Media Coverage, Volatility Overreaction, and Option Returns *

Lihai Yang†

April 2025

Abstract

This paper documents strong media coverage effect in option market. Firms with higher media coverage tend to have lower straddle returns in the next month, which is not explained by other option predictors in literature. I provide supporting evidence for the volatility overreaction in option market: higher media coverage is associated with both higher future realized stock variance and higher current implied variance, while the implied variance responds too much to be reconciled by future realized variance. Alternative explanations, including the investor recognition hypothesis in Fang and Peress (2009), extreme past stock returns, and earnings announcement, cannot fully explain the media coverage effect in option market. Detailed analysis of news topics reveals the five most important topics that drive the media coverage effect, including earnings, revenues, equity actions, acquisition and mergers, and stock prices.

Keywords: media coverage, volatility overreaction, option returns

*I'm grateful for comments from Yakov Amihud, Shuaiqi Li, Junbo Wang, and Di Wu. I thank Shuaiqi Li for providing the option straddle return data.

†City University of Hong Kong, lihaiyang2-c@my.cityu.edu.hk, Hong Kong Kowloon Tong, 852-51783844

1 Introduction

Media report provides information and attracts investor attention, which contributes to investor trading and increases stock volatility (Peress (2014)). As information is incorporated in prices and attention fades, volatility usually declines afterwards. In this paper, I study whether option market investors fully recognize this volatility pattern. If option investors overreact to the increasing stock volatility induced by media report, the option implied volatility will be too high to be reconciled by future realized stock volatility. This would lead to lower returns on volatility assets such as option straddles.

Using RavenPack data from 2000 to 2020, I find that media coverage negatively predicts future option straddle returns, consistent with the volatility overreaction in option market. I measure media coverage at firm level by the number of news articles mentioning the firm in past one month and purge out the main covariates that are known to affect media coverage. In particular, I follow Hillert, Jacobs, and Müller (2014) to obtain residual media coverage by running the month-by-month cross-sectional regression of media coverage on firm size, NASDAQ dummy, S&P 500 dummy, and analyst coverage. On the other hand, I compute monthly option returns using daily-delta-hedged at-the-money (ATM) straddles. By delta-hedging at daily frequency, straddles are insensitive to underlying stock price movements during the holding period and have returns that can proxy for variance shocks (defined as future realized variance divided by implied variance minus 1) much better than straddles without delta-hedging (Li and Yang (2024)). Therefore, delta-hedged straddle returns have a clean economic interpretation as volatility information, which suits this study well.

I find that firms with higher residual media coverage in the past one month tend to have lower future straddle returns. A long-short strategy based on residual media coverage generates an average monthly profit of 5.04% with a high T-statistic of 15.84 and a monthly Sharpe ratio of 1.173. The magnitude is larger than most well-known option return predictors in literature. The strategy is profitable in every 5-year subsample, and has limited crash risk. The predictive ability for the raw media coverage without purging out the main covariates is weaker, but still highly significant and economically large.

Media coverage captures distinct information from option return predictors in previous literature. The alphas of media-based strategy are highly significant and economically large when adjusting for option factor models in literature. It also remains significant in Fama-Macbeth regression after we control for other common option return predictors, including firm size, option momentum in Heston, Jones, Khorram, Li, and Mo (2023), the idiosyncratic volatility in Cao and Han (2013), implied volatility term spread in Vasquez (2017), implied volatility smirk, the volatility spread in Goyal and Saretto (2009), and the volatility of volatility in Cao, Vasquez, Xiao, and Zhan (2023).

I conduct several robustness tests. First, I show that the media coverage effect exists not only for the delta-hedged straddle return in main analysis, but also for the static straddle return that is held to maturity without delta hedging. I also examine the predictive ability of media coverage for the return on ATM put and call options separately and find similarly strong result, suggesting that media coverage does not contain specific directional information. Second, I experiment with different filters about the RavenPack data, including different relevance score restrictions and news type filters. The results are quite similar to the baseline. Finally, I vary the way of calculating residual media coverage by changing the set of explanatory variables for media coverage and find robust results for these alternative specifications. Interestingly, simply purging out the most important covariate that influences media coverage, i.e., the firm size, produces similarly strong results with the baseline specification of residual media coverage.

To examine the impact of option bid-ask spread, I consider different proportions of quoted spread following Muravyev and Pearson (2020). After reasonable transaction cost management, I show that the media-based option strategy remains profitable when traders pay as high as 51.6% of the quoted option bid-ask spread.

To explore the economic source of media coverage effect, I decompose the predictive ability of residual media coverage into two components related to future realized variance and current implied variance. This decomposition is motivated by the observation that delta-hedged straddle return is related to the variance swap return, which captures the difference between future realized variance and current option-implied variance. I find that residual media coverage positively predicts next-month realized stock variance and also positively correlates with the implied variance in the current month. However, the coefficient for the implied variance case is much larger than that for the realized variance case, suggesting that option investors overreact to the media report in the contemporaneous month. The overreaction in option market contributes to the overpriced straddles, which leads to lower future returns and explains the media coverage effect. In addition, I show that the overreaction of option investors to media reports is likely driven by their overestimation about the persistence of media-induced high realized variance.

I also explore other alternative explanations for media coverage effect. Fang and Peress (2009) propose the "Investor Recognition" hypothesis to explain the media coverage effect in stock market, which states that firms with less media coverage have a smaller investor base and therefore should offer higher expected return to compensate its holder for being imperfectly diversified (Merton (1987)). Using various proxies for investor recognition, I do find stronger media coverage effect in the subsample with low investor recognition, consistent with Fang and Peress (2009). However, the media-based strategy remains highly profitable even in the subsample with high investor recognition, suggesting that "Investor Recognition" cannot fully explain the media coverage effect in option market.

Other alternative explanations are related to the endogeneity of media coverage: it is likely that we omit important controls that drive both firm-level media reports and option returns. A potential variable is extreme stock return. News media usually reports firms with extreme past stock returns, which triggers investor attention and leads to overpricing in the option market (Choy and Wei (2023)). Another important variable is the earnings announcement. As perhaps the most important firm event, earnings announcement attracts much attention and is heavily reported by news media. It is also well documented in literature that the uncertainty elevated before earnings announcement while quickly declines afterwards, contributing to the lower option return after earnings announcement (Chung and Louis (2017)). Therefore, high-media-coverage firms are likely those firms with recent earnings announcement, whose volatility declines contribute to lower future option returns.

I carefully design tests to show that the media coverage effect in option market is not simply driven by the extreme stock return and earnings announcement. First, among the subsample of firms without extreme past stock returns and firms without recent earnings announcement, I still find strong negative relation between media coverage and future option returns. Second, the media coverage remains highly significant after controlling for the dummy variables about extreme stock return and earnings announcement in Fama-Macbeth regression. These suggest that media coverage is a much broader concept that is not restricted to reports about extreme return and earnings announcement.

To get a sense of what news topics covered by media reports contribute most to the media coverage effect, I utilize the news topic classification by RavenPack to construct residual media coverage from news articles belonging to different news topics and compare their predictive ability. The top 10 topics selected by mentioning frequency account for about 86% of all news articles in my sample period. Detailed comparison of media coverage of these 10 topics shows that the media coverage effect is mainly driven by the coverage of five key topics, including earnings, revenues, equity actions (such as trading status change, buybacks, public offering, ownership changes, etc.), acquisition and mergers, and stock prices (such as the stock price changes that make news headlines).

To further address the endogeneity concern of media coverage, I follow Dai, Parwada, and

Zhang (2015) and Dai, Shen, and Zhang (2021) to use geographic proximity to news outlets (Dow Jones' offices) as instrument variable for media coverage. The logic for this instrument is that (1) firms located near new outlets are likely to have higher media coverage since it is costly for journalists to cover distant firms, and (2) geographic proximity is unlikely to directly affect straddle returns. A two-stage instrumental regression confirms the negative effect of media coverage on straddle return, suggesting that the media coverage effect is likely causal.

I also provide some further results with respect to media sentiment. First, I decompose the media coverage into positive media coverage and negative media coverage by news sentiment and find that media coverage of positive news contributes more to the media coverage effect. Second, I directly compare the predictive ability of media sentiment and media coverage, and find that media coverage exhibits much stronger predictive ability for straddle returns than media sentiment. Finally, I examine the persistence of media coverage effect by constructing residual media coverage during different past windows. As the past window increases, the predictive ability of media coverage quickly drops compared to the baseline case of past one month, suggesting the short-term predictive ability of media coverage in option market.

This paper contributes to several strands of literature. Firstly, it adds to the recent literature using information from news media to predict option returns. Filippou and Garcia-Ares (2020) find that firms with high media pessimism have lower future delta-hedged call option returns than firms with low media pessimism. Cao, Han, Li, Yang, and Zhan (2022) apply machine learning to news media content and extract textual signals related to option returns. Different from these two papers, I study the information contained in media coverage, without specific focus on media sentiment or textual content. Studying media coverage allows us to have a better cross-sectional firm coverage, including both firms with and without media coverage. In contrast, the media sentiment study in Filippou and Garcia-Ares (2020) only focuses on firms with negative media sentiment, and the extraction of textual information in

Cao, Han, Li, Yang, and Zhan (2022) requires firms to have a reasonable amount of media coverage, which reduces the sample size. Moreover, I take the endogeneity concern of media coverage seriously and perform an additional instrumental analysis. This is missing in Filippou and Garcia-Ares (2020) and Cao, Han, Li, Yang, and Zhan (2022) possibly because it is difficult to find instrument for media sentiment or media textual content. Furthermore, I show that media sentiment is not a robust predictor for option returns and media coverage exhibits much stronger predictive ability.

More broadly, this paper relates to the literature about media coverage and financial market. Fang and Peress (2009) find that stocks with high media coverage earn significantly lower returns than those without media coverage. Gao, Wang, Wang, Wu, and Dong (2020) extend it to the corporate bond market and show that firms with high media coverage have lower bond offering yield. Hillert, Jacobs, and Müller (2014) use media coverage to explain momentum and find that stocks with high media coverage exhibit stronger momentum and subsequent reversal, suggesting that media coverage exacerbates investor bias. Tetlock (2007) and Tetlock, Saar-Tsechansky, and Macskassy (2008) show that the sentiment and tones of media reports predict stock returns. To the best of my knowledge, this paper is the first to study media coverage effect in option market. My findings in option market echo the stock market results in Fang and Peress (2009), but they differ on the economic interpretation: I show that the "Investor Recognition" hypothesis proposed by Fang and Peress (2009) cannot fully explain the media coverage effect in option market; I propose and provide evidence for the volatility overreaction explanation specific to option market. In addition, I provide richer results missing in Fang and Peress (2009), such as media coverage of different news topics, instrumental variable analysis, and comparison with media sentiment.

Finally, my paper contributes to the growing literature about option return predictability by proposing a new predictor. Existing option return predictors are based on stock characteristics (Zhan, Han, Cao, and Tong (2022); Cao and Han (2013)), option information (Goyal and Saretto (2009); Vasquez (2017); Heston, Jones, Khorram, Li, and Mo (2022, 2023); Li and Yang (2024)), or machine learning (Bali, Beckmeyer, Moerke, and Weigert (2023)). This paper is the first to use media coverage to predict cross-sectional option returns.

Section 2 describes the data and variables used in this paper. Section 3 documents the predictive ability of media coverage for option straddle returns and checks its robustness. Section 4 explores the economic explanations for the media coverage effect in option market. Section 5 conducts an instrumental variable analysis for the media coverage effect. Section 6 provides some further results. Section 7 concludes.

2 Data and Variables

I obtain option prices from the OptionMetrics Ivy DB database, which provides daily closing bid and ask quotes for the U.S. equity options. I use the RavenPack database to obtain information about media coverage. Information about stock prices is extracted from CRSP.

2.1 Option Returns

I compute the return on at-the-money (ATM) straddle comprised of one put and one call with the same strike price that is closest to the current underlying stock price for common equities. The sample period starts from January 2000, due to the availability of RavenPack data, and ends at December 2020. I form the straddle on the third Friday of each month and hold it to expiration, which is typically the third Friday of the following month. Thus, the straddle returns are at monthly frequency. To maintain delta-neutral throughout the holding period, I hedge the straddle with the deltas given by OptionMetrics at daily frequency and reinvest the daily delta-hedged payoff until expiration.[1] Then I compute excess return by

[1] In some cases, the Black-Scholes delta for an option is missing in the Optionmetrics. When this occurs, I estimate a delta using the stock price on that day and the most recent non-missing implied volatility for the same option.

subtracting off the risk-free rate obtained from Kenneth French's website. All returns in this paper are excess returns, but I just call them 'returns' for brevity. Straddle return is measured as following:

$$\frac{Straddle(T) - \sum_{n=0}^{N-1} \triangle_{t_n}[S_{t_{n+1}} - S_{t_n} e^{r_f(t_{n+1}-t_n)}] e^{r_f(T-t_{n+1})}}{Straddle(t)} - R_{f,t\to T},$$

where $Straddle(T)$ is the final payoff of straddle at expiration T; $Straddle(t)$ is the initial value of straddle at formation day t; Delta-hedged positions are rebalanced at each of the business days $t_n, n = 0, 1, ..., N-1$, with $t_0 = t, t_N = T$; $\triangle_{t_n}$ is the delta of the straddle at day t_n; S_{t_n} is the stock price at day t_n; r_f is the daily log risk-free rate; $R_{f,t\to T}$ is the risk-free rate from day t to T.

To ensure the tradability of straddles, I apply the following filters: Firstly, I remove options with 0 open interest or bid price; Secondly, I delete options whose ask prices are lower than bid prices, and those with prices violating arbitrage bounds; Thirdly, I discard options with missing implied volatilities or deltas; Lastly, I exclude options whose underlying stock prices are lower than 5 at the formation day, and those with stock-splits during the holding period. Option contracts passing the above filters tend to be actively traded and have reliable returns.

2.2 Media Coverage

I construct measures of media coverage using data from RavenPack, which collects and analyzes real-time, firm-level business news from leading news providers, including Dow Jones Newswires, the Wall Street Journal (all editions), Barron's, etc. As RavenPack data is only available after 2000, the sample period is from January 2000 to December 2020. For each news article, I obtain from RavenPack its publishing date, news type, the firms mentioned in the article, and the relevance score about the mentioned firms ranging from 0 to 100. The news type includes ‘FULL-ARTICLE’, ‘HOT-NEWS-FLASH’, ‘NEWS-FLASH’, ‘PRESS-RELEASE’, and ‘TABULAR-MATERIAL’. To make sure the relevance of firm mentioning, I apply the following filters: (1) relevance score to be at least 75 and (2) news type does not equal ‘TABULAR-MATERIAL’.

On the third Friday of each month (option date), for each firm, I count the number of news articles mentioning the firm during the past 1 month ending at option date. I do not count the news articles published at the most recent option date to avoid using information after market close. *Media Coverage* is defined as the natural log of one plus the number of news articles. For firms without any news article mentioning, their media coverage measure is set to 0.

I also construct the residual media coverage to account for the fact that media coverage is correlated with various firm characteristics. An obvious covariant is firm size, with the intuition that large firms tend to have higher media coverage. Table 2 explores the determinates of media coverage, following the specifications in Hillert, Jacobs, and Müller (2014). Consistent with Hillert, Jacobs, and Müller (2014), firm size alone explains a large proportion of media coverage, with the average adjusted R square reaching 26.8%. Adding NASDAQ and S&P 500 dummies and analyst coverage slightly increases the explanatory power. The explanatory power increases further when adding other firm characteristics and industry dummies, but for parsimonious purpose, we follow Hillert, Jacobs, and Müller (2014) to use specification 3 in Table 2 to calculate residual media coverage. Focusing on the four well-known determinates (firm size, the NASDAQ dummy, the S&P 500 dummy, and the analyst coverage) of media coverage in literature helps to alleviate concerns about data snooping. In addition, adding more firm characteristics as in specification 4 and 5 reduces the cross-sectional firm number.

Specifically, I obtain the *Residual Media Coverage* for each firm at each month by running the following month-by-month cross-sectional regression of media coverage on the firm size, the NASDAQ dummy, the S&P 500 dummy, and the analyst coverage

$$\begin{aligned}\ln(1 + \textit{number of news articles}) =& \alpha + \beta_1 \ln(\textit{Market Cap}) + \beta_2 NASDAQ + \beta_3 S\&P500 + \\ & \beta_4 \ln(1 + \textit{number of analyst coverage}) + \epsilon_{media}\end{aligned} \tag{1}$$

where *Market Cap* is firm's stock market capitalization, $NASDAQ$ is the indicator for NASDAQ listing, $S\&P500$ is the indicator for S&P 500 membership, and *analyst coverage* is the number of analysts who issue at least one FY1 or FY2 earnings forecast for the stock during past 12 months. All the variables in Equation (1) are measured at the option date of each month, and since the regression is run at each cross-section, there is no look-ahead bias for *Residual Media Coverage*. In the robustness tests, I also examine alternative controls in the construction of residual media coverage.

2.3 Control Variables

The choice of control variables largely follows Heston, Jones, Khorram, Li, and Mo (2023): I use the log market capitalization on the third Friday; Following Cao and Han (2013), I calculate idiosyncratic volatility as the standard deviation (annualized) of the residuals from a 22-day rolling regression using Fama-French 3 factors; To measure the slope of implied volatility term structure (IV term spread) in Vasquez (2017), I take the difference between the 60-day and 30-day ATM implied volatilities extracted from the Volatility Surface database; I measure the smirk of the implied volatility curve (IV smirk) as the difference between the implied volatility of a 30-day call with delta of 0.3 and that of a 30-day put with delta of -0.3; To get the volatility difference in Goyal and Saretto (2009) (IV-HV), I subtract the log of ATM straddle implied volatility by the rolling one-year historical volatility (in log) computed from daily stock returns. In addition to above option return predictors, I also control for the volatility of volatility (VOV) in Cao, Vasquez, Xiao, and Zhan (2023) and the option momentum in Heston, Jones, Khorram, Li, and Mo (2023). Specifically, VOV is the standard deviation of the daily percentage changes of 30-day at-the-money implied volatility over the past 22 trading days, and option momentum is the past 1 month straddle return. When calculating the option momentum, I do not require straddles to have positive open interest and bid price to increase the cross-sectional coverage, as in Heston, Jones, Khorram, Li, and Mo (2023) and Heston, Jones, Khorram, Li, and Mo (2022).

Table 1 shows the summary statistics. We have 388911 observations in total over 252 months from January 2000 to December 2020, with 1543 firms each month on average. The average firm is mentioned in 24 news articles every month. Figure 1 compares the time-series of total firm number and the firm number without media coverage in our sample. As shown, the no-media-coverage firms account for around one-third of total firms at the beginning of the sample, but this proportion declines very quickly to an almost negligible level.

3 Media Coverage and Option Returns

In this section, I present the main findings about media coverage effect in option market. I first establish the effect with univariate portfolio sort and Fama-Macbeth regression. Then, I examine the robustness of media coverage effect, the after-cost profit of media-based strategy, and how long the effect lasts. Finally, I investigate whether the effect is driven by positive media coverage or negative media coverage.

3.1 The Media Coverage Effect

To begin with, I perform univariate portfolio sorts to see the relation between media coverage and future straddle returns. Specifically, at the third Friday of each month, I sort firms into five groups by their *Residual Media Coverage* or *Media Coverage*, and then report the mean straddle returns for each portfolio over the next period and the corresponding t-statistics. I form equally weighted quintile portfolios for the continuous variable of *Residual Media Coverage*, while for the discrete variable of *Media Coverage*, I follow Fang and Peress (2009) to separate out the no media coverage group and then divide the rest into four groups by their *Media Coverage*. Throughout the paper, I adjust standard errors according to Newey and West (1987) with 5 lags.[2] Portfolios are rebalanced every month.

Table 3 shows the result. In both panels, the portfolio return monotonically declines as media coverage increases. The predictability based on *Residual Media Coverage* is quite strong, both in terms of return spread and statistical significance: the high-minus-low return sorted by *Residual Media Coverage* is -5.04% per month on average with a t-statistic of -15.84. The negative effect of media coverage is somewhat weaker when sorting by the raw *Media Coverage*, but still quite significant. Firms with high media coverage have significantly lower option straddle returns over the next period than those with no media coverage, with an average high-minus-no return of -3.18% and a t-statistic of -6.68. The weaker result of *Media Coverage* compared to *Residual Media Coverage* is likely because the raw media coverage is correlated with some covariates that pollute its predictive power, which is purged out in residual media coverage.

In order to assess the stability of media coverage effect, I examine 5-year moving averages of return spread sorted by media coverage. Figure 2 shows that the moving average returns based on *Residual Media Coverage* are consistently negative during the whole sample period. For the case of *Media Coverage*, the moving average returns are almost always significantly negative, although its profitability decreases in recent periods.

To better evaluate the economic magnitude of media-based option strategies, I compare its investment performance with existing option return predictors in literature. We adjust

[2] The number of lags is calculated according to the formula of $4 \times (T/100)^{2/9}$ with $T = 252$, following the suggestion in Bali, Engle, and Murray (2016).

the sign of all investment signals such that their long-short spreads are positive to facilitate comparison. Table 4 presents the investment performance evaluation. The most competitive media-based option strategy is based on *Residual Media Coverage* with an average monthly return of 5.04% and a monthly Sharpe ratio of 1.173. This Sharpe ratio is slightly lower than that of the strategy based on IV-HV (1.181) but much higher than any other strategies. In addition, the strategy formed by *Residual Media Coverage* has a maximum drawdown (the largest fraction by which the cumulative value of a portfolio falls below its prior maximum) of only 10.36%, which is the lowest among all strategies, suggesting limited crash risk. For the strategy based on *Media Coverage*, its Sharpe ratio is 0.400, which, although weaker than the *Residual Media Coverage*, is still higher than those strategies formed by firm size, idiosyncratic volatility, and IV smirk.

In Table 5, I further investigate the alphas of media-based option strategies relative to common option factors. Same as Table 4, the strategy buys straddles with low or no media coverage while selling straddles with high media coverage to make its profit positive. The choice of option factors follows Heston, Jones, Khorram, Li, and Mo (2023): (1) the 4-factor model: the long-short returns sorted by market capitalization, idiosyncratic volatility and IV – HV, and the excess straddle return of S&P 500 index; (2) the 7-factor model from the 4 factors plus the long-short returns sorted by IV term spread and IV smirk, and the cross-sectional equally weighted average excess straddle return. We also form a 8-factor model by adding the long-short returns sorted by VOV to the 7-factor model. The strategy alphas for both *Residual Media Coverage* and *Media Coverage* are positive and highly significant, suggesting that the media coverage effect in option market is not explained by other common option factors in literature. An important observation is that the strategy based on *Media Coverage* has significantly negative loadings on the option factor related to firm size, which is almost not present for the strategy based on *Residual Media Coverage*. This negative loading amplifies both the magnitude and the significance level of the strategy alpha compared to those in Panel B of Table 3. This again suggests that purging out important covariates such as firms size from the raw *Media Coverage* is important.

In order to control for other option return predictors simultaneously, I run Fama and MacBeth (1973) regressions and use *Residual Media Coverage* and *Media Coverage* to predict straddle returns in the next month. Table 6 presents the results. The highly significant negative coefficients on *Residual Media Coverage* in columns (1) and (2), and *Media Coverage* in columns (3) and (4) further confirm the existence of media effect in option market. Compared to column (1), the addition of other option predictors in column (2) barely changes the coefficient estimate and t-statistic on *Residual Media Coverage*. This suggests that *Residual Media Coverage* captures distinct information from other predictors. In contrast, the coefficient on *Media Coverage* becomes more negatively significant after adding control variables in column (4), especially the firm size. This again suggests that purging out the main covariates that explain media coverage is important to boost its predictive ability. As an additional test, in Internet Appendix A1, I also perform double sorts and show that the media coverage effect is not explained by other option predictors one by one.

In short, the result in this section suggests the strong negative predictive ability of media coverage for future straddle returns and the superior performance of *Residual Media Coverage* relative to *Media Coverage*. Therefore, in the following analysis, I focus on *Residual Media Coverage*.

3.2 Robustness

Table 7 conducts robustness tests. In Panel A, I examine alternative option returns. Instead of using the dynamic straddle return that is daily delta-hedged until expiration in the main analysis, I look at (1) 'static straddle return', which is the return on a static position of monthly held-to-maturity at-the-money straddle without any delta-hedge, (2) 'dynamic call return', which is the return on at-the-money call option daily-delta-hedged until expiration, and (3) 'dynamic put return', which is the return on at-the-money put option daily-delta-hedged until expiration. The result shows that *Residual Media Coverage* still generates significantly negative high-minus-low spread for static straddle returns, although the result is weaker compared to the dynamic straddle returns. For dynamic put and dynamic call returns, we also see the strong negative predictive ability of *Residual Media Coverage*. In addition, the economic magnitude of the high-minus-low return spread and its significance level are quite similar for put and call, suggesting that *Residual Media Coverage* does not contain specific directional information.

In Panel B, I consider varying the threshold of news relevance score. RavenPack provides the relevance score ranging from 0 to 100 to indicate the relevance of firms mentioned in the article, with 100 as the most relevant. In the baseline analysis, I follow the instruction of RavenPack to use 75 as the threshold. Here, I experiment with thresholds of 80, 85, and 90. We see that both the economic magnitude and the significance level for the high-minus-low return spread are quite close to the baseline result.

Panel C considers different filters about news type. The news type in RavenPack includes the followings: 'FULL-ARTICLE', 'PRESS-RELEASE', 'NEWS-FLASH', 'HOT-NEWS-FLASH', and 'TABULAR-MATERIAL'. In the baseline analysis, I exclude the type about 'TABULAR-MATERIAL', which may simply list the firm in the tables without specific meaning. Here, I consider including it back. I also consider excluding type about 'PRESS-RELEASE', which is mainly about firm announcement. The result shows that these different news type filters barely affect the effect.

Panel D investigates the robustness in the regression specification to get residual media coverage. In the main analysis, I obtain the residual media coverage from the cross-sectional regression of media coverage on firm size, NASDAQ dummy, S&P 500 dummy, and analyst coverage as in Equation (1). Here, I vary the set of right-hand side explanatory variables from the minimum of only firm size to the maximum of all controls plus industry dummies in the specification 5 of Table 2. We see that controlling for firm size only is enough to generate strong predictive ability comparable to the baseline result, consistent with the result in Table 2 that firm size alone captures much of the variation in media coverage. As we increase the control variable list, the effect gets stronger first and then declines, but the results of all these specifications are strongly significant. This suggests that the main conclusion is robust to the regression specification to get residual media coverage, addressing the concern about data mining.

3.3 Adjusting for Transaction Costs

I now examine if the media-based option strategy is still profitable after accounting for transaction cost. So far, the results assume that options can be bought and sold at the midpoint of bid and ask quotes. To evaluate how bid-ask spreads can affect the profits, I consider costs that are a fixed fraction of the quoted half-spread. The choice of fractions follows Heston, Jones, Khorram, Li, and Mo (2023), which is based on the results in Muravyev and Pearson (2020): 'Algo' traders pay 20.3% of the half-spread; 'Adjusted' and 'Effective' are 51.6% and 75.8%, respectively, of the half-spread; 'Quoted' refers to the case in which traders pay full quoted spread and buy options at ask and sell at bid. As an attempt to reduce the impact of bid-ask spreads, I also consider long-short strategies with more extreme deciles, instead of quintiles as in previous analyses, and with low-cost sample consisting of straddles whose percentage bid-ask spreads are in the lowest quartile each month after the portfolio sorting.

Table 8 presents the average monthly returns of the residual-media-coverage-based long-short strategy under different spread ratios. The strategy is still highly profitable under the "Algo" case. Even for the "Adjusted" case where traders pay 51.6% of quoted spread, the strategy still yields significant profit after using low-cost sample and deciles sorting. Therefore, transaction cost does not fully render the media-based option strategy unprofitable.

4 Economic Source

In this section, I investigate the mechanism behind the media coverage effect in option market. I first propose and provide supporting evidence for a channel about volatility overreaction of option investors to the media report. Then I explore other alternative explanations and show that they cannot fully explain the media coverage effect: I show that the traditional "Investor Recognition" hypothesis proposed by Fang and Peress (2009) is unlikely to fully explain the media coverage effect in option market; I also show that the media coverage effect in option market is not driven by extreme stock returns or earnings announcement effect. At last, I delve into the different news topics of media coverage to get a sense of what news topics drive the media coverage effect.

4.1 Volatility Overreaction

Since delta-hedged straddle return is highly correlated with variance swap return (Li and Yang (2024)), which measures the difference between future realized variance (RV) and current option-implied variance (IV), the fact that *Residual Media Coverage* predicts future straddle returns suggests that *Residual Media Coverage* must predict future RV in a different way from how it affects IV. To this end, I decompose the predictive ability of *Residual Media Coverage* into two components related to RV and IV separately. I first use *Residual Media Coverage* to predict future RV in the cross section, where RV is calculated as the sum of daily squared underlying stock returns between option expiration dates from month to month. To check how option market responds to media reports, I regress IV on *Residual Media Coverage*, where IV is the average implied volatility of the put and call in the straddle.

To further examine whether option market overreacts or underreacts to the media reports, I use *Residual Media Coverage* to predict the log difference between RV and IV, which is equivalent to the log variance swap return.

Table 9 presents the result. Column (1) shows that *Residual Media Coverage* positively predicts next-period realized variance, with a significant coefficient of 0.036. When regressing the implied variance on *Residual Media Coverage* in column (2), we obtain a coefficient of 0.071, which is highly significant. So option market responds to media report, but the fact that the predictive coefficient in column (2) is much larger than that in column (1) suggests that option market overreacts to the media report in the contemporaneous month. This is further confirmed when we predict the log variance swap return in column (3), where the coefficient on *Residual Media Coverage* is strongly negative.

But why do option investors overreact to media reports? I propose that it is because they overestimate the persistence of the media-induced high realized variance. To support this argument, I regress the realized variance over the contemporaneous period on the *Residual Media Coverage*. Column (4) shows that high media coverage is associated with high realized stock variance over the contemporaneous period, with a highly significant coefficient of 0.226. This coefficient is much higher than the coefficient for the next-month realized variance in column (1), suggesting the declines of realized variance over the next period after reaching the peak during high media coverage period. This is confirmed in column (5), where the *Residual Media Coverage* significantly and negatively predicts the changes of realized variance. More importantly, the coefficient of 0.226 is much higher than the coefficient for implied variance in column (2), suggesting that option investors do realize the volatility decline pattern following high media coverage. However, the high coefficient in column (2) relative to column (1) suggests that they still overestimate the persistence of the media-induced high realized variance.

In sum, I provide evidence for the volatility overreaction channel for the media coverage effect in option market: high media coverage drives high realized stock variance contemporaneously and option investors overestimate the persistence of this media-induced high realized variance. This leads to the overreaction in implied variance that can not be justified by future realized variance, contributing to overpriced straddles and lower future returns.

4.2 Investor Recognition

Next, I examine an alternative explanation about "Investor Recognition" and show that it cannot fully explain the media coverage effect. The "Investor Recognition" hypothesis states that firms with less media coverage have a smaller investor base and therefore should offer higher expected return to compensate its holder for being imperfectly diversified (Merton (1987)). This hypothesis applies to both stocks and options.

To explore the "Investor Recognition" hypothesis, I follow Fang and Peress (2009) to use analyst coverage, stock idiosyncratic volatility, and stock institutional ownership to proxy for firm's information environment and perform subsample analysis. Analogously, I also add firm size and stock institutional number as two more proxies. Small firms, firms with low analyst following, high idiosyncratic volatility, low stock institutional ownership, and low institutional number, face most severe information problems, where the information dissemination role of media coverage is more important. Therefore, if the "Investor Recognition" hypothesis holds, the media coverage effect in option market is supposed to be stronger in these subsample.

Table 10 presents the results of univariate quintile sorts by *Residual Media Coverage* under different subsample. We see that the media-based option strategy is indeed more profitable under the small firm subsample, and firms with low analyst coverage, high idiosyncratic volatility and low stock institutional number, consistent with the "Investor Recognition" hypothesis. However, the pattern for stock institutional ownership is strangely opposite.

Despite of the stronger result in the subsample with bad information environment, the media-based option strategy still generates highly significant and sizable profits under the subsample with relatively good information environment, i.e., large firms, firms with high analyst coverage, low idiosyncratic volatility, and high institutional number. This suggests that the “Investor Recognition” hypothesis is unlikely to fully explain the media coverage effect in option market.

4.3 Media Coverage VS Extreme Stock Returns

In this section, I distinguish the media coverage effect documented in this paper from the attention effect triggered by extreme stock returns in Choy and Wei (2023). Choy and Wei (2023) find that firms with extreme daily stock returns in the past month have lower future option returns. Their explanation is that news media usually reports firms with extreme stock returns, which attracts investor attention and leads to the overpricing in option market. Although related, the media coverage effect that this paper documents is a much broader concept, which does not restrict to the media coverage of extreme stock returns. Table 11 designs two tests to show that our media coverage effect is not captured by the extreme stock return effect.

In Panel A, I perform the univariate sort for the "Never Winners or Losers" sample consisting of those firms without extreme past stock returns. More specifically, I following Choy and Wei (2023) to divide firms into (1) "Losers": those stocks that are among the daily top 80 losers at least once (but never among the daily 80 winners) during the past one month ending at option date, (2) "Winners": those stocks that are among the daily top 80 winners at least once (but never among the daily 80 losers) during the past one month ending at option date, (3) "Winners & Losers": those stocks that happen to be a winner and a loser at least once for each during the past one month ending at option date, and (4) "Never

Winners or Losers": the rest of the firms. The daily top 80 winners and losers are ranked by daily stock returns in the common stock universe listed in NYSE/AMEX/NASDAQ.

For the "Never Winners or Losers" sample, they are unlikely to be reported by news media for their past extreme stock returns. Therefore, the attention channel induced by extreme stock return in Choy and Wei (2023) is shut down. Nonetheless, we still observe strong media coverage effect in this subsample, with the high-minus-low return sorted by *Residual Media Coverage* to be both economically and statistically significant.

In Panel B, I directly compare the predictive ability of *Residual Media Coverage* with the three dummies about extreme stock returns in Choy and Wei (2023), which are 0/1 indicators for the subsample of "Losers", "Winners", and "Winners & Losers". I run Fama-Macbeth regression in a horse-race form. The result suggests that *Residual Media Coverage* is not subsumed by the three extreme return dummies. In fact, the significance level for *Residual Media Coverage* is much higher than those for extreme return dummies, although all of them are significant when put together.

All in all, the media coverage effect proposed in this paper captures different aspects about option return prediction from the extreme stock return effect in Choy and Wei (2023).

4.4 Media Coverage VS Earnings Announcements

In this section, I investigate if the media coverage effect in option market is driven by firms' earnings announcement. Earnings announcement usually attracts much attention and is heavily reported by news media. Meanwhile, literature has documented elevated uncertainty before earnings announcement, which quickly dies out after earnings announcement. The decline of volatility after earnings announcement leads to the lower return of volatility assets such as option straddles (Chung and Louis (2017)). Therefore, it is possible that earnings announcement drives both high media coverage and lower future straddle returns, contributing to the media coverage effect in option market.

To address this concern, I obtain data about firms' earnings announcement date (the 'rdq' field) from Compustat quarterly file. At the third Friday of each month, I identify the "Earnings Announcement" sample, i.e., those firms who just had earnings announcement in the past one month ending at option date (this covers the same period when measuring media coverage). The rest of firms are classified as "No Earnings Announcement" sample. I also define a dummy variable I_{EA} to indicate the "Earnings Announcement" sample.

In Panel A of Table 12, I perform univariate sort for the "No Earnings Announcement" sample, which by design is unlikely to be reported by news media for their past earnings announcement. Therefore, the earnings announcement effect is shut down. Nonetheless, we still observe strong media coverage effect in this subsample, with the high-minus-low return sorted by *Residual Media Coverage* to be both statistically significant and economically large.

In Panel B, I directly run Fama-Macbeth regression in a horse-race form to see if the predictive ability of *Residual Media Coverage* is subsumed by the dummy variable about earnings announcement. The result shows that *Residual Media Coverage* remains highly significant after controlling for the earnings announcement dummy. In fact, both of them are highly significant when put together. The negative coefficient on earnings announcement dummy is consistent with Chung and Louis (2017) that option returns tend to be low after earnings announcement.

In short, the media coverage effect in option market is not simply driven by firms' earnings announcement.

4.5 Media Coverage of Different News Topics

To better reveal the underlying economic sources, I now examine in detail which news topics covered by the media contribute most to the media coverage effect. For this exercise, I utilize the news topics summarized by RavenPack at the 'Group' level.[3] Panel A of Table 13 lists the top 10 groups in my sample period and the corresponding types within each group, ranked by mentioning frequency. They include earnings, insider-trading, order-imbalances, revenues, product-services, labor-issues, analyst-ratings, equity-actions (such as trading status change, buybacks, public offering, ownership changes, etc.), acquisitions-mergers, and stock-prices (such as the stock price changes that make news headlines). Together, these top 10 groups account for around 86% of all news articles with valid 'Group' label.

At the option straddle formation date of each month, I count for each firm the number of news articles belonging to each of these 10 groups during the past one month and then apply Equation (1) to get the corresponding 10 residual media coverages. Panel B of Table 13 compares the predictive ability of these 10 residuals. From the first 10 columns, we see that 7 out of the 10 residual media coverages significantly predict straddle returns when tested individually. In addition, these seven significant coefficients are all negative, consistent with the overall residual media coverage in the main result.

When all the 10 residuals are added together in column (11), only five of them remain significant, including earnings, revenues, equity actions, acquisition and mergers, and stock prices. This is confirmed in column (12), where only the five significant groups are put together. Other topics like insider trading, order imbalance, product services, labor issues, and analyst ratings, although accounting for a large proportion of media reports, contribute little to the media coverage effect.

Furthermore, when I construct the residual media coverage from news articles belonging to these five significant groups as a whole, its predictive ability is even stronger than the residual media coverage from all groups, as shown in the last three columns. Overall, the

[3] RavenPack provides the classification of news articles at different levels, including '*Topic*', '*Group*', '*Type*', '*Sub_type*', '*Property*', and '*Category*'. Among them, '*Topic*' is the coarsest classification while '*Category*' is the finest. Since the '*Topic*' of news articles in my sample almost exclusively rests on 'Business', I use the classification at '*Group*' level. Classifications at finer levels produce too much categories.

results in this section imply that the media coverage effect is mainly driven by the coverage of five key topics, e.g., earnings, revenues, equity actions, acquisition and mergers, and stock prices.

5 Instrumental Variable Analysis

To further address the endogeneity concern about media coverage, I follow prior literature to construct two instrumental variables for media coverage based on the geographic proximity of firm's headquarter to Dow Jones' offices. The eight Dow Jones offices in the US are in Boston, Chicago, Minneapolis, New York, Princeton, San Francisco, Waltham, and Washington.[4] For firm's headquarter location, I obtain it from Compustat.

The first instrument *DJ Location* comes from Dai, Parwada, and Zhang (2015), which is a categorical variable that takes a value of two (one) if the firm headquarter is located in the same city (state) as one of Dow Jones' offices and zero otherwise. The second instrument *Travel Time* is introduced by Dai, Shen, and Zhang (2021), which is the minimum value of the number of minutes taken for trips between the headquarter of a firm and Dow Jones offices, scaled by 100.[5] The relevance condition for these instruments is that firms located near news outlets should have higher media coverage since it is costly for journalists to cover distant firms. Meanwhile, the geographic proximity between firms and news outlets is unlikely to directly affect option straddle returns, possibly satisfying the exogeneity condition.

I then conduct a two-stage instrumental variable analysis to predict one-period-ahead

[4] See http://www.dowjones.com/contact.

[5] My definition of *Travel Time* is slightly different from Dai, Shen, and Zhang (2021), which uses the median value of the number of minutes taken for trips between the headquarter of a firm and Dow Jones offices. The concern for using median travel time is that for firms located in the same city as one of Dow Jones offices, their median travel time is not necessarily the least. I thank the authors for sharing the travel time data on their website, which covers the period from 2000 to 2014. To fit my sample period, I further assume that the data for year 2014 remains constant until 2020.

straddle returns with fitted value of residual media coverage from instrumental variable:

$$Residual\ Media\ Coverage_{i,t} = \alpha + \beta_{IV} Instrument_{i,t} + \beta_{control} Control_{i,t} + Fixed\ Effect + \epsilon_{i,t}$$
$$R_{i,t+1} = \alpha + \beta_{media} Residual\ \widehat{Media}\ Coverage_{i,t} + \beta_{control} Control_{i,t} + Fixed\ Effect + \epsilon_{i,t+1} \tag{2}$$

where control variables include full set of controls in Table 6. I include year-month fixed effect and cluster standard errors at firm level.

Table 14 reports the result. From columns (1) and (3), we can see that in the first stage regression, *DJ Location* is positively related to residual media coverage, while *Travel Time* is negatively related to residual media coverage, both are very significant. This is consistent with the prior that geographic proximity to news outlets is associated with higher media coverage. In addition, the first-stage F-statistics for these two instruments are quite high, rejecting the null hypothesis of weak instrument. The second-stage regression coefficients on fitted residual media coverage in columns (2) and (4) are both significantly negative. This suggests that the negative effect of media coverage on future straddle returns is likely causal, strengthening the main conclusion of this paper.

6 Further Results

6.1 Positive VS Negative Media Coverage

Investors may react differently to news with positive tones and news with negative tones. To explore whether the media coverage effect is driven by positive news coverage or negative news coverage, I classify news articles by their sentiment into positive news (CSS>50) and negative news (CSS<50) using the composite sentiment score (CSS) in RavenPack. I then count the number of positive news and negative news separately for each firm during the past one month ending at option date and apply Equation (1) to get the two corresponding residual media coverages.

Table 15 shows the main empirical results for these two types of residual media coverage. In the univariate portfolio sort of Panel A, the high-minus-low returns sorted by both types are highly significant and negative. However, the positive type of residual media coverage exhibits stronger predictive ability. This is confirmed in the Fama-Macbeth regression of Panel B, where the coefficient on the positive type is much more significant than the negative type in a horse-race comparison. This suggests that the media coverage effect is mainly driven by positive news coverage.

6.2 Media Coverage VS Media Sentiment

Next, I examine the predictive ability of media sentiment for option returns and compare it with media coverage. To this end, I construct two firm-level measures of media sentiment from all news articles mentioning the firm during the past one month ending at option date. The first is simply the average CSS score of news articles. The second measure is the difference between the number of news articles with positive tones (CSS>50) and that with negative tones (CSS<50), divided by total number of new articles. Higher values of these measures indicate more positive tones of news articles.

Table 16 reports the Fama-Macbeth regression results for media sentiment. We see that higher media sentiment tends to be followed by lower option straddle returns, but the predictive ability is much weaker compared to media coverage. This suggests that media coverage better forecasts option returns than media sentiment.

6.3 Persistence of Media Coverage Effect

To assess the persistence of media coverage effect in option market, Table 17 examines the main empirical result under alternative news windows, including univariate sort in Panel A and Fama-Macbeth regression in Panel B. The residual media coverage is now the residual of media coverage measured over different past windows from 3 months to 12 months, instead of the past one month in the main analysis. We can see that when increasing the past news window, the predictive ability of residual media coverage quickly drops. In fact, for the past 9-month and 12-month windows, the Fama-Macbeth regression coefficients on residual media coverage are only marginally significant or insignificant after adding control variables.

In a similar fashion, Figure 3 plots the Fama-Macbeth regression coefficient of straddle return on residual media coverage at different lags. The result also suggests that the first lag exhibits the strongest predictive ability.

$$R_{i,t} = a_{n,t} + b_{n,t} Residual\ Media\ Coverage_{i,t-n} + \epsilon_{i,t} \tag{3}$$

In a nutshell, the media coverage effect in option market is likely short-term.

6.4 Subperiod Analysis

At last, I check the predictive ability of residual media coverag during different subperiods. I look at the sample period division by the time dimension (first vs. second half of the sample), by the median of investor sentiment in Baker and Wurgler (2006) (low vs. high), by the median of uncertainty index VIX (low vs. high), and by business cycle (normal vs. crisis periods). For the business cycle, I follow the NBER definition and identify two crisis periods during my sample period: January 2008 to June 2009 (both inclusive) and March 2020 to April 2020 (both inclusive). The subperiod division is based on the straddle formation month.

Table 18 presents the Fama-Macbeth regression result over different subperiods. As shown, the coefficient on residual media coverage is significantly negative for almost all subperiods, suggesting the stability of its predictive ability. The weaker result for the crisis period is likely due to the small sample size.

7 Conclusion

This paper documents strong media coverage effect in option market. Firms with higher media coverage in the past one month tend to have lower straddle returns. A long-short strategy based on residual media coverage, which purges out the main covariates that affect media coverage, generates an average monthly profit of 5.04% with a high T-statistic of 15.84 and a monthly Sharpe ratio of 1.173. The magnitude is larger than most well-known option return predictors. The strategy is profitable in every 5-year subsample, and has limited crash risk. It is also quite robust to different specifications, including alternative option returns, alternative relevance score filters, alternative news type filters, and alternative regression specifications to get residual media coverage. After reasonable transaction cost management, the media-based option strategy remains profitable when traders pay as high as 51.6% of the quoted option bid-ask spread.

The media coverage effect cannot be explained by option return predictors in previous literature. In particular, the media-based strategy has highly significant alphas that are not captured by option factor models in literature. It also remains significant after controlling for other common option return predictors in Fama-Macbeth regressions.

In terms of the economic sources underlying the media coverage effect, I find evidence consistent with the volatility overreaction of option investors to the media report. Since straddle return is related to the variance swap return, which measures the difference between future realized variance and current option-implied variance, I decompose the predictive ability of residual media coverage into two components related to future realized variance and current implied variance. I find that residual media coverage positively predicts next-month realized stock variance and also positively correlates with the implied variance in the current month. However, the coefficient for the implied variance case is much larger than that for realized variance case, suggesting that option market overreacts to the media report in the contemporaneous month. The overreaction in option market contributes to the overpricing of straddles, which leads to lower future returns and explains the media coverage effect. I also show that the overreaction of option investors to media reports is likely driven by their overestimation about the persistence of media-induced high realized variance.

I also explore alternative explanations for media coverage effect but fail to find satisfactory results. The traditional “Investor Recognition” hypothesis proposed by Fang and Peress (2009) states that firms with less media coverage have a smaller investor base and therefore should offer higher expected return to compensate its holder for being imperfectly diversified (Merton (1987)). However, I show that the media-based strategy remains highly profitable even in the subsample with high investor recognition, suggesting that “Investor Recognition” cannot fully explain the media coverage effect in option market. The other alternatives are related to the endogeneity of media coverage: firms with extreme past stock return or recent earnings announcement have lower future straddle returns (Choy and Wei (2023); Chung and Louis (2017)) and at the same time they are heavily reported by news media. Inconsistent with these explanations, we still observe a strong media coverage effect after controlling for extreme stock return and earnings announcement.

To get a sense of what news topics covered by media reports contribute most to the media coverage effect, I construct residual media coverage from news articles belonging to different news topics and compare their predictive ability. The results show that the media coverage effect is mainly driven by the coverage of five key topics, including earnings, revenues, equity actions (such as trading status change, buybacks, public offering, ownership changes, etc.), acquisition and mergers, and stock prices (such as the stock price changes that make news headlines).

To further address the endogeneity concern of media coverage, I follow prior literature to use geographic proximity to news outlets as instrumental variable for media coverage and find that the negative effect of media coverage on straddle return survives under two-stage instrumental regression. This suggests that the media coverage effect is likely causal.

I also provide some further results with respect to media sentiment. First, I decompose the media coverage into positive and negative media coverage by news sentiment. Although both types of media coverage exhibit strong predictive ability for straddle returns, the positive media coverage contributes more to the media coverage effect. Second, I directly compare the predictive ability of media sentiment and media coverage, and find that media coverage is a much stronger predictor for straddle returns.

In addition, I examine the persistence of media coverage effect by constructing residual media coverage during different past windows. As the past window increases, the predictive ability of media coverage quickly drops compared to the baseline case of past one month, suggesting the short-term predictive ability of media coverage in option market. At last, I show that the predictive ability of media coverage remains relatively stable across different subperiods.

Overall, this paper contributes to the literature by documenting the media coverage effect in option market and proposing a volatility overreaction explanation specific to option market.

References

Bali, T. G., H. Beckmeyer, M. Moerke, and F. Weigert, 2023, "Option return predictability with machine learning and big data," *The Review of Financial Studies*, 36(9), 3548–3602.

Bali, T. G., R. F. Engle, and S. Murray, 2016, *Empirical asset pricing: The cross section of stock returns.* John Wiley & Sons.

Bandarchuk, P., and J. Hilscher, 2013, "Sources of momentum profits: Evidence on the irrelevance of characteristics," *Review of Finance*, 17(2), 809–845.

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

Cao, J., B. Han, G. Li, R. Yang, and X. Zhan, 2022, "Forecasting option returns with news," *Available at SSRN.*

Cao, J. J., A. Vasquez, X. Xiao, and X. E. Zhan, 2023, "Why Does Volatility Uncertainty Predict Equity Option Returns?," *The Quarterly Journal of Finance*, 13(01), 2350005.

Choy, S. K., and J. Wei, 2023, "Investor attention and option returns," *Management Science*, 69(8), 4845–4863.

Chung, S. G., and H. Louis, 2017, "Earnings announcements and option returns," *Journal of Empirical Finance*, 40, 220–235.

Dai, L., J. T. Parwada, and B. Zhang, 2015, "The governance effect of the media's news dissemination role: Evidence from insider trading," *Journal of Accounting Research*, 53(2), 331–366.

Dai, L., R. Shen, and B. Zhang, 2021, "Does the media spotlight burn or spur innovation?," *Review of Accounting Studies*, 26, 343–390.

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

Fang, L., and J. Peress, 2009, "Media coverage and the cross-section of stock returns," *The journal of finance*, 64(5), 2023–2052.

Filippou, I., and P. A. Garcia-Ares, 2020, "Media Sentiment and the Cross-Section of Option Returns," *Available at SSRN 4054872*.

Gao, H., J. Wang, Y. Wang, C. Wu, and X. Dong, 2020, "Media coverage and the cost of debt," *Journal of Financial and Quantitative Analysis*, 55(2), 429–471.

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

Heston, S. L., C. S. Jones, M. Khorram, S. Li, and H. Mo, 2022, "Seasonal momentum in option returns," *Available at SSRN 4167231*.

———, 2023, "Option momentum," *The Journal of Finance*, 78(6), 3141–3192.

Hillert, A., H. Jacobs, and S. Müller, 2014, "Media makes momentum," *The Review of Financial Studies*, 27(12), 3467–3501.

Li, S., and L. Yang, 2024, "Peer Option Momentum," *Available at SSRN 4964637*.

Merton, R. C., 1987, "A Simple Model of Capital Market Equilibrium with Incomplete Information," *Journal of Finance*, 42(3), 483–510.

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

Newey, W. K., and K. D. West, 1987, "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–708.

Peress, J., 2014, "The media and the diffusion of information in financial markets: Evidence from newspaper strikes," *the Journal of Finance*, 69(5), 2007–2043.

Tetlock, P. C., 2007, "Giving content to investor sentiment: The role of media in the stock market," *The Journal of finance*, 62(3), 1139–1168.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy, 2008, "More than words: Quantifying language to measure firms' fundamentals," *The journal of finance*, 63(3), 1437–1467.

Vasquez, A., 2017, "Equity volatility term structures and the cross section of option returns," *Journal of Financial and Quantitative Analysis*, 52(6), 2727–2754.

Zhan, X., B. Han, J. Cao, and Q. Tong, 2022, "Option return predictability," *The Review of Financial Studies*, 35(3), 1394–1442.

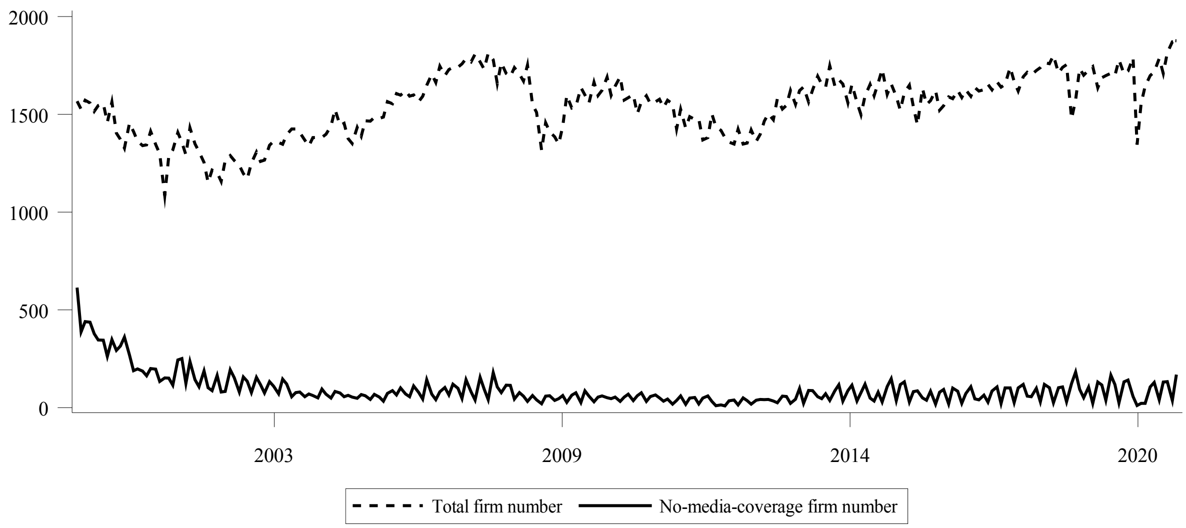

Figure 1: Time-series of Firm Coverage. This figure plots the time-series of total firm number in the sample (dashed line) and the firm number without media coverage (solid line). The sample is restricted to firms with straddle returns.

Figure 2: Five-year Moving Average of Media-based Return Spread. This figure plots the rolling 5-year averages of the high-minus-low return sorted by residual media coverage (top panel) and the high-minus-no return sorted by media coverage (bottom panel), along with the 95% confidence intervals in dashed lines. Standard errors are Newey-West adjusted with 5 lags.

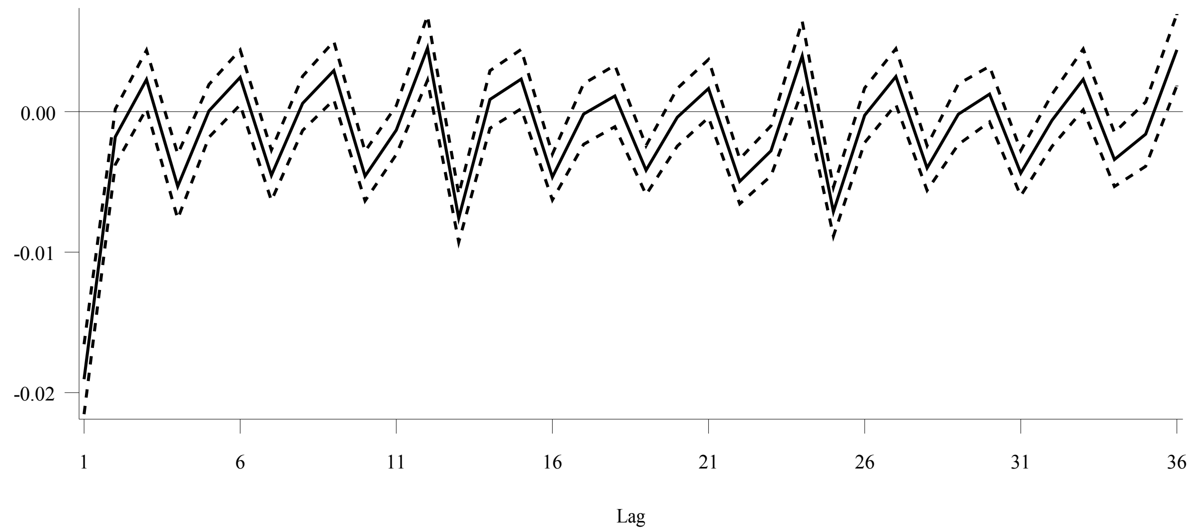

Figure 3: Slope Coefficients of Fama-Macbeth Regressions. This figure plots the slope coefficients (solid line) and its 95% confidence intervals (dash line) from Fama-Macbeth regressions in which monthly straddle returns are regressed on the past residual media coverage, with the month lag in the horizontal axis. I winsorize independent variables at 0.5% and 99.5% levels on a monthly basis. Standard errors are Newey-West adjusted with 5 lags.

Table 1: Summary Statistics

This table reports summary statistics. Sample period is from January 2000 to December 2020, due to the availability of RavenPack data. *Straddle return* is at monthly frequency formed at the third Friday of each month, and it is the return on the ATM straddle daily-delta-hedged until expiration. The excess return is relative to risk-free rate over the option period. *Number of news articles* is the number of news articles mentioning the firm in past 1 month ending at option date. *Media Coverage* equals ln(1+number of news articles). *Residual Media Coverage* is the residual media coverage after controlling for firm size, NASDAQ dummy, S&P 500 dummy, and analyst coverage as in Equation (1). *Option momentum* is the straddle return in past 1 month. *Market capitalization* is a firm's stock market capitalization in billions of dollars. *Idiosyncratic volatility* is the standard deviation (annualized) of the residuals from a 22-day rolling regression using Fama-French 3 factors following Cao and Han (2013). *IV term spread* is the slope of implied volatility term structure (Vasquez (2017)) defined as the difference between 60-day and 30-day ATM implied volatilities. *IV smirk* is the smirk of implied volatility curve defined as the difference between the implied volatilities of the 30-day call with delta of 0.3 and the 30-day put with delta of -0.3. *IV-HV* is the log difference between ATM straddle implied volatility and historical volatility as in Goyal and Saretto (2009). *VOV* is the standard deviation of the daily percentage changes of 30-day at-the-money implied volatility over the past 22 trading days in Cao, Vasquez, Xiao, and Zhan (2023).

	Observation	Mean	StdDev	P10	Median	P90
Number of firms per month	252 months	1543.30	158.26	1342	1565	1740
Excess straddle return (%)	388911	-4.60	39.54	-40.95	-9.55	32.71
Number of news articles	388911	24.43	68.82	1	13	48
Media coverage	388911	2.526	1.178	0.693	2.639	3.892
Residual media coverage	388911	0.000	0.937	-1.235	0.081	1.078
Option momentum	380453	-0.043	0.392	-0.413	-0.095	0.342
Market capitalization (billions)	388911	10.628	36.062	0.390	2.200	21.291
Idiosyncratic volatility	388730	0.348	0.267	0.130	0.277	0.642
IV term spread	388899	-0.005	0.062	-0.058	0.001	0.042
IV smirk	388899	-0.014	0.128	-0.090	-0.023	0.058
IV-HV	378299	-0.011	0.273	-0.324	-0.011	0.306
VOV	387806	0.073	0.075	0.030	0.054	0.120

Table 2: Determinates of Media Coverage

This table reports the coefficients of Fama-Macbeth regressions in which media coverage is regressed on firm characteristics, following the specifications in Table 2 of Hillert, Jacobs, and Müller (2014). Firm size is the log of firm's stock market capitalization. NASDAQ and S&P 500 dummies are the indicators for listing in NASDAQ and being a member of S&P 500, respectively. Analyst coverage is the log of one plus number of analyst coverage in the past 12 months. Book-to-market is the log of firm's book equity to market capitalization ratio. Turnover is the log of firm's average daily turnover over the past 1 month. Idiosyncratic volatility is the standard deviation (annualized) of the residuals from a 22-day rolling regression using Fama-French 3 factors. Momentum strength measures the salience of firm's past 1 month return relative to the market median as in Bandarchuk and Hilscher (2013). I winsorize all independents at 0.5% and 99.5% percentile on a monthly basis except for the dummy variables. T-statistics in parentheses are Newey-West adjusted with 5 lags.

	(1)	(2)	(3)	(4)	(5)
Firm Size	0.376***	0.325***	0.284***	0.408***	0.413***
	(43.73)	(67.58)	(44.56)	(78.83)	(92.58)
NASDAQ Dummy		0.022*	0.004	0.039**	-0.033**
		(1.91)	(0.34)	(2.10)	(-2.31)
S&P 500 Dummy		0.237***	0.221***	0.196***	0.193***
		(10.38)	(9.94)	(9.71)	(10.87)
Analyst coverage			0.153***	0.027***	0.024***
			(16.38)	(2.75)	(2.62)
Book-to-market				0.063***	0.093***
				(12.26)	(19.08)
Turnover				0.161***	0.172***
				(22.40)	(29.89)
Idiosyncratic Volatility				1.218***	1.279***
				(25.60)	(25.64)
Momentum Strength				0.326***	0.335***
				(9.01)	(10.44)
Fama-French 48 industry dummies	No	No	No	No	Yes
Adj. R^2	26.8%	27.5%	28.1%	34.8%	38.1%
$Avg.\#Firms$	1543	1543	1543	1339	1273

Table 3: Univariate Portfolio Sort

This table presents means and t-statistics from univariate portfolio sorts based on *Residual Media Coverage* (Panel A) and *Media Coverage* (Panel B). In Panel A, firms are sorted into quintiles by *Residual Media Coverage*, while in Panel B, I separate out the no media coverage group and then divide the rest into four groups by their *Media Coverage*, as in Fang and Peress (2009). I form equally weighted portfolios on the third Friday of each month and rebalance them every month. All returns are monthly and in percentage. T-statistics in parentheses are Newey-West adjusted with 5 lags.

Panel A: Portfolios sorted by *Residual Media Coverage*

	Low Coverage	2	3	4	High Coverage	High-Low
Excess straddle return	-2.39	-3.47	-4.65	-5.70	-7.44	-5.04***
	(-2.82)	(-4.06)	(-5.05)	(-6.12)	(-8.82)	(-15.84)

Panel B: Portfolios sorted by *Media Coverage*

	No Coverage	2	3	4	High Coverage	High-No
Excess straddle return	-2.60	-3.16	-4.82	-5.46	-5.79	-3.18***
	(-2.86)	(-4.05)	(-5.90)	(-5.97)	(-5.66)	(-6.68)

Table 4: Investment Performance Evaluation

This table compares the investment performance of media-based option strategy with other common option strategies in literature. Portfolios of straddle returns are equally weighted and rebalanced each month. All performance metrics are applied to the long-short returns at monthly frequency. (%) after a variable means that the variable reported is in percentage. I adjust the signs of all investment signals such that their long-short spreads are positive to facilitate comparison. T-statistics in parentheses are Newey-West adjusted with 5 lags.

	Mean (%)	T-stats.	StdDev (%)	Sharpe ratio	Skewness	Excess kurtosis	Max drawdown (%)
Residual Media Coverage	5.04***	(15.84)	4.30	1.173	0.047	1.523	10.36
Media Coverage	3.18***	(6.68)	7.95	0.400	-0.142	5.838	51.92
Option Momentum	2.89***	(5.89)	5.10	0.567	0.941	2.538	30.12
Market Cap	3.27***	(4.31)	9.95	0.329	1.808	8.309	52.42
Idiosyncratic Volatility	3.27***	(4.19)	9.79	0.334	3.114	27.051	76.89
IV term spread	4.86***	(9.41)	6.68	0.728	2.239	16.898	22.33
IV smirk	0.66*	(1.82)	4.43	0.149	-0.470	1.689	62.90
IV-HV	9.10***	(12.98)	7.71	1.181	0.864	2.637	17.64
VOV	4.18***	(8.14)	5.96	0.700	0.427	1.192	37.95

Table 5: Alphas of Media-based Strategy

This table shows the alphas of media-based strategy by regressing its long-short return on option factors listed in the rows. I consider the 4-factor and 7-factor models used in Heston, Jones, Khorram, Li, and Mo (2023), and also augment the 7-factor model with an additional factor that is the long-short return sorted by volatility of volatility. I adjust the signs of all investment signals such that their long-short spreads are positive. The alphas are in percentage. T-statistics in parentheses are Newey-West adjusted with 5 lags.

	by Residual Media Coverage			by Media Coverage		
	(1)	(2)	(3)	(4)	(5)	(6)
Alpha	5.64***	5.43***	4.47***	5.78***	5.42***	4.75***
	(8.89)	(11.67)	(9.59)	(8.31)	(9.07)	(7.59)
Market Cap.	-0.10*	-0.08	-0.10**	-0.48***	-0.47***	-0.48***
	(-1.66)	(-1.56)	(-2.02)	(-7.67)	(-7.97)	(-8.71)
Idiosyncratic Vol.	0.21**	0.25***	0.22***	0.09	0.16**	0.14**
	(2.21)	(3.73)	(3.95)	(1.14)	(2.49)	(2.14)
IV-HV	-0.11*	-0.00	-0.01	-0.15***	-0.01	-0.02
	(-1.83)	(-0.09)	(-0.21)	(-2.63)	(-0.24)	(-0.29)
Short SPX straddle	0.00	0.00	0.01	0.01	0.01	0.01
	(0.48)	(0.33)	(1.31)	(0.78)	(0.46)	(0.90)
IV term spread		-0.20***	-0.16***		-0.26***	-0.24***
		(-3.13)	(-3.13)		(-3.71)	(-3.26)
IV smirk		-0.05	-0.06		0.00	-0.00
		(-0.58)	(-0.76)		(0.04)	(-0.01)
Short EW straddle		0.02	0.00		0.03	0.02
		(0.41)	(0.10)		(0.57)	(0.39)
VOV			0.23***			0.16***
			(4.91)			(2.68)
$Adj.R^2$	12.2%	16.0%	25.4%	35.9%	37.9%	39.0%
$\#Months$	252	252	252	252	252	252

Table 6: Fama-Macbeth Regression

This table reports the coefficients of Fama-Macbeth regressions in which one-period-ahead straddle return is regressed on media-based variables and other characteristics. I winsorize all independents at 0.5% and 99.5% percentile on a monthly basis. T-statistics in parentheses are Newey-West adjusted with 5 lags.

	(1)	(2)	(3)	(4)
Residual Media Coverage	-0.019***	-0.018***		
	(-14.99)	(-14.16)		
Media Coverage			-0.009***	-0.016***
			(-6.47)	(-12.61)
Option Momentum		0.045***		0.045***
		(9.72)		(9.61)
Ln(Market Cap)		-0.003*		0.003**
		(-1.76)		(2.00)
Idiosyncratic Vol.		-0.101***		-0.102***
		(-7.67)		(-7.78)
IV term spread		0.167***		0.165***
		(5.68)		(5.63)
IV smirk		-0.001		-0.001
		(-0.10)		(-0.08)
IV-HV		-0.156***		-0.156***
		(-15.59)		(-15.56)
VOV		-0.141***		-0.143***
		(-5.25)		(-5.35)
$Adj.R^2$	0.3%	3.1%	0.3%	3.0%
$Avg.\#Firms$	1543	1472	1543	1472

Table 7: Robustness of Media-based Option Strategy

This table shows the means and t-statistics from univariate quintile sorts by residual media coverage under different robustness tests. Panel A examines the predictive ability of residual media coverage for alternative option returns, including the 'static straddle' return, which is the return on a static position of monthly held-to-maturity ATM straddle without any delta-hedge, and 'dynamic call' ('dynamic put') return, which is the return on ATM call (put) daily-delta-hedged until expiration. Panel B considers alternative residual media coverage measures by applying different relevance score filters in news mention. Panel C investigates the influence of news type filters. Panel D varies the set of firm characteristics controls in the regression to get residual media coverage. All returns are monthly and in percentage. T-statistics in parentheses are Newey-West adjusted with 5 lags.

Panel A: Alternative option returns

	Low	2	3	4	High	High-Low
Excess static straddle return	-1.75	-2.66	-3.82	-4.82	-5.28	-3.54***
	(-0.93)	(-1.35)	(-1.81)	(-2.28)	(-2.66)	(-5.87)
Excess dynamic call return	-0.35	-0.49	-0.63	-0.81	-1.07	-0.72***
	(-3.42)	(-4.82)	(-6.05)	(-7.78)	(-9.48)	(-15.16)
Excess dynamic put return	-0.64	-0.76	-0.87	-1.04	-1.37	-0.72***
	(-6.27)	(-7.39)	(-8.37)	(-10.02)	(-12.98)	(-19.89)

Panel B: Different restrictions about article relevance score

	Low	2	3	4	High	High-Low
Relevance score $\geq$ 80	-2.40	-3.44	-4.67	-5.70	-7.45	-5.05***
	(-2.83)	(-4.02)	(-5.08)	(-6.13)	(-8.83)	(-15.72)
Relevance score $\geq$ 85	-2.39	-3.45	-4.67	-5.69	-7.46	-5.07***
	(-2.82)	(-4.02)	(-5.07)	(-6.10)	(-8.86)	(-15.77)
Relevance score $\geq$ 90	-2.40	-3.52	-4.63	-5.69	-7.42	-5.03***
	(-2.83)	(-4.10)	(-5.03)	(-6.11)	(-8.77)	(-15.50)

Panel C: Different news type filters

	Low	2	3	4	High	High-Low
Exclude news type 'PRESS-RELEASE'	-2.30	-3.63	-4.70	-5.70	-7.32	-5.02***
	(-2.68)	(-4.20)	(-5.22)	(-6.15)	(-8.68)	(-15.60)
Include news type 'TABULAR-MATERIAL'	-2.53	-3.60	-4.51	-5.64	-7.37	-4.85***
	(-3.00)	(-4.23)	(-4.85)	(-6.24)	(-8.51)	(-15.51)

Panel D: Alternative explanatory variables in calculating residual media coverage

	Low	2	3	4	High	High-Low
Firm size	-2.38	-3.55	-4.73	-5.65	-7.36	-4.98***
	(-2.79)	(-4.12)	(-5.22)	(-6.05)	(-8.66)	(-14.06)
Firm size, NASDAQ, S&P 500	-2.37	-3.64	-4.68	-5.60	-7.36	-4.99***
	(-2.80)	(-4.17)	(-5.15)	(-6.02)	(-8.78)	(-14.97)
Main specification + further controls	-2.47	-3.65	-4.59	-5.55	-6.87	-4.40***
	(-2.93)	(-4.32)	(-5.11)	(-6.06)	(-7.43)	(-13.16)
Main specification + further controls + industry	-2.32	-3.52	-4.57	-5.15	-6.52	-4.20***
	(-2.70)	(-4.11)	(-4.83)	(-5.24)	(-6.93)	(-12.96)

Table 8: Media-based Option Strategy: Adjusting Transaction Cost

This table reports the average returns of residual-media-coverage-based option strategy after adjusting for option bid-ask spreads. I report results for the baseline strategy by quintiles and for strategies that reduce transactions costs by using extreme deciles and by using low-cost sample consisting of only straddles with percentage bid-ask spreads in the lowest quartile each month after the portfolio sort. I consider sizes of transactions costs ranging from zero to the full quoted half-spread. Intermediate cases include the algo, adjusted, and effective spreads, which are 20.3%, 51.6%, and 75.8%, respectively, as large as the quoted spreads (following Muravyev and Pearson (2020)). T-statistics in parentheses are Newey-West adjusted with 5 lags.

	None	Algo	Adjusted	Effective	Quoted
Main sample, quintiles	5.04***	1.09***	-5.13***	-10.25***	-16.31***
	(15.84)	(3.25)	(-11.01)	(-15.85)	(-16.53)
Main sample, deciles	6.04***	2.16***	-3.91***	-8.87***	-14.63***
	(17.21)	(5.95)	(-8.18)	(-13.84)	(-15.85)
Low-cost sample, deciles	3.88***	2.79***	1.12**	-0.17	-1.46**
	(7.16)	(5.11)	(2.02)	(-0.30)	(-2.53)

Table 9: Variance Prediction

This table reports the Fama-Macbeth regression results of variance prediction. In the first three columns, *Residual Media Coverage* is used to predict the log of realized variance over the next period ($\ln(RV_t)$), the log of implied variance at the option date of month t-1 ($\ln(IV_{t-1})$), and the log variance swap return ($\ln(RV_t/IV_{t-1})$). In the last two columns, I use *Residual Media Coverage* to predict the log of realized variance in the contemporaneous period ($\ln(RV_{t-1})$) and the changes of log realized variance ($\ln(RV_t/RV_{t-1})$). Both realized variance and implied variance are monthly variance. I winsorize all independents at 0.5% and 99.5% percentile on a monthly basis. T-statistics in parentheses are Newey-West adjusted with 5 lags.

	(1) $Ln(RV_t)$	(2) $Ln(IV_{t-1})$	(3) $Ln(RV_t/IV_{t-1})$	(4) $Ln(RV_{t-1})$	(5) $Ln(RV_t/RV_{t-1})$
$Residual\ Media\ Coverage_{t-1}$	0.036***	0.071***	-0.035***	0.226***	-0.191***
	(6.09)	(14.41)	(-14.87)	(26.92)	(-17.71)
$Adj.R^2$	0.4%	1.0%	0.4%	4.7%	4.5%
$Avg.\#Firms$	1543	1543	1543	1543	1543

Table 10: Investor Recognition Hypothesis

This table shows the univariate sorts by residual media coverage under different subsample differing by investor recognition. Firms are first sorted into three groups based on firm size, analyst coverage, idiosyncratic volatility, institutional number, or institutional ownership. Then, within each group, I further sort firms into quintiles by residual media coverage. The table reports the mean returns for each portfolio and their t-statistics. All returns are monthly and in percentage. T-statistics in parentheses are Newey-West adjusted with 5 lags.

	Low Coverage	2	3	4	High Coverage	High-Low
			by firm size			
Firm size (Low)	-2.49	-4.62	-5.70	-7.24	-9.62	-7.13***
	(-3.19)	(-6.12)	(-7.37)	(-9.04)	(-12.67)	(-16.05)
Firm size (Median)	-2.51	-3.21	-4.18	-4.90	-7.19	-4.68***
	(-2.89)	(-3.67)	(-4.86)	(-5.29)	(-8.11)	(-12.46)
Firm size (High)	-2.30	-3.00	-4.03	-4.68	-5.28	-2.98***
	(-2.25)	(-2.74)	(-3.50)	(-3.91)	(-4.91)	(-6.71)
			by analyst coverage			
Analyst coverage (Low)	-2.93	-4.15	-5.74	-7.21	-9.34	-6.41***
	(-3.58)	(-4.81)	(-6.78)	(-7.97)	(-11.78)	(-17.17)
Analyst coverage (Median)	-2.54	-3.37	-4.66	-5.23	-7.69	-5.15***
	(-2.94)	(-4.02)	(-5.00)	(-5.71)	(-8.69)	(-12.14)
Analyst coverage (High)	-1.89	-2.88	-3.93	-4.20	-5.38	-3.50***
	(-1.92)	(-2.85)	(-3.81)	(-3.98)	(-5.51)	(-7.22)
			by idiosyncratic volatility			
Idiosyncratic volatility (Low)	-2.36	-3.07	-4.14	-4.52	-4.97	-2.61***
	(-2.35)	(-3.07)	(-3.93)	(-3.77)	(-4.57)	(-6.87)
Idiosyncratic volatility (Median)	-1.94	-3.15	-4.17	-4.99	-6.39	-4.44***
	(-2.28)	(-3.64)	(-4.38)	(-5.30)	(-6.68)	(-10.10)
Idiosyncratic volatility (High)	-3.18	-4.85	-5.98	-7.80	-9.42	-6.24***
	(-4.15)	(-5.99)	(-7.27)	(-9.26)	(-12.05)	(-16.58)
			by institutional number			
Institutional number (Low)	-3.28	-5.01	-6.07	-7.61	-9.53	-6.25***
	(-4.19)	(-6.59)	(-7.81)	(-9.36)	(-12.52)	(-13.95)
Institutional number (Median)	-2.13	-2.82	-3.87	-4.97	-7.18	-5.06***
	(-2.41)	(-3.36)	(-4.27)	(-5.47)	(-8.05)	(-12.00)
Institutional number (High)	-1.97	-2.90	-4.11	-4.38	-5.10	-3.13***
	(-1.88)	(-2.68)	(-3.66)	(-3.63)	(-4.75)	(-6.65)
			by institutional ownership			
Institutional ownership (Low)	-3.51	-4.69	-5.55	-6.77	-7.59	-4.08***
	(-3.83)	(-5.30)	(-5.93)	(-7.59)	(-8.75)	(-9.60)
Institutional ownership (Median)	-1.57	-2.66	-3.92	-4.95	-6.78	-5.21***
	(-1.75)	(-2.73)	(-4.01)	(-4.91)	(-7.42)	(-12.23)
Institutional ownership (High)	-2.09	-3.26	-4.44	-5.26	-7.91	-5.82***
	(-2.53)	(-4.02)	(-5.08)	(-5.67)	(-9.22)	(-13.36)

Table 11: Media Coverage VS Extreme Stock Returns

This table examines if the media coverage effect simply manifests the information about extreme stock returns in Choy and Wei (2023). Panel A shows the means and t-statistics from univariate quintile sorts by residual media coverage under the sample whose stocks have never been a winner or a loser in the past one month ending at option date. Panel B reports the coefficients of Fama-Macbeth regressions of one-period-ahead straddle returns on residual media coverage, the three extreme return dummy variables in Choy and Wei (2023) about stock losers (I_{Losers}), winners ($I_{Winners}$) and both losers and winners ($I_{Winners\&Losers}$), and other characteristics. I_{Losers} is set to 1 if the stock is among the daily top 80 losers at least once (but never among the daily 80 winners) during the past one month ending at option date and 0 otherwise. $I_{Winners}$ is set to 1 if the stock is among the daily top 80 winners at least once (but never among the daily 80 losers) during the past one month ending at option date and 0 otherwise. $I_{Winners\&Losers}$ is set to 1 if the stock happens to be a winner and a loser at least once for each during the past one month ending at option date and 0 otherwise. I winsorize all independents at 0.5% and 99.5% percentile on a monthly basis except for dummy variables. T-statistics in parentheses are Newey-West adjusted with 5 lags.

Panel A: Portfolio sorts in the "Never Winners or Losers" Sample

	Low	2	3	4	High	High-Low
Excess straddle return	-2.02	-3.06	-3.97	-4.57	-5.71	-3.69***
	(-2.28)	(-3.48)	(-4.24)	(-4.58)	(-5.99)	(-11.26)

Panel B: Fama-Macbeth regression in horse-race form

	(1)	(2)	(3)	(4)
Residual Media Coverage	-0.019***		-0.016***	-0.017***
	(-14.99)		(-12.87)	(-13.85)
I_{Losers}		-0.034***	-0.026***	-0.019***
		(-8.45)	(-6.65)	(-5.40)
$I_{Winners}$		-0.031***	-0.024***	-0.023***
		(-6.34)	(-5.09)	(-7.06)
$I_{Winners\&Losers}$		-0.048***	-0.039***	-0.015**
		(-6.16)	(-5.00)	(-2.56)
Option Momentum				0.048***
				(10.15)
Ln(Market Cap)				-0.003*
				(-1.77)
Idiosyncratic Vol.				-0.075***
				(-5.46)
IV term spread				0.173***
				(5.88)
IV smirk				0.001
				(0.08)
IV-HV				-0.157***
				(-15.57)
VOV				-0.136***
				(-5.18)
$Adj.R^2$	0.3%	0.4%	0.6%	3.2%
$Avg.\#Firms$	1543	1543	1543	1471

Table 12: Media Coverage VS Earnings Announcement

This table examines if the media coverage effect is driven by earnings announcement. Panel A shows the means and t-statistics from univariate quintile sorts by residual media coverage under the sample whose stocks did not have earnings announcement in the past one month ending at option date. Panel B reports the coefficients of Fama-Macbeth regressions of one-period-ahead straddle returns on residual media coverage, the dummy variable about earnings announcement (I_{EA}), and other characteristics. I_{EA} is set to 1 if the firm had earnings announcement in the past one month ending at option date and 0 otherwise. I winsorize all independents at 0.5% and 99.5% percentile on a monthly basis except for the dummy variable. T-statistics in parentheses are Newey-West adjusted with 5 lags.

Panel A: Portfolio sorts in the "No Earnings Announcement" Sample

	Low	2	3	4	High	High-Low
Excess straddle return	-1.29	-2.04	-2.58	-3.69	-5.75	-4.46***
	(-1.51)	(-2.35)	(-3.23)	(-4.52)	(-7.09)	(-13.44)

Panel B: Fama-Macbeth regression in horse-race form

	(1)	(2)	(3)	(4)
Residual Media Coverage	-0.019***		-0.015***	-0.012***
	(-14.99)		(-11.84)	(-10.98)
I_{EA}		-0.046***	-0.032***	-0.040***
		(-11.27)	(-6.90)	(-11.40)
Option Momentum				0.046***
				(9.74)
Ln(Market Cap)				-0.002
				(-1.35)
Idiosyncratic Vol.				-0.098***
				(-7.52)
IV term spread				0.202***
				(7.05)
IV smirk				0.001
				(0.06)
IV-HV				-0.162***
				(-16.69)
VOV				-0.118***
				(-4.43)
$Adj.R^2$	0.3%	0.4%	0.6%	3.3%
$Avg.\#Firms$	1543	1543	1543	1472

Table 13: Media Coverage of Different News Topics

This table examines which types of news topic drive the media coverage effect. Panel A shows the details of the top 10 news topics by mentioning frequency in our sample. Panel B reports the coefficients of Fama-Macbeth regressions of one-period-ahead straddle returns on residual media coverage constructed from different news topics, with full set of control variables. I winsorize all independents at 0.5% and 99.5% percentile on a monthly basis except for the dummy variable. T-statistics in parentheses are Newey-West adjusted with 5 lags.

Panel A: Details of each news topic

<table>
<tr><th>News Topics</th><th>Frequency</th><th>Cum. Freq.</th><th>Types|Freq. within each topic</th></tr>
<tr><td>Earnings</td><td>21.31%</td><td>21.31%</td><td>earnings|53.7%, earnings-per-share|29.3%, earnings-guidance|5.4%, earnings-per-share-guidance|5.4%, earnings-estimate|2.8%, operating-earnings|1.8%, ebitda|0.6%, ebitda-guidance|0.4%, pretax-earnings|0.3%, earnings-revision|0.2%, operating-earnings-guidance|0.2%, ebit|0.0%, ebita|0.0%, ebita-guidance|0.0%</td></tr>
<tr><td>Insider-trading</td><td>19.79%</td><td>41.10%</td><td>sell-registration|31.4%, insider-sell|31.2%, insider-buy|22.0%, insider-surrender|13.4%, insider-gift|2.1%, insider-trading-lawsuit|0.0%</td></tr>
<tr><td>Order-imbalances</td><td>8.48%</td><td>49.58%</td><td>buy-moc|39.2%, sell-moc|32.1%, no-moc|20.4%, sell-moo|2.6%, buy-moo|2.2%, delay-imbalance|1.4%, buy-imbalance|1.3%, sell-imbalance|0.8%, no-imbalance|0.0%</td></tr>
<tr><td>Revenues</td><td>8.17%</td><td>57.75%</td><td>revenue|63.4%, revenue-guidance|23.7%, same-store-sales|6.4%, revenue-volume|5.0%, same-store-sales-guidance|1.3%, revenue-estimate|0.3%</td></tr>
<tr><td>Products-services</td><td>6.41%</td><td>64.16%</td><td>business-contract|49.5%, product-release|28.3%, regulatory-product-approval|4.6%, clinical-trials|4.3%, award|2.0%, supply|1.9%, product-price|1.8%, government-contract|1.5%, product-discontinued|1.0%, product-recall|0.9%, clinical-trials-patient-enrollment|0.7%, supply-guidance|0.5%, regulatory-product-application|0.5%, market-entry|0.5%, demand|0.4%, market-share|0.3%, orphan-drug-designation|0.2%, business-combination|0.2%, product-outage|0.2%, fast-track-designation|0.1%, demand-guidance|0.1%, regulatory-product-review|0.1%, grant|0.1%, market-guidance|0.1%, product-resumed|0.1%, project-abandoned|0.0%, regulatory-product-warning|0.0%, production-outlook|0.0%, product-catastrophe|0.0%</td></tr>
<tr><td>Labor-issues</td><td>5.73%</td><td>69.89%</td><td>executive-appointment|70.5%, executive-resignation|17.6%, layoffs|5.0%, executive-salary|3.4%, hirings|0.8%, executive-compensation|0.5%, executive-death|0.5%, executive-firing|0.5%, union-pact|0.4%, executive-shares-options|0.3%, workers-strike|0.3%, executive-health|0.1%, workforce-salary|0.1%, executive-scandal|0.1%</td></tr>
<tr><td>Analyst-ratings</td><td>4.58%</td><td>74.47%</td><td>analyst-ratings-change|71.7%, analyst-ratings-set|25.5%, analyst-ratings-history|2.8%</td></tr>
<tr><td>Equity-actions</td><td>4.37%</td><td>78.84%</td><td>trading[a]|32.4%, buybacks|15.7%, public-offering|12.9%, ownership|9.1%, expenses-guidance|4.3%, expenses|3.9%, capex-guidance|3.2%, fundraising|2.7%, reorganization|2.4%, investment|2.3%, stock-splits|1.2%, initial-public-offering-price|1.1%, private-placement|1.1%, initial-public-offering|1.1%, capex|0.9%, spin-off|0.8%, shelf-registration|0.7%, reverse-stock-splits|0.7%, shareholder-rights-plan|0.7%, savings-guidance|0.6%, rights-issue|0.5%, name-change|0.4%, reorganization-unit|0.4%, capital-increase|0.4%, going-private|0.2%, savings|0.1%, initial-public-offering-unit|0.1%, initial-public-offering-issuance|0.0%, ipo-regulatory-scrutiny|0.0%, ipo-regulatory-approval|0.0%</td></tr>
<tr><td>Acquisitions-mergers</td><td>4.16%</td><td>83.00%</td><td>acquisition|56.3%, unit-acquisition|24.1%, stake|9.0%, merger|6.6%, acquisition-regulatory-approval|2.0%, merger-regulatory-approval|1.9%, unit-acquisition-regulatory-approval|0.1%, acquisition-regulatory-scrutiny|0.1%, merger-regulatory-scrutiny|0.0%, unit-acquisition-regulatory-scrutiny|0.0%</td></tr>
<tr><td>Stock-prices</td><td>2.51%</td><td>85.51%</td><td>stock[b]|100.0%</td></tr>
</table>

[a] The 'trading' type includes listing, delisting, trading-halt, and trading-resume

[b] The 'stock' type include the stock price increases or decreases that make news headlines

Table 13: Continued

Panel B: Comparing media coverage of different topics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Earnings	-0.024***										-0.019***	-0.019***			
	(-14.17)										(-8.41)	(-8.45)			
Insider trading		0.029									0.031				
		(0.36)									(0.40)				
Order imbalance			0.044								0.047				
			(1.00)								(1.04)				
Revenue				-0.023***							-0.006***	-0.006***			
				(-14.25)							(-2.72)	(-2.77)			
Product					-0.005***						0.000				
					(-3.17)						(0.13)				
Labor						-0.004***					-0.000				
						(-2.92)					(-0.31)				
Analyst rating							0.013				0.018				
							(0.52)				(0.73)				
Equity action								-0.021***			-0.013***	-0.013***			
								(-12.17)			(-7.87)	(-7.40)			
Merger & Acquisition									-0.011***		-0.009***	-0.009***			
									(-7.50)		(-5.74)	(-5.64)			
Stock price										-0.018***	-0.005**	-0.006***			
										(-8.16)	(-2.58)	(-2.78)			
Significant-5 topics													-0.022***		-0.020***
													(-19.02)		(-14.23)
All topics														-0.018***	-0.004***
														(-15.44)	(-2.71)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Adj.R^2$	3.1%	2.9%	2.9%	3.0%	2.8%	2.9%	2.9%	2.9%	2.9%	2.9%	3.3%	3.2%	3.2%	3.0%	3.2%
Avg.#Firms	1472	1472	1472	1472	1472	1472	1472	1472	1472	1472	1472	1472	1472	1472	1472

Table 14: Two-Stage Instrumental Variable Regression

This table presents the panel regression results that predict straddle return with instrumental variable for media coverage. In the first stage, *Residual Media Coverage* is regressed on the instrumental variable, with full set of control variables and year-month fixed effect. In the second stage, the one-period-ahead straddle returns are regressed on the fitted value of residual media coverage from the first stage. The first instrument *DJ Location* is a categorical variable that takes a value of two (one) if the firm headquarter is located in the same city (state) as one of Dow Jones' offices and zero otherwise. The second instrument *Travel Time* is the minimum value of the number of minutes taken for trips between the headquarter of a firm and Dow Jones offices. T-statistics shown in parentheses are based on standard errors clustered at firm level.

	(1) 1st stage Residual Coverage	(2) 2nd stage R_{t+1}	(3) 1st stage Residual Coverage	(4) 2nd stage R_{t+1}
DJ Location	0.093***			
	(5.43)			
Travel Time			-0.078***	
			(-5.57)	
$\widehat{\textit{Residual Coverage}}$		-0.048***		-0.050***
		(-3.20)		(-3.10)
Controls	Yes	Yes	Yes	Yes
Year-Month Fixed Effect	Yes	Yes	Yes	Yes
1st stage Cragg-Donald F-stats	1930.10		1368.71	
1st stage Kleibergen-Paap F-stats	29.45		31.04	
Observations	364,660	364,660	342,881	342,881

Table 15: Positive Media Coverage VS Negative Media Coverage

This table studies if the media coverage effect is driven by positive news or negative news. I separate news articles into positive and negative types by the news sentiment score, and then calculate the corresponding two types of residual media coverage by replacing the number of news articles in Equation (1) with number of positive or negative news articles. Panel A shows the means and t-statistics from univariate quintile sorts by the two types of residual media coverage. Panel B reports the coefficients of Fama-Macbeth regressions of one-period-ahead straddle returns on the two types of residual media coverage and other characteristics. I winsorize all independents at 0.5% and 99.5% percentile on a monthly basis. T-statistics in parentheses are Newey-West adjusted with 5 lags.

Panel A: Portfolio sorts

	Low	2	3	4	High	High-Low
Residual of Positive Media Coverage	-2.37	-3.80	-4.56	-5.56	-7.36	-4.99***
	(-2.73)	(-4.42)	(-5.09)	(-6.27)	(-8.46)	(-17.14)
Residual of Negative Media Coverage	-3.15	-3.63	-5.15	-5.46	-6.27	-3.12***
	(-3.43)	(-4.29)	(-6.01)	(-6.03)	(-7.16)	(-10.91)

Panel B: Fama-Macbeth regression in horse-race form

	(1)	(2)	(3)	(4)
Residual of Positive Media Coverage	-0.019***		-0.017***	-0.016***
	(-17.33)		(-16.17)	(-16.78)
Residual of Negative Media Coverage		-0.014***	-0.007***	-0.005***
		(-11.79)	(-5.90)	(-4.88)
Option Momentum				0.045***
				(9.65)
Ln(Market Cap)				-0.003*
				(-1.71)
Idiosyncratic Vol.				-0.100***
				(-7.73)
IV term spread				0.163***
				(5.50)
IV smirk				-0.000
				(-0.00)
IV-HV				-0.157***
				(-15.63)
VOV				-0.139***
				(-5.26)
$Adj.R^2$	0.2%	0.2%	0.3%	3.1%
$Avg.\#Firms$	1543	1543	1543	1472

Table 16: Media Coverage VS Media Sentiment

This table compares the predictive ability between media coverage and media sentiment. It reports the coefficients of Fama-Macbeth regressions of one-period-ahead straddle returns on residual media coverage, media sentiment measures, and other characteristics. Average CSS is the average sentiment score of news articles in the past one month ending at option date, using the composite sentiment score (CSS) in RavenPack. (Pos-Neg)/Total is the difference of news article numbers with positive tones (CSS>50) and negative tones (CSS<50), divided by total number of news articles in the past one month ending at option date. I winsorize all independents at 0.5% and 99.5% percentile on a monthly basis except for the dummy variable. T-statistics in parentheses are Newey-West adjusted with 5 lags.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Residual Media Coverage	-0.019***		-0.023***	-0.021***		-0.023***	-0.021***
	(-14.99)		(-15.29)	(-16.15)		(-15.06)	(-16.02)
Average CSS		0.000	-0.001	-0.002***			
		(0.62)	(-1.10)	(-4.05)			
(Pos-Neg)/Total					-0.003	-0.007***	-0.010***
					(-1.19)	(-3.07)	(-4.62)
Option Momentum				0.043***			0.043***
				(9.42)			(9.43)
Ln(Market Cap)				-0.002			-0.002
				(-1.50)			(-1.59)
Idiosyncratic Vol.				-0.095***			-0.095***
				(-7.14)			(-7.06)
IV term spread				0.164***			0.164***
				(5.67)			(5.72)
IV smirk				0.003			0.003
				(0.26)			(0.24)
IV-HV				-0.160***			-0.160***
				(-14.75)			(-14.77)
VOV				-0.133***			-0.133***
				(-4.96)			(-4.93)
$Adj.R^2$	0.3%	0.1%	0.4%	3.2%	0.0%	0.4%	3.2%
$Avg.\#Firms$	1543	1450	1450	1387	1450	1450	1387

Table 17: Alternative News Windows

This table shows the main results of residual media coverage during alternative windows, which is the regression residuals of media coverage measured during different past windows on the four firm characteristics in Equation (1) (firm size, NASDAQ and S&P 500 dummies, and analyst coverage). Panel A presents means and t-statistics from univariate quintile sorts by residual media coverage, while Panel B reports the Fama-Macbeth regression coefficients on residual media coverage, with and without control variables. T-statistics in parentheses are Newey-West adjusted with 5 lags.

Panel A: Univariate portfolio sort

	Low	2	3	4	High	High-Low
Past 3 months	-3.47	-4.24	-4.53	-5.17	-6.24	-2.77***
	(-3.88)	(-4.66)	(-5.01)	(-5.97)	(-7.60)	(-8.68)
Past 6 months	-3.92	-4.16	-4.76	-4.99	-5.82	-1.90***
	(-4.23)	(-4.56)	(-5.52)	(-5.86)	(-7.00)	(-6.28)
Past 9 months	-4.09	-4.38	-4.60	-5.00	-5.58	-1.49***
	(-4.37)	(-4.72)	(-5.47)	(-5.90)	(-6.64)	(-4.87)
Past 12 months	-4.29	-4.50	-4.58	-4.85	-5.43	-1.15***
	(-4.47)	(-4.88)	(-5.65)	(-5.73)	(-6.41)	(-3.65)

Panel B: Fama-Macbeth regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Past 3 months	-0.011***	-0.008***						
	(-7.59)	(-6.13)						
Past 6 months			-0.006***	-0.004***				
			(-4.75)	(-3.15)				
Past 9 months					-0.004***	-0.002*		
					(-3.27)	(-1.78)		
Past 12 months							-0.003**	-0.001
							(-2.11)	(-0.83)
Control	No	Yes	No	Yes	No	Yes	No	Yes
$Adj.R^2$	0.1%	2.9%	0.1%	2.9%	0.1%	2.9%	0.0%	2.9%
$Avg.\#Firms$	1543	1472	1543	1472	1543	1472	1543	1472

Table 18: Subperiod Analysis

This table reports the Fama-Macbeth regression results during different subperiods in which the one-period-ahead straddle return is regressed on the residual media coverage and control variables. For the division of sample period, I look at first/second half of the sample period, period with low/high sentiment of Baker and Wurgler (2006), low/high uncertainty by VIX, and normal/crisis period. All the independent variables are winsorized at 0.5% and 99.5% levels on a monthly basis. T-statistics in parentheses are Newey-West adjusted with 5 lags.

	First vs. Second half		Sentiment		VIX		Normal vs. Crisis	
	First (1)	Second (2)	Low (3)	High (4)	Low (5)	High (6)	Normal (7)	Crisis (8)
Resid. Media Coverage	-0.014***	-0.021***	-0.020***	-0.015***	-0.021***	-0.014***	-0.018***	-0.009***
	(-12.07)	(-10.68)	(-11.87)	(-10.36)	(-12.95)	(-10.04)	(-14.59)	(-3.49)
Option Momentum	0.061***	0.030***	0.035***	0.056***	0.041***	0.050***	0.044***	0.068***
	(10.04)	(5.62)	(5.60)	(8.79)	(5.49)	(8.85)	(9.02)	(6.51)
Ln(Market Cap)	-0.005**	-0.001	-0.005***	-0.000	-0.002	-0.003	-0.003*	-0.003
	(-2.01)	(-0.36)	(-2.76)	(-0.05)	(-0.92)	(-1.54)	(-1.79)	(-0.35)
Idiosyncratic Vol.	-0.094***	-0.109***	-0.090***	-0.112***	-0.152***	-0.051***	-0.107***	-0.034**
	(-4.28)	(-7.49)	(-7.06)	(-5.05)	(-8.00)	(-3.98)	(-7.71)	(-2.64)
IV term spread	0.275***	0.059**	0.110***	0.224***	0.121***	0.213***	0.169***	0.143
	(7.16)	(2.04)	(2.90)	(5.37)	(3.30)	(5.52)	(5.47)	(1.72)
IV smirk	-0.006	0.004	-0.005	0.002	-0.011	0.009	0.003	-0.049
	(-0.35)	(0.40)	(-0.33)	(0.16)	(-1.10)	(0.53)	(0.32)	(-0.98)
IV-HV	-0.175***	-0.138***	-0.149***	-0.164***	-0.134***	-0.179***	-0.155***	-0.171***
	(-10.02)	(-15.89)	(-15.30)	(-9.11)	(-16.64)	(-10.91)	(-14.90)	(-4.83)
VOV	-0.121**	-0.160***	-0.148***	-0.134***	-0.204***	-0.077*	-0.156***	0.043
	(-2.45)	(-7.78)	(-5.59)	(-3.07)	(-7.36)	(-1.78)	(-5.86)	(0.48)
$Adj.R^2$	3.4%	2.7%	2.8%	3.3%	2.4%	3.7%	2.9%	5.3%
$\#Months$	126	126	126	126	126	126	232	20

Internet Appendix

In this appendix, I provide some additional results to complement the findings in the main text. Table A1 presents the double sort result by residual media coverage and control variables.

Table A1: Double Sorts by Residual Media Coverage and Control Variables

This table presents the results of bivariate sort by *Residual Media Coverage* and control variables. On the third Friday of each month, I first sort firms into quintiles based on the control variable, and then within each group, firms are further sorted into quintiles based on *Residual Media Coverage*. The returns for each media coverage group is the averages across the five control groups. The table reports the means and t-statistics for each media coverage group, as well as the high-minus-low portfolio. All returns are monthly and in percentage. T-statistics in parentheses are Newey-West adjusted with 5 lags.

Control for #	Low Coverage	2	3	4	High Coverage	High-Low	T-stats.
Option Momentum	-2.40	-3.39	-4.62	-5.62	-7.43	-5.04***	(-16.33)
Market Capitalization	-2.52	-3.56	-4.56	-5.73	-7.29	-4.77***	(-14.68)
Idiosyncratic Vol.	-2.59	-3.80	-4.68	-5.72	-6.85	-4.26***	(-13.98)
IV term spread	-2.31	-3.44	-4.68	-5.74	-7.49	-5.18***	(-16.75)
IV smirk	-2.39	-3.60	-4.64	-5.64	-7.39	-5.00***	(-15.49)
IV-HV	-2.15	-3.33	-4.60	-5.78	-7.54	-5.39***	(-16.42)
VOV	-2.54	-3.58	-4.69	-5.63	-7.20	-4.65***	(-16.14)