Center of Volume Mass: Does Aggregate Option Market Activity Predict Stock Returns?*

Gennaro Bernile, *Singapore Management University*

Fei Gao, *Singapore Management University*

Jianfeng Hu, *Singapore Management University*

March 2017

*We are grateful to Ekkehart Boehmer, Andrey Ermolov, Fangjian Fu, Dashan Huang, Roger Loh, Steve Raymar, Qing Tong, Baolian Wang, Rong Wang, Yihui Wang, and seminar participants at Fordham University, NYU Shanghai, and Singapore Management University for comments.

Gennaro Bernile can be reached at: *Phone:* +65 6808 5478; *email:* `gbernile@smu.edu.sg`. Fei Gao can be reached at *email:* `fei.gao.2013@pbs.smu.edu.sg`. Jianfeng Hu can be reached at: *Phone:* +65 6808 5477; *email:* `jianfenghu@smu.edu.sg`. Please address all correspondence to Jianfeng Hu at: Singapore Management University, Lee Kong Chian School of Business, 50 Stamford Road, Level 4, Room 4073, Singapore 178899

Center of Volume Mass: Does Aggregate Options Market Activity Predict Stock Returns?

Abstract – We uncover a novel stock return predictor from the options market, the volume-weighted strike-spot price ratio ($VWKS$) across all traded option contracts. High (low) $VWKS$ indicates that the mass of options volume on an underlying stock centers at the out-of-the-money region of call (put) options. Empirically, $VWKS$ has positive and robust predictive ability for underlying returns after controlling for a long list of variables including known return predictors from the options market, stock illiquidity, and past stock returns, and has more persistent and stronger predictive power for stocks with higher information asymmetry and arbitrage costs. We also find that $VWKS$ exhibits abnormal run-ups and becomes more informative before permanent but not transitory price jumps, suggesting that options traders exploit only fundamental information.

Keywords: Options; Volume; Information; Center of mass

JEL Codes: G11, G12, G14, G17

1 Introduction

The relation between total options trading volumes and underlying stock price dynamic has been well documented by Roll, Schwartz, Subrahmanyam (2009, 2010), Johnson and So (2012), and Ge, Lin, and Pearson (2016). However, the heterogeneity in trading activity across options contracts is largely under-investigated in the literature. Prior studies such as Easley, O'Hara, and Srinivas (1998) and Pan and Poteshman (2006) acknowledge only the difference between call and put options. An important feature of an option contract, the moneyness, is often analyzed at a crude granularity by using in-the-money (ITM), at-the-money (ATM), and out-of-the-money (OTM) regions in prior studies. In this article, we examine the options trading volume distribution across all available contracts with different moneyness to extract information about future stock prices. To the best of our knowledge, such analysis is novel to the financial literature.

We use a simple measure as our main variable to capture the variation in the options volume distribution across individual stocks, the volume-weighted strike-spot ratio. We first calculate the ratio of an option's strike price (K) and the underlying stock price (S) for each option at the end of each trading day. The K/S ratio is a proxy for the option's moneyness. Call (put) options are out-of-the-money when K/S is above (below) one. We then calculate the weighted K/S of all options on the same underlying stock by the number of lots traded on the option contract on the same day. After normalizing the variable by subtracting one, we obtain the volume-weighted strike-spot ratio, which we term $VWKS$. $VWKS$ is high when more OTM calls or ITM puts are traded, whose K/S is high, and vice versa. Intuitively, $VWKS$ captures the center of mass in the options volume distribution along strike prices.

The variation in the location of this place in the cross section can potentially provide important information about future stock prices because sophisticated investors with advanced information may prefer the options market to the underlying market as argued by Black (1975), Figlewski and Webb (1993), and Easley, O'Hara, and Srinivas (1998) among others. An informed trader has multiple trading strategies in the options market for the same type of information given usually hundreds of options contracts available. For example, a trader with positive information about a stock can either buy call options or sell put options on the stock at different strike prices and maturities. The optimal decision of the informed trader depends on the inherent leverage provided by the options contracts, margin requirement, options liquidity, and information duration and precision, as well as subsequent litigation risk if the informed trader is also an insider. In this article, we do not address the optimization problem of the informed. Rather, we rely on some stylized facts to relate $VWKS$ to informed trading. The first thing to note is that OTM options provide higher leverage and are more liquid than ATM and ITM options, hence more attractive to the informed. Secondly, writing options can require more capital due to options exchanges' margin requirement than purchasing options to gain the same exposure to the underlying stock price.[1] Therefore, informed trading in the options market is most likely to occur as purchasing OTM options. In a recent study, Augustin, Brenner, Grass, and Subrahmanyam (2016) also reach similar conclusions from both theoretical and empirical analysis. Practically, informed traders are likely to use a portfolio of options rather than a single contract due to liquidity and revelation concerns. Nonetheless, profitability as the first order of consideration should posit informed options trading activity to center at the OTM region according to the type of information, i.e. at high strike prices when the information is positive and vice versa. If informed traders are active in the options market, we expect that the center of options volume mass is skewed toward the optimal trading region of the informed, and is informative about future stock returns. Specifically, $VKWS$ should rise to where call options are out-of-the-money before positive information shocks and fall to where put options are out-of-the-money before negative shocks, resulting in a positive relation between $VWKS$ and future underlying returns. Alternatively, if informed trading in the options market is trivial and the $VWKS$ captures only noisy dynamic in the trading activity or hedging demand, then $VWKS$ should not have permanent predictive ability for future stock returns.[2]

[1] See Chicago Board Options Exchange's (CBOE) margin requirement at *http://www.cboe.com/framed/pdfframed.aspx?content=/learncenter/pdf/margin2-00.pdf.*

[2] Options trading can also be motivated by volatility dynamics or expected jumps. Classic volatility trading strategies such as straddles and strangles do not create a skewed volume distribution without

Admittedly, the proposed measure of $VWKS$ is a noisy proxy for informed trading activity in the options market. The biggest challenge comes from the fact that there are both call and put options at the same strike price, and the trading volume reflects only aggregate activity but not necessarily directions of risk taking. For example, when $VKWS$ is high, it indicates that trading of high strike price options is more active. It could be due to purchase of OTM calls by informed traders with good news. But it is also possible that sale of OTM calls and both purchase and sale of ITM puts cause the skewness in the volume distribution, hence introducing measurement error to our proxy and biasing the results against our story. Empirically, this concern is mitigated by the evidence in Hu (2014) that customers are net buyers of OTM options but net sellers of ITM options.[3] As a result, trading volumes of the put and call options with the same strike price are likely to point to the same risk exposure to the underlying stock price. For example, a strike price higher than the spot price leads to an OTM call and an ITM put. Unconditionally, the net demand for the call is likely to be positive in this moneyness region, reflecting customers seeking positive *delta* exposure by purchasing calls. At the same time, the net demand for the put is likely to be negative because the put is ITM. The sale of put options by customers again leads to positive *delta* exposure, consistent with the activity in the call. Similarly, options volume on a strike price lower than the spot price is likely to represent selling of *delta* of the underlying stock.

To test whether the center of options volume mass contains stock price information, we compute $VWKS$ at daily level from January 1, 1996 to August 31, 2013. On average, our sample contains 1400 stocks each trading day. A value-weighted investment strategy that are long stocks of the highest $VWKS$ decile and short stocks in the lowest decile

a companion view of the price movement direction. Our normalization of $VWKS$ sufficiently excludes the impact of such trading. Trading on price jumps, on the other hand, is largely consistent with our interpretation of $VWKS$ as a proxy for net demand of the underlying stock because it is also directional.

[3] Hu's (2014) results are based on signed options tick data from the Options Price Reporting Authority across all options exchanges in the US in a recent sample period from 2008 to 2010. The net demand is more balanced between 1990 and 2001 as reported by Pan and Poteshman (2006) using daily aftermarket data from the CBOE. However, Pan and Poteshman find the same pattern of net demand as in Hu (2014) for position-opening transactions, which are more informative about future returns. Between 1996 and 2001, Garleanu, Petersen, and Poteshman (2009) report negative net demand for all types of options on single-name stocks using the same data from CBOE though.

generates an annualized abnormal return of 46.1% with a t-statistic of 7.25. The alpha with respect to Fama-French (1993) factors as well as the liquidity and momentum factors reaches 46.5% and is statistically significant at the 1% level. We find this trading strategy is also profitably at the weekly frequency. We confirm that the return predictability is not driven by other known predictors from the options market. In a double sorting analysis, we examine the profitability of the $VWKS$ strategy controlling for the effects from the put-call ratio similar to Pan and Poteshman (2006), option-to-stock ratio of Roll, Schwartz, Subrahmanyam (2010) and Johnson and So (2012), deviation from the put-call parity as in Cremers and Weinbau (2010), implied skewness as in Xing, Zhang, and Zhao (2010), and implied volatility as in Guo and Qiu (2014). We find that in all of the conditioning quintile portfolios, the abnormal returns as well as the five-factor alphas from the $VWKS$ strategy range from 4.8 bp to 20.5 bp a day with t-statistics all above 2.15. To control for the effects of all options market return predictors at the same time as well as well-known stock return reversals and liquidity effects from the bid-ask spread and trading volume turnover, we run Fama-MacBeth (1973) regressions at the daily level with five lags of explanatory variables. We find that $VWKS$ has positive coefficient estimates at all five lags. Even with the full set of control variables, the coefficient of the first lag of $VWKS$ is 0.053 with a t-statistic of 3.78. A one-standard deviation increase in $VWKS$ is associated with an increase of 0.6 bp in the underlying stock price on the following day, and 1.9 bp in the underlying stock price on the following five days. Moreover, the coefficients of $VWKS$ are statistically significant at the 5% level at the second and third lags, and at the 10% level at the fifth lag. To gauge the cumulative price impact in a longer horizon, we use the five-day moving averages (MA) of independent variables in the regression and find that $VWKS$ has a positive and permanent price impact with a coefficient estimate of 0.201 on the MA term and a t-statistic of 6.01. In robustness tests, we find that $VWKS$ is able to predict both raw returns and mid quote returns, and the predictability is qualitatively the same when we use log transformation to normalize the variable, use option deltas to measure moneyness instead of K/S, or calculate $VWKS$ using lagged stock price to eliminate the effect of return reversal. The strong return predictive ability of $VWKS$ is consistent with our conjecture that $VWKS$ captures informed trading in the options market.

To better understand the nature of the return predictability, we analyze $VWKS$ calculated using different types of options and subsamples conditioning on stock characteristics. When we separate call and put options, we find that $VWKS$ calculated using both call and put options contain valuable information. The positive coefficient estimates on both call $VWKS$ and put $VWKS$ are consistent with the information explanation. When separating options by maturity, we find that both short-term (expiring within 30 days) and long-term (expiring beyond 30 days) options are informative as $VWKS$ calculated using either maturity group is able to predict subsequent returns. We also find that the predictive ability of $VWKS$ comes from both positive and negative $VWKS$ although the pricing effect is stronger from positive $VKWS$. When exploring nonlinear pricing impact of $VKWS$ by including a singed quadratic term into the regression, we find that the quadratic term has strong predictive ability that subsumes the predictive ability of $VKWS$, suggesting that the effect of $VWKS$ mainly comes from the tails. Conjecturing that informed trading should be more common for less transparent stocks, we divide our sample into subsamples based on measures of information asymmetry including firm size, analyst coverage, book-to-market ratio, probability of informed trading (PIN), idiosyncratic volatility, illiquidity, and institutional ownership as well as the time period. Our subsample analysis generates two important findings. First, the $VWKS$ effect is statistically significant throughout all the subsamples. Second, we find that $VWKS$ has more persistent predictive power for stocks with higher level of information asymmetry, i.e. those with low market capitalization, little analyst coverage, high options volume, large PIN, high volatility, low liquidity, and low institutional ownership as well as in the first half of the sample period. These results reinforce our claim that $VWKS$ is an information measure.

There might be concerns that the documented return predictability invites alternative explanations. To establish a concrete and unambiguous link between our variable and information flow, we conduct an event study of $VWKS$ around earnings announcements, non-earnings 8-K filings, and price jumps due to other reasons. While earnings announcements are usually pre-scheduled and have been examined extensively in the literature, non-earnings 8-K filings are mostly unscheduled information events. The permanent price jumps represent other information shocks not captured by the mandatory SEC filings, and are largely unscheduled. Therefore, our sample of information shocks is more comprehensive than those typically examined in such event studies. Our results show that $VWKS$ significantly increases (decreases) before positive (negative) information events, consistent with the hypothesis that $VWKS$ contains private information. Moreover, the predictive ability becomes greater before permanent price jumps too. Finally, when we examine transitory price jumps as a placebo test, we do not find similar pre-event run-ups of $VWKS$ or enhanced predictive ability, suggesting that either options traders are not better at predicting such transitory jumps than stock traders, or options traders voluntarily do not exploit such opportunities.

The contributions of this paper to the financial literature are mainly twofold. First, the center of options volume mass across strike prices is a new variable to measure private information from the options market. We find that this variable has strong predictive ability for future stock returns not subsumed by known return predictors. Second, we show that the predictive ability from options trading is due to arrival of fundamental information. Although many existing studies have also used scheduled events such as earnings announcements, we are among the first finance studies that use 8-K filings to identify unscheduled information flow. Moreover, our analysis using permanent and transitory price jumps directly compares the pricing effects of $VWKS$ in a real information shock and a placebo test. Because $VWKS$ responds to only corporate events and permanent price jumps but not transitory jumps, the result suggests that options traders exploit only fundamental information, which is not documented in existing studies.

The rest of the paper is organized as follows. Section 2 reviews the literature on the price impact of option trading and develops testable hypotheses. Section 3 described the data and sample selection. Section 4 reports the empirical results. Section 5 concludes.

2 Related literature and hypotheses

The financial literature on lead-lag analysis across the options market and underlying stock market has two important findings in consensus. First, options prices have limited contribution to price discovery compared to the underlying stock price as evidenced by e.g., Stephan and Whaley (1990), Chan, Chung, and Johnson (1993), Chan, Chung, and Fong (2002), Chakravarty, Gulen, and Mayhew (2004), and Muravyev, Pearson, and Broussard (2013). Second, options volumes contain valuable information about future stock returns and volumes as documented by Anthony (1988), Easley, O'Hara, and Srinivas (1998), Pan and Poteshman (2006), Johnson and So (2012), Hu (2014), and Ge, Lin, and Pearson (2016) among others. These two seemingly inconsistent findings can be reconciled by the derivative nature of options. On the one hand, options can be priced according to the non-arbitrage framework started from Black and Scholes (1973) and Merton (1973), hence endogenous to the underlying asset price. While any change in the underlying price can lead to changes in quoted options prices, options price updates are still governed by pricing bounds derived from the non-arbitrage rule even if private information reaches the options market before the stock market. Therefore, it is not surprising that the options prices lag the stock price on average. On the other hand, the options market can attract informed traders because of the higher leverage embedded in derivative contracts, relaxing of short-sale constraints, and ability to hide behind multiple contracts available. Options volumes can be informative as a result of informed trading in the options market. In the stock market without perfect liquidity, informed stock volume will lead to informative stock price inevitably such as in Kyle (1985) and Glosten and Milgrom (1985). However, the same is not true in the options market because the options market makers are no longer exposed to information risk after they hedge the options positions. Therefore, even if options trading can be motivated by private information, options prices are not necessarily informative but options volumes.

Unlike most studies that use intraday data to compute options order flow, Roll, Schwartz, Subrahmanyam (2009) calculate a total options volume to stock volume ratio as a proxy for information efficiency and show that this O/S ratio is positively related to firm value. Roll, Schwartz, Subrahmanyam (2010) and Johnson and So (2012) find that the same ratio contains price information about the underlying stock. A desirable feature of the O/S ratio is that the calculation requires only daily aftermarket data, hence avoiding the use of intraday options data. However, by treating all options trades equally, the O/S ratio fails to capture the heterogeneity in options contracts while such heterogeneity can be important in the options market. Although an option contract is defined by many parameters, we focus on its strike price in this study given its close relation to an option's moneyness. In the US equity options market, the strike price (K) interval is normally set to be 2.50 points for stocks under \$25, 5 points for stocks selling over \$25 per share, and 10 points (or greater) is acceptable for stocks over \$200 per share. K contains little information by itself since it is preset by the exchange. However, the volume distribution across options contracts with different K can be informative. As argued earlier, informed traders should prefer to purchase OTM call (put) options upon acquiring positive (negative) information about the underlying stock. If the private information is strong enough, the informed trader can place more aggressive bets on deep OTM options. As a result of their trading, the distribution of options volumes will be skewed toward the region that informed traders prefer. In other words, the center of volume mass approaches the optimal strike price for the informed traders. Therefore, we use the volume-weighted strike price to extract private information embedded in the options market. To normalize the strike prices across different stocks, we divide the strike price by the spot price (S) of the underlying stock and subtract one from the ratio. We term the normalized volume-weighted strike-spot price ratio as $VWKS$. Given its potential relation to informed trading in the options market, we have the following hypothesis.

Hypothesis 1. The volume-weighted strike-spot price ratio, $VWKS$, has a positive and permanent price impact on the underlying stock.

To test the above hypothesis, we need to control for the pricing effects of known return predictors from the options market at the daily or longer horizons. In addition to the O/S ratio discussed earlier, Pan and Poteshman (2006) find that a put-call ratio (PC) negatively predicts future returns because informed traders are likely to purchase calls ahead of good news and puts ahead of bad news. Although intraday studies show that options prices have marginal contribution to price discovery, several studies find that the options implied volatility surface can exhibit significant return predictive ability at longer horizons. For example, Cremers and Weinbaum (2010) find that a measure of deviation from put-call parity (DEV) positively predicts returns at the weekly horizon and they interpret the results as evidence of mispricing in the stock market. Xing, Zhang, and Zhao (2010) find that an options implied skewness measure ($SKEW$) predicts the underlying stock return up to six months. Guo and Qiu (2014) confirm the idiosyncratic volatility puzzle using the options implied volatility ($IVOL$). In our empirical tests, we control for all of these options market variables as well as stock return reversals, bid-ask spread, and turnover ratio. Since we mainly use daily data from the Options Metrics database to form a long and representative sample in our empirical analysis, we do not include other volume variables constructed from intraday data such as options order flow in our analysis.

In the cross section of stocks, the effectiveness of an information measure can depend on a stock's information environment. Holding everything else the same, informed trading is less likely to happen to transparent stocks. Therefore, we have the following hypothesis.

Hypothesis 2. The volume-weighted strike-spot price ratio, $VWKS$, has stronger return predictive ability for stocks with higher levels of information asymmetry.

Our empirical tests of this hypothesis rely on several proxies for information asymmetry. It is well-known that many return predictors are less effective for large stocks and stocks well covered by analysts. Hong, Lim and Stein (2000) find that even for momentum strategies, the profitability declines sharply with firm size and analyst coverage. Fang and Peress (2009) suggest that the breadth of information dissemination affects stock returns. Their results are more pronounced among small stocks and stocks with high individual ownership, low analyst following, and high idiosyncratic volatility. Spiegel and Wang (2005) find that stock returns increase with the level of idiosyncratic risk and decrease with a stock's liquidity. Chen, Hong, and Stein (2002), D'Avolio (2002), Asquith, Pathak, and Ritter (2005) and Nagel (2005) show that institutional ownership signals short-sale constraints and thus affects stock returns. We examine the predictive power of $VWKS$ in subsamples based on all the firm characteristics mentioned above to separate the effects on transparent and opaque stocks.

The return prediction test is important to understand the effect of options trading. However, the results are often subject to alternative explanations. For example, Cremers and Weinbaum (2010) and Goncalves-Pinto et al. (2016) attribute the predictive ability of options variables to mispricing of the underlying stocks at longer horizons. To relate the information proxies from the options market to fundamental information flow, prior studies often examine the pricing effects around earnings announcements as in Pan and Poteshman (2006), Roll, Schwartz, Subrahmanyam (2010), Xing, Zhang, and Zhao (2010), Johnson and So (2012), and Ge, Lin, and Pearson (2016). While such analysis relies on unambiguous information shocks from earnings announcements, the anticipated nature of such announcements can greatly increase expected volatility in the event window and potentially affect the relation between stock returns and options trading activities as shown by Cremers, Fodor, and Weinbaum (2015). In this regard, testing the behavior of an information proxy around unscheduled corporate events can significantly complement the analysis. Such studies of informed options trading before unscheduled corporate events include Cao, Chen, and Griffin (2005), Chan, Ge, and Lin (2015), and Augustin, Brenner, and Subrahmanyam (2016) for mergers and acquisitions, Hayunga and Lung (2014) for analyst revisions, Augustin, Brenner, Hu, and Subrahmanyam (2015) for spinoffs, Gharghori, Maberly, and Nguyen (2015) for stock splits, and Ge, Hu, Humphery-Jenner, and Lin (2016) for bankruptcies. To include all types of unscheduled announcements in the analysis, we choose to use the SEC's 8-K filings unrelated to earnings announcements. The majority of such 8-K forms are not pre-scheduled. Therefore it can serve as a more comprehensive set of information events where informed trader are likely to exploit their information advantage, as documented by Thompson and Sale (2003), Brochet (2010), Skaife, Veenman and Wangerin (2013), and Zhao (2016). There could be other information shocks not captured by earnings and 8-K filings. Based on the notion that significant information should lead to price adjustment, we use extreme price jumps unrelated to earnings news and 8-K filings to identify additional information shocks. We divide these jumps into transitory jumps, which reverse quickly, and permanent jumps otherwise. As argued by Boehmer and Wu (2013), by definition, permanent jumps involve new information while transitory jumps do not. Lee and Mykland (2008) show that jumps are associated with scheduled earnings announcements and other company-specific news events. Jiang and Yao (2013) find that jumps account for portfolio value premium in equity market. Although fundamental information is associated with permanent jumps only, we choose to examine transitory jumps as well because these jumps serve as a placebo test. If the predictive ability of $VWKS$ is due to informed trading, we expect that $VWKS$ will exhibit abnormal behavior and become more informative before permanent jumps but not transitory jumps. Therefore, we have the following hypothesis.

Hypothesis 3. The volume-weighted strike-spot price ratio, $VWKS$, becomes larger ahead of earnings announcements, 8-K filings and permanent jumps in the same direction as the subsequent announcement return, and the return predictability strengthens.

Similar to our approach, several recent studies also analyze options trading around unscheduled events. For example, Jin, Livnat, and Zhang (2012) examine the predictive ability of options implied skewness and deviation from put-call parity around unscheduled events in the Key Development database of Capital IQ. Cremers, Fodor, and Weinbaum (2015) use Reuters news release to study the choice of options traders before scheduled and unscheduled events and similar analysis is performed by Augustin, Brenner, Grass, and Subrahmanyam (2016) using Dow Jones news release processed by RavenPack. Unlike those studies, we use the 8-K filings and price jumps in this paper. One advantage of using the 8-K filings and price jumps is that it covers the entire sample period of Option Metrics, from which we extract the options market information. News databases used in the other studies start much later. Another advantage is that the majority of 8-K forms are not pre-scheduled. According to Securities and Exchange Commission (SEC) Accounting Series Release NO. 306, the report "plays a critical role in the periodic reporting system, which is intended to provide investors with a continuous stream of corporate information". Whisenant, Sankaraguruswamy, and Raghunandan (2003) found that reportable events disclosed in 8-K filings are considered by investors to have information content, evidenced by the substantial cumulative abnormal return over the period of disclosure. The 8-K filings is a more complete universe of information events where insiders can take advantage, documented by Skaife, Veenman and Wangerin (2013), Brochet (2010) and Thompson and Sale (2003). Therefore, if $VWKS$ does capture a more complete set of information, it would signal before the event. Zhao (2016) uses the intensity of 8-K filings to study the risk premium associated with information asymmetry. Goldstein and Wu (2015) and

Niessner (2015) use 8-K filings to study disclosure timing by managers. Our study, on the other hand, uses 8-K filings to complement scheduled earnings announcements and form a comprehensive sample of information events. Moreover, price jumps largely completes all possible information shocks. And the placebo test of transitory jumps is novel in this strand of the literature.

3 Data, sample selection and variable construction

We collect options data from OptionMetrics, which provides daily option trading volumes, strike prices, expiration date of the option, delta of option, call and put flags starting from 1996. Equity returns, bid and ask spread, trading volume and shares outstanding data are from the Center for Research in Security Prices (CRSP). Earnings announcement date and analyst coverage are from the Institutional Brokers Estimate System (IBES), and 8K Filing data are from SEC Analytics Suite. Our sample ends at August 2013. We exclude all indexes, units, ADRs, REITs, closed end funds, ETFs, and foreign firms and focus on common stocks only (CRSP share codes 10 and 11). We follow Jegadeesh and Titman (2001) and many others to exclude stocks whose closing prices are below \$5. After merging, our final sample has 3837 unique firms. On each trading day, there are on average 1400 firms.

We propose the volume-weighted strike-spot ratio for firm i on day t, $VWKS_{i,t}$, as the center of mass in the options volume distribution along strike prices:

$$VWKS_{i,t} = \frac{\sum_{j=1}^{n} volume_{i,t,j}(\frac{K_{i,t,j}}{S_{i,t}} - 1)}{\sum_{j=1}^{n} volume_{i,t,j}}, \tag{1}$$

where $K_{i,t,j}$ is the strike price for contract j, $volume_{i,t,j}$ is the trading volume of contracts j, n is the total number of contracts with unique strike price, and $S_{i,t}$ is the underlying equity price. Our variable captures the mean of strike-spot ratio whose density function is described by the options trading volume. If we replace the volume-weighted strike price $\frac{\sum_{j=1}^{n} volume_{i,t,j} K_{i,t,j}}{\sum_{j=1}^{n} volume_{i,t,j}}$ for firm i to its future equity price $S_{i,t+1}$, then $VWKS$ can be viewed as $\frac{S_{i,t+1}}{S_{i,t}} - 1$. Therefore, $VWKS$ can be interpreted as the aggregated views from option market on the future equity prices. If there are no options traded on a particular day, $VWKS$ is set to be missing.

Table 1 gives the summary statistics for variables from both the options and equity markets. Both options and equity trading volumes are counted as number of shares traded. All the variables are winsorized at 0.5 and 99.5 percentiles at daily level to mitigate the effect of potential outliers.

[Table 1 about here]

On average, $VWKS$ is slightly above zero, with a mean of 0.022 and a standard deviation of 0.111. It is skewed to the right, implying more volumes on OTM calls and ITM puts than on ITM calls and OTM puts.

The put-call ratio (PC) is calculated as the logarithm of put volume over call volume. We use $(put + 0.001)/(call + 0.001)$ to avoid zero volume in call or put. The mean of PC is -2.013 and standard deviation is 4.210, suggesting that in general there are more puts traded than calls. OS is calculated as the logarithm of option volume over underlying equity volume. Similarly, we use volume+0.001 to avoid zero trading volume in options or underlying equities. The average OS is -4.939 since the underlying equity volume is larger than the options trading volume. We follow Cremers and Weinbaum (2010) to compute the deviation from put-call parity DEV as the average difference between the call option implied volatility and put option implied volatility at all strike prices. Consistent with the findings of Cremers and Weinbaum, the mean of DEV is negative (-0.008). We follow Xing, Zhang, and Zhao (2010) to calculate options implied skewness measure $SKEW$ as the difference between the OTM (with K/S between 0.8 and 0.95) put option implied volatility and ATM all option implied volatility. Consistent with the findings of Xing, Zhang, and Zhao and many other studies reporting a volatility smile or smirk, the mean of $SKEW$ is positive (0.037), suggesting the ATM implied volatility is lower. We follow Quo and Qiu (2014) to compute options implied volatility $IVOL$ as the average implied volatility of ATM call and put options. It is a froward-looking measure of conditional variance and the mean is 0.478. Using CRSP data, we compute the percentage bid-ask spread ($SPREAD$) as the close ask minus close bid scaled by the midpoint of the bid and ask prices, and turnover ratio ($TURN$) as the total trading volume over the number of shares outstanding. RET is raw stock return in CRSP. V is the squared raw returns. $QRET$ is mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends, and $QRET1$ is the mid quote return on the following day. All the variables are at daily level unless otherwise specified.

Table 2 provides time-series averages of cross-sectional correlations. The contemporaneous correlation between $VWKS$ and stock return is large and negative (-0.16 for raw returns and -0.159 for mid quote returns). This is not surprising given the fact that the stock price enters calculation of $VWKS$ as a denominator. Mechanically, high stock returns can lead to a lower $VWKS$ for the same options volume distribution. This also raises a concern about potentially contaminated pricing effect by return reversals. In one of our robustness checks, we change the definition of $VWKS$ to get rid of such impact. $VWKS$ is positively correlated with the next day's mid quote return, $QRET1$ with a correlation coefficient of 1.3%. Because OTM options are traded more frequently than ITM options, $VWKS$ is negatively correlated with PC and the correlation reaches -0.175. The correlations between $VWKS$ and OS and DEV are less than 5%. The correlation between $VWKS$ and $SKEW$ is -0.093. It is possible that imbalanced options trading can create a price impact on options implied volatility as argued by Bollen and Whaley (2004) and Galeanu, Petersen, and Poteshman (2009). If high (low) $VKWS$ reflects positive (negative) risk taking, it could drive up the left (right) tail in the implied volatility curve, leading to lower (higher) implied skewness. Finally, $VKWS$ is also highly correlated with the level of implied volatility, percentage bid-ask spread, and stock turnover with correlations of 0.251, 0.115, and 0.08, respectively.

[Table 2 about here]

4 Empirical results

4.1 Portfolio analysis

We first perform an investment analysis to gauge the economical significance of the return predictability. On each trading day (week), the value-weighted decile portfolios are formed based on each return predictors from the options market ($VWKS$, PC, OS, DEV, $SKEW$ and $IVOL$). We use market capitalization as weights. Investment strategies are formed using high decile minus low decile. We report the returns as well as alphas with respect to Fama-French factors and liquidity and momentum factors. The t-statistic reported in the table are using Newey-West (1987) standard errors. Table 3 gives both daily and weekly reblanced portfolio sorted by an options market predictor.

[Table 3 about here]

Panel A reports the average annualized mid quote returns of daily rebalanced decile portfolios. Besides PC and OS, all the other predictors from option market see more than 23% returns, even after controlling for Fama-French factors and liquidity and momentum factors. The t-statistics are above 3.5 for daily strategies using $VWKS$, DEV, $SKEW$ and $IVOL$. From 1996 to 2013, a portfolio using high $VWKS$ decile minus low $VWKS$ decile as the trading signal has generated an annualized return of 46.1% before transaction cost, with a t-statistic of 7.25. The long leg of the strategy has a larger contribution to the return differential than the short leg of the strategy as the average annualized return of the high (low) decile portfolio is 0.302 (-0.159). The alpha with respect to Fama-French factors as well as the liquidity and momentum factors reaches 46.5% and is statistically significant at the 1% level. The 46% returns from the daily strategy might appear high, but we expect a substantial drag from the transactions cost. The second best performance comes from $IVOL$, with 36.5% alphas with respect to Fama-French factors and liquidity and momentum factorsm with t-statistic of 3.52. Panel B reports weekly rebalanced strategies with a week starting on Wednesday and ending on the coming Tuesday. As a lower turnover and more practical strategy, a portfolio using high $VWKS$ decile minus low $VWKS$ decile as the trading signal has generated a more modest 14.2% annualized return after controlling Fama-French factors and liquidity and momentum factors. $IVOL$ individually generate 23.8% return with a t-statistic of 2.46, followed by DEV with a 16.6% annualized return and t-statistic of 4.15. The t-statistics decreases in the weekly portfolio but are still significant. Although $IVOL$ generates extremely decent returns in both daily and weekly rebalanced portfolios, the signal reverse at monthly rebalanced portfolio. Portfolio returns using PC and OS are not economically nor statistically significant in weekly rebalanced portfolios. The findings in this table show that the return predictability from $VWKS$ is economically significant.

To visualize the returns generated, we plot the cumulative return from the daily rebalanced $VWKS$ strategy on a log scale in Figure 1. It is clear that the strategy achieves stable and persistent performance during the entire sample period.

[Figure 1 about here]

The strategy is the same as the first column in panel A table 3. Each day we sort the sample into 10 deciles by $VWKS$, buy all the stocks in the highest decile and sell all stocks in the lowest decile, weighted by market capitalization. We accumulate the next trading day's mid quote return and reported on a log scale. From 1996 to 2013, we consistently observe economically significant wealth growth even through the financial crisis.

The sound performance might be driven by other known predictors from the options market which are correlated to $VWKS$. To rule out this possibility, we examine the profitability of the $VWKS$ strategy controlling the effects from PC, OS, DEV, $SKEW$ and $IVOL$. The results of the double sorting analysis are presented in table 4.

[Table 4 about here]

In panel A, the sample is first sorted into 5 quintiles based on PC. Within each PC quintile, the subsample is further sorted into 5 quintiles by $VWKS$. Investment strategies are formed using high $VWKS$ quintile minus low $VWKS$ quintile and the returns are calculated using the market capitalization as weights. We report the annualized returns as well as alphas with respect to Fama-French factors and liquidity and momentum factors. The t-statistic reported in the table are using Newey-West (1987) standard errors. Similarly, we first sort the sample into 5 quintiles each day by OS (panel B), $SKEW$ (panel C), $IVOL$ (panel D) and DEV (panel E). Within each quintile, we further sorted the portfolio into 5 quintiles by VWKS. For each panel, all the five high minus low $VWKS$ portfolios are generating significant returns economically and statistically. The most profitable strategy comes from panel E. In the high $IVOL$ quintile, the portfolio alpha reaches 52.1%, with a t-statistic of 6.17. The second best performance is from high OS quintile, with a portfolio alpha of 39.3%. In low PC quintile, the portfolio alpha reaches 36.4%. In low $SKEW$ quintile, the portfolio alpha is 30.9%. In high DEV quintile, the portfolio alpha is 24.4%. All the alphas are significant at 1%. The double sorting analysis confirm the previous findings that the return predictability from $VWKS$ is economically significant.

4.2 Multivariate regression analysis

The multivariate regression analysis tests our first hypothesis that $VWKS$ has a positive and permanent price impact on the underlying stock returns. The standard Fama-MacBeth regression has two stages. We first estimate the following regression in cross section for each trading day t:

$$QRET_{i,t} = \alpha + \sum_{l=1}^{5} \beta_l VWKS_{i,t-l} + \sum_{l=1}^{5} \theta_l X_{i,t-l} + \epsilon, \tag{2}$$

where $QRET_{i,t}$ is the mid quote returns for firm i on day t. Five lags of all the explanatory variables for firm i on trading day $t-1$, $t-2$, $t-3$, $t-4$ and $t-5$, including $VWKS$, PC, OS, DEV, $SKEW$, $IVOL$, $QRET$, $SPREAD$, $TURN$ and V, are controlled. After obtaining a time series of the slope coefficients, we then examine these coefficients using Newey-West (1987) adjustment, allowing for autocorrelation structures. For ease of reporting, the dependent variables $QRET1$ and $RET1$ from table 5 to table 15 are expressed as percentages. Table 5 contains the main results of the Fama-MacBeth regression.

[Table 5 about here]

The first model presents the regression results from using one lag of $VWKS$, which has an average slope coefficient of 0.205, statistically significant at the 1% level (t-statistic = 7.82). A one-standard deviation increase in $VWKS$ is associated with an increase of 2.28 bp in the underlying stock price on the following date. Adding four more lags, the second model shows that all five lags of $VWKS$ are significant, and the first four lags are significant at 1% level. The coefficient of $VWKS$ in the second model is 0.125, statistically significant at the 1% level (t-statistic = 6.75). The third model controls five lags of all return predictors from option markets, i.e. $VWKS$, PC, OS, DEV, $SKEW$ and $IVOL$, and the first three lags of $VWKS$ are significant at 1% level. The coefficient of $VWKS$ in the third model is 0.101, statistically significant at the 1% level (t-statistic = 6.45). After further controlling for $QRET$, $SPREAD$, $TURN$ and their lags, $VWKS$ in the fourth model has a coefficient of 0.057 with t-statistic of 4.05, statistically significant at the 1% level. The first three lags of $VWKS$ are significant in the third and fourth model. The full model presents the regression results from using all return predictors from both equity and options market. $VWKS$ has a coefficient of 0.053 with t-statistic of 3.78, statistically significant at the 1% level. A one-standard deviation increase in $VWKS$ is associated with an increase of 0.6 bp in the underlying stock price on the following day, and 1.9 bp in the underlying stock price on the following five days. Moreover, the coefficients of $VWKS$ are statistically significant at the 5% level at the second and third lags, and at the 10% level at the fifth lag. With more control variables added into the Fama-MacBeth regression, both economic value and t-statistics decrease for $VWKS$ and its lags. Throughout the five models, the first lag of $VWKS$ is consistently statistically significant at 1% level, followed by its second and third lags. The adjusted R-square is 4.7% in the third model, controlling return predictors from options market only. It reaches 8.2% in the full model.

Since most return predictors ($VWKS$, PC, OS, $SKEW$, $IVOL$, $QRET$, $SPREAD$, $TURN$ and V) are statistically significant at multiple lags in the full model, we examine the aggregate effect by taking the moving averages of the five past trading days for all the variables. The empirical test is based on the following Fama-MacBeth regression:

$$QRET_{i,t} = \alpha + \beta VWKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon, \tag{3}$$

where

$$VWKS\_MA5_{i,t-1} = \sum_{l=1}^{5} VWKS_{i,t-l}/5, \tag{4}$$

and

$$X\_MA5_{i,t-1} = \sum_{l=1}^{5} X_{i,t-l}/5. \tag{5}$$

The X_MA5 is a set of control variables observable on day $t-1$. Estimating the coefficients of the MA terms in Equation (3) is equivalent to estimating the aggregate coefficients of all lags in Equation (2) under the constraint that the coefficients are the same for all lags of the same independent variable.

[Table 6 about here]

Throughout the four models in Table 6, $VWKS\_MA5$ is statistically significant at the 1% level. The first model only controls $VWKS\_MA5$, which has an average slope coefficient of 0.457 for $VWKS\_MA5$ (t-statistic = 8.22). The second model controls return predictors from options market. $VWKS\_MA5$ has a coefficient of 0.322 and t-statistic of 8.90. The third model controls PC_MA5, OS_MA5, DEV_MA5, $SKEW\_MA5$, $IVOL\_MA5$, $QRET\_MA5$, $SPREAD\_MA5$ and $TURN\_MA5$. The coefficient of $VWKS\_MA5$ is 0.197 and the t-statistic is 5.86. The fourth model further adds V_MA5. The coefficient of $VWKS\_MA5$ is 0.201 and its t-statistic is 6.01. All the variables in the full model are significant at 1% level except $IVOL\_MA5$ (significant at 10%) and V_MA5 (insignificant). This is due to the fact that their estimators reverse within five lags in Table 5. The adjusted R-square is 3.9% in the second model, controlling return predictors from options market only. It reaches 5.4% in the full model.

Using moving averages of the past five trading days, Table 6 captures the permanent price impact from the examined return predictors, which is more relevant for our question of information arrival to financial markets. The table is also easier to read, and the results are consistent with those in table 5. Both tables confirmed the first hypothesis that $VWKS$ has a positive and permanent price impact on the underlying stock returns. In the following Fama-MacBeth regression analyses, we report the results using 5-day moving averages and the tables using individual lags are available upon request.

4.3 Robustness tests

We conduct robustness tests to strengthen our first hypothesis about the return predictability of $VWKS\_MA5$, the aggregate effect of five lags of $VWKS$. Table 7 contains regression results on four alternative measures of $VWKS\_MA5$ in predicting future returns using full controls (model 4 in table 6).

[Table 7 about here]

In the first Fama-MacBeth regression, we test the return predictability of $VWKS\_MA5$ on raw returns ($RET1$) instead of mid quote returns by estimating the following equation:

$$RET_{i,t} = \alpha + \beta VWKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon, \tag{6}$$

where $RET_{i,t}$ is raw returns on day t and is expressed as a percentage. The X_MA5 is a set of control variables observed on day $t-1$, including PC_MA5, OS_MA5, DEV_MA5, $SKEW\_MA5$, $IVOL\_MA5$, RET_MA5, $SPREAD\_MA5$, $TURN\_MA5$ and V_MA5. Since RET is highly correlated with $QRET$, the result is consistent with the full model in table 6. $VWKS\_MA5$ has a coefficient of 0.165 with t-statistic of 4.88, statistically significant at 1% level.

In the second regression, we test the return predictability of $VWLNKS$, the log transformation of $VWKS$, defined as follows:

$$VWLNKS_{i,t} = \frac{\sum_{j=1}^{n} volume_{i,t,j}(log(K_{i,t,j}) - log(S_{i,t}))}{\sum_{j=1}^{n} volume_{i,t,j}}. \tag{7}$$

If we replace the volume-weighted log strike price $\frac{\sum_{j=1}^{n} volume_{i,t,j} log(K_{i,t,j})}{\sum_{j=1}^{n} volume_{i,t,j}}$ for firm i to its future log equity price $log(S_{i,t+1})$, $VWLNKS$ can be viewed as $log(S_{i,t+1})-log(S_{i,t})$. Same as $VWKS$, $VWLNKS$ can be interpreted as the aggregated views from option market on the future equity prices. Given the similarity in their definitions, $VWLNKS$ is expected to perform as well as $VWKS$ in return prediction using the following equation:

$$QRET_{i,t} = \alpha + \beta VWLNKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon, \tag{8}$$

where

$$VWLNKS\_MA5_{i,t-1} = \sum_{l=1}^{5} VWLNKS_{i,t-l}/5. \tag{9}$$

$VWLNKS\_MA5$ has a coefficient of 0.202 with t-statistic of 5.49, statistically significant at the 1% level. Compared to $VWKS\_MA5$, $VWLNKS\_MA5$ has a higher economic value in return prediction and is more significant statistically. This is mainly due to the logarithm transformation. PC and OS are both computed using the logarithm since their return predictability increases after the transformation. Qualitatively, the predictability of $VWLNKS$ is the same as $VWKS$.

While $VWKS$ measures option moneyness, another closely related measure is option *delta*. In the third Fama-MacBeth regression, we test the return predictability from the trading center of mass of *delta*:

$$QRET_{i,t} = \alpha + \beta VWDELTA\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon, \tag{10}$$

where

$$VWDELTA_{i,t} = \frac{\sum_{j=1}^{n} volume_{i,t,j} DELTA_{i,t,j}}{\sum_{j=1}^{n} volume_{i,t,j}}, \tag{11}$$

and

$$VWDELTA\_MA5_{i,t-1} = \sum_{l=1}^{5} VWDELTA_{i,t-l}/5. \tag{12}$$

Since $DELTA$ is negative for put options and positive for call options, we set put $DELTA = DELTA + 1$ and call $DELTA = DELTA$ such that a large $VWDELTA_{i,t}$ implies a small

$VWKS$. Note that the strike price is inversely related to the call option *delta*. Therefore, the value-weighted $VWDELTA$ is negatively correlated with $VWKS$ and should negatively predict future returns. The coefficient of $VWDELTA\_MA5$ is -0.099 with t-statistic of -3.49, statistically significant at the 1% level, supporting our hypothesis that the center of volume mass along options strike prices predicts returns positively.

As previously discussed, we are concerned that the predicting power of $VWKS$ is mechanical from return reversals. When stock price drops, without any changes in trading center of K, $VWKS$ will increase. Assuming return reversal, even if there is no informed trading in options, the previous result will look the same: high $VWKS$ predicts a higher return. Therefore, besides controlling the lagged returns, we create $VWKLS$ to address the concern in the following regression analysis:

$$QRET_{i,t} = \alpha + \beta VWKLS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon, \tag{13}$$

where

$$VWKLS_{i,t} = \frac{\sum_{j=1}^{n} volume_{i,t,j}(\frac{K_{i,t,j}}{S_{i,t-1}} - 1)}{\sum_{j=1}^{n} volume_{i,t,j}}, \tag{14}$$

and

$$VWKLS\_MA5_i = \sum_{l=1}^{5} VWKLS_{i,t-l}/5. \tag{15}$$

Stock price S_t is replaced by lagged price S_{t-1} when computing $VWKLS$. The coefficient of $VWKLS\_MA5$ is 0.129 with t-statistic of 5.58, statistically significant at the 1% level. Although the economic value decreases by 0.043, the statistically significance is even stronger. The result implies that $VWKS$ is not mechanical from return reversals. All four regressions have adjusted R-squares at around 5%.

The results in table 7 further confirms the first hypothesis. In this section, we find that $VWKS$ is able to predict both raw returns and mid quote returns. The predictability is qualitatively the same when we use $VWLNKS$, log transformation to normalize the variable; use $VWDELTA$, option deltas to measure moneyness instead of K/S; or use $VWKLS$, lagged stock price to eliminate the effect of return reversal. The strong return predictive ability of $VWKS$ is consistent with our conjecture that $VWKS$ captures informed trading in the options market.

4.4 Separating calls and puts

To better understand the nature of the return predictability, we analyze $VWKS$ using different types of options. We first compute volume weighted call options strike price over underlying stock price $VWKSCALL$, and volume weighted put options strike price over underlying stock price $VWKSPUT$. The empirical test is based on the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSCALL\_MA5_{i,t-1} + \beta_2 VWKSPUT\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon, \tag{16}$$

where $VWKSCALL\_MA5$ and $VWKSPUT\_MA5$ is the 5-day moving averages of $VWKSCALL$ and $VWKSPUT$ respectively. The X_MA5 is the same set of control variables on day $t-1$ defined earlier.

[Table 8 about here]

While $VWKSPUT\_MA5$ is slightly more significant than $VWKSCALL\_MA5$, both of them are positive and significant at 1% level in table 8. $VWKSCALL\_MA5$ alone has a coefficient of 0.387. After controlling all return predictors, its coefficient drops to 0.195, with t-statistic of 5.65. $VWKSPUT\_MA5$ alone has a coefficient of 0.438 and drops to 0.195 after controlling all return predictors. Model five combines both $VWKSCALL\_MA5$ and $VWKSPUT\_MA5$ and we see a coefficient of 0.154 (t-statistic = 4.88) for $VWKSCALL\_MA5$ and a coefficient of 0.160 (t-statistic = 5.09) for $VWKSPUT\_MA5$. The result suggests that the center of options volume mass using either call or put options contains stock price information.

4.5 Separating options maturity

We then investigate $VWKS$ using options with different maturities. $VWKSM1$ is $VWKS$ computed options expiring in fewer than or equal to 30 days, and $VWKSM2$ is $VWKS$ computed options expiring in in more than 30 days. The empirical test is based on the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSM1\_MA5_{i,t-1} + \beta_2 VWKSM2\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon, \quad (17)$$

where $VWKSM1\_MA5$ is the 5-day moving average of $VWKSM1$ and $VWKSM2\_MA5$ is the 5-day moving average of $VWKSM2$. The X_MA5 is the same set of control variables on day $t-1$ defined earlier.

[Table 9 about here]

Both $VWKSM1\_MA5$ and $VWKSM2\_MA5$ are significant at 1% level through the five models in table 9. In the first model, $VWKSM1\_MA5$ alone has a coefficient of 0.211 with t-statistic of 5.67. After controlling all the other return predictors from both options and equities market mentioned earlier, the second model reports a coefficient of 0.077 for $VWKSM1\_MA5$ with t-statistic of 3.42. In the third model, $VWKSM2\_MA5$ alone has a coefficient of 0.284 with t-statistic of 7.61. After controlling all the other return predictors from both options and equities market mentioned earlier, the fourth model reports a coefficient of 0.112 for $VWKSM2\_MA5$ with t-statistic of 5.27. When combining $VWKSM1\_MA5$ and $VWKSM2\_MA5$ in model 5, $VWKSM1\_MA5$ has a coefficient of 0.055 (t-statistic = 2.66) and $VWKSM2\_MA5$ has a coefficient of 0.096 (t-statistic = 4.82). While $VWKSM1$ and $VWKSM2$ predict returns positively, we find that more information comes from $VWKSM2$, with more significant result economically and statistically. The result suggests that the center of options volume mass using options with different maturities contains stock price information, and options expiring in more than 30 days contains more information than those expiring in less than or equal to 30 days.

4.6 Asymmetric pricing effects from center of options volume mass

In this section, we divide $VWKS$ into the positive side and negative side to test where the predictive power comes from. Positive $VWKS$ is denoted as $VWKSP = max(VWKS, 0)$ and negative $VWKS$ is denoted as $VWKSN = min(VWKS, 0)$. The empirical test is based on the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSP\_MA5_{i,t-1} + \beta_2 VWKSN\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon, \quad (18)$$

where $VWKSP\_MA5$ is the 5-day moving average of volume weighted strike price over underlying stock price if strike price is larger than stock price, zero otherwise. $VWKSN\_MA5$ is the 5-day moving average of volume weighted strike price over underlying stock price if strike price is smaller than stock price, zero otherwise.

[Table 10 about here]

$VWKSP\_MA5$ is positive and significant at 1% level in table 10. In the first model, $VWKSP\_MA5$ alone has a coefficient of 0.483 with t-statistic of 9.08. After controlling all the other return predictors from both options and equities market mentioned earlier, the second model reports a coefficient of 0.168 for $VWKSP\_MA5$ with t-statistic of 5.34. In the third model, $VWKSN\_MA5$ alone has a coefficient of 0.237 with t-statistic of 3.12. After controlling all the other return predictors from both options and equities market mentioned earlier, the fourth model reports a coefficient of 0.184 for $VWKSN\_MA5$ with t-statistic of 3.86. When combining $VWKSP\_MA5$ and $VWKSN\_MA5$ in model 5, $VWKSP\_MA5$ has a coefficient of 0.147 (t-statistic = 4.72) and $VWKSN\_MA5$ has a coefficient of 0.084 (t-statistic = 1.89), significant at 10% level. The result suggests that while both sides of $VWKS$ have positive impact on stock price, the positive side of $VWKS$ contains more information.

4.7 Non-linear price impact of center of options volume mass

We then test the predicting power from the tails of $VWKS$ distribution. Since $VWKS$ centers around zero, we create variable $VWKSSQ = sign(VWKS) * VWKS^2$, which is the signed $VWKS$ square. An absolute large value of $VWKSSQ$ implies it is at the tails of the distribution, where a positive sign indicates the right tail and a negative sign indicates the left tail. The empirical test is based on the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSSQ\_MA5_{i,t-1} + \beta_2 VWKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon, \quad (19)$$

where $VWKSSQ\_MA5$ is the 5-day moving average of squared volume weighted strike price over underlying stock price. $VWKS\_MA5$ is the 5-day moving average of volume weighted strike price over underlying stock price. Table 11 reports the Fama-MacBeth regression result.

[Table 11 about here]

We see positive and significant $VWKSSQ\_MA5$ at 1% level. $VWKSSQ\_MA5$ by itself (model 1), has a coefficient as large as 0.611, with t-statistic of 10.89. After controlling other return predictors, we see a coefficient of 0.376 for $VWKSSQ\_MA5$ with t-statistic of 8.31 in model 2. Finally, even after controlling $VWKS\_MA5$, we find $VWKSSQ\_MA5$ with coefficient 0.600 and t-statistic 8.13 in model 3. In this model, $VWKS\_MA5$ turns negative, implying that the return predicting power mainly comes from the tail of $VWKS$.

4.8 Shocks to center of options volume mass

We also test the predicting power from the shocks of $VWKS$. We first compute the past 20-day moving average of $VWKS$ as $VWKSMA20 = \sum_{j=1}^{20} VWKS_{t-j-5}/20$. Then we take $VWKS\_MA5$'s deviation from $VWKSMA20$ to obtain $VWKSS20 = VWKS\_MA5 - VWKSMA20$. The empirical test is based on the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSS20_{i,t-1} + \beta_2 VWKSMA20_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon. \quad (20)$$

Table 12 reports Fama-MacBeth regression result of 20-day moving average and its shock, $VWKSMA20$ and $VWKSS20$.

[Table 12 about here]

We find that both $VWKSMA20$ and $VWKSS20$ can predict return positive and significant at 1% level. The shock $VWKSS20$ has a coefficient of 0.275 with t-statistic of 7.41 in the first model by itself. After controlling all the other return predictors from both options and equities market mentioned earlier, the coefficient becomes 0.109 with t-statistic of 4.65 (model 2). In the third model, the 20-day moving average $VWKSMA20$ has a coefficient of 0.383 with t-statistic of 4.61 and after controlling other return predictors from options market, the coefficient becomes 0.175 with t-statistic of 3.15 in the fourth model. Finally, combining both $VWKSMA20$ and $VWKSS20$ and PC_MA5, OS_MA5, DEV_MA5, $SKEW\_MA5$, $IVOL\_MA5$, $QRET\_MA5$, $SPREAD\_MA5$, $TURN\_MA5$ and V_MA5, we see $VWKSS20$ with a coefficient 0.156 (t-statistic = 6.40) and $VWKSMA20$ with a coefficient 0.222 (t-statistic = 3.86). The economic significance for the shock is 0.012. With one standard deviation change in $VWKSS20$, the daily return will move by 1.2%. The economic significance for 20-day moving average of $VWKS$ is 0.015. With one standard deviation change in $VWKSMA20$, the daily return will move by 1.5%. Comparing the shocks and 20-day moving average, we see relatively similarly economic significance.

Both 20-day moving average of $VWKS$ and its deviation positively predict future returns. The significance in the moving average from day $t-25$ to $t-6$ suggests that there is some delayed price response even beyond first five days and the positive sign suggests that there is no reversal. However, a larger impact comes from the shock, due to arrival of new information. Therefore we see a much more significant $VWKSS20$ in the last model.

The even stronger predicting power from the positive part of $VWKS$ ($VWKSP$), tail of $VWKS$ ($VWKSSQ$) as well as shocks of $VWKS$ ($VWKSS20$), further confirms our first hypothesis that $VWKS$ has a positive and permanent price impact on the underlying stock.

4.9 Conditioning on information asymmetry

This subsection tests the second hypothesis 2 that $VWKS$ has stronger return predictability for stocks with higher levels of information asymmetry. Based on the eight proxies for information asymmetry, we divide the sample by size (measured by market capitalization), idiosyncratic volatility, Amihud's (2002) illiquidity, analyst coverage, institutional ownership, probability of informed trading (PIN), options trading volume and sample time period. For each proxy for the information asymmetry, the full sample is divided into two groups: low (<50th percentile) and high (>50th percentile). The slope coefficients and t-statistics (in parentheses) are reported only for five lags of $VWKS$ and $VWKS\_MA5_{i,t-1}$, but the regressions are based on the Fama-Macbeth regression using the following equation:

$$QRET_{i,t} = \alpha + \sum_{l=0}^{4} \beta_l VWKS_{i,t-l} + \sum_{l=0}^{4} \theta_l X_{i,t-l} + \epsilon.$$

To study the aggregate effect of the five lags of $VWKS$, we use the following equation:

$$QRET_{i,t} = \alpha + \beta VWKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon.$$

Table 13 reports Fama-MacBeth regression results for each subsample. We report the five lags of $VWKS$, 5-day moving average $VWKS\_MA5$, and the differences between low and high groups within each subsample.

[Table 13 about here]

The conditioning variable is firm market capitalization ($Size$) in Panel A, idiosyncratic stock volatility ($Idio$) in Panel B, illiquidity measured as in Amihud (2002) in Panel C, analyst coverage ($Analyst$) in Panel D, fraction of institutional ownership ($Ownership$) in Panel E, the probability of informed trading (PIN) as of Easley, Kiefer, O'Hara, and Paperman (1996) in Panel F, total options trading volume ($Volume$) in Panel E, and sample period ($Year$) in Panel F.

We find $VWKS$ is statistically significant throughout all the subsamples. For large firms, we see a more significant $VWKS$ at the first lag, whose coefficient is 0.093 with t-statistics of 4.01. For small firms, while the first $VWKS$ has a coefficient of 0.055 with t-statistics of 2.73, the second, third and fifth lags are all significant. Therefore we see a more persistent predictive power for stocks in smaller firms. At the aggregate level, $VWKS\_MA5$ for smaller firms have a coefficient of 0.316 (t-statistics = 7.83) while $VWKS\_MA5$ for larger firms have a coefficient of 0.118 (t-statistics = 2.31). We then test the differences of $VWKS\_MA5$ between small and large firms and find $VWKS\_MA5$ is statistically more significant in smaller firms with t-value of 4.2, at 0.1% level.

Firms with more idiosyncratic risks see more significant and persistent $VWKS$. $VWKS$ has a coefficient of 0.029 with t-statistics of 1.72 for firms with lower idiosyncratic risks. For firms with more idiosyncratic risks, the first, second, third and fifth lags of $VWKS$ are all significant, each with a coefficient of 0.05 (t-statistics = 2.22), 0.058 (t-statistics = 2.67), 0.046 (t-statistics = 2.04) and 0.040 (t-statistics = 1.73). $VWKS\_MA5$ is only significant for high idiosyncratic risks firms. The differences are quite obvious with a t-value of -5.79, significant at 0.1%.

More liquid firms have higher significance in $VWKS$, but less persistent over multiple lags. High illiquidity firms have statistically significant $VWKS$ for all its five lags. While the first and fourth lags are significant at 5% level, the second, third and fifth lags are significant at 1% level. For low illiquidity firms, the first lag of $VWKS$ has a coefficient of 0.080 with t-statistics of 3.19. The second, third and fourth lags have negative coefficient estimates. $VWKS\_MA5$ is only significant for less liquid firms. The differences between liquid and illiquid firms are quite obvious with a t-value of -6.4, significant at 0.1%.

For more analyst covered firms, we see a more significant $VWKS$ at the first lag, whose coefficient is 0.097 with t-statistics of 4.21. For less analyst covered firms, while the first $VWKS$ has a coefficient of 0.041 with t-statistics of 2.03, the second lag is also significant with t-statistics of 2.19. Similarly, we find a more persistent predictive power for stocks in less analyst covered firms.

Firms with higher institutional ownership have slightly more significant but less persistent $VWKS$. For low institutional ownership firms, the first lag of $VWKS$ has a coefficient of 0.045 and t-statistics of 2.14. The second lag has a coefficient of 0.038 with t-statistics of 1.81. The third lag has a coefficient of 0.050 with t-statistics of 2.36 and the fifth lag has a coefficient of 0.043 with t-statistics of 1.90. On the other hand, the higher institutional ownership firms has a coefficient of 0.077 (t-statistics = 3.23) at its first lag and a coefficient of 0.042 (t-statistics = 1.88) at the fourth lag. $VWKS\_MA5$ is significant for both samples, the lower institutional ownership sees a coefficient of 0.247 with t-statistics of 5.68 while the higher institutional ownership sample sees a coefficient of 0.186 with t-statistics of 3.82. The differences between the two subsamples are not statistically significant.

For high PIN firms, the first three lags are significant. The first lag has a coefficient of 0.042 and t-statistics of 4.21. The second lag has a coefficient of 0.034 and t-statistics of 1.69. The third lag has a coefficient of 0.063 and t-statistics of 3.05. While only the first lag for low PIN firms is significant, with a coefficient of 0.086 and t-statistics of 3.62. $VWKS\_MA5$ in both high PIN and low PIN firms are significant at 1% and their differences are not significant, with a t-value of -1.34. $VWKS\_MA5$ in both high PIN and low PIN firms are significant at 1% and their differences are not significant, with a t-value of 0.29.

Firms with more options traded see much more significant and persistent predicting power from $VWKS$. The first lag is significant at 1% with a coefficient of 0.092 and the second lag is significant at 5% level with a coefficient of 0.062. Only the third lag of $VWKS$ for firms with fewer options trading volume is significant, with a coefficient of 0.031 and t-statistics of 1.65. $VWKS\_MA5$ in both subsamples are significant at 1%, with coefficient of 0.111 for low volume sample and 0.236 for high volume sample. The t-value for low minus high sample is -2.61, indicating that $VWKS\_MA5$ is stronger in predicting returns for firms with more options traded.

Although $VWKS$ is significant in both 1996-2004 and 2005-13, its economic significance has decreased from 0.064 (t-statistics = 2.60) to 0.049 (t-statistics = 2.71). We find $VWKS\_MA5$ significant at 1% level from 1996 to 2004 while only significant at 5% level from 2005-2013. The differences between the two sample period is significant with t-value of 3.25.

The finding reinforces hypothesis 2 that $VWKS$ has more persistent predictive power for stocks with higher level of information asymmetry, i.e. those with low market capitalization, little analyst coverage, high options volume, large PIN, high volatility, low liquidity and low institutional ownership, and in the first half of the sampler period.

4.10 An event study on information shocks

Having documented the strong return predicting power from $VWKS$, the analysis now turns to examining the underlying mechanism. We establish a concrete and unambiguous link between $VWKS$ and information flow by studying its behavior around earnings announcements, non-earnings 8-K filings and price jumps due to other reasons. We define four types of events. The first type is scheduled events, which is earnings announcements. It has been examined as an instruments of information shocks extensively in the literature. Information revealed in scheduled events is found to be associated with options implied uncertainty by Dubinsky and Johannes (2006). 8-K filings is a more comprehensive set of information events, and are mostly unscheduled. As required by SEC, ´companies must report certain materiel corporate events on a more current basis´, in the form of 8-K. The second type is unscheduled events, defined as 8-K filings unrelated to earnings announcements. We are among the first finance studies that use 8-K filings to identify unscheduled events. There could be other information events not captured by these two types. In the sample, we observe extreme price jumps, which are not related to either scheduled or unscheduled events. Extreme price jumps are those days when the risk-adjusted return is higher than 10% as in Savor (2012) or when the risk-adjusted return is above two standard deviations as in Boehmer and Wu (2013). We further classify the price jumps into transitory and permanent categories. A transitory jump ($tranjump$) would see a return reversal within five days that completely offsets the initial jump. A permanent price jump ($permjump$) survives the subsequent return reversal. The four types of events are mutually exclusive in our analysis. Events are signed by the cumulative abnormal return on the event day ($CAR0$).

To test if $VWKS$ exhibits any abnormal behavior around events, we estimate the following equation using ordinary least squares (OLS) regressions with the firm, year and week fixed effects:

$$VWKS = \alpha + \beta_0 EVENT + \sum_{i=1}^{5} \beta_1^i PREEVENTi + \sum_{i=1}^{5} \beta_2^i POSTEVENTi + \epsilon, \quad (21)$$

where $EVENT$ is a category variable with a value of 1 (-1) if there is a positive (negative) corporate event on the same day t, and zero otherwise; $PREEVENTi$ is a pre-event category variable with a value of 1 (-1) if there is a positive (negative) event on day $t+i$, and zero otherwise; and $POSTEVENTi$ is a post-event category variable with a value of 1 (-1) if there is a positive (negative) event on day $t-i$, and zero otherwise. We cluster standard errors around firms. The results are reported in table 14.

[Table 14 about here]

We find that before positive (negative) scheduled events $VWKS$ increases (decreases) significantly at 1% level. $VWKS$ moves in the same direction as the event's cumulative abnormal return by 0.002 five days before, 0.003 four days before, 0.006 three days before, 0.008 two days before, and 0.010 one day before scheduled event. The significance level is the same for unscheduled events. $VWKS$ moves in the same direction as the event's cumulative abnormal return by 0.001 five days before, 0.002 four days before, 0.003 three days before, 0.003 two days before, and 0.004 one day before unscheduled event. $VWKS$ behaves similarly around permanent jumps. At 1% significance level, $VWKS$ moves in the same direction as the event's cumulative abnormal return by 0.003 five days before, 0.004 four days before, 0.006 three days before, 0.009 two days before, and 0.010 one day before permanent price jumps. However, before transitory price jumps, $VWKS$ does not show significant changes. By combining the previous four events, $VWKS$ becomes significantly larger prior to firms' earnings announcements, 8-K filings, as well as extreme price jumps in the same direction as the subsequent cumulative abnormal returns. Comparing the two types of jumps, we find $VWKS$ is much more significant before permanent price jumps and insignificant before transitory price jumps. Since permanent jumps involve new information while transitory jumps do not (Boehmer and Wu, 2013), the finding establishes a clear link between $VWKS$ and information shocks.

We investigate $VWKS$ price sensitivity prior to different events using Fama-MacBeth regression:

$$QRET_{i,t} = \alpha + \beta VWKS\_MA5_{i,t-1} + dummy + VWKS\_MA5_{i,t-1} * dummy + \theta X\_MA5_{i,t-1} + \epsilon, \tag{22}$$

where the event dummy $SCHEDULED = 1$ if there is an earnings announcement on day t. The event dummy $UNSCHEDULED = 1$ if there is an 8-K filing unrelated to earnings announcement on day t. The event dummy $PERMJUMP = 1$ if there is a permanent price jump but no 8-K filing nor earnings announcement on day t. The event dummy $TRANJUMP = 1$ if there is the price jump is only transitory (return reverse within five trading days) but no 8-K filing nor earnings announcement on day t. Results are reported in the last table.

[Table 15 about here]

In model 1, 2, 3 and 4, we test $VWKS\_MA5$'s sensitivity to each type of event and model 5 combines all the four types. By studying the interactions between $VWKS\_MA5$ and event dummies, we find its interaction with permanent price jump dummy, $KS * PERMJUMP$, is the most sensitive. The coefficient is 1.423 and t-statistics is 3.64, significant at 1% level, suggesting that if there is a permanent price jump on day t, $VWKS\_MA5$ will move in the same direction as the subsequent cumulative abnormal returns significantly. The fact that $VWKS$ is most sensitive to permanent price jumps but not transitory price jumps further confirms our third hypothesis.

5 Conclusion

In this article, we document a novel stock return predictor from the options market, the volume-weighted strike-to-spot price ratio ($VWKS$) across all traded option contracts. Intuitively, $VWKS$ measures the hot spot in the distribution of options volume along strike prices, and could reflect the activity of informed traders. A daily rebalanced investment strategy based on this variable generates an annualized abnormal return of 48% over the past 18 years. Double sorting analysis shows that this abnormal return exists in all quintile portfolios based on other known options market return predictors including the put-call ratio, options-to-stock volume ratio, deviation from put-call parity, implied skewness, and implied volatility. We then examine the cross-sectional pricing effect of $VWKS$ controlling for all the other options market return predictors as well as stock market illiquidity and past returns using Fama-MacBeth regressions. The predictability from $VWKS$ survives in the multivariate regressions, and is robust to alternative measures of the center of options volume mass using log transformation and option *delta*. In additional analysis, we find that the predictive ability of $VWKS$ exists in both call and put options, and in both short-term and long-term options. The predictive power is positive and significant for both positive and negative $VWKS$, and mainly comes from the tail of the distribution of $VWKS$. Consistent with options traders exploiting advanced information, we find that $VWKS$ has more persistent predictive power for stocks with higher levels of information asymmetry proxied by low market capitalization, high options volume, little analyst coverage, large PIN, high volatility, low liquidity, and low institutional ownership as well as in the first half of our sample period. Moreover, $VWKS$ retains significant predictive power in all subsamples. We also document a close link between $VWKS$ and firms' future fundamental news (e.g. earnings, 8-K filings, and permanent price jumps). We find that $VWKS$ exhibits abnormal run-ups before all of these corporate events. However, there is no abnormal run-up in $VWKS$ before transitory price jumps in a placebo test. Moreover, the price sensitivity to $VWKS$ also strengthens before permanent price jumps. The event study supports our hypothesis that $VWKS$ captures the fundamental information flow in options trading. For future research, it would be interesting to study the impact of $VWKS$ on the options implied volatility curve too. We use a simple point, the center of mass, to capture the heterogeneity across options volume distributions. There are other statistical metrics that can potentially shed more light on the information content in the volume distribution. Moreover, adding options maturity into the picture will expand the distribution curve to a surface, which further increases the possibilities as well as challenges in information extraction.

References

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of financial markets, 5(1), pp.31-56.

Anthony, J. H., 1988. The interrelation of stock and options market trading-volume data. Journal of Finance 43, 949-964.

Augustin, P., Brenner, M., Grass, G. and Subrahmanyam, M.G., 2016. How do insiders trade?.

Augustin, P., Brenner, M., Hu, J. and Subrahmanyam, M.G., 2015. Are Corporate Spin-offs Prone to Insider Trading?.

Asquith, P., P. Pathak, and J. Ritter, 2005, Short interest, institutional ownership, and stock returns, 78, Journal of Financial Economics, 243-276.

Black, F., 1975. Fact and fantasy in the use of options. Financial Analysts Journal, 31(4), pp.36-41.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637-654.

Boehmer, E. and Wu, J.J., 2013. Short selling and the price discovery process. Review of Financial Studies, 26(2), pp.287-322.

Brochet, F., 2010. Information content of insider trades before and after the Sarbanes-Oxley Act. The Accounting Review, 85(2), pp.419-446.

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

Cao, C., Chen, Z., and Griffin, J., 2005. Informational content of option volume prior to takeovers, Journal of Business 78, 1073-1109.

Chakravarty, S., Gulen, H., Mayhew, S., 2004. Informed trading in stock and option markets. Journal of Finance 59, 1235-1258.

Chan, K., Chung, Y. P., Fong, W. M., 2002. The informational role of stock and option volume. Review of Financial Studies 14, 1049-1075.

Chan, K., Chung, Y. P., Johnson, H., 1993. Why option prices lag stock prices: A trading-based explanation. Journal of Finance 48, 1957-1967.

Chen, J, H. Hong, and J. Stein, 2002, Breadth of ownership and stock returns, Journal of Financial Economics, 66, 171-205.

Chan, K., Ge, L. and Lin, T.C., 2015. Informational content of options trading on acquirer announcement return. Journal of Financial and Quantitative Analysis, 50(05), pp.1057-1082.

Cremers, M., Fodor, A. and Weinbaum, D., 2015. Where Do Informed Traders Trade First? Option Trading Activity, News Releases, and Stock Return Predictability. Working Paper.

Cremers, M. and Weinbaum, D., 2010. Deviations from put-call parity and stock return predictability.

D'Avolio, G., 2002, The market for borrowing stock, Journal of Financial Economics, 66, 271-306.

Dubinsky, A., and M. Johannes, 2005, Earnings Announcements and Equity Options. Working Paper, Columbia University.

Easly, D., M. O'Hara, and P. Srinivas, 1998, Option Volume and Stock Prices: Evidence on Where Informed Traders Trade. Journal of Finance 53, 431-465.

Fama, E.F. and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), pp.3-56.

Fama, E.F. and MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. The journal of political economy, 607-636.

Fang, L. and Peress, J., 2009. Media coverage and the crosssection of stock returns. The Journal of Finance, 64(5), pp.2023-2052.

Figlewski, S. and Webb, G.P., 1993. Options, short sales, and market completeness. The Journal of Finance, 48(2), pp.761-777.

Garleanu, N., Pedersen, L.H. and Poteshman, A.M., 2009. Demand-based option pricing. Review of Financial Studies, 22(10), pp.4259-4299.

Ge, L., Hu, J., Humphery-Jenner, M. and Lin, T.C., 2016. Informed options trading prior to bankruptcy filings.

Ge, L., Lin, T.C. and Pearson, N.D., 2016. Why does the option to stock volume ratio predict stock returns?. Journal of Financial Economics, 120(3), 601-622.

Gharghori, P., Maberly, E.D. and Nguyen, A., 2015. Informed trading around stock split announcements: Evidence from the option market.

Glosten, L. R., Milgrom, P. R., 1985. Bid, ask, and transaction prices in a specialist market with heterogenously informed traders. Journal of Financial Economics 14, 71-100.

Goldstein, I. and Wu, D., 2015. Disclosure timing, information asymmetry, and stock returns: Evidence from 8-K filing texts. Working Paper.

Goncalves-Pinto, L., Grundy, B.D., Hameed, A., van der Heijden, T. and Zhu, Y., 2016. Why Do Option Prices Predict Stock Returns? The Role of Price Pressure in the Stock Market.

Guo, H. and Qiu, B., 2014. Options-implied variance and future stock returns. Journal of Banking & Finance, 44, 93-113.

Hayunga, D., and Lung, P., 2014. Trading in the options market around financial analysts consensus revisions, Journal of Financial and Quantitative Analysis 48, 725-747.

Hu, J., 2014, Does Option Trading Convey Stock Price Information? Journal of Financial Economics 111, 625-645.

Hong, H., Lim, T. and Stein, J.C., 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. The Journal of Finance, 55(1), pp.265-295.

Jiang, G.J. and Yao, T., 2013. Stock price jumps and cross-sectional return predictability. Journal of Financial and Quantitative Analysis, 48(05), pp.1519-1544.

Jin, W., Livnat, J. and Zhang, Y., 2012. Option Prices Leading Equity Prices: Do Option Traders Have an Information Advantage?. Journal of Accounting Research, 50(2), pp.401-432.

Johnson, T.L. and So, E.C., 2012. The option to stock volume ratio and future returns. Journal of Financial Economics, 106(2), 262-286.

Kyle, A. S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315-1336.

Lee, S.S. and Mykland, P.A., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. Review of Financial studies, 21(6), pp.2535-2563.

Merton, R. C., 1973. An intertemporal asset pricing model. Econometrica 41, 867-887.

Muravyev, D., Pearson, N. D., Broussard, J. P., 2013. Is there price discovery in equity options?. Journal of Financial Economics 107(2), 259-283.

Nagel, S., 2005, Short sales, institutional investors, and the cross-section of stock returns, Journal of Financial Economics, 78, 277-309.

Niessner, M., 2015. Strategic disclosure timing and insider trading. Working Paper.

Newey, W.K. and West, K.D., 1987. Hypothesis testing with efficient method of moments estimation. International Economic Review, 777-787.

Pan, J. and Poteshman, A.M., 2006. The information in option volume for future stock prices. Review of Financial studies, 19(3), 871-908.

Spiegel, M.I. and Wang, X., 2005. Cross-sectional variation in stock returns: Liquidity and idiosyncratic risk. Working Paper.

Roll, R., Schwartz, E. and Subrahmanyam, A., 2009. Options trading activity and firm valuation. Journal of Financial Economics, 94(3), pp.345-360.

Roll, R., Schwartz, E. and Subrahmanyam, A., 2010. O/S: The relative trading activity in options and stock. Journal of Financial Economics, 96(1), 1-17.

Savor, P.G., 2012. Stock returns after major price shocks: The impact of information. Journal of financial Economics, 106(3), pp.635-659.

Skaife, H.A., Veenman, D. and Wangerin, D., 2013. Internal control over financial reporting and managerial rent extraction: Evidence from the profitability of insider trading. Journal of Accounting and Economics, 55(1), pp.91-110.

Stefan, J., and R. Whaley, 1990, Intraday Price Change and Trading Volume Relations in the Stock and Stock Option Markets, Journal of Finance 45, 191-220.

Thompson, R.B. and Sale, H.A., 2003. Securities fraud as corporate governance: Reflections upon federalism. Vand. L. Rev., 56, p.859.

Whisenant, J.S., Sankaraguruswamy, S. and Raghunandan, K., 2003. Market reactions to disclosure of reportable events. Auditing: A Journal of Practice & Theory, 22(1), pp.181-194.

Xing, Y., Zhang, X. and Zhao, R., 2010. What Does Individual Option Volatility Smirks Tell Us about Future Equity Returns. In Journal of Financial and Quantitative Analysis, 641-662.

Zhao, X., 2016. Does Information Intensity Matter for Stock Returns? Evidence from Form 8-K Filings. Management Science.

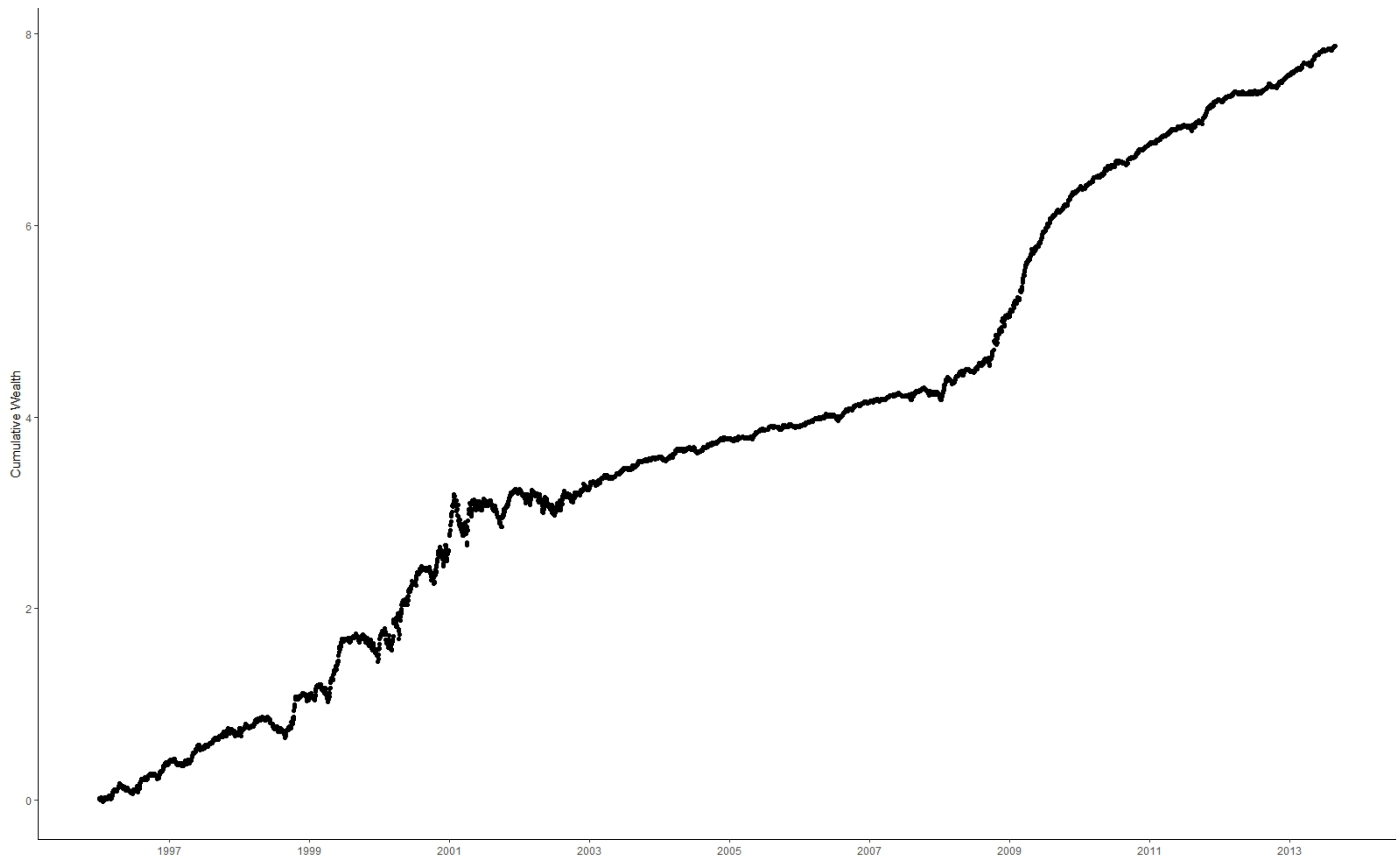

Figure 1: Cumulative return from the $VWKS$ strategy

This figure plots the cumulative return of the daily rebalanced strategy based on the volume-weighted strike-to-spot ratio ($VWKS$) from 1996 to 2013. Each day, we sort the sample into decile portfolios by $VWKS$ and buy all the stocks in the highest decile and sell all the stocks in the lowest decile. The total cumulative return calculated using the mid quote return is reported on a log scale.

Table 1: Summary statistics

This table reports descriptive statistics of the main variables in the analysis. We obtain daily stock and options data from CRSP and OptionMetrics between January 1, 1996 and August 31, 2013. Only common stocks with CRSP security code of 10 and 11 are included. We also exclude those stocks with prices below five dollars. There are 3837 unique firms in the sample with on average 1400 firms per day. $VWKS$ is the options-volume weighted strike price over underlying price minus one. PC is the put-call ratio, calculated as the logarithm of put options volume over call options volume. We use (put+0.001)/(call+0.001) to avoid zero volume in call or put. OS is the logarithm of total options volume over underlying stock volume. Similarly, we use volume+0.001 to avoid zero trading volume in options or underlying equities. DEV is the deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options across all pairs. $SKEW$ is the options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts (strike-to-price ratio lower than 0.95 but higher than 0.80) and at-the-money calls. $IVOL$ is the options-implied volatility, calculated as the average implied volatility of at-the-money call and put options. RET is raw stock return in CRSP. $QRET$ is mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends. $SPREAD$ is the percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices times 100. $TURN$ is the turnover ratio calculated as the total trading volume over the number of shares outstanding times 100. All variables except $QRET$ are winsorized at the 0.5% and 99.5% levels.

Variable	N	Mean	Std Dev	Minimum	Maximum
$VWKS$	6187703	0.022	0.111	-0.490	2.532
PC	6273113	-2.013	4.210	-12.054	10.043
OS	6273113	-4.939	3.368	-13.830	-0.461
DEV	6273113	-0.008	0.070	-2.292	1.338
$SKEW$	6273113	0.037	0.092	-1.327	2.228
$IVOL$	6273113	0.478	0.231	0.071	2.291
RET	6273042	0.001	0.031	-0.411	0.425
$QRET$	6149871	0.001	0.030	-0.333	0.436
$SPREAD$	6024194	0.580	0.926	0.005	100.388
$TURN$	6273113	4.204	0.949	0.171	7.704

Table 2: Correlations

This table reports the time-series averages of cross-sectional correlations between January 1996 and August 2013. $VWKS$ is the options-volume weighted strike price over underlying price minus one. PC is the put-call ratio, calculated as the logarithm of put options volume over call options volume. OS is the logarithm of total options volume over underlying stock volume. DEV is the deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW$ is the options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL$ is the options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. RET is raw stock return in CRSP. $QRET$ is mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends, and $QRET1$ is the mid quote return on the following day. $SPREAD$ is the percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices. $TURN$ is the turnover ratio calculated as the total trading volume over the number of shares outstanding.

	QRET1	*QRET*	*RET*	*VWKS*	*PC*	*OS*	*DEV*	*SKEW*	*IVOL*	*SPREAD*	*TURN*
QRET1	1.000										
QRET	0.002	1.000									
RET	0.004	0.989	1.000								
VWKS	0.013	-0.159	-0.160	1.000							
PC	-0.009	-0.107	-0.106	-0.175	1.000						
OS	-0.008	0.020	0.020	-0.004	-0.015	1.000					
DEV	0.011	-0.092	-0.121	0.045	-0.015	-0.043	1.000				
SKEW	-0.015	0.056	0.069	-0.093	0.054	0.090	-0.460	1.000			
IVOL	0.012	-0.020	-0.020	0.251	-0.072	0.131	-0.093	0.136	1.000		
SPREAD	0.011	-0.001	0.000	0.115	-0.032	-0.027	-0.063	-0.015	0.253	1.000	
TURN	0.010	0.047	0.047	0.080	-0.044	0.054	-0.017	0.069	0.493	-0.020	1.000

Table 3: Univariate portfolio analysis

Panel A reports the average annualized returns of daily rebalanced decile portfolios sorted by an options market predictor individually, as well as the return differentials between the top and bottom deciles and the alphas with respect to the Fama-French (1993) factors and liquidity and momentum factors. $VWKS$ is the options-volume weighted strike price over underlying price minus one. PC is the put-call ratio, calculated as the logarithm of put options volume over call options volume. OS is the logarithm of total options volume over underlying stock volume. DEV is the deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW$ is the options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL$ is the options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. The stock returns are calculated using the midpoint of the bid and ask prices at market close adjusted for stock splits and dividends. Portfolio returns are calculated using the market capitalization as weights. Panel B reports the results of weekly rebalanced strategies with a week starting on Wednesday and ending on the coming Tuesday. The weekly options market signals are calculated as the means of their daily values. Corresponding t-statistics with Newey-West (1987) standard errors are reported in squared brackets.

	Panel A: Value-weighted daily returns						*Panel B:Value-weighted weekly returns*					
Portfolio	*VWKS*	*PC*	*OS*	*DEV*	*SKEW*	*IVOL*	*VWKS*	*PC*	*OS*	*DEV*	*SKEW*	*IVOL*
low	-0.159	0.140	0.152	0.083	0.275	0.028	0.086	0.138	0.138	0.102	0.192	0.073
2	-0.008	0.037	0.139	0.043	0.169	0.071	0.085	0.139	0.150	0.100	0.116	0.108
3	0.003	0.107	0.154	0.016	0.164	0.132	0.141	0.140	0.131	0.101	0.156	0.123
4	0.087	0.108	0.154	0.021	0.095	0.090	0.106	0.130	0.140	0.090	0.123	0.171
5	0.134	0.171	0.108	0.036	0.090	0.128	0.123	0.125	0.138	0.066	0.077	0.134
6	0.164	0.113	0.122	0.096	0.084	0.109	0.141	0.124	0.141	0.102	0.129	0.170
7	0.171	0.104	0.127	0.146	0.090	0.134	0.166	0.107	0.122	0.163	0.139	0.172
8	0.192	0.047	0.104	0.160	0.033	0.196	0.164	0.134	0.110	0.160	0.116	0.216
9	0.225	0.056	0.097	0.239	0.042	0.215	0.162	0.114	0.113	0.189	0.128	0.172
high	0.302	0.040	0.020	0.315	0.021	0.385	0.214	0.132	0.119	0.243	0.103	0.313
high-low	0.461	-0.100	-0.132	0.232	-0.254	0.357	0.128	-0.007	-0.019	0.141	-0.088	0.239
	[7.25]	[-3.97]	[-3.64]	[5.33]	[-4.29]	[3.31]	[2.15]	[-0.24]	[-0.51]	[3.66]	[-1.70]	[2.46]
FF5 alpha	0.465	-0.101	-0.121	0.237	-0.250	0.365	0.142	-0.007	-0.004	0.166	-0.100	0.238
	[7.20]	[-4.10]	[-3.46]	[5.43]	[-4.05]	[3.52]	[2.41]	[-0.26]	[-0.10]	[4.15]	[-1.93]	[2.46]

Table 4: Double sorting analysis

This table reports the average annualized returns of daily rebalanced double sorted portfolios, as well as the return differentials between the top and bottom quintiles and the alphas with respect to the Fama-French (1993) factors and liquidity and momentum factors. On each day, we first sort all stocks into quintile portfolios based on a known options market predictor. Within each quintile portfolio, we further sort stocks into quintile portfolios based on $VWKS$, the options-volume weighted strike price over underlying price minus one. PC is the put-call ratio, calculated as the logarithm of put options volume over call options volume. OS is the logarithm of total options volume over underlying stock volume. DEV is the deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW$ is the options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL$ is the options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. The stock returns are calculated using the midpoint of the bid and ask prices at market close adjusted for stock splits and dividends. Portfolio returns are calculated using the market capitalization as weights. Corresponding t-statistics with Newey-West (1987) standard errors are reported in squared brackets.

VWKS	low	2	3	4	high
			Panel A: PC		
low	-0.110	-0.053	0.022	-0.042	-0.025
2	0.118	0.105	0.125	0.036	0.074
3	0.140	0.134	0.170	0.108	0.043
4	0.190	0.199	0.208	0.206	0.085
high	0.258	0.231	0.332	0.217	0.101
high-low	0.368	0.284	0.310	0.260	0.126
	[7.66]	[5.15]	[4.40]	[3.92]	[2.15]
FF5 alpha	0.364	0.282	0.322	0.256	0.125
	[8.00]	[5.25]	[4.68]	[3.85]	[2.25]
			Panel B: OS		
low	0.110	0.050	0.017	-0.013	-0.174
2	0.150	0.118	0.064	0.066	0.043
3	0.113	0.173	0.138	0.191	0.124
4	0.174	0.206	0.227	0.159	0.143
high	0.203	0.249	0.342	0.231	0.219
high-low	0.093	0.199	0.326	0.244	0.393
	[2.28]	[4.17]	[5.24]	[3.91]	[5.38]
FF5 alpha	0.097	0.205	0.325	0.245	0.393
	[2.44]	[4.21]	[5.48]	[3.88]	[5.46]

Table 4 (continued):

VWKS	low	2	3	4	high
			Panel C: DEV		
low	-0.017	-0.097	-0.051	0.010	0.160
2	0.028	-0.036	0.029	0.120	0.226
3	0.072	0.109	0.088	0.194	0.292
4	0.144	0.082	0.212	0.188	0.313
high	0.210	0.129	0.149	0.299	0.415
high-low	0.227	0.226	0.200	0.290	0.255
	[3.72]	[3.81]	[3.54]	[4.57]	[3.80]
FF5 alpha	0.237	0.225	0.198	0.293	0.244
	[3.67]	[3.97]	[3.35]	[4.79]	[3.68]
			Panel D: SKEW		
low	0.103	-0.021	-0.011	-0.019	-0.066
2	0.144	0.078	0.040	-0.003	-0.004
3	0.243	0.145	0.127	0.134	0.074
4	0.264	0.209	0.126	0.092	0.133
high	0.409	0.204	0.196	0.271	0.199
high-low	0.306	0.226	0.208	0.290	0.265
	[5.04]	[4.04]	[3.99]	[4.81]	[3.84]
FF5 alpha	0.309	0.226	0.221	0.280	0.262
	[5.08]	[4.26]	[4.13]	[4.70]	[3.89]
			Panel E: IVOL		
low	-0.116	0.006	-0.03	-0.085	0.015
2	-0.003	0.017	0.062	0.189	0.256
3	0.094	0.162	0.123	0.137	0.293
4	0.118	0.182	0.195	0.237	0.406
high	0.137	0.192	0.272	0.242	0.547
high-low	0.253	0.186	0.302	0.327	0.532
	[8.52]	[3.94]	[5.00]	[4.30]	[6.19]
FF5 alpha	0.248	0.193	0.292	0.326	0.521
	[8.10]	[4.22]	[4.98]	[4.15]	[6.17]

Table 5: Multivariate Fama-MacBeth regressions

This table investigates daily mid quote return predictability from $VWKS$, the options-volume weighted strike price over underlying price minus one. Presented are Fama-MacBeth regression results of the following equation:

$$QRET_{i,t} = \alpha + \sum_{l=1}^{5} \beta_l VWKS_{i,t-l} + \sum_{l=1}^{5} \theta_l X_{i,t-l} + \epsilon,$$

where $QRET_{i,t}$ is the mid quote returns on day t for firm i scaled to basis points; $X_{i,t-l}$ is a set of control variables on day t for firm i, including five lags of PC, OS, DEV, $SKEW$, $IVOL$, $QRET$, $SPREAD$, $TURN$ and V. PC is the put-call ratio, calculated as the logarithm of put options volume over call options volume. OS is the logarithm of total options volume over underlying stock volume. DEV is the deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW$ is the options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL$ is the options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. $QRET$ is mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends, and $QRET1$ is the mid quote return on the following day. $SPREAD$ is the percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices. $TURN$ is the turnover ratio calculated as the total trading volume over the number of shares outstanding. V is the squared raw stock returns in CRSP. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5 (continued):

	QRET1	*QRET1*	*QRET1*	*QRET1*	*QRET1*
Intercept	0.063***	0.054***	-0.058***	-0.217***	-0.186***
	[2.93]	[2.65]	[-3.35]	[-9.41]	[-8.02]
L1VWKS	0.205***	0.125***	0.101***	0.057***	0.053***
	[7.82]	[6.75]	[6.45]	[4.05]	[3.78]
L2VWKS		0.163***	0.121***	0.040***	0.035**
		[8.67]	[7.71]	[2.84]	[2.54]
L3VWKS		0.075***	0.047***	0.032**	0.032**
		[4.27]	[3.09]	[2.24]	[2.26]
L4VWKS		0.048***	0.019	0.022	0.023
		[2.67]	[1.24]	[1.50]	[1.55]
L5VWKS		0.040**	0.009	0.031**	0.029*
		[2.21]	[0.58]	[2.06]	[1.94]
L1PC			-0.003***	-0.004***	-0.003***
			[-8.40]	[-10.66]	[-10.23]
L2PC			-0.001	-0.002***	-0.002***
			[-1.45]	[-4.70]	[-4.72]
L3PC			0	-0.001**	-0.001***
			[-1.19]	[-2.55]	[-2.67]
L4PC			0	-0.001	-0.001*
			[-0.72]	[-1.60]	[-1.75]
L5PC			0	0	0
			[-0.32]	[-0.17]	[-0.29]
L1OS			-0.001	0	0
			[-1.46]	[-0.91]	[-0.69]
L2OS			-0.001**	-0.001***	-0.001**
			[-2.13]	[-2.73]	[-2.45]
L3OS			-0.001**	-0.002***	-0.002***
			[-2.47]	[-3.47]	[-3.50]
L4OS			-0.001	-0.001**	-0.001**
			[-1.10]	[-2.41]	[-2.41]
L5OS			0	-0.001**	-0.001**
			[-0.52]	[-2.17]	[-2.17]
L1DEV			-0.022	-0.001	-0.006
			[-0.52]	[-0.03]	[-0.15]
L2DEV			0.046	0.064*	0.081**
			[1.16]	[1.67]	[2.12]
L3DEV			0.074*	0.04	0.042
			[1.86]	[1.02]	[1.08]
L4DEV			-0.068*	-0.061	-0.059
			[-1.75]	[-1.57]	[-1.54]
L5DEV			0.046	0.045	0.035
			[1.26]	[1.25]	[0.97]
L1SKEW			-0.336***	-0.275***	-0.263***
			[-8.77]	[-7.43]	[-7.18]
L2SKEW			-0.135***	-0.089**	-0.091**
			[-3.51]	[-2.37]	[-2.44]
L3SKEW			0.067*	0.034	0.032
			[1.83]	[0.94]	[0.90]

Table 5 (continued):

	QRET1	*QRET1*	*QRET1*	*QRET1*	*QRET1*
L4SKEW			-0.037	-0.056	-0.049
			[-0.96]	[-1.47]	[-1.30]
L5SKEW			0.012	-0.018	-0.015
			[0.36]	[-0.56]	[-0.47]
L1IVOL			0.541***	0.322***	0.283***
			[9.09]	[5.90]	[5.21]
L2IVOL			0.098	-0.023	-0.016
			[1.62]	[-0.40]	[-0.27]
L3IVOL			-0.180***	-0.069	-0.067
			[-2.91]	[-1.19]	[-1.15]
L4IVOL			-0.056	-0.009	-0.013
			[-0.95]	[-0.15]	[-0.23]
L5IVOL			-0.203***	-0.116**	-0.104**
			[-3.99]	[-2.42]	[-2.19]
L1QRET				-0.654***	0.049
				[-4.83]	[0.34]
L2QRET				-1.430***	-1.300***
				[-11.21]	[-9.65]
L3QRET				-0.745***	-0.880***
				[-6.67]	[-7.02]
L4QRET				-0.682***	-0.788***
				[-6.19]	[-6.56]
L5QRET				-0.373***	-0.449***
				[-3.49]	[-3.92]
L1SPREAD				0.046***	0.047***
				[4.41]	[4.42]
L2SPREAD				0.031***	0.031***
				[3.11]	[3.05]
L3SPREAD				0.035***	0.032***
				[3.45]	[3.16]
L4SPREAD				0.020**	0.016*
				[2.02]	[1.65]
L5SPREAD				0.019*	0.016
				[1.89]	[1.60]
L1TURN				0.095***	0.101***
				[24.62]	[26.85]
L2TURN				-0.025***	-0.023***
				[-7.10]	[-6.74]
L3TURN				-0.004	-0.010***
				[-1.17]	[-2.83]
L4TURN				-0.002	-0.004
				[-0.68]	[-1.31]
L5TURN				-0.020***	-0.023***
				[-6.01]	[-6.90]
L1V					-2.937
					[-1.43]
L2V					-3.259*
					[-1.88]
L3V					4.879***
					[2.99]
L4V					3.143*
					[1.85]
L5V					6.668***
					[4.30]
adj. R^2	0.003	0.009	0.047	0.075	0.082
Obs	5575626	5575626	5575626	5575626	5575626

Table 6: Multivariate Fama-MacBeth regressions using moving averages

This table reports the time-series averages of the cross-sectional coefficients of the following equation:

$$QRET_{i,t} = \alpha + \beta VWKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $QRET_{i,t}$ is the mid quote returns on day t for firm i scaled to basis points; $VWKS\_MA5$ is the 5-day moving average (MA) of options-volume weighted strike price over underlying price minus one calculated on day $t-1$; and X_MA5 is a set of control variables on day $t-1$. PC_MA5 is the 5-day MA of put-call ratio, calculated as the logarithm of put options volume over call options volume. OS_MA5 is the 5-day MA of logarithm of total options volume over underlying stock volume. DEV_MA5 is the 5-day MA of deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW\_MA5$ is the 5-day MA of options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL\_MA5$ is the 5-day MA of options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. $QRET\_MA5$ is 5-day MA of mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends. $SPREAD\_MA5$ is the 5-day MA of percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices. $TURN\_MA5$ is the 5-day MA of turnover ratio calculated as the total trading volume over the number of shares outstanding. V_MA5 is the 5-day MA of squared raw stock returns in CRSP. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	QRET1	*QRET1*	*QRET1*	*QRET1*
Intercept	0.054***	-0.060***	-0.212***	-0.189***
	[2.59]	[-3.39]	[-8.82]	[-8.19]
VWKS_MA5	0.457***	0.322***	0.197***	0.201***
	[8.22]	[8.90]	[5.86]	[6.01]
PC_MA5		-0.004***	-0.007***	-0.007***
		[-5.02]	[-8.72]	[-8.66]
OS_MA5		-0.004***	-0.006***	-0.006***
		[-3.29]	[-5.75]	[-5.70]
DEV_MA5		0.150***	0.133***	0.138***
		[2.90]	[2.75]	[2.86]
SKEW_MA5		-0.447***	-0.422***	-0.419***
		[-12.32]	[-12.60]	[-12.58]
IVOL_MA5		0.197***	0.094**	0.077*
		[4.69]	[2.28]	[1.93]
QRET_MA5			-3.955***	-3.689***
			[-12.22]	[-11.24]
SPREAD_MA5			0.155***	0.148***
			[9.25]	[8.66]
TURN_MA5			0.039***	0.036***
			[8.83]	[8.08]
V_MA5				5.324
				[1.63]
adj. R^2	0.007	0.039	0.051	0.054
Obs	6121440	6121440	6121440	6121440

Table 7: Alternative measures of returns and center of volume mass in return prediction
This table presents Fama-Macbeth regression estimates of daily return prediction using alternative measures. The first model reports regression results using raw returns rather than mid quote returns as the dependent variable:

$$RET_{i,t} = \alpha + \beta VWKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $RET_{i,t+1}$ is the raw stock returns in CRSP on day t. The second model reports regression results using the following equation:

$$QRET_{i,t} = \alpha + \beta VWLNKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $VWLNKS\_MA5$ is the 5-day moving average (MA) of the volume weighted log strike price over underlying stock price. The third model reports regression using the following equation:

$$QRET_{i,t} = \alpha + \beta VWDELTA\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $VWDELTA\_MA5$ is the 5-day MA of the volume weighted $DELTA$ (put $DELTA = DELTA + 1$). And the fourth model reports the following regression:

$$QRET_{i,t} = \alpha + \beta VWKLS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $VWKLS\_MA5$ is defined as the 5-day MA of volume weighted strike price over previous day's underlying stock price. $QRET_{i,t}$ is the mid quote returns on day t. The X_MA5 is a set of control variables on day $t-1$. PC_MA5 is the 5-day MA of put-call ratio, calculated as the logarithm of put options volume over call options volume. OS_MA5 is the 5-day MA of logarithm of total options volume over underlying stock volume. DEV_MA5 is the 5-day MA of deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW\_MA5$ is the 5-day MA of options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL\_MA5$ is the 5-day MA of options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. $QRET\_MA5$ is 5-day MA of mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends. RET_MA5 is 5-day MA of raw returns in CRSP. $SPREAD\_MA5$ is the 5-day MA of percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices. $TURN\_MA5$ is the 5-day MA of turnover ratio calculated as the total trading volume over the number of shares outstanding. V_MA5 is the 5-day MA of squared raw stock returns. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 7 (continued):

	RET1		*QRET1*		*QRET1*		*QRET1*
Intercept	-0.157***	*Intercept*	-0.190***	*Intercept*	-0.143***	*Intercept*	-0.187***
	[-6.87]		[-8.23]		[-5.53]		[-7.93]
VWKS_MA5	0.165***	*VWLNKS_MA5*	0.202***	*VWDELTA_MA5*	-0.099***	*VWKLS_MA5*	0.129***
	[4.88]		[5.49]		[-3.49]		[5.58]
PC_MA5	-0.007***	*PC_MA5*	-0.007***	*PC_MA5*	-0.006***	*PC_MA5*	-0.007***
	[-9.12]		[-8.67]		[-7.95]		[-8.58]
OS_MA5	-0.005***	*OS_MA5*	-0.006***	*OS_MA5*	-0.005***	*OS_MA5*	-0.006***
	[-5.01]		[-5.45]		[-5.13]		[-5.19]
DEV_MA5	1.102***	*DEV_MA5*	0.135***	*DEV_MA5*	0.139***	*DEV_MA5*	0.180***
	[19.39]		[2.80]		[2.89]		[3.54]
SKEW_MA5	-0.352***	*SKEW_MA5*	-0.421***	*SKEW_MA5*	-0.439***	*SKEW_MA5*	-0.425***
	[-10.47]		[-12.62]		[-13.05]		[-12.13]
IVOL_MA5	0.082**	*IVOL_MA5*	0.084**	*IVOL_MA5*	0.098**	*IVOL_MA5*	0.087**
	[2.03]		[2.11]		[2.43]		[2.10]
RET_MA5	-4.681***	*QRET_MA5*	-3.702***	*QRET_MA5*	-3.820***	*QRET_MA5*	-3.693***
	[-13.66]		[-11.30]		[-11.63]		[-11.18]
SPREAD_MA5	0.150***	*SPREAD_MA5*	0.149***	*SPREAD_MA5*	0.148***	*SPREAD_MA5*	0.162***
	[8.71]		[8.71]		[8.69]		[8.62]
TURN_MA5	0.032***	*TURN_MA5*	0.036***	*TURN_MA5*	0.036***	*TURN_MA5*	0.035***
	[7.35]		[8.13]		[8.11]		[7.78]
V_MA5	8.446**	*V_MA5*	5.558*	*V_MA5*	5.273	*V_MA5*	4.169
	[2.54]		[1.70]		[1.61]		[1.27]
adj. R^2	0.055	adj. R^2	0.054	adj. R^2	0.053	adj. R^2	0.054
Obs	6121440	Obs	6121440	Obs	6121440	Obs	6121440

Table 8: Center of volume mass in call and put options

This table reports Fama-Macbeth estimates of daily mid quote return predictions using the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSCALL\_MA5_{i,t-1} + \beta_2 VWKSPUT\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $QRET_{i,t}$ is the mid quote returns on day t. $VWKSCALL\_MA5$ is the 5-day moving average (MA) of volume weighted call options strike price over underlying stock price. $VWKSPUT\_MA5$ is the 5-day MA of volume weighted put options strike price over underlying stock price. The X_MA5 is a set of control variables on day $t-1$. PC_MA5 is the 5-day MA of put-call ratio, calculated as the logarithm of put options volume over call options volume. OS_MA5 is the 5-day MA of logarithm of total options volume over underlying stock volume. DEV_MA5 is the 5-day MA of deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW\_MA5$ is the 5-day MA of options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL\_MA5$ is the 5-day MA of options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. $QRET\_MA5$ is 5-day MA of mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends. $SPREAD\_MA5$ is the 5-day MA of percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices. $TURN\_MA5$ is the 5-day MA of turnover ratio calculated as the total trading volume over the number of shares outstanding. V_MA5 is the 5-day MA of squared raw stock returns. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	QRET1	*QRET1*	*QRET1*	*QRET1*	*QRET1*
Intercept	0.048**	-0.195***	0.073***	-0.185***	-0.187***
	[2.45]	[-8.44]	[3.30]	[-8.02]	[-8.16]
VWKSCALL_MA5	0.387***	0.195***			0.154***
	[6.29]	[5.65]			[4.88]
VWKSPUT_MA5			0.438***	0.195***	0.160***
			[8.89]	[5.65]	[5.09]
PC_MA5		-0.007***		-0.007***	-0.007***
		[-9.63]		[-9.24]	[-9.20]
OS_MA5		-0.007***		-0.005***	-0.006***
		[-6.22]		[-4.67]	[-5.67]
DEV_MA5		0.129***		0.154***	0.145***
		[2.69]		[3.18]	[3.03]
SKEW_MA5		-0.433***		-0.419***	-0.415***
		[-12.93]		[-12.55]	[-12.50]
IVOL_MA5		0.077*		0.093**	0.075*
		[1.92]		[2.30]	[1.88]
QRET_MA5		-3.688***		-3.747***	-3.626***
		[-11.24]		[-11.42]	[-11.08]
SPREAD_MA5		0.152***		0.146***	0.148***
		[8.94]		[8.57]	[8.79]
TURN_MA5		0.035***		0.036***	0.036***
		[8.02]		[8.15]	[8.14]
V_MA5		5.454*		5.1	4.858
		[1.67]		[1.56]	[1.49]
adj. R^2	0.009	0.054	0.004	0.054	0.055
Obs	6121440	6121440	6121440	6121440	6121440

Table 9: Center of volume mass in long-term and short-term options

This table reports Fama-Macbeth regression using the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSM1\_MA5_{i,t-1} + \beta_2 VWKSM2\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $QRET_{i,t}$ is the mid quote returns on day t. $VWKSM1\_MA5$ is the 5-day moving average (MA) of volume weighted strike price over underlying stock price using options expiring in fewer than or equal to 30 days. $VWKSM2\_MA5$ is the 5-day MA of volume weighted strike price over underlying stock price using options expiring more than 30 days. The X_MA5 is a set of control variables on day $t-1$. PC_MA5 is the 5-day MA of put-call ratio, calculated as the logarithm of put options volume over call options volume. OS_MA5 is the 5-day MA of logarithm of total options volume over underlying stock volume. DEV_MA5 is the 5-day MA of deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW\_MA5$ is the 5-day MA of options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL\_MA5$ is the 5-day MA of options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. $QRET\_MA5$ is 5-day MA of mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends. $SPREAD\_MA5$ is the 5-day MA of percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices. $TURN\_MA5$ is the 5-day MA of turnover ratio calculated as the total trading volume over the number of shares outstanding. V_MA5 is the 5-day MA of squared raw stock returns. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	QRET1	*QRET1*	*QRET1*	*QRET1*	*QRET1*
Intercept	0.065***	-0.194***	0.054***	-0.188***	-0.188***
	[3.00]	[-8.33]	[2.59]	[-8.15]	[-8.15]
VWKSM1_MA5	0.211***	0.077***			0.055***
	[5.67]	[3.42]			[2.66]
VWKSM2_MA5			0.284***	0.112***	0.096***
			[7.61]	[5.27]	[4.82]
PC_MA5		-0.008***		-0.007***	-0.007***
		[-10.41]		[-8.71]	[-9.35]
OS_MA5		-0.006***		-0.006***	-0.006***
		[-5.24]		[-5.42]	[-5.43]
DEV_MA5		0.137***		0.138***	0.138***
		[2.83]		[2.88]	[2.87]
SKEW_MA5		-0.432***		-0.428***	-0.420***
		[-12.95]		[-12.78]	[-12.65]
IVOL_MA5		0.095**		0.082**	0.079**
		[2.32]		[2.06]	[1.98]
QRET_MA5		-3.781***		-3.768***	-3.722***
		[-11.50]		[-11.47]	[-11.37]
SPREAD_MA5		0.151***		0.148***	0.148***
		[8.79]		[8.62]	[8.63]
TURN_MA5		0.036***		0.035***	0.036***
		[8.11]		[8.02]	[8.07]
V_MA5		5.627*		5.531*	5.397*
		[1.72]		[1.69]	[1.65]
adj. R^2	0.004***	0.053***	0.006***	0.053***	0.054***
Obs	6121440	6121440	6121440	6121440	6121440

Table 10: Asymmetric price impact from center of options volume mass

This table presents Fama-Macbeth regression using the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSP\_MA5_{i,t-1} + \beta_2 VWKSN\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $QRET_{i,t}$ is the mid quote returns on day t. $VWKSP\_MA5$ is the 5-day moving average (MA) of volume weighted strike price over underlying stock price if strike price is larger than stock price, zero otherwise. $VWKSN\_MA5$ is the 5-day MA of volume weighted strike price over underlying stock price if strike price is smaller than stock price, zero otherwise. The X_MA5 is a set of control variables on day $t-1$. PC_MA5 is the 5-day MA of put-call ratio, calculated as the logarithm of put options volume over call options volume. OS_MA5 is the 5-day MA of logarithm of total options volume over underlying stock volume. DEV_MA5 is the 5-day MA of deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW\_MA5$ is the 5-day MA of options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL\_MA5$ is the 5-day MA of options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. $QRET\_MA5$ is 5-day MA of mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends. $SPREAD\_MA5$ is the 5-day MA of percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices. $TURN\_MA5$ is the 5-day MA of turnover ratio calculated as the total trading volume over the number of shares outstanding. V_MA5 is the 5-day MA of squared raw stock returns. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	QRET1	*QRET1*	*QRET1*	*QRET1*	*QRET1*
Intercept	0.038*	-0.188***	0.076***	-0.190***	-0.188***
	[1.86]	[-7.98]	[3.28]	[-8.05]	[-7.97]
VWKSP_MA5	0.483***	0.168***			0.147***
	[9.08]	[5.34]			[4.72]
VWKSN_MA5			0.237***	0.184***	0.084*
			[3.12]	[3.86]	[1.89]
PC_MA5		-0.007***		-0.007***	-0.007***
		[-8.59]		[-8.94]	[-8.56]
OS_MA5		-0.005***		-0.006***	-0.006***
		[-4.89]		[-5.55]	[-5.21]
DEV_MA5		0.182***		0.183***	0.185***
		[3.56]		[3.59]	[3.64]
SKEW_MA5		-0.424***		-0.440***	-0.421***
		[-12.08]		[-12.46]	[-12.03]
IVOL_MA5		0.080*		0.110***	0.085**
		[1.93]		[2.62]	[2.03]
QRET_MA5		-3.692***		-3.746***	-3.662***
		[-11.15]		[-11.32]	[-11.09]
SPREAD_MA5		0.160***		0.167***	0.161***
		[8.52]		[8.89]	[8.64]
TURN_MA5		0.035***		0.034***	0.035***
		[7.88]		[7.64]	[7.85]
V_MA5		3.973		4.288	3.991
		[1.22]		[1.31]	[1.22]
adj. R^2	0.006	0.054	0.002	0.053	0.054
Obs	5880724	5880724	5880724	5880724	5880724

Table 11: Nonlinear pricing impact from center of options volume mass

This table presents Fama-Macbeth regression using the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSSQ\_MA5_{i,t-1} + \beta_2 VWKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $QRET_{i,t}$ is the mid quote returns on day t. $VWKSSQ\_MA5$ is the 5-day moving average (MA) of squared volume weighted strike price over underlying stock price. $VWKS\_MA5$ is the 5-day MA of volume weighted strike price over underlying stock price. The X_MA5 is a set of control variables on day $t-1$. PC_MA5 is the 5-day MA of put-call ratio, calculated as the logarithm of put options volume over call options volume. OS_MA5 is the 5-day MA of logarithm of total options volume over underlying stock volume. DEV_MA5 is the 5-day MA of deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW\_MA5$ is the 5-day MA of options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL\_MA5$ is the 5-day MA of options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. $QRET\_MA5$ is 5-day MA of mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends. $SPREAD\_MA5$ is the 5-day MA of percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices. $TURN\_MA5$ is the 5-day MA of turnover ratio calculated as the total trading volume over the number of shares outstanding. V_MA5 is the 5-day MA of squared raw stock returns. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	QRET1	*QRET1*	*QRET1*
Intercept	0.064***	-0.188***	-0.187***
	[2.86]	[-7.90]	[-7.95]
VWKSSQ_MA5	0.611***	0.376***	0.600***
	[10.89]	[8.31]	[8.13]
VWKS_MA5			-0.067*
			[-1.82]
PC_MA5		-0.007***	-0.007***
		[-9.00]	[-8.89]
OS_MA5		-0.006***	-0.006***
		[-5.14]	[-5.19]
DEV_MA5		0.184***	0.185***
		[3.59]	[3.63]
SKEW_MA5		-0.425***	-0.423***
		[-12.06]	[-12.07]
IVOL_MA5		0.088**	0.088**
		[2.09]	[2.13]
QRET_MA5		-3.706***	-3.770***
		[-11.15]	[-11.47]
SPREAD_MA5		0.161***	0.158***
		[8.59]	[8.44]
TURN_MA5		0.035***	0.035***
		[7.81]	[7.87]
V_MA5		3.952	3.943
		[1.21]	[1.20]
adj. R^2	0.002	0.054	0.055
Obs	5880724	5880724	5880724

Table 12: Long-term moving average of and shock to center of options volume mass
This table reports Fama-Macbeth regression using the following equation:

$$QRET_{i,t} = \alpha + \beta_1 VWKSS20_{i,t-1} + \beta_2 VWKSMA20_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon,$$

where $QRET_{i,t}$ is the mid quote returns on day t. $VWKSMA20$ is the past 20-day moving average (MA) of volume weighted strike price over underlying stock price from day $t-25$ to day $t-6$. $VWKSS20$ is the difference between 5-day MA of $VWKS$ and $VWKSMA20$. The X_MA5 is a set of control variables on day $t-1$. PC_MA5 is the 5-day MA of put-call ratio, calculated as the logarithm of put options volume over call options volume. OS_MA5 is the 5-day MA of logarithm of total options volume over underlying stock volume. DEV_MA5 is the 5-day MA of deviation from put-call parity, calculated as the average difference in implied volatilities between call options and put options. $SKEW\_MA5$ is the 5-day MA of options implied skewness, calculated as the difference between the implied volatilities of out-of-the-money puts and at-the-money calls. $IVOL\_MA5$ is the 5-day MA of options-implied volatility, calculated as the average implied volatility of at-the-money call options and at-the-money put options. $QRET\_MA5$ is 5-day MA of mid quote returns calculated using closing bid-ask prices and adjusted for stock splits and dividends. $SPREAD\_MA5$ is the 5-day MA of percentage bid-ask spread calculated as the ask minus bid scaled by the midpoint of the bid and ask prices. $TURN\_MA5$ is the 5-day MA of turnover ratio calculated as the total trading volume over the number of shares outstanding. V_MA5 is the 5-day MA of squared raw stock returns. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	QRET1	*QRET1*	*QRET1*	*QRET1*	*QRET1*
Intercept	0.062***	-0.168***	0.049**	-0.165***	-0.162***
	[2.74]	[-6.81]	[2.34]	[-6.79]	[-6.70]
VWKSS20	0.275***	0.109***			0.156***
	[7.41]	[4.65]			[6.40]
VWKSMA20			0.383***	0.175***	0.222***
			[4.61]	[3.15]	[3.86]
PC_MA5		-0.006***		-0.006***	-0.006***
		[-7.25]		[-7.34]	[-6.87]
OS_MA5		-0.005***		-0.006***	-0.006***
		[-4.11]		[-4.65]	[-4.47]
DEV_MA5		0.247***		0.257***	0.256***
		[4.10]		[4.25]	[4.24]
SKEW_MA5		-0.441***		-0.419***	-0.406***
		[-11.06]		[-10.47]	[-10.21]
IVOL_MA5		0.103**		0.075*	0.067
		[2.34]		[1.77]	[1.58]
QRET_MA5		-3.614***		-3.861***	-3.669***
		[-10.50]		[-11.32]	[-10.76]
SPREAD_MA5		0.172***		0.169***	0.168***
		[7.79]		[7.57]	[7.53]
TURN_MA5		0.030***		0.030***	0.031***
		[6.37]		[6.51]	[6.62]
V_MA5		2.362		1.623	1.507
		[0.69]		[0.48]	[0.45]
adj. R^2	0.003	0.056	0.009	0.058	0.059
obs	4982833	4982833	4982833	4982833	4982833

Table 13: Subsample analysis

In each panel, the full sample is divided into two groups based on a proxy for information asymmetry: low (<50th percentile) and high (>50th percentile). The slope coefficients and t-statistics (in parentheses) are reported only for five lags of volume-weighted strike-to-spot price ratio ($VWKS$) from the Fama-Macbeth regression using the following equation:

$$QRET_{i,t} = \alpha + \sum_{l=0}^{4} \beta_l VWKS_{i,t-l} + \sum_{l=0}^{4} \theta_l X_{i,t-l} + \epsilon,$$

and coefficients of a 5-day moving average (MA) of $VWKS$ ($VWKS\_MA5_{i,t-1}$) from the regression using the following equation:

$$QRET_{i,t} = \alpha + \beta VWKS\_MA5_{i,t-1} + \theta X\_MA5_{i,t-1} + \epsilon.$$

The conditioning variable is firm market capitalization ($Size$) in Panel A, idiosyncratic stock volatility ($Idio$) in Panel B, illiquidity measured as in Amihud (2002) in Panel C, analyst coverage ($Analyst$) in Panel D, fraction of institutional ownership ($Ownership$) in Panel E, the probability of informed trading (PIN) as of Easley, Kiefer, O'Hara, and Paperman (1996) in Panel F, total options trading volume ($Volume$) in Panel E, and sample period ($Year$) in Panel F. $X_{i,t-l}$ is a set of control variables on day t for firm i, and $X\_MA5_{i,t-l}$ is the 5-day MA of $X_{i,t-l}$, and are defined the same as before. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. We further compare the coefficient estimates on $VWKS\_MA5$ between each pair of subsamples using the unpaired t-test.

	Size			*Idio*		
	low	*high*	*low-high*	*low*	*high*	*low-high*
L1VWKS	0.055***	0.093***		0.029*	0.050**	
	[2.73]	[4.01]		[1.72]	[2.22]	
L2VWKS	0.056***	0.015		0.021	0.058***	
	[2.94]	[0.62]		[0.79]	[2.67]	
L3VWKS	0.068***	0.002		0.02	0.046**	
	[3.41]	[0.08]		[0.59]	[2.04]	
L4VWKS	0.026	0.036		-0.033	0.034	
	[1.23]	[1.52]		[-1.27]	[1.45]	
L5VWKS	0.053**	-0.016		0.005	0.040*	
	[2.53]	[-0.59]		[0.28]	[1.73]	
VWKS_MA5	0.316***	0.118**	0.198***	0.023	0.271***	-0.248***
	[7.83]	[2.31]	[4.20]	[0.63]	[5.82]	[-5.79]

Table 13 (continued):

	Illiquidity			*Analyst*		
	low	*high*	*low-high*	*low*	*high*	*low-high*
L1VWKS	0.080***	0.051**		0.041**	0.097***	
	[3.19]	[2.48]		[2.03]	[4.21]	
L2VWKS	-0.017	0.068***		0.042**	0.093	
	[-0.67]	[3.51]		[2.19]	[1.35]	
L3VWKS	-0.080**	0.073***		0.03	0.025	
	[-1.99]	[3.67]		[1.50]	[1.03]	
L4VWKS	-0.013	0.047**		0.03	0.024	
	[-0.59]	[2.18]		[1.44]	[0.86]	
L5VWKS	0.014	0.061***		0.016	-0.025	
	[0.45]	[2.88]		[0.77]	[-0.40]	
VWKS_MA5	0.044	0.350***	-0.306***	0.205***	0.192***	0.013
	[0.84]	[8.67]	[-6.40]	[5.19]	[3.92]	[0.29]

	Owner			*PIN*		
	low	*high*	*low-high*	*low*	*high*	*low-high*
L1VWKS	0.045**	0.077***		0.086***	0.042**	
	[2.14]	[3.23]		[3.62]	[2.03]	
L2VWKS	0.038*	0.026		-0.042	0.034*	
	[1.81]	[1.15]		[-0.49]	[1.69]	
L3VWKS	0.050**	0.018		0.007	0.063***	
	[2.36]	[0.76]		[0.11]	[3.05]	
L4VWKS	0.022	0.042*		0.021	0.026	
	[0.99]	[1.88]		[0.62]	[1.18]	
L5VWKS	0.043*	0		0.025	0.032	
	[1.90]	[-0.02]		[1.01]	[1.45]	
VWKS_MA5	0.247***	0.186***	0.061	0.172***	0.237***	-0.062
	[5.68]	[3.82]	[1.36]	[3.45]	[5.66]	[-1.34]

	Volume			*Year*		
	low	*high*	*low-high*	*96-04*	*05-13*	*low-high*
L1VWKS	-0.016	0.083***		0.064***	0.049***	
	[-0.89]	[3.45]		[2.60]	[2.71]	
L2VWKS	0.006	0.065***		0.049**	0.015	
	[0.34]	[2.67]		[2.08]	[0.76]	
L3VWKS	0.03	0.007		0.047*	0.008	
	[1.63]	[0.28]		[1.95]	[0.42]	
L4VWKS	0.023	0.007		0.042	0.004	
	[1.18]	[0.31]		[1.62]	[0.23]	
L5VWKS	0.02	0.027		0.044*	0.015	
	[1.04]	[1.05]		[1.66]	[0.79]	
VWKS_MA5	0.111***	0.236***	-0.125***	0.202***	0.066**	0.136**
	[2.90]	[4.79]	[-2.61]	[5.49]	[2.39]	[3.25]

Table 14: Center of options volume mass around corporate events

For each type of corporate event in each column, we present the pooled ordinary least squares results of the following equation:

$$VWKS = \alpha + \beta_0 EVENT + \sum_{i=1}^{5} \beta_1^i PREEVENTi + \sum_{i=1}^{5} \beta_2^i POSTEVENTi + \theta X + \epsilon,$$

where $EVENT$ is a category variable with a value of 1 (-1) if there is a positive (negative) corporate event on the same day t, and zero otherwise. Events are signed by the cumulative abnormal return on the event day ($CAR0$). $PREEVENTi$ is a pre-event category variable with a value of 1 (-1) if there is a positive (negative) event on day $t+i$, and zero otherwise; and $POSTEVENTi$ is a post-event category variable with a value of 1 (-1) if there is a positive (negative) event on day $t-i$, and zero otherwise. We include the firm, year and week fixed effects (FE). Standard errors are clustered by firms. Scheduled events are those from earnings announcements. Unscheduled events are 8-K filings that are not related to earnings news. Jumps are identified if the risk-adjusted return is higher than 10% based on Savor (2012) or if the risk-adjusted return is above two standard deviations by Boehmer and Wu (2013), and are not related to either 8-K filings or earnings announcements. *permjump* are price jumps whose $CAR0$ has the same sign as the cumulative abnormal return on the following one to five days ($CAR5$). *tranjump* are price jumps that reverse within the following five trading days ($sign(CAR0 * CAR5) < 0$). Associated t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

VWKS	*scheduled*	*unscheduled*	*permjump*	*tranjump*	*all events*
EVENT	-0.005***	-0.004***	-0.007***	-0.018***	-0.005***
	[-6.28]	[-8.74]	[-6.02]	[-15.62]	[-15.97]
PREEVENT1	0.010***	0.004***	0.010***	0	0.005***
	[11.93]	[9.58]	[8.16]	[0.28]	[15.65]
PREEVENT2	0.008***	0.003***	0.009***	-0.001	0.004***
	[10.12]	[6.83]	[7.58]	[-0.79]	[10.99]
PREEVENT3	0.006***	0.003***	0.006***	-0.001	0.003***
	[7.50]	[6.33]	[5.51]	[-1.08]	[9.07]
PREEVENT4	0.003***	0.002***	0.004***	-0.001	0.002***
	[4.46]	[4.15]	[4.18]	[-1.27]	[5.73]
PREEVENT5	0.002***	0.001***	0.003***	0	0.001***
	[2.72]	[3.92]	[3.49]	[-0.51]	[5.10]
POSTEVENT1	-0.022***	-0.006***	-0.003***	-0.010***	-0.009***
	[-26.85]	[-15.06]	[-2.63]	[-8.35]	[-26.14]
POSTEVENT2	-0.012***	-0.004***	-0.005***	-0.007***	-0.005***
	[-15.47]	[-9.99]	[-4.14]	[-6.52]	[-16.52]
POSTEVENT3	-0.010***	-0.003***	-0.004***	-0.005***	-0.004***
	[-12.71]	[-7.33]	[-3.46]	[-4.89]	[-13.01]
POSTEVENT4	-0.007***	-0.002***	-0.005***	-0.004***	-0.003***
	[-9.22]	[-6.06]	[-4.66]	[-3.80]	[-10.75]
POSTEVENT5	-0.005***	-0.002***	-0.004***	0	-0.002***
	[-7.60]	[-4.97]	[-4.08]	[0.40]	[-8.01]
Year & Week FE	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES
Intercept	0.028***	0.028***	0.028***	0.028***	0.028***
	[583.46]	[582.85]	[583.12]	[582.97]	[583.61]
adj. R^2	0.367	0.367	0.367	0.367	0.367
Obs	5130705	5130705	5130290	5130113	5129698

Table 15: Price sensitivity to center of options volume mass around corporate events
This table studies price sensitivity of $VWKS\_MA5$ to different types of events by Fama-Macbeth regression:

$$QRET_{i,t} = \alpha + \beta VWKS\_MA5_{i,t-1} + dummy + VWKS\_MA5_{i,t-1} * dummy + \theta X\_MA5_{i,t-1} + \epsilon,$$

where the event dummy $SCHEDULED$ takes the value of one if there is an earnings announcement on day t, and zero otherwise; the event dummy $UNSCHEDULED$ takes the value of one if there is an 8-K filing unrelated to earnings announcement on day t, and zero otherwise; the event dummy $PERMJUMP$ takes the value of one if there is a permanent price jump unrelated to 8-K filing or earnings announcement on day t, and zero otherwise; and the event dummy $TRANJUMP$ takes the value of one if there is a transitory price jump (with a complete return reversal within five trading days) unrelated to 8-K filing or earnings announcement on day t, and zero otherwise. Standard errors are calculated with the Newey-West adjustment to four lags. Associated t-statistics are reported in parentheses ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	QRET1	*QRET1*	*QRET1*	*QRET1*	*QRET1*
Intercept	-0.190***	-0.189***	-0.189***	-0.191***	-0.192***
	[-8.19]	[-8.17]	[-8.19]	[-8.27]	[-8.32]
VWKS_MA5	0.135***	0.132***	0.129***	0.132***	0.117***
	[5.80]	[5.64]	[5.59]	[5.68]	[5.09]
SCHEDULED	14.096				14.113
	[1.01]				[1.01]
*KS*SCHEDULED*	-76.818				-76.883
	[-0.79]				[-0.79]
UNSCHEDULED		0.033**			0.039**
		[2.01]			[2.30]
*KS*UNSCHEDULED*		0.715			0.749
		[1.49]			[1.53]
PERMJUMP			0.155***		0.158***
			[5.54]		[5.65]
*KS*PERMJUMP*			1.423***		1.433***
			[3.64]		[3.66]
TRANJUMP				0.242***	0.245***
				[8.44]	[8.51]
*KS*TRANJUMP*				-0.868	-0.857
				[-0.48]	[-0.47]
PC_MA5	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***
	[-9.05]	[-8.99]	[-9.04]	[-9.13]	[-9.23]
OS_MA5	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***
	[-5.43]	[-5.35]	[-5.38]	[-5.47]	[-5.64]
DEV_MA5	0.133***	0.136***	0.135***	0.135***	0.123**
	[2.77]	[2.82]	[2.81]	[2.80]	[2.57]
SKEW_MA5	-0.424***	-0.423***	-0.422***	-0.422***	-0.424***
	[-12.76]	[-12.69]	[-12.72]	[-12.70]	[-12.85]
IVOL_MA5	0.079**	0.082**	0.082**	0.081**	0.077*
	[1.98]	[2.03]	[2.04]	[2.01]	[1.92]
QRET_MA5	-3.715***	-3.733***	-3.739***	-3.741***	-3.728***
	[-11.32]	[-11.36]	[-11.39]	[-11.41]	[-11.40]
SPREAD_MA5	0.146***	0.147***	0.148***	0.148***	0.146***
	[8.59]	[8.58]	[8.63]	[8.67]	[8.61]
TURN_MA5	0.036***	0.035***	0.035***	0.036***	0.035***
	[8.08]	[8.04]	[8.06]	[8.13]	[8.02]
V_MA5	5.741*	5.522*	5.837*	6.033*	6.478**
	[1.75]	[1.69]	[1.79]	[1.85]	[1.99]
adj. R^2	0.057	0.055	0.056	0.056	0.064
Obs	6121440	6121440	6121440	6121440	6121440