See Bernales et al. (2019) for a good summary of investors' behavioral biases and anomalies in equity option returns.

findings. Section 3.5 presents discussions on robustness. Section 3.6 concludes.

3.2 Gambling activity measure: CallMoney

Extensive literature discusses gambling behaviours in capital market. De Long et al. (1990) suggest that there exist noise traders who are drawn to financial markets to take risks, and are less concerned about fundamental values. Brown et al. (2018) show that sensation-seeking fund managers take on more risk that is unrelated to performance. Clark et al. (2018) demonstrate that pension-plan participants who show gambling-like behaviours are drawn to the market when there is higher market volatility, which is not compensated by commensurate return. This strand of literature indicates that gamblers enter the market for risk-based excitement and behave as noise traders, but they do not generate higher returns. Such behaviour stands in clear contrast to the implications of traditional models of risk and return.

Another strand of literature suggests that gamblers tend to be those who overweight small probability events, and overpay for assets with lottery features (Barberis and Huang, 2008; Kumar, 2009*b*; Doran et al., 2011). Following these studies, we consider lottery preference as the main feature to identify gamblers. OTM call positions are ideal arenas for gamblers with lottery preference to enter, as they offer substantially high ex-ante return skewness (Boyer and Vorkink, 2014). Several studies propose that deep OTM single stock calls resemble overpriced lottery-like securities (Barberis and Huang, 2008; Doran et al., 2011; Boyer and Vorkink, 2014; Félix et al., 2019). This provides economic intuition of using OTM calls for the construction of our gambling activity measure.

We propose a new gambling measure by jointly considering the information of open interest and moneyness, which is motivated by Lakonishok et al. (2006) and Boyer and Vorkink (2014). Lakonishok et al. (2006) find that open interest as a percentage of shares outstanding for OTM calls increased substantially during the dot-com bubble, and they attribute this increase to the speculation demand from investors in options market during the bubble period. Boyer and Vorkink (2014) show that call options with larger moneyness provide substantially higher skewness, implying that call options with higher moneyness will be more attractive to gamblers with lottery preference. By jointly considering open interest and moneyness of OTM individual equity call options, we propose a new activity-based gambling measure.

Our new gambling activity measure $CallMoney$, is defined at time t as:

$$CallMoney_t \equiv \sum_{j=1}^{N} \frac{100 \times OI_{j,t}}{SO_t} M_{j,t} \times 100,$$

where $OI_{j,t}$ is the open interest associated to the jth moneyness level, SO_t is the number of total outstanding shares; moneyness level $M_{j,t} = \frac{K_{j,t}}{S_t}$ with S_t being the current stock price and $K_{j,t}$ being jth strike price; N is the number of different strike prices of OTM calls. The factor of 100 in the numerator converts the open interest into an equivalent number of shares of the underlying stock (100 shares per option contract). The final factor of 100 converts the quantity into a percentage.

Comparing to Bergsma et al. (2020)'s $aveMoney$ measure, our construction has two major differences. First, we only use OTM call option moneyness. By doing that, our measure captures more on lottery features of options and traders' gambling preference on such features. Second, while *aveMoney* uses dollar volume as weight, we use open interest versus shares outstanding as weight. Using trading volume as weight reflects more on traders' information, and using normalised open interest as weight reflects more on (OTM call) option demands. Combining option demands with lottery features of OTM calls, *CallMoney* captures gambling demands that mainly reflects traders' optimism of the underlying stock.

OTM call options are those with moneyness higher than 1.05. Our empirical results are robust when we choose 1.025 as an alternative moneyness threshold. Since short-term options provide the highest return skewness (Boyer and Vorkink, 2014), and therefore being more attractive to gamblers, we use options with short-term maturity (between 7 to 40 days) to construct this gambling measure. We average daily $CallMoney_t$ in a month to create a monthly *CallMoney* measure. One thing that the readers should be cautious when interpreting our results is that since we use the OTM call options' information to measure the gambling activity, our measure can only capture gambling activity in stocks with available OTM call option data.

The open interest on call options comprises both long and short positions. Lakonishok et al. (2006) find that written calls are comprised mainly of covered calls so both long and short positions in calls represent positive exposure on the underlying stock price. Furthermore, Andreou et al. (2018) show in a theoretical framework that option trader's optimal choice of moneyness monotonically increases with optimism level, irrespective of the selection of long or short option positions. In our work, we draw upon the results of these two studies and do not distinguish between long and short positions in OTM call's open interest.

To provide intuitive evidence of $CallMoney$ being a gambling activity measure, we examine the relation between $CallMoney$'s historical time series patterns and financial bubbles. Figure 3.1 illustrates the yearly time series of market and industry-wide median $CallMoney$ from 1996 to 2017.[2].

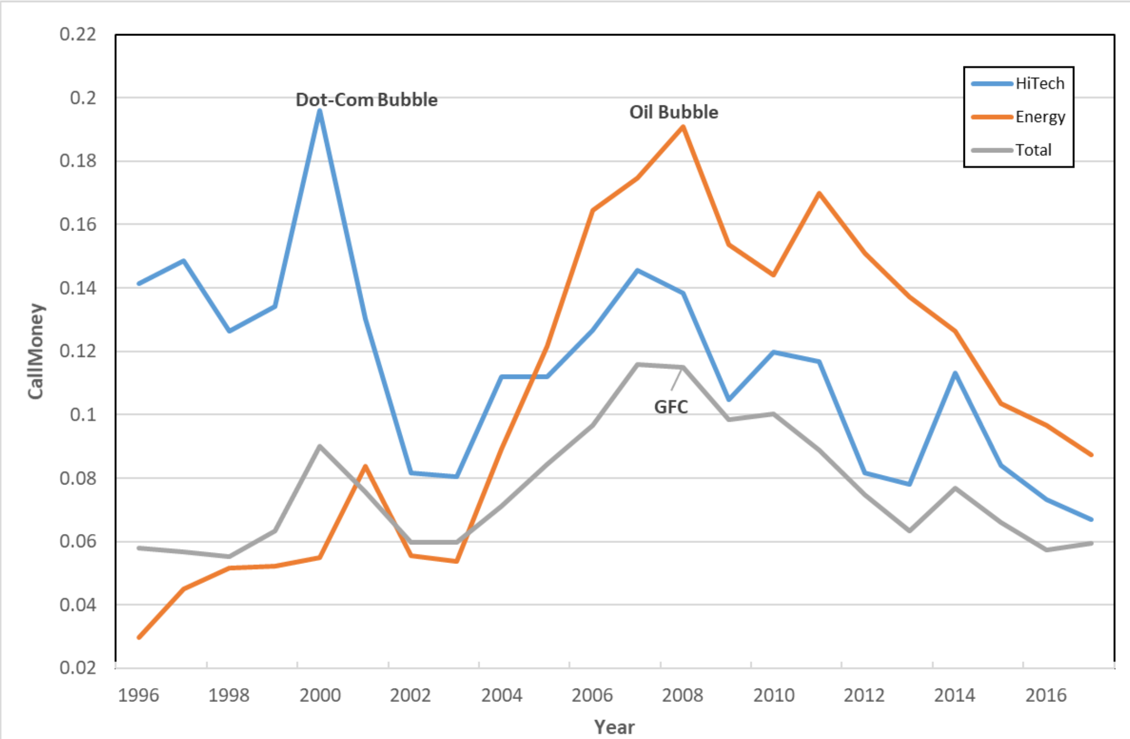

Figure 3.1: $CallMoney$ time Series: 1996-2017

The figure illustrates the time series of yearly median $CallMoney$ measures for HiTech industry, energy industry and the whole sample from 1996 to 2017. A Stock is classified into HiTech industry if its Compustat SIC code is within 7370-7391 or 8730-8734, and into energy industry if its Compustat SIC code is within 1200-1399 or 2900-2999.

The market-wide $CallMoney$ time series have two spikes, one in 2000-2001 (the dot-com bubble) and the other in 2007-2008 (the pre-GFC stock market bubble). HiTech industry $CallMoney$ time series also have two spikes, but the spike in the dot-com bubble is more pronounced than the pre-GFC one. Energy industry $CallMoney$ time series have an opposite pattern, with a more pronounced spike in 2008 comparing

[2]Quarterly and monthly $CallMoney$ time series also exhibit similar (but noisier) patterns

to the dot-com one. In 2007-2008, oil price bubble and pre-GFC stock market bubble together drive the excessive speculation in engergy sector which is successfully captured by industry-wide $CallMoney$.

Market and industry-wide $CallMoney$ time series show that our new measure captures investors' excessive optimism and gambling activity during the bubble periods. Comparing to the valuation-based bubble measures (such as Shiller PE ratio[3]), $CallMoney$ has the potential to serve as a bubble warning measure from the angle of options market trading activity.

3.3 Data

We collect data from several sources. We obtain stock trading data from the Center for Research in Security Prices (CRSP). The accounting data are from Compustat. The Fama-French-Carhart common risk factors and the risk-free rate are from Kenneth French's website. The analyst coverage and forecast data are from I/B/E/S. The institutional ownership data are from Thomson Reuters (13F) database.

Options data are obtained from OptionMetrics from January 1996 to December 2017. The data include the daily closing and ask quotes, trading volume, strike price, open interest and implied volatility. Closing option prices are calculated as the midpoint of the closing bid and ask prices. We apply a series of data filters to minimize the impact of recording errors in the options market following previous literature. First we eliminate all observations for which the ask price is lower than

[3]The adjusted stock market price-earnings ratio defined by Shiller (2015). Shiller PE ratio peaks around 1929, 2000 and 2007, in all major market bubbles.

the bid price. Second we eliminate options with zero or missing bid price. Third we eliminate options with zero open interest. Fourth we eliminate options with nonstandard settlement (settlement flag to be nonzero). Fifth we eliminate options with missing or zero implied volatility. Sixth we eliminate options with delta to be below -1 or above $+1$. Our final sample contains 235,224 option-month observations for OTM calls and 236,520 for ATM calls. We present summary statistics of gambling measures and their correlations in Table 3.1. These measures are defined in details in Appendix A.

Table 3.1: Summary statistics and correlations of gambling measures

The table reports the summary statistics and correlations for the gambling measures: $CallMoney$, maximum daily return (MAX), idiosyncratic volatility ($IVOL$), idiosyncratic skewness ($ISKEW$), and option ex-ante skewness ($OSKEW$). At the end of each month, lower quartile (P25), median, upper quartile (P75) and standard deviation (SD) of each variable and correlation matrix among variables are calculated cross-sectionally, resulting in a time-series of summary statistics. The time-series summary statistics are then averaged and reported in Panel A. Panel B reports the time-series average of cross-sectional correlations. The sample period is from January 1996 to December 2017.

Panel A: Summary statistics of gambling measures				
	P25	Median	P75	SD
CallMoney	0.022	0.075	0.242	2.321
MAX	0.032	0.047	0.070	0.046
IVOL	0.016	0.022	0.031	0.014
ISKEW	-0.243	0.204	0.695	1.153
OSKEW	1.316	1.667	2.108	6.699
Panel B: Correlation matrix				
	MAX	IVOL	ISKEW	OSKEW
CallMoney	0.132	0.189	0.022	0.080
MAX		0.403	0.037	0.180
IVOL			0.074	0.229
ISKEW				-0.102

Table 3.1 shows that $CallMoney$ has very large standard deviation (2.321) and skews to the right as expected. Summary statistics for indirect gambling measures are generally in line with previous literature. The correlations between $CallMoney$ and indirect gambling measures are positive and in general small, indicating that our

new measure captures different set of information.

In this paper, we consider both option returns and stock returns. For option returns, we focus on delta-hedged call option returns as they are generally immune from the underlying stock price movement. In line with Cao et al. (2017) and Ruan (2020*b*), we define the delta-hedged option gain as the change in value of a self-financing portfolio consisting of a long call position, hedged by a short position in delta units of underlying stock, with the proceeds earning the risk-free return.

Specifically, if a self-financing portfolio is rebalanced N times over a period $[t, t+\tau]$, its delta-hedged gain is:

$$\Pi(t, t+\tau) = C_{t+\tau} - C_t - \sum_{n=0}^{N-1} \Delta_{C,t_n}[S(t_{n+1}) - S(t_n)] - \sum_{n=0}^{N-1} \frac{\alpha_n r_{t_n}}{365}[C(t_n) - \Delta_{C,t_n} S(t_n)],$$

where Δ_{C,t_n} is the delta of the option C at date t_n; r_{t_n} is the annualised risk-free rate at date t_n; and α_n is the number of calendar days between t_n and t_{n+1}. Then scale the dollar return $\Pi(t, t+\tau)$ by $\Delta_t S_t - C_t$ to obtain delta-hedged call option return as follows:

$$R(t, t+\tau) = \frac{\Pi(t, t+\tau)}{\Delta_t S_t - C_t}.$$

In subsequent empirical tests, all option and option portfolio returns are delta-hedged returns. We use one OTM call (moneyness lower than 1.20 and higher than 1.05, and closest to 1.10) and one ATM call (moneyness lower than 1.05 and higher than 0.95, and closest to 1) for each stock when we examine option returns. The moneyness range selection of OTM call and ATM call options follows the option literature (Xing et al., 2010; Doran et al., 2011; Boyer and Vorkink, 2014).

3.4 Empirical analysis

3.4.1 Direct and indirect gambling measures

In Section 3.2, we show that market-wide $CallMoney$ captures investors' excessive optimism and gambling activity during the bubble periods. In this section, we examine the cross-sectional relation between gambling activity measure $CallMoney$ and stocks' lottery features. We conduct both univariate and multivariate tests to examine whether $CallMoney$ is higher for lottery-like stocks. We adopt lottery-like payoff characteristics (i.e. indirect gambling measures) as identified by previous literature: MAX, $IVOL$ and $ISKEW$.[4]

We first sort our stock-month observations into quintiles based on each of the indirect gambling measures observed at the previous month-end. We then report time-series average $CallMoney$ across quintiles. Panel A of Table 3.2 reports the results of the analysis.

[4] We do not investigate $OSKEW$ in this section as it is an option-specific variable.

Table 3.2: CallMoney and lottery characteristics

This table reports $CallMoney$ across each of the three lottery characteristics (Panel A) and the results of Fama-MacBeth regressions of month $t+1$ $CallMoney$ on month t lottery characteristics and control variables (Panel B). The lottery characteristics include maximum daily return (MAX), idiosyncratic volatility ($IVOL$), and idiosyncratic skewness ($ISKEW$). The control variables include stock return in the past month (REV), cumulative stock return from past 11 months (MOM), log market capitalization (log($SIZE$)), beta ($BETA$), illiquidity ($ILLIQUIDITY$), and book-to-market ratio (BTM). In Panel A, stock-month observations are first sorted by each of the lottery characteristics during the prior month. Then time-series average $CallMoney$ across quintiles, and the difference between extreme quintiles and the associated t-statistics are reported. Panel B presents the time series averages of the monthly cross-sectional regression coefficients and the associated t-statistics. Also reported are the average adjusted R-square. The sample period is from January 1996 to December 2017.

Panel A: CallMoney across lottery characteristics			
	MAX_{t-1}	$IVOL_{t-1}$	$ISKEW_{t-1}$
Q1	0.125	0.090	0.244
Q2	0.165	0.146	0.266
Q3	0.217	0.208	0.238
Q4	0.329	0.294	0.259
Q5	0.457	0.555	0.286
Q5-Q1	0.332***	0.465***	0.043**
	(15.49)	(11.64)	(2.28)
Panel B: Fama-MacBeth regressions of CallMoney			
	Model 1	Model 2	Model 3
MAX	1.715***		
	(6.16)		
IVOL		16.218***	
		(2.82)	
ISKEW			0.007**
			(2.33)
REV	-0.294***	-0.075	-0.039
	(-3.04)	(-0.84)	(-0.50)
MOM	-0.044	-0.121	-0.055
	(-0.73)	(-1.34)	(-0.88)
log(SIZE)	-0.090***	-0.043***	-0.104***
	(-2.85)	(-2.67)	(-3.47)
BETA	0.087***	0.066***	0.099***
	(6.65)	(6.70)	(7.79)
ILLIQUIDITY	-0.312***	-0.336***	-0.310***
	(-3.55)	(-3.42)	(-3.44)
BTM	-0.024*	-0.021	-0.028*
	(-1.68)	(-1.12)	(-1.82)
INTERCEPT	YES	YES	YES
Adjusted R^2 (%)	5.99%	6.82%	4.33%

Panel A shows that $CallMoney$ increases across MAX, $IVOL$, and $ISKEW$ quintiles in the previous month. The difference between extreme quintiles (Q5-Q1) are all statistically significant. These results indicate that $CallMoney$ depends on the lottery features of stocks observed during the previous month.

We then conduct multivariate tests to determine whether $CallMoney$ is higher for each of these indirect gambling measures controlling for other firm-specific characteristics. Following Kumar (2009*b*), we conduct Fama-MacBeth (1973) regressions, where the dependent variable is $CallMoney$, and the key independent variables are the indirect gambling measures (MAX, $IVOL$, and $ISKEW$) observed at the previous month-end. The control variables include stock return in the past month (REV), cumulative stock return from past 11 months (MOM), log market capitalization (log($SIZE$)), beta ($BETA$), illiquidity ($ILLIQUIDITY$), and book-to-market ratio (BTM). These variables are defined in details in Appendix A. We report the time-series averages of the coefficients of the monthly cross-sectional regressions, along with their t-statistics in Panel B of Table 3.2.

In Model 1, we estimate the multivariate regressions with the key independent variable being MAX. The coefficient on MAX is 1.715 with a t-statistic of 6.16, confirming that MAX can positively predict future $CallMoney$. We obtain qualitatively similar results in Models 2 and 3 with the key independent variable being $IVOL$ and $ISKEW$ respectively. We show a robust positive relation between $CallMoney$ and the lottery-like characteristics. This result indicates that investors' preference for lottery stocks is reflected in higher $CallMoney$ value. We also look into the coefficients on the other control variables. The coefficients on $log(SIZE)$ and $ILLIQUIDITY$ are consistently negative, indicating that $CallMoney$ is higher for smaller-size firms and stocks with higher liquidity. The coefficient on $BETA$ is significantly positive, indicating that $CallMoney$ is higher for firms with higher systematic risk.

3.4.2 Single-sort portfolio analysis and stochastic dominance

We provide some intuitive evidence that $CallMoney$ negatively predicts option returns and high $CallMoney$ value captures option traders' gambling activity. We sort OTM/ATM call options into 10 decile portfolios based on their previous month's $CallMoney$ levels, where decile 10 (decile 1) option portfolio is the one with the highest (lowest) $CallMoney$. We equally weight delta-hedged monthly option returns to compute decile portfolio's returns.[5] Table 3.3 reports the average monthly returns and their corresponding Fama-French-Carhart (1997) four-factor alphas of the decile portfolios; together with the spread between the highest and lowest $CallMoney$ decile portfolio returns ("10-1") and its associated four-factor alpha.

Single-sort portfolio results in Table 3.3 show that decile option portfolio returns monotonically decrease when $CallMoney$ increases. For OTM call options, the "10-1" return spread is -1.321% per month and its four-factor alpha is -1.285% per month. While for ATM options, the "10-1" return spread and its alpha are -0.831% and -0.778% per month. The spreads and alphas are all significant at 1% level. The larger return spread for OTM calls may be attributed to the higher return skewness of OTM options, which endear them to gamblers more than ATM options.

We further employ a non-parametric test based on stochastic dominance to examine whether the negative option return predictability by $CallMoney$ is driven by risk factors that are not captured by Fama-French-Carhart four factors. The stochastic dominance approach provides an utility-based framework to compare choices under uncertainty. Previous finance literature applies stochastic dominance approach to

[5]We use the equal-weighted portfolio return because all the option returns are delta-hedged returns.

Table 3.3: Portfolios single sorted on $CallMoney$

In each month, all OTM call options are single sorted into equally weighted quintile portfolios based on $CallMoney$ (1 = Low $CallMoney$, 5 = High $CallMoney$). Panel A reports the average monthly delta-hedged returns and four-factor alphas for portfolios thus obtained. The single sort results for ATM calls are reported in Panel B. The sample period is from January 1996 to December 2017. The t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level. All returns are expressed as percentages.

Panel A: OTM calls												
	1	2	3	4	5	6	7	8	9	10	10-1	
Average Return	0.448	0.323	0.221	0.232	0.116	0.096	-0.160	-0.166	-0.409	-0.873	-1.321***	(-9.96)
FF4 Alpha	0.455	0.347	0.242	0.249	0.138	0.124	-0.144	-0.136	-0.372	-0.830	-1.285***	(-11.09)
Panel B: ATM calls												
	1	2	3	4	5	6	7	8	9	10	10-1	
Average Return	0.232	0.213	0.112	0.089	0.055	0.001	-0.080	-0.185	-0.268	-0.599	-0.831***	(-8.67)
FF4 Alpha	0.150	0.141	0.048	0.019	-0.022	-0.062	-0.139	-0.233	-0.309	-0.628	-0.778***	(-9.33)

examine return puzzles and anomalies (e.g. Monday effect, January effect, and momentum) to rule out the possibility of omitted risk factors (Seyhun, 1993; Fong et al., 2005; Cho et al., 2007; Fong, 2010).

Following these studies, we conduct a stochastic dominance analysis on option returns. We divide OTM call options into low $CallMoney$ and high $CallMoney$ groups, based on their $CallMoney$ values relative to the median $CallMoney$ level. Figure 3.2 illustrates the empirical cumulative distribution functions (CDFs) of low and high $CallMoney$ groups' equal-weighted monthly returns.

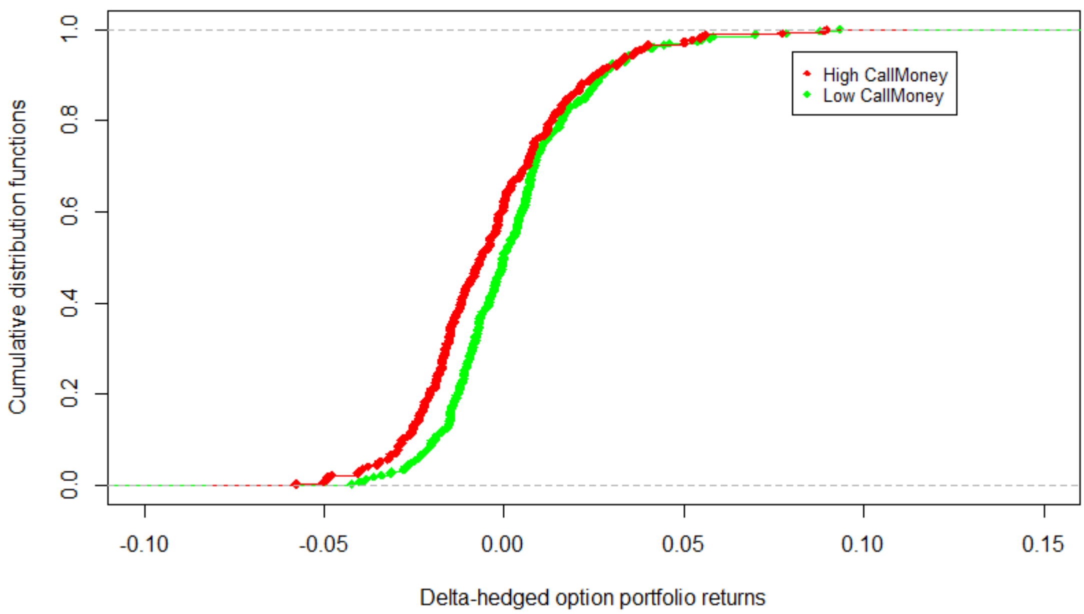

Figure 3.2: Stochastic dominance: cumulative distribution functions of high and low $CallMoney$ groups' returns

OTM call options are divided into two (High and Low $CallMoney$) groups based on $CallMoney$ values relative to the median $CallMoney$ level. The figure illustrates the cumulative distribution functions of the equal-weighted monthly delta-hedged returns of High and Low $CallMoney$ option groups. The sample period is from January 1996 to December 2017.

Using the test proposed by Davidson and Duclous (2000), we find that low $CallMoney$

group's returns dominate those of high $CallMoney$ group by the second order stochastic dominance, suggesting all risk-averse investors will prefer low $CallMoney$ group to high $CallMoney$ group.[6] The stochastic dominance analysis shows that $CallMoney$'s option return predictability cannot be explained by omitted risk factors. The trading activity in high $CallMoney$ option group is more likely to be initiated by gamblers, who enter the market for risk-based excitement.

3.4.3 CallMoney versus lottery-like-payoffs based measures: double-sort analyses

$CallMoney$ provides a direct gauge on investors' gambling activity. By contrast, the indirect gambling measures use lottery-like-payoffs as a proxy of optimism-induced gambling. We compare the performance on option return predictability of our gambling measure $CallMoney$ to the existing lottery-like-payoffs based (indirect) gambling measures (MAX, $IVOL$, $ISKEW$, and $OSKEW$) by conducting double-sort analyses. Table 3.4 reports the average equal-weighted monthly OTM call portfolio returns based on double sorts of direct-indirect gambling measures.[7]

[6] There is no first-order stochastic dominance because the two CDFs cross. See Appendix B for details on Davidson and Duclous (2000) test.

[7] We obtain qualitatively similar results for ATM call option returns. The results are available from the authors upon request.

Table 3.4: Portfolio double sorted on $CallMoney$ and indirect gambling measure

This table reports the OTM call option returns double sorted by $CallMoney$ and the indirect gambling measures. In each month, all OTM call options are double sorted in two ways. In one double sort, options are first sorted into quintiles based on one of the four indirect gambling measures: maximum daily return (MAX), idiosyncratic volatility ($IVOL$), idiosyncratic skewness ($ISKEW$), and option ex-ante skewness ($OSKEW$). Then within each quintile, options are further sorted into five equally weighted portfolios by their $CallMoney$ (1 = Low $CallMoney$, 5 = High $CallMoney$). Panel A reports the average monthly delta-hedged returns for portfolios thus obtained. The second double sort reverses the sort order, and the corresponding results are reported in Panel B. The sample period is from January 1996 to December 2017. The t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level. All returns are expressed as percentages.

Panel A: Sorted by CallMoney controlling other measures				
	MAX	IVOL	ISKEW	OSKEW
1 (Low CallMoney)	0.279	0.367	0.263	0.243
2	0.133	0.170	0.153	0.107
3	-0.030	0.004	-0.007	-0.002
4	-0.221	-0.140	-0.191	-0.175
5 (High CallMoney)	-0.702	-0.533	-0.746	-0.520
5-1	-0.982***	-0.900***	-1.009***	-0.763***
	(-10.63)	(-8.13)	(-9.47)	(-7.68)
5-1 FF4 Alpha	-0.973***	-0.895***	-0.982***	-0.726***
	(-11.01)	(-8.48)	(-10.22)	(-8.16)
Panel B: Sorted by other measures controlling CallMoney				
	MAX	IVOL	ISKEW	OSKEW
1 (Low)	-0.146	-0.166	-0.161	0.056
2	-0.097	-0.088	-0.026	-0.171
3	-0.067	0.103	-0.066	-0.121
4	-0.044	0.100	-0.103	-0.144
5 (High)	-0.183	-0.093	-0.172	0.036
5-1	-0.037	0.073	-0.011	-0.020
	(-0.28)	(0.62)	(-0.17)	(-0.14)
5-1 FF4 Alpha	0.053	0.132	-0.012	-0.001
	(0.47)	(1.15)	(-0.20)	(-0.01)

In Table 3.4 Panel A, we sort OTM calls on $CallMoney$ after controlling for each of the four indirect gambling measures. For example, we control for MAX by first forming quintile portfolios ranked based on MAX of the underlying stock. Then, within each MAX-ranked portfolio, we sort options into quintile portfolios ranked based on $CallMoney$ so that quintile 1 (quintile 5) contains options with the lowest (highest) $CallMoney$. The first column of Table 3.4 Panel A presents monthly option returns for each $CallMoney$ quintile averaged across the MAX quintiles.

After controlling for MAX, the "5-1" portfolio spread is -0.982% per month with a t-statistic of -10.63, the associated four-factor alpha is -0.973% per month with a t-statistic of -11.01. Using the same approach, the second to fourth columns in Panel A show that after controlling for $IVOL$, $ISKEW$, and $OSKEW$, the return spreads and the corresponding alphas between the highest and lowest $CallMoney$ quintiles are all negative and significant at 1% level. Results in Panel A imply that none of the indirect gambling measures can explain the high (low) returns of low (high) $CallMoney$ call options.

In Table 3.4 Panel B, we perform double sorts on the reversed order, sorting options on each of the four gambling measures after controlling for $CallMoney$. The results show that the portfolio return explanation power of MAX, $IVOL$, $ISKEW$ and $OSKEW$, after controlling for $CallMoney$, are insignificant; as portfolio spreads and the corresponding four-factor alphas are statistically insignificant. These results indicate that $CallMoney$ is better than the indirect gambling measures at predicting future option returns.

3.4.4 Fama-MacBeth regressions

In this section, we use Fama-MacBeth (1973) regressions to formally test the negative relation between $CallMoney$ and cross-sectional future option returns by controlling indirect gambling measures and other variables.

The dependent variable is delta-hedged OTM/ATM call returns. The control variables include REV, MOM, log($SIZE$), $BETA$, BTM, stock turnover ($TURNOVER$), $ILLIQUIDITY$, realized-implied volatility spread (VRP), volatility of volatility

(VOV), and option bid-ask spread (PBA). Extant literature shows that option implied volatility contains private information of option traders (Xing et al., 2010; Cremers and Weinbaum, 2010; Jin et al., 2012; An et al., 2014). Following An et al. (2014), we also control for change in ATM call implied volatilities (CIV) and change in ATM put implied volatilities (PIV) to rule out the possibility that the result is driven by informed trading in options market. The variables are defined in details in Appendix A.

Fama-MacBeth regressions are extensively used in the literature to examine the cross-sectional return predictability of the interested predictors (see e.g. Xing et al. (2010), Cao and Han (2013), Byun and Kim (2016), and Atilgan et al. (2020*b*)). However, there are several limitations to the regression methodology. First, it imposes linearity on the relation between the dependent variable and the independent variable. To solve this, we have conducted portfolio sorting analysis to examine the relation between $CallMoney$ and call option returns in Sections 3.4.2 and 3.4.3. Furthermore, the second stage of Fama-MacBeth regressions is to conduct inference on the time-series of the coefficients by assuming the coefficients over time are independent and identically distributed. However, there may exist the autocorrelation structure of the time series of coefficients. To adjust for the time-series correlation of the coefficients, we use the Newey-West (1987) adjustments to calculate the t-statistics of the coefficients.

We report the time-series averages of the coefficients, along with their Newey-West adjusted t-statistics in Table 3.5 (for OTM calls) and in Table 3.6 (for ATM calls). Both tables include 7 models.

Table 3.5: Fama-MacBeth regression analyses: OTM call return predictability

The table presents the results of Fama and MacBeth (1973) regressions of month $t+1$ delta-hedged OTM call option returns on month t $CallMoney$ and control variables. The table presents the time series averages of the monthly cross-sectional regression coefficients. Also reported are the average adjusted R-square. The sample period is from January 1996 to December 2017. Newey and West (1987) t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
CallMoney	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***
	(-7.08)	(-5.95)	(-5.99)	(-5.53)	(-5.87)	(-5.82)	(-5.58)
MAX			-0.032***				-0.029***
			(-2.67)				(-2.60)
IVOL				-0.076*			-0.050
				(-1.75)			(-1.15)
ISKEW					0.000		0.000
					(-0.84)		(-0.67)
OSKEW						0.000	0.000
						(-0.95)	(-1.43)
REV		0.004	0.009*	0.004	0.004	0.004	0.009*
		(0.99)	(1.88)	(1.03)	(0.95)	(1.11)	(1.78)
MOM		-0.003**	-0.002**	-0.002	-0.002**	-0.002**	-0.002
		(-2.17)	(-2.03)	(-1.59)	(-2.04)	(-2.03)	(-1.48)
log(SIZE)		0.001	0.000	0.000	0.001	0.000	0.000
		(1.62)	(0.95)	(0.76)	(1.58)	(1.40)	(0.34)
BETA		0.001*	0.001**	0.001**	0.001*	0.001*	0.001**
		(1.94)	(2.10)	(2.09)	(1.93)	(1.88)	(2.20)
BTM		-0.000	-0.000	0.000	-0.000	0.000	0.001
		(-0.09)	(-0.10)	(0.14)	(-0.11)	(0.17)	(0.37)
TURNOVER		0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
		(2.79)	(3.37)	(2.65)	(2.72)	(2.77)	(3.04)
ILLIQUIDITY		-16.813***	-15.597**	-15.334**	-16.876***	-16.886***	-14.813**
		(-2.62)	(-2.41)	(-2.41)	(-2.61)	(-2.59)	(-2.26)
CIV		-0.051***	-0.048***	-0.050***	-0.051***	-0.052***	-0.047***
		(-4.93)	(-4.47)	(-4.87)	(-4.92)	(-4.96)	(-4.43)
PIV		0.006	0.003	0.004	0.006	0.006	0.001
		(0.57)	(0.31)	(0.41)	(0.54)	(0.54)	(0.12)
VRP		-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
		(-3.90)	(-4.36)	(-3.07)	(-3.86)	(-3.82)	(-3.48)
VOV		-0.039***	-0.037***	-0.037***	-0.039***	-0.038***	-0.036***
		(-8.05)	(-7.83)	(-7.97)	(-7.92)	(-7.98)	(-7.63)
PBA		0.017***	0.017***	0.016***	0.017***	0.018***	0.017***
		(5.29)	(5.15)	(4.96)	(5.28)	(5.18)	(4.81)
INTERCEPT	YES	YES	YES	YES	YES	YES	YES
Adjusted R^2 (%)	0.62	5.22	5.54	5.57	5.26	5.38	6.01

Table 3.6: Fama-MacBeth regression analyses: ATM call return predictability

The table presents the results of Fama and MacBeth (1973) regressions of month $t+1$ delta-hedged ATM call option returns on month t $CallMoney$ and control variables. The table presents the time series averages of the monthly cross-sectional regression coefficients. Also reported are the average adjusted R-square. The sample period is from January 1996 to December 2017. Newey and West (1987) t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
CallMoney	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
	(-5.18)	(-7.11)	(-7.08)	(-6.62)	(-7.08)	(-7.06)	(-6.58)
MAX			-0.039***				-0.033***
			(-6.63)				(-6.02)
IVOL				-0.141***			-0.116***
				(-5.81)			(-4.75)
ISKEW					-0.000		-0.000
					(-1.25)		(-0.52)
OSKEW						-0.000	-0.000
						(-1.52)	(-0.20)
REV		-0.002	0.004	-0.002	-0.002	-0.002	0.004
		(-1.12)	(1.46)	(-0.89)	(-1.13)	(-1.21)	(1.60)
MOM		-0.000	-0.000	0.000	-0.000	-0.000	0.000
		(-0.53)	(-0.63)	(0.53)	(-0.46)	(-0.62)	(0.69)
log(SIZE)		0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
		(5.91)	(4.51)	(3.34)	(5.89)	(5.91)	(2.74)
BETA		-0.000	0.000	0.000	-0.000	-0.000	0.000
		(-0.35)	(0.11)	(0.51)	(-0.36)	(-0.35)	(0.71)
BTM		0.001	0.001	0.001	0.001	0.001	0.001
		(1.04)	(1.02)	(1.03)	(1.02)	(1.08)	(1.18)
TURNOVER		0.000*	0.000***	0.000**	0.000*	0.000**	0.000***
		(1.90)	(3.50)	(2.57)	(1.89)	(2.10)	(3.84)
ILLIQUIDITY		-7.715***	-6.070***	-6.221***	-7.655***	-7.791***	-5.107**
		(-3.38)	(-2.63)	(-2.66)	(-3.38)	(-3.42)	(-2.18)
CIV		-0.044***	-0.043***	-0.045***	-0.044***	-0.044***	-0.043***
		(-8.50)	(-8.12)	(-8.61)	(-8.44)	(-8.46)	(-8.09)
PIV		0.011**	0.011*	0.010*	0.011**	0.011*	0.009
		(2.02)	(1.92)	(1.82)	(2.01)	(1.93)	(1.64)
VRP		-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
		(-4.78)	(-5.69)	(-4.63)	(-4.76)	(-4.86)	(-5.57)
VOV		-0.014***	-0.012***	-0.013***	-0.014***	-0.014***	-0.012***
		(-7.31)	(-6.59)	(-7.07)	(-7.29)	(-7.46)	(-6.49)
PBA		0.014***	0.014***	0.013***	0.014***	0.016***	0.015***
		(4.65)	(4.68)	(4.34)	(4.66)	(4.79)	(4.39)
INTERCEPT	YES	YES	YES	YES	YES	YES	YES
Adjusted R^2 (%)	1.00	6.09	6.53	6.65	6.11	6.35	7.22

In Model 1, we estimate univariate regressions of option returns on $CallMoney$. The coefficient on $CallMoney$ is -0.005 (-0.004) and with a t-statistic of -7.08 (-5.18) for OTM (ATM) call regressions, confirming that $CallMoney$ can negatively predict future option returns.

In Model 2, we estimate the multivariate regressions by adding control variables. The coefficients on $CallMoney$ remain negative and significant at 1% level. The $CallMoney$'s option return predictability is robust after controlling for option and stock characteristics. The coefficients on $ILLIQUIDITY$, VRP, and VOV are negative and significant, in line with findings of Goyal and Saretto (2009) and Ruan (2020*b*). While log($SIZE$), $BETA$, and PBA have positive and significant coefficients, consistent with the studies of Ruan (2020*b*), Cao and Han (2013) and Choy (2015). The coefficient on CIV is significantly negative, while the coefficient on PIV is significantly positive in Table 3.6. This is consistent with the argument that informed option traders' demand will influence option implied volatility level and option prices (Garleanu et al., 2008).

From Model 3 to Model 7, we further control for each and then all of the indirect gambling measures (MAX, $IVOL$, $ISKEW$, and $OSKEW$) . The negative and significant option return predictability of $CallMoney$ holds consistently across models. This confirms that $CallMoney$'s option return predictability cannot be explained by any of the indirect gambling measures. By contrast, the coefficient on $IVOL$ in Table 3.5 and the coefficients on $ISKEW$ and $OSKEW$ are insignificant.

Tables 3.5 and 3.6 show that $CallMoney$ is better at capturing the gambling activity in the options market and its option return predicability is robust after controlling indirect gambling measures and other variables.

3.4.5 Anchoring and gambling behaviour: 52-week high and capital gain overhang

Investors' trading behaviour tends to anchor on past return performance. We analyse how investors' anchoring behaviour interacts with gambling activity in options market. If investors' behavioural biases such as anchoring induce more risk-taking activity, then these biases should increase the return predictability of our gambling measure.

The 52-week high stock price is one of the most readily available information of past stock price performance. Acting as a resistance level in technical analysis, 52-week high represents a psychological price barrier to investors. George and Hwang (2004) find that 52-week high price explains a large portion of the momentum investing profits. Driessen et al. (2011) find that option implied volatilities decrease when the stock prices approach 52-week highs, suggesting option treaders' optimism is negatively associated to the nearness to the 52-week high price. Following previous studies, we define nearness to 52-week high (denoted as NH) for each stock at the end of month t as:

$$NH_t = \frac{P_t}{52HIGH_t},$$

where P_t is the price of stock at the end of month t, and $52HIGH_t$ is the highest daily closing price from the beginning of month $t-11$ to the end of month t.

Another behavioural anchoring concept is capital gain overhang. Grinblatt and Han (2005) find that investors tend to be risk-seeking (risk-averse) as they experience capital losses (gains). An et al. (2020) find that the lottery-related anomalies are more pronounced when stocks experience capital losses. They argue that in order to break even, investors are more likely to gamble after capital loss. Following Grinblatt and Han (2005), we calculate capital gain overhang (denoted as CGO) as the normalized difference between the current stock price and the reference price. The reference price, a proxy for stock purchase cost, is defined as:

$$RP_t = \frac{1}{k}\sum_{n=1}^{T}\Big(V_{t-n}\prod_{\tau=1}^{n-1}(1-V_{t-n+\tau})\Big)P_{t-n},$$

where P_t is the stock price at the end of week t, V_t is week t's stock turnover, $T = 260$, the number of weeks in the previous 5 years, and $k = \sum_{n=1}^{T}\Big(V_{t-n}\prod_{\tau=1}^{n-1}(1-V_{t-n+\tau})\Big)$ is a constant that normalise the weights on past prices sum to one. The CGO at week t is defined as:

$$CGO_t = \frac{P_{t-1} - RP_t}{P_{t-1}}.$$

Following Wang et al. (2017), we use the last-week's CGO in each month as monthly CGO. We conduct independent double-sort analysis on $CallMoney$ and NH/CGO by sorting call options into five groups based on $CallMoney$, and then sorting independently into five groups based on their underlying stocks' NH or CGO. We next construct 25 portfolios defined by the intersections of this 5×5 sort, and we eqaul-weight the delta-hedged option returns when computing option portfolio returns. Table 3.7 presents the independent double-sort results.

Panel A of Table 3.7 reports the independent double-sort results of sorting by $CallMoney$ and NH. For OTM calls, when stock price is far from its 52-week high ("Low NH"), return spread between high and low $CallMoney$ quintile portfolios ("5-1") is -1.103% per month, with the corresponding four-factor alpha being -1.122%

Table 3.7: Portfolios double sorted independently on $CallMoney$ and anchoring measure

In each month, all OTM/ATM call options are double sorted in two ways. In one double sort, options are independently sorted into 5 × 5 equally weighted portfolios based on $CallMoney$ (1 = Low $CallMoney$, 5 = High $CallMoney$) and nearness to 52-week high measure NH (1 = Low NH, 5 = High NH). Panel A reports the average monthly delta-hedged returns for portfolios thus obtained. In the second double sort, options are independently sorted into 5 × 5 equally weighted portfolios based on $CallMoney$ (1 = Low $CallMoney$, 5 = High $CallMoney$) and capital gain overhang measure CGO (1 = Low CGO, 5 = High CGO). Panel B reports the average monthly delta-hedged returns for portfolios thus obtained. The sample period is from January 1996 to December 2017. The t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level. All returns are expressed as percentages.

Panel A: CallMoney and NH													
	OTM calls							ATM calls					
	1 (Low NH)	2	3	4	5 (High NH)	N5-N1		1 (Low NH)	2	3	4	5 (High NH)	N5-N1
1 (Low CallMoney)	0.465	0.373	0.355	0.289	0.457		1 (Low CallMoney)	0.356	0.116	0.190	0.106	0.233	
2	0.538	0.327	0.202	0.315	0.336		2	0.146	0.154	0.065	0.084	0.198	
3	0.122	0.178	-0.025	0.056	0.182		3	-0.022	0.039	-0.040	-0.012	0.059	
4	-0.110	-0.194	-0.081	-0.060	0.105		4	-0.143	-0.185	-0.161	-0.115	0.018	
5 (High CallMoney)	-0.638	-0.730	-0.699	-0.635	-0.222		5 (High CallMoney)	-0.546	-0.504	-0.376	-0.377	-0.107	
5-1	-1.103***	-1.104***	-1.054***	-0.924***	-0.679***	0.424*	5-1	-0.902***	-0.620***	-0.566***	-0.483***	-0.340***	0.562***
	(-6.25)	(-8.32)	(-8.36)	(-6.37)	(-5.14)	(1.89)		(-6.95)	(-5.87)	(-6.82)	(-6.43)	(-3.94)	(3.52)
5-1 FF4 Alpha	-1.122***	-1.129***	-1.027***	-0.889***	-0.660***	0.462**	5-1 FF4 Alpha	-0.890***	-0.582***	-0.537***	-0.435***	-0.293***	0.597***
	(-7.29)	(-8.95)	(-8.61)	(-5.94)	(-4.68)	(2.27)		(-6.51)	(-5.85)	(-6.38)	(-5.38)	(-3.22)	(3.62)
Panel B: CallMoney and CGO													
	OTM calls							ATM calls					
	1 (Low CGO)	2	3	4	5 (High CGO)	C5-C1		1 (Low CGO)	2	3	4	5 (High CGO)	C5-C1
1 (Low CallMoney)	0.687	0.495	0.413	0.350	0.411		1 (Low CallMoney)	0.313	0.258	0.197	0.181	0.241	
2	0.601	0.434	0.287	0.194	0.334		2	0.172	0.116	0.104	0.070	0.206	
3	0.197	0.186	0.038	0.012	0.115		3	-0.093	0.088	-0.004	-0.027	0.164	
4	-0.030	-0.020	-0.287	-0.118	-0.050		4	-0.254	-0.059	-0.156	-0.206	-0.050	
5 (High CallMoney)	-0.684	-0.659	-0.620	-0.649	-0.603		5 (High CallMoney)	-0.593	-0.451	-0.502	-0.428	-0.367	
5-1	-1.371***	-1.154***	-1.033***	-0.999***	-1.014***	0.357**	5-1	-0.905***	-0.710***	-0.699***	-0.609***	-0.608***	0.297***
	(-9.18)	(-10.05)	(-8.57)	(-6.63)	(-7.77)	(2.16)		(-7.90)	(-9.16)	(-7.65)	(-6.60)	(-6.82)	(2.61)
5-1 FF4 Alpha	-1.393***	-1.161***	-0.976***	-0.991***	-1.001***	0.392**	5-1 FF4 Alpha	-0.914***	-0.689***	-0.655***	-0.573***	-0.564***	0.350***
	(-10.03)	(-10.32)	(-8.17)	(-7.07)	(-7.94)	(2.30)		(-7.35)	(-9.22)	(-7.33)	(-6.84)	(-6.62)	(3.00)

per month, both significant at 1% level. In contrast, when stock price is near its 52-week high ("High NH"), the "5-1" return spread is -0.679%, and the corresponding four-factor alpha is -0.660%. Column "N5-N1" reports the difference between the above return spreads ($1.103\% - 0.679\%$) and alphas ($1.122\% - 0.660\%$). The difference in return spreads is 0.424%, significant at 10% level, and the difference in alphas is 0.462% per month, significant at 5% level.

For ATM calls, when stock price is far from its 52-week high ("Low NH"), the "5-1" return spread is -0.902% per month, with the corresponding four-factor alpha being -0.890% per month, both significant at 1% level. In contrast, when stock price is near its 52-week high ("High NH"), the "5-1" return spread is -0.340%, and the corresponding four-factor alpha is equal to -0.293%, only about one-third of the low NH's alpha spread. Column "N5-N1" reports the difference between the above return spreads ($0.902\% - 0.340\%$) and alphas ($0.890\% - 0.293\%$). This difference in return spreads is 0.562% and the difference in alphas is 0.597% per month, both significant at 1% level.

Results in Panel A indicate that $CallMoney$'s option return predictability is stronger when the underlying stock is further from its 52-week high, suggesting that option traders gamble more heavily when the current stock price is further from its psychological barrier. All the "5-1" return spreads are negative in Panel A, indicating a consistently negative option return predictability across different NH states.

Panel B of Table 3.7 reports the results of sorting by $CallMoney$ and CGO. For OTM calls, within the low CGO quintile, the return spread between high and low $CallMoney$ quintile portfolios is -1.371% per month, with the corresponding four-factor alpha being -1.393% per month. In contrast, within the high CGO quintile, the return spread is -1.014%, and the corresponding four-factor alpha is equal to -1.001%. Column "C5-C1" reports the difference between the above return spreads $(1.371\% - 1.014\%)$ and alphas $(1.393\% - 1.001\%)$. The difference in return spreads is 0.357%, and the difference in alphas is 0.392% per month, both significant at 5% level. For ATM calls, the difference "C5-C1" in return spreads (alphas) is 0.297% (0.350%) per month, both significant at 1% level.

Results in Panel B indicate that $CallMoney$'s option return predictability is stronger when capital gain overhang is lower, suggesting that option traders gamble more heavily when the desire to cover previous losses and break even is stronger. Again, all the "5-1" return spreads are negative in Panel B, indicating a consistently negative option return predictability across different CGO states.

Overall, our findings in Table 3.7 suggest that the interaction between gambling activity and past stock performance anchoring amplifies the negative option return predictability of $CallMoney$ when stock price is further from its 52-week high or investors have suffered capital loss.

3.4.6 Information uncertainty and gambling

Information plays an important role of rational investors' decision making process. Previous literature (Hirshleifer, 2001; Jiang et al., 2005; Kumar, 2009*a*; Zhang, 2006*b*) shows that investors exhibit stronger behavioural bias when there is greater information uncertainty. We examine whether $CallMoney$'s option return predictability is weakened or accentuated by information uncertainty of the underlying stock. If driven by investors' behavioural bias such as gambling, the predictability should accentuate when information uncertainty is high.

Following previous literature (Zhang, 2006*b*; Jiang et al., 2005), we use two proxies of information uncertainty: analysts' coverage and analysts forecast dispersion. A firm with larger analysts' coverage is likely to have more information available and is therefore less uncertain in its valuation. We measure analyst coverage ($ANALYSTS$) as the number of analysts following the firm. A larger analysts forecast dispersion reflects less of consensus among analysts, implying a higher degree of information uncertainty. We calculate analysts' forecast dispersion ($DISP$) as the standard deviation of analysts forecasts scaled by the mean analysts' forecast.

To analyse the interaction between gambling activity and information uncertainty, we conduct independent double-sort on $CallMoney$ and each of the information uncertainty proxies. We independently sort call options into quintiles based on $CallMoney$, and quintiles based on their underlying stocks' information uncertainty proxies and then calculate the equal-weighted delta-hedged call option returns for each of the 25 resulting portfolios. The results are reported in Table 3.8.

Panel A of Table 3.8 reports the results on sorting by $CallMoney$ and $ANALYSTS$. For OTM calls,when analysts coverage is low ("Low $ANALYSTS$"), the return spread between high $CallMoney$ and low $CallMoney$ quintiles ("5-1") is -1.591% per month, and the corresponding four-factor alpha is -1.577% per month. When analysts coverage is high ("High $ANALYSTS$"), the "5-1" return spread is -0.566% per month, and the corresponding four-factor alpha is -0.506% per month. The "A5-A1" column reports the differences between the above return spreads and alphas.

Table 3.8: Portfolios double sorted independently on $CallMoney$ and information uncertainty measure

In each month, all OTM/ATM call options are double sorted in two ways. In one double sort, options are independently sorted into 5 × 5 equally weighted portfolios based on $CallMoney$ (1 = Low $CallMoney$, 5 = High $CallMoney$) and analysts coverage measure $ANALYSTS$ (1 = Low ANALYSTS, 5 = High ANALYSTS). Panel A reports the average monthly returns for portfolios thus obtained. In the second double sort, options are independently sorted into 5 × 5 equally weighted portfolios based on lagged $CallMoney$ (1 = Low $CallMoney$, 5 = High $CallMoney$) and analysts' forecast dispersion measure $DISP$ (1 = Low DISP, 5 = High DISP). Panel B reports the average monthly returns for portfolios thus obtained. The sample period is from January 1996 to December 2017. The t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level. All returns are expressed as percentages.

Panel A: CallMoney and ANALYSTS

OTM calls	1 (Low ANALYSTS)	2	3	4	5 (High ANALYSTS)	A5-A1
1 (Low CallMoney)	0.306	0.300	0.429	0.462	0.465	
2	0.171	0.244	0.323	0.316	0.352	
3	-0.031	0.038	0.199	0.201	0.185	
4	-0.368	-0.204	0.012	0.066	0.091	
5 (High CallMoney)	-1.285	-0.788	-0.510	-0.295	-0.102	
5-1	-1.591***	-1.088***	-0.938***	-0.757***	-0.566***	1.025***
	(-11.49)	(-8.55)	(-7.03)	(-4.82)	(-3.82)	(5.65)
5-1 FF4 Alpha	-1.577***	-1.089***	-0.937***	-0.753***	-0.506***	1.071***
	(-11.23)	(-9.52)	(-7.65)	(-5.68)	(-3.73)	(5.69)

ATM calls	1 (Low ANALYSTS)	2	3	4	5 (High ANALYSTS)	A5-A1
1 (Low CallMoney)	0.199	0.172	0.197	0.256	0.230	
2	0.084	0.152	0.171	0.150	0.127	
3	-0.113	-0.043	0.065	0.133	0.114	
4	-0.376	-0.213	-0.125	-0.021	0.077	
5 (High CallMoney)	-1.012	-0.517	-0.259	-0.129	-0.129	
5-1	-1.211***	-0.689***	-0.457***	-0.385***	-0.360***	0.851***
	(-13.52)	(-6.75)	(-4.93)	(-3.66)	(-3.48)	(7.76)
5-1 FF4 Alpha	-1.174***	-0.653***	-0.422***	-0.341***	-0.320***	0.854***
	(-13.02)	(-7.02)	(-5.52)	(-3.77)	(-3.60)	(8.25)

Panel B: CallMoney and DISP

OTM calls	1 (Low DISP)	2	3	4	5 (High DISP)	D5-D1
1 (Low CallMoney)	0.287	0.463	0.418	0.587	0.542	
2	0.180	0.207	0.368	0.367	0.556	
3	0.006	0.073	0.219	0.218	0.176	
4	-0.165	-0.080	-0.025	0.064	-0.077	
5 (High CallMoney)	-0.461	-0.286	-0.425	-0.508	-0.506	
5-1	-0.748***	-0.749***	-0.843***	-1.095***	-1.048***	-0.300**
	(-5.64)	(-4.69)	(-6.61)	(-7.71)	(-8.24)	(-2.30)
5-1 FF4 Alpha	-0.717***	-0.773***	-0.863***	-1.079***	-1.025***	-0.308**
	(-5.55)	(-5.28)	(-6.88)	(-8.90)	(-8.61)	(-2.36)

ATM calls	1 (Low DISP)	2	3	4	5 (High DISP)	D5-D1
1 (Low CallMoney)	0.131	0.232	0.227	0.337	0.161	
2	0.096	0.170	0.173	0.111	0.139	
3	0.009	0.027	0.083	0.092	0.123	
4	-0.145	-0.018	-0.086	-0.079	-0.086	
5 (High CallMoney)	-0.255	-0.176	-0.258	-0.346	-0.488	
5-1	-0.386***	-0.409***	-0.485***	-0.683***	-0.648***	-0.262**
	(-4.48)	(-3.97)	(-4.07)	(-6.24)	(-6.57)	(-2.12)
5-1 FF4 Alpha	-0.350***	-0.391***	-0.464***	-0.650***	-0.604***	-0.254**
	(-3.65)	(-4.28)	(-4.60)	(-6.46)	(-6.65)	(-2.00)

Both differences (1.025% and 1.071% per month) are significant at 1% level, showing that more analysts' coverage decreases $CallMoney$'s predictability on option returns.

For ATM calls, when analysts coverage is low, the "5-1" return spread between high $CallMoney$ and low $CallMoney$ quintiles -1.211% per month, and the corresponding four-factor alpha is -1.174% per month. When analysts coverage is high, the "5-1" return spread is -0.360% per month, and the corresponding four-factor alpha is -0.320% per month. The "A5-A1" difference in return spreads is 0.851% per month and the difference in alphas is 0.854% per month, both are significant at 1% level. Panel A shows that option return predictability of $CallMoney$ is stronger when analysts' coverage is lower. More analysts' coverage leads to less information uncertainty and less overpricing of OTM/ATM call options.

Panel B of Table 3.8 reports the independent double-sort results of sorting by $CallMoney$ and $DISP$. For OTM calls, when analysts forecast disperion is low ("Low DISP"), the "5-1" return spread between high $CallMoney$ and low $CallMoney$ quintiles is -0.748% per month, and the corresponding four-factor alpha is -0.717% per month. When analysts forecast disperion is high ("High DISP"), the "5-1" return spread is -1.048% per month, and the corresponding four-factor alpha is -1.025% per month. Column "D5-D1" reports the difference between the above return spreads (-0.300%), and the difference in four-factor alphas (-0.308%), with 5% significance level. For ATM calls, the sorting results show similar pattern. The "D5-D1" difference in return spreads is -0.262% and difference in alphas is -0.254%, both are significant at 5% level. Panel B shows that option return predictability of $CallMoney$ stronger when analysts forecast dispersion is higher.

Overall, results in Table 3.8 support that $CallMoney$'s option return predictability comes from option traders' gambling activity. Traders gamble more when company information is more uncertain and gambling activity leads to overpricing of OTM/ATM call options. Also, In Table 3.8, all "5-1" spreads are negative, suggesting $CallMoney$'s consistent negative option return predictability across all information uncertainty levels.

3.4.7 CallMoney and stock returns

If $CallMoney$ successfully captures gambling activity in options market driven by investors' optimism in the underlying stock, then this optimism will also cause overvaluation of the stock and should negatively predict stock returns. In this section, we test whether $CallMoney$ has stock return predictability.

We run Fama-MacBeth regressions with monthly future stock returns as dependent variable. We consider six regression models and Table 3.9 reports the results.

Table 3.9: Fama-MacBeth regression analyses: stock return predictability

The table presents the results of Fama and MacBeth (1973) regressions of month $t+1$ stock returns on month t *CallMoney* and control variables. The table presents the time series averages of the monthly cross-sectional regression coefficients. Also reported are the average adjusted R-square. The sample period is from January 1996 to December 2017. Newey and West (1987) t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
CallMoney	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***
	(-3.67)	(-4.17)	(-4.17)	(-3.83)	(-4.13)	-3.83
MAX			-0.012			0.001
			(-0.71)			(0.10)
IVOL				-0.222***		-0.218***
				(-3.01)		(-3.19)
ISKEW					-0.001**	-0.001**
					(-2.41)	(-2.47)
REV		-0.014**	-0.012*	-0.013**	-0.014**	-0.013**
		(-2.25)	(-1.84)	(-2.15)	(-2.27)	(-2.02)
MOM		0.004	0.004	0.005*	0.004	0.005**
		(1.54)	(1.64)	(1.89)	(1.64)	(2.05)
log(SIZE)		0.000	0.000	-0.001	0.000	-0.001
		(0.39)	(0.27)	(-1.02)	(0.34)	(-0.96)
BETA		0.000	0.000	0.001	0.000	0.001
		(0.16)	(0.21)	(0.40)	(0.18)	(0.38)
BTM		-0.003	-0.003	-0.003	-0.003	-0.003
		(-1.10)	(-1.11)	(-1.28)	(-1.08)	(-1.27)
ILLIQUIDITY		8.900***	9.153***	9.635***	8.917***	9.827***
		(2.78)	(2.91)	(2.89)	(2.79)	(2.99)
CIV		0.008	0.008	0.005	0.007	0.004
		(0.68)	(0.67)	(0.49)	(0.62)	(0.37)
PIV		-0.010	-0.009	-0.010	-0.009	-0.009
		(-1.06)	(-0.94)	(-1.04)	(-1.01)	(-0.92)
INTERCEPT	YES	YES	YES	YES	YES	YES
Adjusted R^2 (%)	0.55	8.71	9.16	9.40	8.77	9.78

In Model 1, we estimate univariate regressions of future stock returns on $CallMoney$. The coefficient is negative (-0.006) and statistically significant (with a t-statistic of -3.67), confirming that $CallMoney$ can negatively predict the future stock returns.

In Model 2, we estimate the multivariate regressions after controlling for the stock characteristics. The control variables are those typically used in the empirical asset pricing literature (e.g. Bali et al. (2011)). We also control for the option informed trading measures CIV and PIV. The coefficients on $CallMoney$ remain negative and significant at 1% level. The coefficients on the other variables have the signs consistent with the literature (Bali et al., 2011; Andreou et al., 2018).

From Model 3 to Model 6, we further control for each or all of the indirect gambling measures (MAX, $IVOL$, and $ISKEW$).[8] The negative and significant stock return predictability of $CallMoney$ holds consistently across these models. In Model 3, the coefficient on MAX is insignificant. In Model 4, the coefficient on $IVOL$ is negative and significant at 1% level. In Model 5, the coefficient on $ISKEW$ is negative and significant at 5% level. In Model 6, when all indirect gambling measures are included, the coefficient on MAX is insignificant.

Across all models, the magnitude of $CallMoney$'s coefficient persists at -0.006, indicating a unique set of information content that cannot be explained by either control variables or indirect gambling measures. $CallMoney$ performs better than indirect gambling measures at predicting stock returns.

We conduct one additional test in this section to provide evidence that the gambling activity measure $CallMoney$ captures irrational speculative trading behaviours. As mentioned in Section 3.2, one strand of literature defines gamblers as those who enter the market for risk-based excitement and are less concerned about fundamental values. We test whether gambling activity as measured by $CallMoney$ is more active in stocks with higher volatility but are expected to have worse fundamental performance. We sort stock-month observations into quintiles based on the standard deviation of stock daily return VOL observed in the previous month. We then report time-series average $CallMoney$ across quintiles. Column 1 of Table 3.10 shows

[8] We omit $OSKEW$ because it is option contract dependent measure.

that $CallMoney$ increases across increasing VOL quintiles. The difference between extreme quintiles (Q5-Q1) is 0.320 with a t-statistic of 10.41. This indicates that $CallMoney$ increases on the stock return volatility during the previous month.

Table 3.10: $CallMoney$ and stocks' risk and mispricing characteristics

This table reports $CallMoney$ across quintiles of stock volatility and stock mispricing level. Stock volatility VOL is the standard deviation of stock daily return during the previous month. Stock mispricing level $Mispricing$ is calculated following Stambaugh and Yuan (2017). At the beginning of each month, we sort stocks into quintiles based on each of the two proxies observed at the previous month-end, where Q5 (Q1) is the quintile with the highest (lowest) level of the proxy. Then we calculate and report the time-series average of the mean $CallMoney$ in each quintile. We also report the $CallMoney$ spread between Q5 and Q1. The t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level.

	Sort by VOL_{t-1}	Sort by $Mispricing_{t-1}$
Q1	0.280	0.189
Q2	0.211	0.184
Q3	0.240	0.217
Q4	0.299	0.245
Q5	0.600	0.360
Q5-Q1	0.320***	0.171***
t-stat	(10.41)	(5.17)

Though we show that a higher $CallMoney$ is associated with a lower future stock return, our result is based on realized returns. We would like to know whether the risk-seeking investors are pursuing "worse-performance" stocks ex ante. Stambaugh and Yuan (2017) estimate a mispricing level for each stock using the accounting-based information, and identify stocks that are relatively overpriced (high mispricing level) and underpriced (low mispricing level). Stocks with high (low) mispricing level are expected to underperform (outperform) the counterparties, and the accounting information is publicly available ex ante. We therefore estimate an accounting-based mispricing level following Stambaugh and Yuan (2017), and we examine the relation between the mispricing level of stocks in month $t-1$ and $CallMoney$ in month t. We sort stock-month observations into quintiles based on the mispricing level of stock observed at the previous month-end. We then report time-series average $CallMoney$

across quintiles. Column 2 of Table 3.10 shows that $CallMoney$ increases across increasing mispricing quintiles. The difference between extreme quintiles (Q5-Q1) is 0.171 with a t-statistic of 5.17.

Combining the above results, the positive relations between $CallMoney$ and VOL, and between $CallMoney$ and mispricing level indicate that the traders in stocks with high $CallMoney$ are trading securities with higher risk but lower expected return. This is consistent with our evidence from the stochastic dominance analysis, where the high $CallMoney$ investors are risk-seeking.

3.4.8 Subperiod analysis

In this section, we conduct subperiod analysis. We first examine whether the negative option and stock return predictability of $CallMoney$ is more pronounced during high market volatility periods, as Clark et al. (2018) show that gamblers prefer a volatile stock market to express their predisposition to take risks. Kumar (2009b) also argues that gamblers are willing to undertake more risk as they believe that the extreme return events observed in the past are more likely to be realized again when volatility is high.

We define a high (low) market volatility month as one in which the average VIX level in the month is above (below) the sample median, and we examine the option and stock return predictability of $CallMoney$ using Fama-MacBeth (1973) regressions following high market volatility months ($High\_VIX$) and low market volatility months (Low_VIX) respectively. We report the time-series averages of the coefficients, along with their Newey-West adjusted t-statistics in Panel A of Table

3.11.

Table 3.11: Fama-MacBeth regression analyses: different market volatility and sentiment subperiods

This table presents the results of Fama-MacBeth (1973) regressions for high/low market volatility periods (High_VIX/Low_VIX) in Panel A and for high/low market sentiment periods (High_sent/Low_sent) in Panel B. The dependent variables are month $t+1$ delta-hedged OTM call option returns (Panels A1 and B1), delta-hedged ATM call option returns (Panels A2 and B2), and stock returns (Panels A3 and B3). The independent variables are $CallMoney$ measures and the control variables in month t. High (low) market volatility month is the month in which the average VIX level in the month is above (below) the sample median. High (low) market sentiment month is the month in which the Baker and Wugler sentiment index level in the month is above (below) the sample median. The time series average coefficients on $CallMoney$ of the monthly cross-sectional regressions are reported. Also reported are the average adjusted R-square. The sample period is from January 1996 to December 2017. Newey and West (1987) t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level.

Panel A: Market volatility						
	Panel A1: OTM_call		Panel A2: ATM_call		Panel A3: Stock	
	High_VIX	Low_VIX	High_VIX	Low_VIX	High_VIX	Low_VIX
CallMoney	-0.006***	-0.003***	-0.006***	-0.003***	-0.009***	-0.003***
	(-5.15)	(-3.75)	(-6.61)	(-4.61)	(-3.89)	(-2.67)
CONTROL	YES	YES	YES	YES	YES	YES
INTERCEPT	YES	YES	YES	YES	YES	YES
Adjusted R^2 (%)	6.13	4.29	6.80	5.37	10.80	6.59
Panel B: Market sentiment						
	Panel B1: OTM_call		Panel B2: ATM_call		Panel B3: Stock	
	High_sent	Low_sent	High_sent	Low_sent	High_sent	Low_sent
CallMoney	-0.006***	-0.004***	-0.005***	-0.003***	-0.007***	-0.004***
	(-4.97)	(-4.05)	(-6.59)	(-4.28)	(-3.18)	(-3.14)
CONTROL	YES	YES	YES	YES	YES	YES
INTERCEPT	YES	YES	YES	YES	YES	YES
Adjusted R^2 (%)	5.88	4.59	6.46	5.75	8.64	6.79

In Panel A1, for regressions of OTM call returns, the coeffcient on $CallMoney$ is -0.006 with a t-statistic of -5.15 for $High\_VIX$ periods, and -0.003 with a t-statistic of -3.75 for Low_VIX periods. In Panel A2, for regressions of ATM call returns, the coeffcient on $CallMoney$ is -0.006 with a t-statistic of -6.61 for $High\_VIX$ periods, and -0.003 with a t-statistic of -4.61 for Low_VIX periods. In Panel A3, for regressions of stock returns, the coeffcient on $CallMoney$ is -0.009 with a t-statistic of -3.89 for $High\_VIX$ periods, and -0.003 with a t-statistic of -2.67 for Low_VIX periods. Overall, the results show that the negative option/stock return predictability of $CallMoney$ is more pronounced when the market volatility is high, indicating that gamblers are more active during high market volatility periods.

We argue that $CallMoney$ measures investors' gambling activity by showing that it captures excessive investor optimism during the bubble periods, and it negatively predicts future option/stock returns. We would like to know whether such gambling-induced overpricing of options/stocks differs in periods with different investor optimism levels. To determine the market optimism level, we use the well-known Baker and Wurgler (2006) market sentiment index (denoted as BW).

We define a high-sentiment (low-sentiment) month as the one with the BW index above (below) the sample median value in the month, and we examine the option and stock return predictability of $CallMoney$ using Fama-MacBeth (1973) regressions following high market sentiment months ($High\_sent$) and low market sentiment months (Low_sent) respectively. Panel B of Table 3.11 reports the time-series averages of the coefficients, along with their Newey-West adjusted t-statistics.

In Panel B1, for regressions of OTM call returns, the coeffcient on $CallMoney$ is -0.006 with a t-statistic of -4.97 for $High\_sent$ periods, and -0.004 with a t-statistic of -4.05 for Low_sent periods. In Panel B2, for regressions of ATM call returns, the coeffcient on $CallMoney$ is -0.005 with a t-statistic of -6.59 for $High\_sent$ periods, and -0.003 with a t-statistic of -4.28 for Low_sent periods. In Panel B3, for regressions of stock returns, the coeffcient on $CallMoney$ is -0.007 with a t-statistic of -3.18 for $High\_sent$ periods, and -0.004 with a t-statistic of -3.14 for Low_sent periods.

Overall, the results suggest that the negative option/stock return predictability of $CallMoney$ exists in both high and low market sentiment periods. Nevertheless, the negative option/stock return predictability of $CallMoney$ is more pronounced when the market sentiment is high, indicating that there is more pronounced overpricing of options and stocks by the gamblers during high market sentiment periods.

3.5 Discussions

3.5.1 Alternative constructions of CallMoney

In this section, we check the robustness of $CallMoney$ under alternative constructions. We test whether $CallMoney_t$ observed at the last trading day of each month and the change of $CallMoney$ ($\Delta CallMoney$) also have option and stock return predictability. We calculate $\Delta CallMoney$ as the change in $CallMoney$ relative to the average of past 6-months' $CallMoney$ as follows:

$$\Delta CallMoney = \frac{CallMoney - \overline{CallMoney}}{\overline{CallMoney}},$$

where $\overline{CallMoney}$ is the average $CallMoney$ for the firm over the previous 6 months.

We perform Fama-MacBeth regressions of delta-hedged OTM call option returns and stock returns on these two alternative $CallMoney$ measures. Table 3.12 presents the regression results.

Table 3.12: Fama-MacBeth regression analyses: alternative $CallMoney$ measures

The table presents the results of Fama and MacBeth (1973) regressions. The dependent variables are month $t+1$ delta-hedged option returns for OTM call options and stock returns. The independent variables are one of the alternative $CallMoney$ measures and month t control variables. The alternative $CallMoney$ measures are: 1) $CallMoney_t$ estimated on the last trading day of month t; and 2) $\Delta CallMoney$, the change in monthly $CallMoney$ relative to the 6-month average of past $CallMoney$. The table presents the time series averages of the monthly cross-sectional regression coefficients. Also reported are the average adjusted R-square. The sample period is from January 1996 to December 2017. Newey and West (1987) t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level.

Panel A: OTM calls			Panel B: Stock		
$CallMoney_t$	-0.005***		$CallMoney_t$	-0.006***	
	(-6.27)			(-4.54)	
$\Delta CallMoney$		-0.003***	$\Delta CallMoney$		-0.001***
		(-8.33)			(-2.67)
REV	0.003	0.004	REV	-0.013*	-0.013**
	(0.89)	(0.99)		(-1.93)	(-2.05)
MOM	-0.003**	-0.002*	MOM	0.003	0.003
	(-2.31)	(-1.84)		(1.23)	(1.35)
log(SIZE)	0.001*	0.001	log(SIZE)	0.001	0.001
	(1.70)	(1.57)		(0.83)	(0.84)
BETA	0.002**	0.001*	BETA	0.000	0.000
	(2.07)	(1.91)		(0.26)	(-0.08)
BTM	-0.000	-0.000	BTM	-0.004	-0.003
	(-0.12)	(-0.02)		(-1.14)	(-1.05)
TURNOVER	0.000**	-0.000	ILLIQUIDITY	18.342***	10.068***
	(2.54)	(-0.55)		(3.34)	(3.11)
ILLIQUIDITY	-19.169**	-18.744***	CIV	0.002	0.007
	(-2.21)	(-2.84)		(0.12)	(0.59)
CIV	-0.049***	-0.049***	PIV	-0.004	-0.011
	(-4.30)	(-4.76)		(-0.28)	(-1.18)
PIV	0.007	0.012	INTERCEPT	YES	YES
	(0.56)	(1.14)	Adjusted R^2 (%)	8.81	8.38
VRP	-0.000***	-0.000***			
	(-4.38)	(-3.94)			
VOV	-0.043***	-0.037***			
	(-7.30)	(-7.45)			
PCR	0.017***	0.019***			
	(5.34)	(5.68)			
INTERCEPT	YES	YES			
Adjusted R^2 (%)	5.78	5.00			

Panel A and B of Table 3.12 report the results for regressions on future OTM call option returns and stock returns respectively. The coefficients on both $CallMoney_t$ and $\Delta CallMoney$ in both panels are negative and significant at 1% level, after controlling for option and stock characteristics. This results indicate that the negative $CallMoney$-return predictability is robust to alternative specifications for $CallMoney$. Compared to the indirect gambling measures that depend on time-series observations for construction, $CallMoney$ only needs point observation to capture investors' gambling demand.

3.5.2 Gambling activity and option demand

The open interest normalised by either trading volume or shares outstanding is an important proxy of option demands. The option overpricing may come from option demand as option dealers charge a higher option premium when the option demands are higher (Cao and Han, 2013). We show that $CallMoney$'s return predictability cannot be explained by option demands of both call and put options, highlighting the importance of considering OTM call options to capture gambling demands. We calculate option demand OD as the ratio of total number of option contracts (calls and puts) that are open to shares outstanding.

Then we double sort option (stock) portfolios in two ways We first form decile option (stock) portfolios ranked based on OD, and within each OD decile, we sort options (stocks) again into decile portfolios based on $CallMoney$. And we also conduct the reverse sort by sorting on OD after controlling $CallMoney$. Table 3.13 report the double-sort results.

Table 3.13: Portfolios double sorted on $CallMoney$ and option demand

In each month, all OTM call options (stocks) are double sorted in two ways. In one double sort, options are first sorted into deciles based on option demand measure OD. Then within each decile, options (stocks) are further sorted into ten equally weighted portfolios based on $CallMoney$ (1 = Low $CallMoney$, 10 = High $CallMoney$). Panel A reports the average monthly returns for portfolios thus obtained. The second double sort reverses the sort order, and the corresponding results are reported in Panel B. The sample period is from January 1996 to December 2017. The t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level. All returns are expressed as percentages.

Panel A: Sorted by CallMoney controlling for OD			Panel B: Sorted by OD controlling for CallMoney		
	OTM call	Stock		OTM call	Stock
1 (Low CallMoney)	0.401	0.817	1 (Low OD)	-0.377	0.553
2	0.316	0.830	2	-0.164	0.737
3	0.228	0.787	3	-0.150	0.520
4	0.206	0.845	4	0.014	0.593
5	0.082	0.610	5	0.011	0.539
6	0.006	0.502	6	0.031	0.517
7	-0.025	0.454	7	0.117	0.541
8	-0.241	0.436	8	0.085	0.493
9	-0.367	0.298	9	0.104	0.457
10 (High CallMoney)	-0.767	-0.089	10 (High OD)	0.157	0.571
10-1	-1.168***	-0.906***	10-1	0.533***	0.018
	(-11.27)	(-3.98)		(6.83)	(0.14)
10-1 FF4 Alpha	-1.122***	-0.962***	10-1 FF4 Alpha	0.495***	0.077
	(-11.84)	(-5.84)		(6.53)	(0.68)

Panel A of Table 3.13 shows the option and stock decile portfolio returns sorted by $CallMoney$ after controlling for OD.When OD is controlled for in Panel A, the return spread between decile 10 and decile 1 option portfolios ("10-1") is -1.168% per month (t-statistic -11.27). The magnitude is only slightly lower than the portfolio return spread reported in Table 3.3. The spread between decile 10 and decile 1 stock portfolios ("10-1") is -0.906% per month (t-statistic -3.98). Resutls in Panel A suggest that option demand cannot explain $CallMoney$'s option (stock) return predictability.

In Panel B, we sort options into deciles based on OD after controlling for $CallMoney$. The spread between decile 10 and decile 1 portfolios ("10-1") is significantly *positive* at 0.533% per month. Cao and Han (2013) show that option demands negatively predict delta-hedge option returns. After controlling $CallMoney$ the negative predictability of option demands flips to be a positive one. The spread and four-factor alpha between "10-1" OD deciles for stock returns are both insignificant.

Table 3.13 shows that $CallMoney$'s option (stock) return predictability cannot be explained by overpricing caused by option demands.[9]

3.5.3 Dot-Com Bubble

We show in Figure 3.1 that aggregate $CallMoney$ successfully captures the bubble events in the market. In this section, we formally test the relation between $CallMoney$ and the market bubble level. We focus on the bubble events in the Nasdaq index, as the literature uses the aggregate market to book ratio of the Nasdaq index to capture the overvaluation of the Nasdaq index and the Nasdaq bubble (e.g. the Dot-Com bubble) (Pástor and Veronesi, 2006; Li and Xue, 2009). Following these studies, we calculate the aggregate market to book ratio of the Nasdaq index (M/B), and test the relation between M/B and $CallMoney$ of the HiTech/non-HiTech firms.[10]

Following Clark et al. (2018), we perform the following time-series OLS regressions to examine the contemporaneous relation between $CallMoney$ and M/B:

$$CallMoney_{t,i} = \beta_{0,i} + \beta_{1,i}\, M/B_t + \epsilon_{t,i},$$

where $CallMoney_{t,i}$ is the median $CallMoney$ of HiTech firms or non-HiTech firms

[9] We conduct Fama-MacBeth regressions by regressing future option and stock returns on orthogonalized $CallMoney$, OD and all control variables and we find orthogonalized $CallMoney$ still has a significantly negative coefficient. The orthogonalized $CallMoney$ is the residual term via regressions of $CallMoney$ on OD. We also conduct robustness check by considering put-call trading volume ratio in Fama-MacBeth regressions, $CallMoney$'s negative return predictability still robustly hold. These results can be obtained from authors upon request.

[10] We define non-HiTech firms as those firms that are not classified into the HiTech industry.

in quarter t, and M/B_t is the aggregate market to book ratio of the Nasdaq index in quarter t. We also perform an alternative test to examine the relation between change in $CallMoney$ and change in M/B:

$$\%CallMoney_{t,i} = \gamma_{0,i} + \gamma_{1,i}\,\%M/B_t + \epsilon_{t,i},$$

where $\%CallMoney_{t,i}$ ($\%M/B_t$) is the percentage change in $CallMoney_{t,i}$ (M/B_t) from quarter $t-1$ to quarter t.

We expect to observe significantly positive β_1 and γ_1 for HiTech firms if $CallMoney$ successfully captures HiTech investors' optimism during the Nasdaq bubble. We are also interested in whether the coefficients are significant for the non-HiTech firms.

Table 3.14 reports the coefficients on M/B_t and $\%M/B_t$ and the adjusted R^2 for Model 1 (2), where the dependent variable is $CallMoney_{t,i}$ ($\%CallMoney_{t,i}$), and the independent variable is M/B_t ($\%M/B_t$).

Table 3.14: Nasdaq bubble and $CallMoney$

This table presents the OLS regression results, where the dependent variable is the gambling activity measure for HiTech (Non-HiTech) firms in Panel A (B), and the independent variable is the Nasdaq Index valuation proxy. In Model 1, the dependent variable is $CallMoney_t$, the median $CallMoney$ in quarter t, and the independent variable is M/B_t, the aggregate market to book ratio of the Nasdaq Index in quarter t. In Model 2, the dependent variable is $\%CallMoney_t$, the percentage change of $CallMoney_t$ from quarter $t-1$ to quarter t, and the independent variable is $\%M/B_t$, the percentage change of M/B_t from quarter $t-1$ to quarter t. The sample period is from January 1996 to December 2017. The t-statistics are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level.

	Panel A: HiTech firms		Panel B: Non-HiTech firms	
	Model 1	Model 2	Model 1	Model 2
β_1/γ_1	0.033**	0.187***	0.002	0.058
t-stat	(2.30)	(2.97)	(0.18)	(1.24)
INTERCEPT	YES	YES	YES	YES
Adjusted R^2 (%)	4.69	8.31	0.04	0.63

Panel A presents the results for the HiTech firms. The coefficient on M/B_t is 0.033

(t-stat=2.30) in Model 1, and the coefficient on $\%M/B_t$ is 0.187 (t-stat=2.97) in Model 2. The significantly positive coefficients in both models indicate that the gambling activity in HiTech firms is more active when the Nasdaq index is more overvalued. Panel B presents the results for the non-HiTech firms. The coefficients in both models are insignificant. Collectively, these results imply that during the Nasdaq bubble, the market participants gamble actively in the HiTech firms instead of the non-HiTech firms.

3.6 Conclusion

We propose a new gamling activity measure, $CallMoney$, by jointly considering open interest and moneyness of OTM call options. The time series of $CallMoney$ captures the excessive optimism during the previous financial bubbles. $CallMoney$ negatively and robustly predicts cross-sectional delta-hedged call option returns and stock returns, suggesting the overpricing of call options and stocks when gambling activity is high. Empirical study shows that $CallMoney$ generally outperforms existing indirect gambling measures with respect to option (stock) return predictability.

This negative $CallMoney$ return predictability is more pronounced when the stock price is further far from its' 52-week high, stock's capital gain overhang is lower, and when information uncertainty of stock is higher, further suggesting that $CallMoney$ captures gambling activity in options market.

Bibliography

Acharya, V. and Naqvi, H. (2019), 'On reaching for yield and the coexistence of bubbles and negative bubbles', *Journal of Financial Intermediation* **38**, 1–10.

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

An, B.-J., Ang, A., Bali, T. G. and Cakici, N. (2014), 'The joint cross section of stocks and options', *Journal of Finance* **69**(5), 2279–2337.

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

Andreou, P. C., Kagkadis, A., Philip, D. and Tuneshev, R. (2018), 'Differences in options investors' expectations and the cross-section of stock returns', *Journal of Banking & Finance* **94**, 315–336.

Ang, A., Hodrick, R. J., Xing, Y. and Zhang, X. (2006), 'The cross-section of volatility and expected returns', *Journal of Finance* **61**(1), 259–299.

Atilgan, Y., Bali, T. G., Demirtas, K. O. and Gunaydin, A. D. (2020*a*), 'Left-tail momentum: Underreaction to bad news, costly arbitrage and equity returns', *Journal of Financial Economics* **135**(3), 725–753.

Atilgan, Y., Bali, T. G., Demirtas, K. O. and Gunaydin, A. D. (2020*b*), 'Left-tail momentum: underreaction to bad news, costly arbitrage and equity returns', *Journal of Financial Economics* **135**(3), 725–753.

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

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

Bali, T. G., Cakici, N. and 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., Engle, R. F. and Tang, Y. (2016), 'Dynamic conditional beta is alive and well in the cross section of daily stock returns', *Management Science* **63**(11), 3760–3779.

Bali, T. G. and Murray, S. (2013), 'Does risk-neutral skewness predict the cross section of equity option portfolio returns?', *Journal of Financial and Quantitative Analysis* **48**(4), 1145–1171.

Baltussen, G., Van Bekkum, S. and Van Der Grient, B. (2018), 'Unknown unknowns: uncertainty about risk and stock returns', *Journal of Financial and Quantitative Analysis* **53**(4), 1615–1651.

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

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

Bates, D. S. (1991), 'The crash of 87: was it expected?the evidence from options markets', *Journal of Finance* **46**(3), 1009–1044.

Bates, D. S. (1996), 'Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options', *The Review of Financial Studies* **9**(1), 69–107.

Bergsma, K., Csapi, V., Diavatopoulos, D. and Fodor, A. (2020), 'Show me the money: Option moneyness concentration and future stock returns', *Journal of Futures Markets* **40**(5), 761–775.

Berkelaar, A. B., Kouwenberg, R. and Post, T. (2004), 'Optimal portfolio choice under loss aversion', *Review of Economics and Statistics* **86**(4), 973–987.

Bernales, A., Verousis, T., Voukelatos, N. and Zhang, M. (2019), 'What do we know about individual equity options?', *Available at SSRN* .

Billings, M. B. and Jennings, R. (2011), 'The option market's anticipation of information content in earnings announcements', *Review of Accounting Studies* **16**(3), 587–619.

Blau, B. M., Bowles, T. B. and Whitby, R. J. (2016), 'Gambling preferences, options markets, and volatility', *Journal of Financial and Quantitative Analysis* **51**(2), 515–540.

Bollen, N. P. and Whaley, R. E. (2004), 'Does net buying pressure affect the shape of implied volatility functions?', *Journal of Finance* **59**(2), 711–753.

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

Brown, S., Lu, Y., Ray, S. and Teo, M. (2018), 'Sensation seeking and hedge funds', *Journal of Finance* **73**(6), 2871–2914.

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

Byun, S.-J. and Kim, D.-H. (2016), 'Gambling preference and individual equity option returns', *Journal of Financial Economics* **122**(1), 155–174.

Cao, J. and Han, B. (2013), 'Cross section of option returns and idiosyncratic stock volatility', *Journal of Financial Economics* **108**(1), 231–249.

Cao, J., Han, B., Tong, Q. and Zhan, X. (2017), 'Option return predictability', *Rotman School of Management Working Paper* (2698267).

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

Chabi-Yo, F., Ruenzi, S. and Weigert, F. (2018), 'Crash sensitivity and cross-section of expected stock returns', *Journal of Financial and Quantitative Analysis* **53**(3), 1059–1100.

Cho, Y.-H., Linton, O. and Whang, Y.-J. (2007), 'Are there monday effects in stock returns: A stochastic dominance approach', *Journal of Empirical Finance* **14**(5), 736–755.

Choi, Y., Kim, W. and Kwon, E. (2019), 'Are disposition effect and skew preference correlated? evidence from account-level elw transactions', *Journal of Futures Markets* .

Choy, S.-K. (2015), 'Retail clientele and option returns', *Journal of Banking & Finance* **51**, 26–42.

Clark, G. L., Fiaschetti, M., Tufano, P. and Viehs, M. (2018), 'Playing with your future: Who gambles in defined-contribution pension plans?', *International Review of Financial Analysis* **60**, 213–225.

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

Coval, J. D. and Shumway, T. (2001), 'Expected option returns', *The journal of Finance* **56**(3), 983–1009.

Cremers, M., Halling, M. and Weinbaum, D. (2015), 'Aggregate jump and volatility risk in the cross-section of stock returns', *The Journal of Finance* **70**(2), 577–614.

Cremers, M. and Weinbaum, D. (2010), 'Deviations from put-call parity and stock return predictability', *Journal of Financial and Quantitative Analysis* pp. 335–367.

Davidson, R. and Duclos, J.-Y. (2000), 'Statistical inference for stochastic dominance and for the measurement of poverty and inequality', *Econometrica* **68**(6), 1435–1464.

De Long, J. B., Shleifer, A., Summers, L. H. and Waldmann, R. J. (1990), 'Noise trader risk in financial markets', *Journal of Political Economy* **98**(4), 703–738.

Doran, J. S., Jiang, D. and Peterson, D. R. (2011), 'Gambling preference and the new year effect of assets with lottery features', *Review of Finance* **16**(3), 685–731.

Driessen, J., Lin, T.-C. and Van Hemert, O. (2011), 'How the 52-week high and low affect option-implied volatilities and stock return moments', *Review of Finance* **17**(1), 369–401.

Driessen, J., Lin, T.-C. and Van Hemert, O. (2013), 'How the 52-week high and low affect option-implied volatilities and stock return moments', *Review of Finance* **17**(1), 369–401.

Driessen, J., Maenhout, P. J. and Vilkov, G. (2009), 'The price of correlation risk: Evidence from equity options', *Journal of Finance* **64**(3), 1377–1406.

Dubinsky, A., Johannes, M., Kaeck, A. and Seeger, N. J. (2019), 'Option pricing of earnings announcement risks', *The Review of Financial Studies* **32**(2), 646–687.

Fama, E. F. and 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. and MacBeth, J. D. (1973), 'Risk, return, and equilibrium: Empirical tests', *Journal of Political Economy* **81**(3), 607–636.

Félix, L., Kräussl, R. and Stork, P. (2019), 'Single stock call options as lottery tickets: overpricing and investor sentiment', *Journal of Behavioral Finance* **20**(4), 385–407.

Fong, W. M. (2010), 'A stochastic dominance analysis of yen carry trades', *Journal of Banking & Finance* **34**(6), 1237–1246.

Fong, W. M., Wong, W. K. and Lean, H. H. (2005), 'International momentum strategies: a stochastic dominance approach', *Journal of Financial Markets* **8**(1), 89–109.

Frazzini, A. and Pedersen, L. H. (2012), 'Embedded leverage', *NBER Working Paper Series No. 18558* .

Gao, C., Xing, Y. and Zhang, X. (2018), 'Anticipating uncertainty: Straddles around earnings announcements', *Journal of Financial and Quantitative Analysis* **53**(6), 2587–2617.

Garleanu, N., Pedersen, L. H. and Poteshman, A. M. (2008), 'Demand-based option pricing', *The Review of Financial Studies* **22**(10), 4259–4299.

George, T. J. and Hwang, C.-Y. (2004), 'The 52-week high and momentum investing', *Journal of Finance* **59**(5), 2145–2176.

Goyal, A. and Saretto, A. (2009), 'Cross-section of option returns and volatility', *Journal of Financial Economics* **94**(2), 310–326.

Grinblatt, M. and Han, B. (2005), 'Prospect theory, mental accounting, and momentum', *Journal of Financial Economics* **78**(2), 311–339.

Han, B. (2008), 'Investor sentiment and option prices', *Review of Financial Studies* **21**(1), 387–414.

Hirshleifer, D. (2001), 'Investor psychology and asset pricing', *Journal of Finance* **56**(4), 1533–1597.

Hirshleifer, D. and Teoh, S. H. (2003), 'Limited attention, information disclosure, and financial reporting', *Journal of accounting and economics* **36**(1-3), 337–386.

Hong, H., Lim, T. and Stein, J. C. (2000), 'Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies', *Journal of Finance* **55**(1), 265–295.

Jarrow, R. and Zhao, F. (2006), 'Downside loss aversion and portfolio management', *Management Science* **52**(4), 558–566.

Jegadeesh, N. (1990), 'Evidence of predictable behavior of security returns', *Journal of Finance* **45**(3), 881–898.

Jiang, G., Lee, C. M. and Zhang, Y. (2005), 'Information uncertainty and expected returns', *Review of Accounting Studies* **10**(2-3), 185–221.

Jin, W., Livnat, J. and Zhang, Y. (2012), 'Option prices leading equity prices: Do option traders have an information advantage?', *Journal of Accounting Research* **50**(2), 401–432.

Kahneman, D. and Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* **47**(2), 263–292.

Kelly, B. and Jiang, H. (2014), 'Tail risk and asset prices', *The Review of Financial Studies* **27**(10), 2841–2871.

Kelly, B., Lustig, H. and Van Nieuwerburgh, S. (2016), 'Too-systemic-to-fail: What option markets imply about sector-wide government guarantees', *American Economic Review* **106**(6), 1278–1319.

Kelly, B., Pástor, L. and Veronesi, P. (2016), 'The price of political uncertainty: Theory and evidence from the option market', *Journal of Finance* **71**(5), 2417–2480.

Kumar, A. (2009*a*), 'Hard-to-value stocks, behavioral biases, and informed trading', *Journal of Financial and Quantitative Analysis* **44**(6), 1375–1401.

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

Lakonishok, J., Lee, I., Pearson, N. D. and Poteshman, A. M. (2006), 'Option market activity', *The Review of Financial Studies* **20**(3), 813–857.

Lee, S. S. and Mykland, P. A. (2008), 'Jumps in financial markets: A new nonparametric test and jump dynamics', *The Review of Financial Studies* **21**(6), 2535–2563.

Lemmon, M. and Ni, S. X. (2014), 'Differences in trading and pricing between stock and index options', *Management Science* **60**(8), 1985–2001.

Li, C. W. and Xue, H. (2009), 'A bayesian's bubble', *Journal of Finance* **64**(6), 2665–2701.

Liu, J. and Pan, J. (2003), 'Dynamic derivative strategies', *Journal of Financial Economics* **69**(3), 401–430.

Lu, Z. and Murray, S. (2019), 'Bear beta', *Journal of Financial Economics* **131**(3), 736–760.

Maheu, J. M. and McCurdy, T. H. (2004), 'News arrival, jump dynamics, and volatility components for individual stock returns', *The Journal of Finance* **59**(2), 755–793.

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

Muravyev, D. (2016), 'Order flow and expected option returns', *Journal of Finance* **71**(2), 673–708.

Muravyev, D. and Pearson, N. D. (2020), 'Options trading costs are lower than you think', *The Review of Financial Studies* **33**(11), 4973–5014.

Newey, W. K. and West, K. D. (1987), 'A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica* **55**(3), 703–708.

Pan, J. (2002), 'The jump-risk premia implicit in options: Evidence from an integrated time-series study', *Journal of financial economics* **63**(1), 3–50.

Pástor, L. and Veronesi, P. (2006), 'Was there a nasdaq bubble in the late 1990s?', *Journal of Financial Economics* **81**(1), 61–100.

Patell, J. M. and Wolfson, M. A. (1979), 'Anticipated information releases reflected in call option prices', *Journal of accounting and economics* **1**(2), 117–140.

Patell, J. M. and Wolfson, M. A. (1981), 'The ex ante and ex post price effects of quarterly earnings announcements reflected in option and stock prices', *Journal of Accounting Research* pp. 434–458.

Ruan, X. (2020*a*), 'Volatility-of-volatility and the cross-section of option returns', *Journal of Financial Markets* **48**, 100492.

Ruan, X. (2020*b*), 'Volatility-of-volatility and the cross-section of option returns', *Journal of Financial Markets* **48**, 100492.

Santa-Clara, P. and Yan, S. (2010), 'Crashes, volatility, and the equity premium:

Lessons from s&p 500 options', *The Review of Economics and Statistics* **92**(2), 435–451.

Seyhun, H. N. (1993), 'Can omitted risk factors explain the january effect? a stochastic dominance approach', *Journal of Financial and Quantitative Analysis* **28**(2), 195–212.

Shiller, R. J. (2015), *Irrational exuberance: Revised and expanded third edition*, Princeton university press.

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

Stambaugh, R. F. and Yuan, Y. (2017), 'Mispricing factors', *The Review of Financial Studies* **30**(4), 1270–1315.

Stilger, P. S., Kostakis, A. and Poon, S.-H. (2016), 'What does risk-neutral skewness tell us about future stock returns?', *Management Science* **63**(6), 1814–1834.

Stoline, M. R. and Ury, H. K. (1979), 'Tables of the studentized maximum modulus distribution and an application to multiple comparisons among means', *Technometrics* **21**(1), 87–93.

Todorov, V. (2010), 'Variance risk-premium dynamics: The role of jumps', *The Review of Financial Studies* **23**(1), 345–383.

Tversky, A. and Kahneman, D. (1991), 'Loss aversion in riskless choice: A reference-dependent model', *Quarterly Journal of Economics* **106**(4), 1039–1061.

Van Oordt, M. and Zhou, C. (2016), 'Systematic tail risk', *Journal of Financial and Quantitative Analysis* **51**(2), 685–705.

Wang, H., Yan, J. and Yu, J. (2017), 'Reference-dependent preferences and the risk–return trade-off', *Journal of Financial Economics* **123**(2), 395–414.

Whaley, R. E. and Cheung, J. K. (1982), 'Anticipation of quarterly earnings announcements: a test of option market efficiency', *Journal of Accounting and Economics* **4**(2), 57–83.

Xing, Y., Zhang, X. and Zhao, R. (2010), 'What does the individual option volatility smirk tell us about future equity returns?', *Journal of Financial and Quantitative Analysis* **45**(3), 641–662.

Yu, J. and Yuan, Y. (2011), 'Investor sentiment and the mean–variance relation', *Journal of Financial Economics* **100**(2), 367–381.

Zhang, X. F. (2006*a*), 'Information uncertainty and analyst forecast behavior', *Contemporary Accounting Research* **23**(2), 565–590.

Zhang, X. F. (2006*b*), 'Information uncertainty and stock returns', *Journal of Finance* **61**(1), 105–137.

Appendix A

Appendix to Chapter 2: Mathematical Proof

Proof of Proposition 1: Proof of the uniqueness of decomposition in Eq. (1).

We construct a volatility factor-mimicking portfolio V and a jump risk-factor mimicking portfolio J using a short-maturity straddle $S1$ and a long maturity straddle $S2$:

$$V = (-(Gamma_{S2}/Gamma_{S1})S1, S2);$$

$$J = (S1, -(Vega_{S1}/Vega_{S2})S2),$$

where straddle $S1$ ($S2$) is composed of one unit of the call option $c1$ ($c2$) and $-\Delta_{c1}/\Delta_{p1}$ ($-\Delta_{c2}/\Delta_{p2}$) units of the put option $p1$ ($p2$). Therefore, V and J are investable portfolios of $c1$, $c2$, $p1$ and $p2$. The construction can be expressed by a 4×2 matrix Y, where the first (second) column contains the composition of V (J), and the first, second, third, and fourth row contain the units of $c1$, $p1$, $c2$ and $p2$

respectively:

$$Y = \begin{bmatrix} -Gamma_{S2}/Gamma_{S1} & 1 \\ Gamma_{S2}/Gamma_{S1} \times \Delta_{c1}/\Delta_{p1} & -\Delta_{c1}/\Delta_{p1} \\ 1 & -Vega_{S1}/Vega_{S2} \\ -\Delta_{c2}/\Delta_{p2} & Vega_{S1}/Vega_{S2} \times \Delta_{c2}/\Delta_{p2} \end{bmatrix}$$

We need to solve portfolio weights a and b in decomposition $S1 = (aV, bJ)$. That is to solve a and b from the following equation system:

$$Y \times \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ -\Delta_{c1}/\Delta_{p1} \\ 0 \\ 0 \end{bmatrix}$$

There are four equations but only two of them are linearly independent, therefore the equation system has a unique solution for a and b:

$$a = \frac{\frac{Vega_{S1}}{Vega_{S2}}}{1 - \frac{Vega_{S1}}{Vega_{S2}} \frac{Gamma_{S2}}{Gamma_{S1}}}$$
$$b = \frac{1}{1 - \frac{Vega_{S1}}{Vega_{S2}} \frac{Gamma_{S2}}{Gamma_{S1}}}.$$

Appendix B

Appendix to Chapter 3: Variable Definitions and Davidson and Duclos test

B.1 Variable definitions

This appendix provides definitions of the variables used in this paper.

Maximum daily return (MAX): Bali et al. (2011) show a negative relation between the maximum daily return over the past month (MAX) and stock return in the future. They argue that this is driven by investors' preference for lottery-like assets. Byun and Kim (2016) also document a negative relation between MAX and option returns. Following their studies, we calculate MAX as the maximum daily return within month t.

Idiosyncratic volatility ($IVOL$): We estimate idiosyncratic volatility ($IVOL$) following Ang et al. (2006). We run time-series regressions of excess stock returns on the Fama and French (1973) three-factors (market, SMB and HML factors) using daily stock returns in each quarter, and obtain the residuals from the model. We calculate $IVOL$ for each stock as the standard deviation of the residuals in the last quarter.

Idiosyncratic skewness ($ISKEW$): Previous studies regard stocks with high idiosyncratic skewness ($ISKEW$) as lottery-like (Barberis and Huang, 2008; Kumar, 2009*b*; Mitton and Vorkink, 2007). Mitton and Vorkink (2007) document a negative relation between $ISKEW$ and stock return; Byun and Kim (2016) document a negative relation between $ISKEW$ and option return. Following Byun and Kim (2016), we calculate $ISKEW$ using daily returns over last quarter and fit the following equation:

$$R_{i,d} - R_{f,d} = \alpha_i + \beta_i(R_{m,d} - R_{f,d}) + \gamma_i(R_{m,d} - R_{f,d})^2 + \epsilon_{i,d},$$

where $R_{i,d}$, $R_{m,d}$ and $R_{f,d}$ are the day d returns on stock i, the market portfolio, and the risk-free rate respectively. $ISKEW$ is the skewness of daily residuals $\epsilon_{i,d}$.

Option Skewness ($OSKEW$): Boyer and Vorkink (2014) construct the ex-ante skewness of option returns ($OSKEW$), and find a negative relation between $OSKEW$ and future option returns. This suggests that investors not only overprice lottery-like stocks, but also overprice lottery-like options. We construct ex-ante skewness measure for the option returns following Boyer and Vorkink (2014). Consider a call option written on stock i with the strike price K and time to maturity τ in years:

$$OSKEW = \frac{Z(\mu_i, \sigma_i, M, \tau) - 3Y(\mu_i, \sigma_i, M, \tau)X(\mu_i, \sigma_i, M, \tau) + 2X(\mu_i, \sigma_i, M, \tau)^3}{(Y(\mu_i, \sigma_i, M, \tau) - X(\mu_i, \sigma_i, M, \tau)^2)^{3/2}},$$

where $M = K/S_t$, μ_i and σ_i are the stock log return and volatility respectively. We estimate μ_i and σ_i using daily stock returns during the quarter prior to the portfolio formation date.

Expressions for $X(\mu_i, \sigma_i, M, \tau)$, $Y(\mu_i, \sigma_i, M, \tau)$, and $X(\mu_i, \sigma_i, M, \tau)$ are given as follows:

$$X(\mu_i, \sigma_i, M, \tau) = exp[(\mu_i + \frac{\sigma_i^2}{2})\tau]N(d_1) - MN(d_2)$$

$$Y(\mu_i, \sigma_i, M, \tau) = exp[(2\mu_i + 2\sigma_i^2)\tau]N(d_3) - 2Mexp[(\mu_i + \frac{\sigma_i^2}{2})\tau]N(d_1) + M^2N(d_2)$$

$$Z(\mu_i, \sigma_i, M, \tau) = exp[(3\mu_i + \frac{9}{2}\sigma_i^2)\tau]N(d_4) - 3Mexp[(2\mu_i + 2\sigma_i^2)\tau]N(d_3) \\ + 3M^2exp[(\mu_i + \frac{\sigma_i^2}{2})\tau]N(d_1) - M^3N(d_2),$$

where

$$d_1(\mu_i, \sigma_i, M, \tau) = (-lnM + (\mu_i + \frac{\sigma_i^2}{2})\tau)/\sigma_i\sqrt{\tau}$$

$$d_2(\mu_i, \sigma_i, M, \tau) = d_1(\mu_i, \sigma_i, M, \tau) - \sigma_i\sqrt{\tau}$$

$$d_3(\mu_i, \sigma_i, M, \tau) = d_1(\mu_i, \sigma_i, M, \tau) + \sigma_i\sqrt{\tau}$$

$$d_4(\mu_i, \sigma_i, M, \tau) = d_1(\mu_i, \sigma_i, M, \tau) + 2\sigma_i\sqrt{\tau}.$$

$N(x)$ denotes the probability that a variable with a standard normal distribution will be less than x.

Short-term reversal (REV): Following Jegadeesh (1990), the reversal for each stock is defined as the return in month t.

Momentum (MOM): Following Jegadeesh (1990), we caculate the momentum variable for each stock in month t as the cumulative return on the stock over the previous 11 months starting from month $t-11$ to month $t-1$.

Firm size ($SIZE$): We calculate firm's size as the firm's market capitalization at the end of month t (stock price multiplied by the number of shares outstanding). We take logarithm of firm's size, log($SIZE$), and use it in our regressions.

Beta ($BETA$): We estimate firm's beta as the slope coefficient from the time-series regressions of excess stock returns on the excess market returns using three years of monthly stock returns including month t. We require a minimum of 12 non-missing monthly observations.

Book-to-market ratio (BTM): We calculate firm's book-to-market ratio as the ratio of book value of common equity to the market capitalization of the equity at the end of the last quarter.

Turnover ($TURNOVER$): We compute the monthly turnover of each stock by averaging daily turnover within month t. Daily turnover is calculated as the ratio of the number of shares traded to the number of shares outstanding (Bali et al., 2016).

Illiquidity ($ILLIQUIDITY$): Following Amihud (2002), we calculate the stock illiquidity in month t as the ratio of the absolute monthly stock return to its dollar trading volume.

Change in call implied volatility (CIV): Following An et al. (2014), we calculate call implied volatility change as the difference between short-term ATM calls' average implied volatilities over month t and month $t-1$. ATM calls have moneyness ranging from 0.975 to 1.025 and maturity between 10 to 60 days.

Change in put implied volatility (PIV): Following An et al. (2014), we calculate put implied volatility change as the difference between short-term ATM puts' average implied volatilities over month t and month $t-1$. ATM puts have moneyness ranging from 0.975 to 1.025 and maturity between 10 to 60 days.

Variance risk premium (VRP): We estimate variance risk premium (VRP) as the difference between ATM calls' and puts' implied volatilities IV and monthly stock volatility. We use 30-day realized volatility in month t as a proxy for monthly stock volatility following Cao and Han (2013).

Volatility-of-volatility (VOV): We calculate volatility of volatility (VOV) following Baltussen et al. (2018). We first calculate the average of ATM calls' and puts' implied volatilities (IV) with moneyness ranging from 0.975 to 1.025 and maturity between 7 to 40 days on each day. We then calculate the standard deviation of the implied volatilities within month t, and scale it by the average implied volatility over the month t.

Option bid-ask spread (PBA): Following Cao and Han (2013), we calculate option bid-ask spread as the ratio of bid-ask spread of option quotes over the midpoint of bid and ask quotes at the end of month t.

Analysts' coverage ($ANALYSTS$): We measure analyst coverage ($ANALYSTS$) as the number of analysts following the firm in month t.

Analysts forecast dispersion ($DISP$): We calculate forecast dispersion ($DISP$) as the standard deviation of analysts forecasts scaled by the mean analysts' forecast in month t.

B.2 Davidson and Duclos (2000) test

This Appendix provides the construction of Davidson and Duclos (2000) non-parametric second stochastic dominance (DD) statistics.

Let y and z be two risky assets, both with N observations, with common support of $[a, b]$, and with the corresponding cumulative distribution functions to be F_y and F_z respectively. Let $D_i^2(x) = \int_a^x F_i(s)ds$, for $i = y, z$.

y is said to dominate z stochastically at second-order if $D_y^2(x_i) \leq D_z^2(x_i)$ for all i, with strictly significant inequality for some i. The DD statistic to test the null hypothesis of the equality of $D_y^2(x_i) = D_z^2(x_i)$ is:

$$T(x) = \frac{\hat{D}_y^2(x) - \hat{D}_z^2(x)}{\sqrt{\hat{V}^2(x)}},$$

where

$$\hat{V}^2(x) = \hat{V}_y^2(x) + \hat{V}_z^2(x) - 2\hat{V}_{y,z}^2(x)$$
$$\hat{D}_y^2(x) = \frac{1}{N}\sum_{i=1}^{N}(x - y_i)_+$$
$$\hat{D}_z^2(x) = \frac{1}{N}\sum_{i=1}^{N}(x - z_i)_+$$
$$\hat{V}_y^2(x) = \frac{1}{N}[\frac{1}{N}\sum_{i=1}^{N}(x - y_i)_+^2 - \hat{D}_y^2(x)^2]$$
$$\hat{V}_z^2(x) = \frac{1}{N}[\frac{1}{N}\sum_{i=1}^{N}(x - z_i)_+^2 - \hat{D}_z^2(x)^2]$$
$$\hat{V}_{y,z}^2(x) = \frac{1}{N}[\frac{1}{N}\sum_{i=1}^{N}(x - y_i)_+(x - z_i)_+ - \hat{D}_y^2(x)\hat{D}_z^2(x)]$$

To test for stochastic dominance, we select a pre-designed finite number of x following Fong et al. (2005). We first partition the data equally into 10 major grids. Then we partition each major grid into 10 equal sub-intervals. For the obtained fixed values of x_1, x_2, ..., x_m and their corresponding statistics $T(x_i)$ for $i = 1, 2, ..., m$, the following hypotheses are investigated:

$H_0 : D_y^2(x_i) = D_z^2(x_i)$ for all i;
$H_A : D_y^2(x_i) \neq D_z^2(x_i)$ for some i;
$H_{A1} : D_y^2(x_i) \geq D_z^2(x_i)$ for all i; $D_y^2(x_i) > D_z^2(x_i)$ for some i;
$H_{A2} : D_y^2(x_i) \leq D_z^2(x_i)$ for all i; $D_y^2(x_i) < D_z^2(x_i)$ for some i;

For risk averters, the null hypothesis is rejected in favor of the alternative hypothesis that y (low $CallMoney$ group) dominates z (high $CallMoney$ group) if none of the DD statistics is significantly positive and at least some of the DD statistics are significantly negative.

Statistical inference is based on the student maximum modulus (SMM) distribution for $K = 10$ and infinite degrees of freedom. The 5% asymptotic critical value of the SMM distribution is 3.254 from Stoline and Ury (1979).