Article:

Zhang, M., Verousis, T. and Kalaitzoglou, I. (2022) Information and the arrival rate of option trading volume. Journal of Futures Markets, 42 (4). pp. 605-644. ISSN: 0270-7314

https://doi.org/10.1002/fut.22299

Information and the arrival rate of option trading volume.

Mengyu Zhang[a], Thanos Verousis[b*], Iordanis Kalaitzoglou[c]

Abstract

In this paper we investigate the interaction between liquidity and information in the options market and its impact on the pricing of the underlying asset. We model option trade duration and volume jointly, for the first time, as a natural measure of options' trading intensity and we associate it with differential degrees of information present in option trades. We report a highly significant association between option trading intensity with contemporaneous and future underlying volatility and returns, which is robust to the presence of other information measures, market factors and structural forms.

JEL classifications: G12, G14

Keywords**:** Options, stocks, trading volume, liquidity, information, conditional duration

* Corresponding author at: Thanos Verousis, Essex Business School, University of Essex, UK.
a. Kent Business School, University of Kent, Medway, UK. Email: M.Zhang-280@kent.ac.uk

b. Essex Business School, University of Essex, Colchester, UK. Email: t.verousis@essex.ac.uk

c. Audencia Business School, Department of Finance, France. Email: ikalaitzoglou@audencia.com

1. Introduction

"*…an investor who wants the action on a stock has two ways of getting it. He can deal directly in the stock, or he can deal in the option*"

Black (1975, p. 37)

Options are redundant securities: their payoff can be fully replicated by the payoff of the underlying securities. That is however, when markets are complete. In incomplete markets, Ross (1976) argued that options have the power to improve market efficiency by expanding the number of contingencies that can be covered by the market. In incomplete markets with information asymmetries, Black (1975) introduced the idea that informed traders could use the options market due to their high leverage and Biais and Hillion (1994) showed that the effect of the introduction of options markets on market completeness and efficiency depends on the type of liquidity orders. Understanding therefore the role of liquidity in options markets is key to understanding the information content of option trades. Against this background, in this paper, we employ a model of intraday trading intensity in options markets that accounts for most of the dimensions of liquidity and classifies option trades according to their inferred level of information.

A crucial contribution of our paper is that we associate information resolution with fluctuations of liquidity in options markets, by modelling an observable proxy: trading intensity. Even though the idea that option trades may reveal important information is not new, previous literature does not account for the multiple ways that information is manifested in the options markets. Easley et al. (1998) investigate the role of option volume in a market microstructure model of two markets in which investors decide to trade in the option or the underlying market (see also Pan and Poteshman, 2006).[1] Hu (2014) shows that option trading and in particular option-induced imbalance significantly predicts future stock returns. More recently, Mkansi and Acheampong (2012) and Chung et al. (2016) investigate the role of option trade duration on the price

[1] Empirically, Kacperczyk and Pagnotta (2019) show that option markets are used extensively by informed traders. Patel et al. (2020) further show that the average level of price discovery of US stocks with active stock options is approximately 29%.

impact of trades.[2] However, none of these studies considers that information might affect all of them simultaneously, or that they might have a combined impact on subsequent price discovery. This is the primary concern of our study.

In particular, we apply a Smooth Transition Autoregressive Conditional Weighted Duration (STM-ACWD) model (see Kalaitzoglou and Ibrahim, 2015) to capture the duality of the impact of both option trade duration and volume. This model employs a time rescaling that focuses on event, rather than on calendar time, which in this case is defined as the waiting time (in seconds) for trading an option contract (i.e., arrival rate of option trading volume). This natural measure of trading intensity exhibits several properties that are essential for investigating the combined effect of volume and time. It combines in one variable two dimensions of trading intensity, namely duration and volume of trading, without an explicit specification of their distributional properties and therefore, a lower degree of parameterization is required. In addition, this measure of trading intensity relates to four out of five recognized dimensions of liquidity[3] and therefore, it can capture latent variations in what can be considered as "normal" liquidity levels; a concept introduced as 'relative' liquidity. This is a particularly relevant attribute to the venture here, because it provides a conditional way to infer the presence of information from 'relative' liquidity fluctuation. Easley and O'Hara (1992) links 'relative' liquidity levels to information and the model employed here, investigates this at a granular, i.e., transaction level.

---

[2] In stock markets, Easley and O'Hara (1992) show that if volume and therefore trading reveals some information, then the absence of volume, and therefore the absence of any trades, reveals a different type of information to market participants. Dufour and Engle (2000) show that duration between trades is strongly and negatively correlated with the presence of informed traders. These two findings combined imply that it is not simply the magnitude of volume that is associated with information, but also its accumulation rate. The empirical findings of Dufour and Engle (2000) suggest that faster realization of trades and thus, faster accumulation of volume, is associated with more information and subsequent higher pricing impact. Our approach here focuses exactly on this point and models the rate at which volume accumulates, i.e., the "arrival rate of volume".

[3] Trading intensity relates to immediacy, resiliency, tightness and depth. It may only indirectly relate to breadth (see Kalaitzoglou and Ibrahim, 2015).

The STM-ACWD framework, employed here, suggests a modelling of this rescaled time as a superposed point process and tries to extract the marginal (latent) point processes from the observed market wide trading activity. This is particularly important in our analysis because with this model we can identify the probability of each individual trade to be initiated by different agents with progressively higher access to price unresolved information. The way the liquidity is linked to various degrees of information is based on the arrival rate of volume and its time variation, which can only be done if volume and duration are modelled jointly. This is a major contribution of our work because, unlike previous literature that infers the 'informativeness' of options trades from the properties of aggregated variables such as imbalances (e.g., Hu, 2014), the O/S ratio (e.g., Ryu and Yang, 2018), volume (e.g., Pan and Poteshman, 2006) and/or duration (e.g., Chung et al., 2016), it focuses on how different traders with differing degrees of access to information might interact with the market. This interaction is captured by the 'arrival rate' of the volume of their trading, which is measured by a precise statistical measure, i.e., the conditional intensity ('hazard rate') of trading intensity, which might take different shapes that are consistent with the behaviour of different agent types (e.g., Kalaitzoglou, 2020, 2021). Consequently, we provide a precise statistical measure for agent classification (based on information) at a granular, i.e., instantaneous, rather than at an aggregated level.

Previous literature that is most closely related to our study focuses mostly on the arrival rate of options (e.g., Chung et al. 2016) or the underlying asset (e.g., Cartea and Meyer-Brandis , 2010) trades (i.e., duration), mainly controlling for, rather than modelling the dimension of volume. They consistently report a positive correlation between more frequent trading (i.e., shorter duration) and the presence of information, which is highly in line with previous literature on market microstructure (e.g., Easley and O'Hara, 1992; Engle, 2000). However, although they account for the effects of trading volume by controlling for confounding effects, they do not explicitly model it as a factor associated to the presence of information (e.g., Easley et al., 2011b, 2012, 2014a, 2014b). Our paper however, instead of controlling for the effect of option trade volume, it explicitly models option trade duration and trading volume. The joint modelling of option trade duration and option trading volume enables us to rank option trades based on the level of informativeness encapsulated in the intensity of trading and subsequently investigate how options' trading intensity predicts the underlying stock price movements.[4]

With respect to trading intensity modelling and information, the empirical findings conform to the initial intuition. Trading intensity is persistent, which implies a prolonged acceleration/deceleration of trading activity and thus, relatively distinct periods of higher or lower liquidity. Our modelling, based on the variations of the arrival rate of options trading volume, is able to differentiate between relatively low and high trading intensity levels and associate them with differential degrees of association with information. In consistency with previous empirical evidence on other assets, we report a direct link between higher options trading intensity and presence of information.

We investigate further the link between the extracted information signal and subsequent price discovery and report a number of significant findings: First, we report a positive and highly significant association between option trading intensity and underlying volatility. This significant relationship holds for contemporaneous volatility as well as next-day's volatility forecasts. The magnitude of the slope coefficient for the option trading intensity measure decreases as the forecasting horizon increases. Second, in line with the predictions made in the excess volatility puzzle literature (see Chordia et al., 2011) (i) relative increases in the presence of informed trading lead to higher volatility and (ii) the effect of private information trading in the volatility of the underlying assets is positive and significant for more than one trading day and this is robust to stricter information measures. Third, our directional information measures successfully predict contemporaneous and next-day underlying stock returns but there is no significant association between our option intensity measures and returns for longer forecasting horizons. This finding implies that our measures capture significant private information in the options market that does not subsequently reverse. Fourth, in line with previous literature, we show that call and put options have differential information roles in predicting underlying stock returns. Fifth, we demonstrate a series of robustness tests to confirm that our results are robust to new announcements effects, moneyness effect, time-of-day effect, different level of market uncertainty, and market factors.

[4] For the futures market, Kalaitzoglou and Ibrahim (2015) and Ryu and Yang (2018) jointly model trade duration and trade volume employing an ACDW and a trade indicator model, respectively.

Our findings have important implications. First, we contribute to the literature that investigates the role of options markets in price discovery (see Patel et al., 2020). We show that option trading reveals significant information in the underlying market. In fact we find that 'relative' options liquidity fluctuations do affect the underlying asset's volatility and returns on the short term, which, considering the role of option in price discovery, we interpret as information inflow from the options market to the underlying asset.[5] Second, we contribute to the explanation of the excess volatility puzzle by investigating the effect of informed trades in the options market to underlying volatility. Finally, we contribute to the duration modelling literature by investigating the validity of modelling the arrival rate of events and its link to information on a "redundant" (in the sense that the main price discovery occurs in another asset) asset.

The remainder of the paper is organised as follows: In Section 2 we introduce the STM-ACWD model. In Section 3 we describe the data and provide a set of descriptive statistics. In Section 4 we present the results of the classification method and in Section 5 we examine the relation between our option trading intensity measures and next-period stock returns. In Section 6, we check the robustness of our results. In Section 7, we conclude the paper.

2. Trading intensity modelling

In this section, we discuss the joint modelling of trade duration and volume as a natural proxy of liquidity. We start by introducing the modelling of trading intensity as a rescaled marked point process that reduces its dimensionality. We then proceed with an empirical specification that enables the empirical investigation of the different levels of relatively low/high trading intensity for the asset of interest. Finally, we link the level of trading intensity to the presence of information, as it is revealed into the differential degrees of the arrival rate of the options trading volume.

[5] The underlying assumption here is that an investor with better access to price relevant information would trade the underlying asset. She might however select to trade in the options market either because the information is ambiguous or for leverage reasons. In both cases, her actions would lead to a 'relative' increase in options liquidity, which we capture with our modelling and associate it (using the shape of the hazard function) to presence of information. We find that these 'relative' liquidity fluctuations do have a consistent short term price impact on the price discovery underlying asset, which we interpret as information inflow. In brief, what we investigate is whether options liquidity fluctuations are priced in a way that is consistent to the presence of information and we report empirical findings that support that higher 'relative' options liquidity is highly likely to be associated with information and that there is a short term transition mechanism to the pricing of the underlying asset.

3. General framework

An important premise of market microstructure models is that the adjustment of prices to new information is a continuously revealing a price equilibrium that is broadly characterised as the price discovery process. The role of liquidity in that process is largely assumed to be that it improves market efficiency by enabling a faster price discovery process. In that respect, liquidity is a quintessential element of price discovery (e.g., O'Hara, 2003; Kalaitzoglou, 2021), both in aggregated terms (i.e., overall liquidity) facilitating price discovery, as well as in relative terms (i.e., relative to a benchmark) acting as an information signal (e.g., Kalaitzoglou and Ibrahim, 2013, 2015). Take for example the actions of informed traders: their information is primarily reflected in higher relative liquidity. In parallel, uninformed traders are generally understood to trade for reasons unrelated to information and therefore, their arrival rate should be expected to be time invariant (e.g., Hujer and Vuletić, 2007). Consequently, the actions of any agent, independently of her motivations for trading, are first revealed by the timing and volume of her trading. We focus on when transactions occur and how large they are, and we reverse engineer the originator of the trade.

For this purpose, we employ a natural proxy for liquidity, namely trading intensity, first introduced by Kalaitzoglou and Ibrahim (2015). According to Engle and Russell (1998), when focusing on intraday events, like the arrival of trades, the time in between them is irregularly spaced and therefore, conventional econometric estimators tend to be biased. Instead, a point process framework that focuses on modelling the arrival rate of this events is deemed more appropriate. Consider a time-series $\{t_0, t_1, \ldots, t_i, \ldots\}$ with $t_0 < t_1 < \cdots < t_i < \cdots, t \in T$ of arrival times t of events i, i.e., transactions. This is a point process $\{t_i\}$ that can be fully described by its counting function, $N(t_i)$, or its conditional, on $\mathcal{F}_i^t := \{t_0, t_1, \ldots, t_i\}$, intensity function, $\lambda(t_i|\mathcal{F}_{i-1}^t) = \lim_{\Delta t \to 0} \frac{P(N(t_i + \Delta t) > N(t_i)|\mathcal{F}_{i-1}^t)}{\Delta t}$.

The specification above focuses solely on arrival times, but in several cases the information set, $\mathcal{F}_i^t$, should be expanded in order to account for additional information, called marks, that is associated with the events that arrive at times t_i. Here, we are interested not simply at the arrival rate of transactions, but also on the volume of these transactions. For this purpose, we shift the focus from solely event time to the arrival of trading volume. This is done by expanding the information set, $\mathcal{F}_i^t \subset \mathcal{F}_{i-1} :=$

$\{t_0, t_1, \ldots, t_i, v_0, v_1, \ldots, v_i\}$, to also include the history of trading volume per transaction, $v \sim N(\bar{v}, \sigma_v^2)$ $v \in V$. Following Kalaitzoglou and Ibrahim (2015), instead of modelling the marginal distributions of time and volume separately, we apply a time rescaling that focuses on the joint distribution.

In more detail, let $d_i = t_i - t_{i-1}$ denote the (raw) duration of transaction i, and x_i its diurnally adjusted transformation (see Engle, 2000). This is a direct measure of trading frequency with the same counting and intensity functions, $N(t_i)$ and $\lambda(t_i|\mathcal{F}_{i-1}^t)$. In addition, consider the function $K(v_i) = exp(-(v_i - \bar{v})/2\sigma_v)$ as a scaling factor for durations. $K(v_i)$ is a normal density kernel that transforms raw volume into a decreasing function, i.e., $K(v_i)' < 0$. Then, duration and volume are combined into a direct measure of trading intensity, z_i

$$z_i = x_i * K(v_i) \tag{1}$$

z_i is a rescaled time variable that transforms the marked point process $\{(t_i, v_i)\}$ into a rescaled point process $(\mathrm{Y}\{t_i,\ v_i\}|F_{i-1}) = (z_i|F_{i-1}) \sim f(z_i|\check{z}_{i-1})$, where $\check{z}_{i-1}$ is the history of z up to time $i-1$ and $f(:)$ is the distribution of z. This is a dimension reduction approach, i.e., $T \times V \to T$, where $(t_i, v_i) \to t_i^* := t_i K(v_i)$ where $K(v_i)$ is a the scaling factor, expressed as a function of the associated mark. The sequence $\{t_0^*, t_1^*, \ldots, t_n^*, \ldots\}$ with $t_0^* < t_1^* < \cdots < t_n^* < \cdots, t^* \in T$ is a temporal point process and refers to the arrival time of a unit quantity rather than a transaction. This re-definition of the event type focuses on a more granular level and the conditional intensity $\lambda(t_i K(v_i)|F_{i-1}) = \lambda(t_i^*|F_{i-1}) = \lambda(t_i^*|t_0^*, t_1^*, \ldots, t_{i-1}^*)$ describes fully the arrival rate of volume and thus, it fully describes $z_i := (\mathrm{Y}\{t_i,\ v_i\}|F_{i-1}) = (z_i|F_{i-1})$. For the purposes of modelling, it is more convenient to work with durations and, therefore, z is defined as scaled durations, which can be heuristically interpreted as the waiting time for a single options contract, or simply its trading intensity.

Following Kalaitzoglou and Ibrahim (2015), z_i is modelled as:

$$z_i = \theta_i \varepsilon_i, where\ \theta_i = \theta(z_{i-1}, \ldots, z_1; \varphi_1) \tag{2}$$

where z_i is the volume-weighted duration from Eq. (1), $\theta_i = E(z_i|\check{z}_{i-1})$ is the expected value of trading intensity assumed to follow an ARMA specification; $\varepsilon_i = z_i/\theta_i$ is the error term, defined as the standardized weighted duration.

The specification in Eq. (2) is conditional on past information and therefore it does account for how past information is incorporated into the accumulation rate of volume and, in a sense, it works as the impulse response function of the dimensions of liquidity captured by trading intensity. The notable advantage of such a formulation is that an event, such as a liquidity and/or a volatility shock, that might have an impact on subsequent liquidity beyond the expected factors, i.e., θ_i, it does so through the innovations ε_i. ε_i affects the magnitude of z_i, which in turn affects the magnitude of z_{i+1}. This way, the liquidity impact of any shock/innovation is accounted for through the impulse response function implied by Eq. (2) and, therefore, this modelling provides an indirect way to account for various types of risks, such as liquidity and volatility risk, as they are translated into innovations in trading intensity.[6]

Furthermore, the main objective of this paper is the identification of the potential presence of information, as it is gradually revealed through deviations from a 'normal' level of trading intensity. Previous literature (e.g. Kyle, 1985), recognizes the presence of various types of traders, the actions of which are reflected on the arrival rate of their trades. As a general classification, previous literature suggest that the majority of market participants trade for exogenous reasons, unrelated to information. These agents are considered to be the 'normal' state of the market and their arrival rate should be expected to be unrelated to the arrival of information and thus, time invariant (e.g., Kalaitzoglou and Ibrahim, 2013). Mathematically this can be represented with a flat conditional intensity, i.e., $\lambda_{uninformed}(t_i^*|F_{i-1}) \longrightarrow \left(\lambda_{uninformed}(t_i^*|F_{i-1})\right)' = 0$, such as the one associated with an exponential distribution. In sharp contrast to the

[6] Eq. (2) is a direct modelling of the arrival rate of volume and, therefore, it can account for how different types of risk are translated into variations of this arrival rate. For example, a liquidity shock can be defined as the inability to trade at the desirable time. This would be translated into a low arrival rate of volume and thus, a limited demand that might render some orders non-executable due to limited depth. The impulse response function of Eq. (2) provides a modelling of the time required to revert back to "normal" liquidity levels and, thus, it does indirectly account for liquidity risk. Furthermore, considering that liquidity is the means by which information is resolved, any information shock, such as a volatility shock, should result in excessive liquidity and, thus, it would be captured by ε_i. Consequently, the impulse response function of Eq. (2) can also indirectly account for the liquidity response of other types of risk.

uninformed traders, there is another group that trades only upon the arrival of private information. These agents, called the informed traders, are the first to poses this information and have the incentive to exploit it before it becomes public information. Consequently, especially when they compete with each other, their probability of trading should be highest when new information arrives and it should be a decreasing function of time, until it reaches zero when the information becomes public. Mathematically, this can be expressed with a downward slope conditional intensity, i.e., $\lambda_{informed}(t_i^*|F_{i-1}) \longrightarrow \left(\lambda_{informed}(t_i^*|F_{i-1})\right)' < 0$.

The presence of these agents, or their relative proportions in the market, is a latent piece of information and it cannot be known ex-post or ex-ante. Our intention is to infer the presence of informed agents, or at least the probability of a transaction to be initiated by an informed agent. For this, we assume that the (observed) arrival rate of all transactions, or else the trading activity of the market as a whole, is a combination the arrival rates of the three groups of traders described above, $\int_s^t \lambda(t_i^*|F_{i-1})dt = \sum_k \int_s^t \lambda_k(t_i^*|F_{i-1})dt$, where $k = (uninformed, informed)$. The integral notation ensures that the superposed process is simple. Following Kalaitzoglou and Ibrahim (2013), since the presence of the agents is not directly detectable, we assume that each trade could have been initiated by any one of the two agent types with a time varying associated probability. We model this using a mixture of distributions that represents the distribution of the standardized durations as a weighted average of the distributions of the standardized durations of each group of agents.

$$\varepsilon_i|J_i \sim i.i.d., \text{with density } f(\varepsilon_i|J_i) \text{ and } E(\varepsilon_i|J_i) = \mathrm{E}(\varepsilon_i) = 1 \qquad (3)$$

J_i is an economically relevant threshold variable that is used in as an observable classification proxy. The general framework in Eq. (3) implies that the conditional density function of all durations changes according to a variable J_i and consequently, the conditional intensity of the superposed process can take different shapes. Our intension is to match the shape of the conditional intensity with the shape of the intensities of each one of the trading types described above. For this purpose, we need a distribution that is flexible enough to accommodate a flat and an increasing intensity function. We select the Weibull distribution for its simplicity and the versatility of the intensity functions that can be generate, and we assign all the variation to its shape parameter, $h(J_i\colon\tau)$:

$$f(z_i|J_i;\tau) =$$

$$(h(J_i\colon\tau)z_i)\left[\frac{z_i\Gamma(1+\frac{1}{h(J_i\colon\tau)})}{\theta_i}\right]^{h(J_i\colon\tau)} \exp\left(-\left[\frac{z_i\Gamma(1+\frac{1}{h(J_i\colon\tau)})}{\theta_i}\right]^{h(J_i\colon\tau)}\right) \quad (4)$$

where,

$$h(J_i\colon\tau) = \gamma_1 * \left(1 - G(J_i\colon g, j)\right) + \gamma_2 * G(J_i\colon g, j) \quad (5)$$

$$G(z_i\colon g, j) = (1 + \exp\{-g * (J_i - j)\})^{-1} \quad (6)$$

The shape parameter, $h(J_i\colon\tau)$, of the mixture of Weibull distribution is a function, h, of the threshold variable, J_i, defined as $\log(\frac{1}{z_i})$, and a vector of parameter coefficients $\tau = (\gamma_1, \gamma_2, g_1, j)$. γ_1 and γ_2 are the shape parameters of the marginal Weibull distributions that define the shape of the intensity function in each one of the two regimes of J_i, as they are classified by the threshold j and the smoothness parameter g. When $J_i < j$ ($J_i > j$) then, $G \to 0$ ($G \to 1$) and $h(J_i\colon\tau) \to \gamma_1$ ($h(J_i\colon\tau) \to \gamma_2$). Heuristically, this classification can be interpreted the following way. When J_i is lower (higher) than the threshold j, the shape parameter of the distribution of the trading intensity of the market is closer to the one described by γ_1 (γ_2) and therefore, there is a higher probability that this transaction is initiated by an agent that belongs to group 1 (2). If the shape of the intensity described by γ_1 (γ_2) matches the theoretical expectations, e.g., flat or decreasing, then we can allocate the trade accordingly. Our expectation is that higher values of J_i that describe a higher trading intensity should be associated with a decreasing intensity function, i.e., $\gamma_2 < 1$, because information is usually linked with more intense trading (e.g., Easley and O'Hara, 1992). In the opposite case, a more quiet market is expected to be associated with less informed trading and thus with a flat intensity function, i.e., $\gamma_1 \approx 1$.[7] Finally, the certainty level of the classification, e.g., the probability of a trade belonging to one or the other group, is determined by the distance from the threshold j and consequently, j is the level under or over which a normal level of trading intensity is defined. We expect higher levels of J_i to be associated with higher probability of informed trading, captured by higher values of the function G. The estimation of all parameters is conducted with maximum likelihood (refer to Kalaitzoglou and Ibrahim (2013) for a detailed derivation of the log-Likelihood function)

4. Presence of information

The STM-ACWD model, presented above, provides an empirical estimate of the probability of a transaction to be initiated by an informed versus an uninformed agent and therefore, it provides a conditional classification. In statistical terms, the shape parameter of the Weibull distribution, $h(J_i: \tau)$, is a function of an observable variable, which determines its magnitude. Following previous literature, values below 1, would be associated with a downward slope conditional intensity function and thus, with informed trading. The lower the value of $h(J_i: \tau)$, the steeper is the slope of the conditional intensity and this is interpreted as a higher probability of the transaction to be informed. We use this statistical property in order to create an information variable, I. In more detail, we observe the threshold variable J_i and its relative magnitude with respect to j and then we create a dummy variable that counts the trades that ca be characterized as informed, i.e., $h(J_i: \tau) < 1$, following different confidence intervals.

---

[7] The shape of the Weibull distribution is fully determined by its shape parameter. When the shape parameter takes the value of 1, then the Weibull distribution collapses to the Exponential distribution, which is characterised by a time invariant, i.e., flat, intensity function. When, the shape parameter is lower than 1, then the intensity function of the Weibull distribution exhibits a decreasing function. Following our theoretical assumptions, the first case captures uninformed trading, while the second captures the presence of informed agents.

5. Data and descriptive statistics

In this section, we first discuss the intraday options data. We then give an overview of the data cleaning criteria and present a set of descriptive statistics.

We obtain an extensive intraday trade and quote dataset for the 30 individual equity options written on the components of the Dow Jones Industrial Average Index (DJI) from January 2012 to June 2014. Given that, the implementation of the conditional duration model on the options markets is not time dependent and that the market structure of the U.S. options markets has remained unchanged, have no reason to believe that a more recent sample would have altered the results.[8] The contracts traded for each option class vary according to strike price, maturity date and contract type (i.e. call or put). The dataset includes information about the option price, strike price, maturity date and volume for every option contract, time-stamped to the nearest millisecond, separately for best asks, best bids and trades.

We estimate moneyness as S/K, where K refers to the strike price of the options contracts and S refers to the underlying concurrent mid-quote price. We exclude Deep-in-the-money (DITM) and deep-out-of-the-money (DOTM) options. We control for expiration effects by focusing on short-term options only; sub-tickers that are between 7 and 37 days from expiration are selected. We delete outliers based on spread and price criteria as follows: all zero volume, zero price and out-of-hours observations are deleted. We drop quotes with negative or zero bid-ask spreads. Also, we control for possible outlying data by dropping quotes with excessively large bid-ask spreads. The cut-off point for the percentage bid-ask spreads is set at 150%. Finally, we diurnally adjust durations to address intraday seasonality as suggested by Kalaitzoglou and Ibrahim (2015).

In Table 1, we report the descriptive statistics for the main variables of interest: option trade duration and option trading volume. Duration refers to the average time difference between two trades (excluding overnight durations), measured in seconds. Volume refers to the average option volume per trade. For calls (puts), the average number of transactions over the sample period is just over 285 (200) thousand. Consistent with the

[8] In Tables A1 and A2 in the Appendix we test the hypothesis that our results are not time-dependent. We employ a bootstrapping resampling scheme, set the number of bootstrap replications to 1000 and re-estimate our baseline model for volatility and returns. Both tables confirm our baseline results (Tables 5 and 6).

differences in transaction frequency, the average duration between trades for calls and puts is 767" (976"), respectively. The average traded volume for calls is 15.72 contracts and 14.31 contracts for puts. As expected the correlation between duration and volume is negative (-0.10 for calls, -0.16 for puts, results not reported to conserve space).

Following Kalaitzoglou and Ibrahim (2015), we model transaction time with the option trade volume v_i.

*** Table 1 ***

In Table 1, we report the average time duration between trades and the average trading volume per contract of 30 components of DJI. Call options have a shorter trade duration than the put options, while not all call options represent a high volume than the puts counterparty. Additionally, trading duration and trading volume vary significantly across the option series.

6. Trading Intensity and Information

In this section we present the first set of empirical results. We start with a discussion of the estimation results of the STM-ACWD model. We then discuss the development of the main and robust information measures and present the results of the classification method. The results provide supportive evidence of the validity of our model in capturing the time variation of information arrival in options.

7. Initial observations

In Table 2, we present the estimation results of the STM-ACWD model on each one of the options' transaction series in our sample. The columns present the estimates of the parameters of the conditional mean specification model parameters (ω, α, and β) and of the mixture of distributions ($\gamma_1, \gamma_2, g_1, j_1$). All results are statistically significant at 1% level, while α and β are consistently very close to, but less than 1 (at more decimal places than reported). This indicates stationarity with high persistence, suggesting prolonged trading intensity shocks (Kalaitzoglou and Ibrahim, 2015). This is an early sign that our initial proposition might have some empirical grounding. Trading intensity shocks seem to have a long lasting impact on the subsequent trading intensity, as this is reflected on the parameter estimates that indicate high persistence, and therefore, a higher or low trading intensity event appears to be prolonged. This is relevant in our analysis, because we link fluctuations in trading intensity with the presence or not of better informed agents. According to previous literature, when informed agents enter the market, this might result in an increase in overall trading volume (we capture this by an increase in trading volume per unit of time). This information has to be price resolved and this process results also in higher volumes of trading. Consequently, this high persistence in trading intensity might, indeed, be linked to the presence of information. A high (lower) trading intensity shock might lead to more (less) intense trading due to (absence of) information resolution and thus, to trading episodes of higher or lower trading activity.

*** Table 2 ***

We identify these episodes and link them to potential presence of information, or not, by distinguishing between a “normal” and a relatively “high” level of trading intensity. This is done by considering a mixture of distributions distinguished by different levels of an observable threshold variable. The threshold value is a parameter that is estimated and it provides a tangible way to identify groups of trades that might belong to a trading episode that is related to information. At a later stage we will investigate whether the identified trades are indeed associated with contemporaneous or subsequent price changes.

More precisely, there is a clear distinction in the shape of the Weibull distribution identified by the two regimes. γ_1 is consistently close to 1, while γ_2 is consistently significantly lower. This, according to our prior analysis, indicates that the trades identified to be in second regime exhibit a downward slope conditional intensity function (as opposed to the flat shape of regime one) and thus the are classified as being informed. The classification is done according to the threshold variable J_i and the threshold value j. j's range is between 0.413 and 0.811 and when $J_i > j$ and $J_i \to \infty$, then $h(J_i\colon \tau) \to \gamma_2$.

8. Main information measures

Consequently, we consider trades where $J_i > j$ as being informed, with an increasing probability as $J_i \to \infty$, because $G \to 1$. Accordingly, we construct the dummy variable $I^{m=1}, m = 1,2,3$ is different specifications of the dummy variable I discussed in the robustness section below, based on j as follows:

$$I_{k,i}^{m=1} \text{ equals one if } J_{k,i} > j_k \text{, otherwise zero;} \quad (7)$$

where $I_{k,i}^{m=1}$ is the information dummy variable for the trade of options written on stock k on time i. j_k refers to the threshold value and $J_{k,i}$ is the trading intensity of trade i of options on stock k. $I_{k,i}^{m=1}$ therefore defines as informed trades, strictly all trades with $J_i > j$.

$I_{k,i}^{m=1}$ identifies the information content of options' trades, as it might be captured by its trading speed (frequency) and its trading volume. As the conditional extraction of latent information based on observable information might be rather noisy at a trade-by-trade level, we aggregate our classification at a lower frequency level. The daily aggregate figures reflect the relative arrival rate of trading volume $RTI_{k,t}^{m}$, as follows:

$$RTI_{k,t}^{m} = \frac{\sum_{i=1}^{\#trades_t} I_{k,i,t}^{m}}{\#trades_t} \quad (8)$$

For the main results, $RTI_{k,t}^{m}$ with $m = 1$, is the proportion of the total number of trades in a day t, $\#trades_t$, identified as informed, $I_{k,i,t}^{m=1} = 1$. In the robustness section, we test the validity of our findings by employing a stricter definitions of $I_{k,i,t}^{m}$ based on the distance from j (denoted as $m = 2$) and according to the magnitude of $h(J_i\colon \tau)$ (denoted as $m = 3$). $RTI_{k,t}^{m}$ identifies the average level of information present in a trading day.

Since the estimates of the parameters that classify each trade as informed, are sample-wide estimates, this measure should be able to capture deviations from a "normal" trading intensity level that is identified in a sample wide manner; thus, it should be less noisy.

Furthermore, previous studies (e.g. Easley et al., 1998; Johnson and So, 2012; Ryu and Yang, 2018) highlight the predictive power of trade direction in the options market on subsequent returns in the underlying asset market, suggesting that aggregated estimates of trading volume are likely to overlook temporary informational content (e.g., Blasco et al., 2010). Consequently, a time, as well as a direction dimension should also be considered. $RTI_{k,t}^{m}$ is a measure that identifies the intensity of the presence of information taking into consideration the time dimension, but largely ignores the direction of the trades.

For this purpose we develop a second set of information measures, the directional $RTI_{k,t}^{m}$, or $DRTI_{k,t}^{m}$, that take into consideration the direction of each trade, $Dir_{k,i,t}$. $Dir_{k,i,t}$ takes the value of +1 (-1) if the trade is buyer (seller) initiated and it is created by applying the Lee and Ready (1991) methodology, followed by the tick rule for the trades that cannot be classified. $DRTI_{k,t}^{m}$ is constructed in a similar fashion to $RTI_{k,t}^{m}$, with the notable difference that the numerator is now the sum of signed information measures.

$$DRTI_{k,t}^{m} = \frac{\sum_{i=1}^{\#trades_t} Dir_{k,i,t} I_{k,i,t}^{m}}{\#trades_t} \tag{9}$$

All these information measures, attempt to capture the presence of information at granular level, by linking fluctuations in the accumulation rate of volume with the presence or absence of information. This is a "zoomed in" version of the well-established measure for the Probability of Informed Trading (PIN, Easley et al., 1996, 1997a, 1997b) (e.g., Kalaitzoglou, 2020), which has the notable advantage of being instantaneous and interval free.[9]

[9] Previous approaches in identifying the presence of information focus on the aggregated properties of the actions of different agent types and classify them according to how far their collective outcome is from what is considered to be "normal" trading activity levels. For example, two widely used measures of the presence of informed traders, namely the PIN (Easley et al., 1996, 1997a, 1997b) and the VPIN

9. Robustness of information measures

As a robustness check, we construct the dummy variable $I^{m=2}$ based on the distance from j (3 standard deviations). In addition, we classify the trades ($I^{m=3}$) according to the magnitude of $h(J_i: \tau)$. The two dummy variables are constructed as follows:

$I_{k,i}^{m=2}$ equals one if $J_{k,i} > j_k + 3 * SE_j$, otherwise zero; (10)

$I_{k,i}^{m=3}$ equals one if $h_{k,i}$ is lower than 30th percentile of the shape parameter $h_{i,t}$, otherwise zero; (11)

where SE_j is the standard error of the estimate of the parameter j.

Relative to $I^{m=1}$, the construction of these three information dummy variables is progressively stricter with $I^{m=1} \supseteq I^{m=2} \supseteq I^{m=3}$. $I^{m=2}$ recognizes that there might be some measurement error and therefore, it defines j considering a three standard deviation interval. $I^{m=1}$ and $I^{m=2}$, though, are constructed in a way that considers a step-wise transition from regime 1 to regime 2; for example if $j = 0.5$ a trade with $J_i = 0.499$ would be allocated as uninformed by $I^{m=1}$, while a trade with $J_i = 0.501$ would be allocated as informed. In order to avoid this, as further robustness, we consider a smooth transition function, G, which is taken into consideration in $I^{m=3}$. More precisely, $I^{m=3}$ classifies the trades as informed according to the distributional properties of $h(J_i: \tau)$, by focusing on the bottom 30th percentile of the shape parameter

(e.g., Easley et al., 2011a, 2011b), derive the probability of informed trading from the distributional properties of trade imbalances or aggregated volume imbalances, respectively. They are based on the idea that excessive trading volumes should be associated with an increased presence of information (e.g., Easley and O'Hara, 1992), the direction of which would yield excessive imbalances, the magnitude of which is considered to be proportional to the presence of informed agents. These interval estimates mitigate the impact of noise present in high-frequency data, but they have also been found to suffer from distributional sampling properties (Easley et al., 2011b, 2014a; Andersen and Bondarenko, 2014), where the grouping/binning mechanism is detrimental to specifying what is considered to be normal and what is considered to be far from it and thus, better informed. The proposition in this paper ventures the idea that focusing on a granular, rather than on an aggregated level, first, it overcomes the issue of an optimal bin selection and, second, it might be more appropriate for shorter investment horizons. This is based on O'Hara's (2015) critique, that in high-frequency trading the timing of possessing price-relevant information is crucial because it can turn a time priority into an information advantage, leading to a situation where "to be uninformed is to be slow" (Haldane, 2012). For further information please refer to Kalaitzoglou (2020).

of the joint distribution.[10] In Equation (7) therefore, $m = (2,3)$ for $I^{m=2}$ and $I^{m=3}$, respectively.

Finally, in order to account for differential levels of the presence of information we develop one more measure based on the distributional properties of the shape parameter, i.e., $DisRTI_{k,t}$. $DisRTI_{k,t}$ is defined as:

$$DisRTI_{k,t} = \sum_{i=1}^{\#trades_t} Dir_{k,i,t} * \frac{1}{h_{k,i,t}} \qquad (12)$$

$h_{k,i,t}$ is the shape parameter function of the joint distribution, $h(J_i : \tau)$, for each stock/day/trade. Lower values of $h_{k,i,t}$ indicate higher presence of information and this leads to higher values of $DisRTI_{k,t}$. The $Dir_{k,i,t}$ variable indicates the direction of the information present.

10. Trading intensity classification results

The information measures presented above capture a progressively stricter definition of relatively "high" trading intensity. In Table 3, we present the results for $RTI_{k,t}^{m}$ for

---

[10] The 30% threshold is arbitrarily selected to represent a smaller proportion of high intensity trades, as they are defined by the shape of the hazard function, i.e., captured by the shape parameter $h(J_i : \tau)$, and not by the level of the threshold variable J. The added benefit of doing so is that it defines as better informed trades the ones that exhibit a steeper slope on the hazard function and thus, according to theory, they might be more likely to be motivated by information. In a sense, the hazard function is the statistical measure we employ to link trading intensity (outcome of informed action) to information (motivation). We claim that $I^{m=1}$ and $I^{m=2}$ capture the presence of information by focusing on the outcome of the actions (trading intensity), while $I^{m=3}$ captures information by focusing on the motive. We appreciate that the 30% threshold is rather arbitrarily selected and thus we test the robustness of this threshold further. We find no significant difference with the threshold being defined at 25% and 20% levels.

progressively stricter levels of I: $I^{m=1} \supseteq I^{m=2} \supseteq I^{m=3}$, separately for each option and option type.

*** Table 3 ***

As expected, the proportion of trades identified as informed $RTI_{k,t}^{m=3}$ is consistently lower than the one identified by $RTI_{k,t}^{m=2}$, which is lower than the ones identified by $RTI_{k,t}^{m=1}$. More precisely, $RTI_{k,t}^{m=3}$ classifies on average 25.47% of call option trades and 25.14% of put option trades as informed. The within stock average daily proportion of informed trades varies from 17.84% (put: TRV) to 44.93% (call: JNJ). In sharp contrast, a significantly higher proportion of trades is identified as informed by $RTI_{k,t}^{m=1}$, i.e., 49.87% of call option trades and 50.26% of put option trades. In purely statistical terms, the STM-ACWD model splits the sample into two parts – lower vs higher trading intensity – which appear to be roughly balanced in terms of number of observations. According to the model, roughly 50% of trades are considered to be of lower/"normal" trading intensity, while the remaining 50% is classified as relatively "high" trading intensity, which we associate with potentially higher presence of information. The relative benefit of the STM-ACWD classification over a simple quantile classification is that it can identify this on a trade by trade basis and therefore, it can vary across stocks and over time. For example, $RTI_{k,t}^{m=1}$, varies from 41.78% in TRV (call) to 69.26% in JNJ (put). In addition, in Figure 1, the proportion of trades classified as informed is time variant.

*** Figure 1 ***

These two features combined, provide some supportive evidence of the versatility of our classification in capturing the time variation of information arrival, as this is revealed in trading intensity. The underlying assumption of our modelling is that an investor might select the options market for informed trades either under ambiguous information or for leverage reasons. According to theory, this is not systematic or time invariant and, therefore, liquidity fluctuations and, consequently, inferred information levels should be expected to be time variant. We link the levels of trading intensity to information and therefore, our measure is also time variant.

In Table 4, we present the average value of the directional information measures $DRTI_{k,t}^{m}$ and $DisRTI_{k,t}$.

*** Table 4 ***

Both measures capture both the magnitude and the direction of the inferred information. The sign of both $DRTI_{k,t}^{m}$ and $DisRTI_{k,t}$indicates whether there is positive (+ sign := good news) or negative (- sign := bad news) information, while the magnitude of the measure indicates how "strong" is the information signal. Under the presence of unambiguous clear information the market is expected to follow the direction of the information. Consequently, in call options, more buys (sells) are expected in good (bad) news days. The presence of information would lead to higher intensity (e.g., Easley and O'Hara, 1992), which the model would classify as increased presence of information in two ways: i) The threshold variable would be higher than the threshold value, which would make $I_{k,i} = 1$ and ii) The higher values of the threshold value would make $h(J_i\colon \tau) \rightarrow \gamma_2$. The magnitude of the information measure $DRTI_{k,t}^{m}$ captures the first, while $DisRTI_{k,t}$ captures the second.

According to the empirical results reported in Table 4, the majority of options exhibit a positive information imbalance. This positive imbalance is consistently larger in magnitude than the average negative imbalance, which implies that when there is information to be price resolved and the traders choose to trade in the options market, usually they align their trades with the direction of the information; call (put) options for good (bad) news. A more comprehensive view on the variability of the directional measures, is provided in Figures 2 and 3, which present the time variation of $DRTI_{k,t}^{m}$ and $DisRTI_{k,t}$. $DRTI_{k,t}^{m}$ varies from around -1.99% to +2.43%, while $DisRTI_{k,t}$ varies from around -8.30% to +11.96%. Again, the directional information imbalance seems to be skewed to the positive side, (higher magnitude of positive extremes), which is in line with the average values in Table 4, confirming that options' traders align their demand with the direction of the information. Finally, stricter classification criteria lead to greater in magnitude. Since fewer proportion of trades identified as informed with stricter classification criteria, the captured directional information may be more imbalanced and difficult to be cancelled out.

*** Figure 2 and Figure 3 ***

In this section, we presented the results of the classification method at varying levels of information classification. In the next section, we empirically test the validity of the information measures in forecasting underlying asset volatility and returns.

11. The informational role of option trading intensity

To form a better picture of the informational role of option trading intensity, in this section we examine the relation between our option trading intensity measures and next-period stock returns and volatility. In Section 5.1 we examine stock predictability using our trading intensity measures using a set of daily panel regressions. Finally, in Section 5.2, we show that our main results hold even when we control for option duration and transaction frequency and the O/S ratio.

Motivated to a large extent by the empirical model of Easley et al. (1998), Pan and Poteshman (2006), and Hu (2014), we investigate the predictive ability of option trading intensity on next-trading day stock returns by estimating the following equations:

$$V_{k,t+\tau} = \alpha \quad + \beta_1 RTI^{m}_{k,t} + F_i + \; M_t + \varepsilon_{k,t+\tau} \,, \tau = 0 \; to \; 10 \tag{13}$$

$$R_{k,t+\tau} = \alpha \quad + \beta_1 IM + F_i + \; M_t + \varepsilon_{k,t+\tau}, \tau = 0 \; to \; 10 \tag{14}$$

$V_{k,t}$ refers to stock volatility, estimated as the absolute value of stock returns. $R_{k,t}$ refers to stock close-to-close return, estimated as the logarithmic change in successive daily closing prices. $RTI^{m}_{k,t}$ refers to the percentage of trades with predictive information in a trading day; $IM = \left(DRTI^{m}_{k,t}, DisRTI_{k,t}\right)'$ refers to the information measures $DRTI^{m}_{k,t}$ and $DisRTI_{k,t}$ of options written on stock *k* on date *t*. We estimate separate regressions for calls and puts and include asset and monthly fixed effects (F_i and M_t). Importantly, in the first set of regressions, we investigate the predictability of $RTI^{m}_{k,t}$ on underlying stock volatility. Given that $DRTI^{m}_{k,t}$ and $DisRTI_{k,t}$ are directional measures, we use them to investigate the predictability of option trading intensity on the underlying stock returns.

12. Option trading intensity and underlying asset volatility

In Table 5, we report the regression results of Equation 13. In Panel A, we present the results for calls and in Panel B for puts, for contemporaneous volatility and up to 10 days in the future. Model 1 refers to $RTI^{m=1}_{k,t}$ and Models 2 and 3 refer to $RTI^{m}_{k,t}$ with m = 2 and m = 3, respectively.

*** Table 5 ***

For call options, we find that $RTI^{m=1}_{k,t}$ successfully predicts underlying asset volatility (see also Sarwar, 2005 and Ni et al., 2008). $RTI^{m=1}_{k,t}$ is positive and highly significant for contemporaneous volatility and remains highly statistically significant for up to 10 days. For contemporaneous volatility, $RTI^{m=1}_{k,t}$ has a slope coefficient of 0.01677 (t-statistic=22.86), indicating that the presence of more informed trading in the options market contributes to an increase in contemporaneous volatility in the underlying market. The magnitude of the slope coefficient for $RTI^{m=1}_{k,t}$ decreases as the forecasting horizon increases. As anticipated, when we adopt a progressively stricter definition of relatively "high" trading intensity, the slope coefficients for $RTI_{k,t}^{m=2}$ and $RTI_{k,t}^{m=3}$ become smaller relative to $RTI_{k,t}^{m=1}$ for the same forecasting horizon. In particular, for contemporaneous volatility, the slope coefficient for $RTI_{k,t}^{m=2}$ is 0.01678 (t-statistic=22.94) and for $RTI_{k,t}^{m=3}$ is 0.00737 (t-statistic=10.13). Both $RTI_{k,t}^{m=2}$ and $RTI_{k,t}^{m=3}$ are positive and highly significant at 1% level for next day's volatility, while only $RTI_{k,t}^{m=2}$ shows a longer forecasting horizon.

For put options, the results remain relatively similar albeit the proportion of volatility explained by $RTI_{k,t}^{m}$ is smaller. In particular, the slope coefficient for $RTI_{k,t}^{m=1}$ and contemporaneous volatility is 0.01237 (t-statistic=20.30) and remains positive and highly statistically significant for up to 10 days. Equally, for contemporaneous volatility, the slope coefficient for $RTI_{k,t}^{m=2}$ is 0.01233 (t-statistic=20.70) and for $RTI_{k,t}^{m=3}$ is 0.00332 (t-statistic=5.99). For next-days volatility, the slope coefficient for only $RTI_{k,t}^{m=2}$ is positive and highly significant at 0.00428 (t-statistic=6.19). Importantly, in consistency with call options, this positively significant relationship remains for up to 10 days.

Overall, the results presented in Table 5 contribute to the excess volatility puzzle found in the literature. In particular, prior literature has shown that variance ratios of stock returns tend to be more variable during trading hours than non-trading hours (see Boudoukh et al., 2019 for a recent study). One side of this literature has attributed this effect on trading on private information (see French and Roll, 1986 and Chordia et al., 2011). Importantly, French and Roll (1986) show that for the private information hypothesis to be a valid explanation of the excess volatility puzzle, the effect of informed trading on stock volatility needs to be persistent. In other words, private information will have to affect returns for more than one trading day. In Table 5, we shed further light to this finding. We show that, in line with this literature, relative increases in the presence of informed trading lead to higher volatility. Importantly, however, we show these results for the interaction between the option and the underlying market. Also, as predicted by the private information hypothesis, the effect of private information trading in the volatility of the underlying assets is positive and significant for more than one trading tray.

13. Option trading intensity and underlying asset returns

Further, we investigate the predictability of option trading intensity on underlying stock returns for up to 10 days in the future. In Table 6 we present set of regression results of Equation 14. Unlike the previous section, in this set of regressions we use the directional option trading intensity measures, $DRTI_{k,t}^{m}$ and $DisRTI_{k,t}$. For Models 1 to 3, the directional information measure refers to $DRTI_{k,t}^{m}$, where m = 1 to m = 3, respectively. For Model 4, the information measure refers to $DisRTI_{k,t}$.

*** Table 6 ***

Our main result is that both directional information measures successfully predict contemporaneous and next-days underlying stock returns. In particular, for call options, we find that the slope coefficient for $DRTI_{k,t}^{m=1}$ for contemporaneous returns is positive and highly significant at 0.00754 (t-statistic = 14.37). As we adopt a progressively stricter definition of relatively "high" trading intensity, the magnitude of the slope coefficient for contemporaneous returns decreases to 0.00756 (t-statistic = 14.28) and 0.00588 (t-statistic = 7.42) for $DRTI_{k,t}^{m=2}$ and $DRTI_{k,t}^{m=3}$, respectively. When we account for the distributional properties of the shape parameter, the slope coefficient for $DisRTI_{k,t}$ and contemporaneous returns is positive and highly significant at 0.00179 (t-statistic = 13.58). Further, $DRTI_{k,t}^{m}$ and $DisRTI_{k,t}$ successfully forecast next-days underlying stock returns. The slope coefficients for $DRTI_{k,t}^{m}$ and $DisRTI_{k,t}$ for *t = 1* range from 0.00284 (t-statistic = 5.15) for $DRTI_{k,t}^{m=1}$ to 0.00068 (t-statistic = 5.30) for $DisRTI_{k,t}$.

For put options, our directional information measures are negatively associated with contemporaneous and next-day stock returns. The slope coefficient for $DRTI_{k,t}^{m=1}$ for contemporaneous returns is negative and highly significant at -0.00470 (t-statistic = -9.66). The equivalent coefficient values for progressingly stricter levels of informed trading, $DRTI_{k,t}^{m=2}$ and $DRTI_{k,t}^{m=3}$, are -0.00470 (t-statistic = -9.72) and -0.00354 (t-statistic = -4.50), respectively. The corresponding value for $DisRTI_{k,t}$ and contemporaneous returns is -0.00116 (t-statistic = -10.27). For next-days stock returns, the results are negative and highly significant across directional information measures. For $DRTI_{k,t}^{m=1}$, the estimated slope coefficient value is -0.00238 (t-statistic = -5.15).

The magnitude of the slope coefficient decreases as m increases. For $DisRTI_{k,t}$, the slope coefficient for next-days returns is -0.00061 (t-statistic = -5.72).

Overall, the results presented in Table 6 have important implications for our hypotheses. Importantly, for both call and put options, our directional information measures predict the stock returns of the next trading day and the slope coefficients for longer forecasting horizons are not statistically significant. This finding implies that $DRTI_{k,t}^{m}$ and $DisRTI_{k,t}$ capture significant private information in the options market that does not subsequently reverse (see also Hu , 2014). Further, we confirm previous research on the directional information of option informed trading. In particular, Du and Fung (2018) use the call-to-stock volume and put-to-stock volume ratios to proxy for directional informed trading in the options market. They show that, for the banking industry, call and put options have differential information roles in predicting underlying stock returns. Along similar lines, Chan et al. (2002) show that positive (negative) call (put) returns have a positive impact on stock returns. Our directional information measures however generalise the findings of Du and Fung (2018), relying solely on option market information. Finally, Tsai et al. (2015) show that trading volume in the VIX option market conveys no significant predictive information for the underlying VIX index. In contrast, our results show that, for the components of the DJIA, our directional information proxies capture significant informed trading in the options market and successfully forecast stock returns of the next trading day.

14. Robustness tests

In the previous section, we show that option trading intensity contains important information about the underlying stock returns and volatility. In this section, we perform a number of additional robustness tests. First, we examine whether our information measures are higher before corporate news announcements. Also, we test whether our baseline results vary across moneyness levels. Second, we test for time-of-day effects in our trading intensity measures. Third, we examine the effect of uncertainty on our findings. Finally, we test whether the predictability of our information measures is robust to other market factors

15. Options trading intensity and information resolution

Informed trading is more active and profitable before corporate news announcements, such as mergers and acquisitions (Chan et al., 2015; Chordia et al., 2019) and split announcements (Gharghori et al., 2017). Given that our trading intensity measures capture the degree of information in the options market, in this subsection, we test whether our information measures are higher before such events. We focus on annual and quarterly corporate releases of financial reports that appear in the U.S. Securities and Exchange Commission website. Further, we create pre- and post-announcement sub-samples for our trading intensity measures, 10 days before and after the event days, respectively. In Table 7, we estimate the test statistics for our trading intensity measures around the announcement days.

*** Table 7 ***

As expected, for both call and put options, our intensity measures ($RTI_{k,t}^{m=i}$) are higher before corporate announcements than after the announcement day (Table 7, Panel A). These differences are significant at 1% level except $RTI_{k,t}^{m=3}$ for put options. In general, our $RTI_{k,t}^{m=i}$ measures capture private information before earnings announcements, in line with Jin et al. (2012) and Chordia et al. (2021b). Further, Roll et al. (2010) and Johnson and So (2012) suggest that option trading rather than trade direction reflect more information in the options market, therefore we expect that $DRTI_{k,t}^{m=i}$ will be insignificant as it is a directional intensity measure. Our results in Table 7, Panel B confirm this hypothesis.

Chakravarty et al. (2004) and Pan and Poteshman (2006) show that informed trading varies across moneyness levels as informativeness is related to leverage, trading volume and spreads. It is therefore reasonable to expect that our trading intensity measures will also vary across moneyness levels. We test this hypothesis in Table 8 below.

*** Table 8 ***

In particular, Table 8 reports the distribution of information measures separately for OTM, ATM and ITM options and separately for calls and puts.[11] Chakravarty et al. (2004) and Pan and Poteshman (2006) show that the most liquid options contain the highest level of information and in line with these results we show that, for both call and put options, informed trading is concentrated in ATM options, followed by OTM and ITM options. Informed OTM put options (approximately 16%) is generally higher than informed OTM call options (approximately 9%), suggesting that informed traders have a higher demand on OTM puts because of short-selling restrictions (see Pan and Poteshman, 2006). In Tables A3 and A4 in the appendix, we replicate our predictability tests across moneyness groups. The results remain qualitatively similar with the full sample results.

16. Time-of-day effects

Gao et al. (2018) and Elaut et al. (2018) show that the first and last trading sessions of the day contain more information than the rest of the trading day. Elaut et al. (2018) argue that this predictive relationship is driven by liquidity demand rather than informed trading. Inspired by these findings, we test for time-of-day effects in our trading intensity measures. In particular, we estimate the following regression models:

$$V_{k,t+\tau} = \alpha \quad + \beta_1 RTI\_F_{k,t}^m + \beta_2 RTI\_M_{k,t}^m + \beta_3 RTI\_L_{k,t}^m + F_i + \ M_t + \varepsilon_{k,t+\tau}\ , \tau = 0\ to\ 10 \qquad (15)$$

$$R_{k,t+\tau} = \alpha \quad + \beta_1 IM\_F + \beta_2 IM\_M + \beta_3 IM\_L + F_i + \ M_t + \varepsilon_{k,t+\tau}, \tau = 0\ to\ 10 \qquad (16)$$

where $RTI\_F_{k,t}^m, RTI\_M_{k,t}^m, and\ RTI\_L_{k,t}^m$ refer to the percentage of trades with predictive information across 30-minute intervals at the beginning, middle, and end of the trading day, respectively. IM_F, IM_F, and IM_L refer to information measures, $DRTI_{k,t}^m$ and $DisRTI_{k,t}$, across 30-minute intervals at the beginning, middle, and end of the trading day, respectively.

[11] We define moneyness as the difference between the stock price and the strike price that is calculated by S/K, where *K* refers to the strike price of the options contracts and *S* refers to the underlying concurrent mid-quote price. DOTM call contracts have moneyness smaller than 0.9, OTM call contracts have moneyness between 0.9 and 0.95, ATM call contracts have moneyness between 0.95 and 1.05, ITM call contracts have moneyness between 1.05 and 1.1, and DITM call contracts have moneyness over 1.1. Put contracts are based on the opposite classification.

**** Table 9 ****

In Table 9, we investigate the relationship between intraday trading intensity and underlying asset volatility. In general, intraday trading intensity measures predict contemporaneous and next-day volatility. For contemporaneous volatility, we find that the magnitude of the slope coefficient for RTI_F^m is generally higher than that for RTI_M^m and RTI_L^m. For instance, for call options, the slope coefficient for $RTI\_F^{m=1}$ is 0.00828 (t-statistic=23.12), while that for $RTI\_M^{m=1}$ and $RTI\_L^{m=1}$ are 0.0033 (t-statistic=11.69) and 0.00361 (t-statistic=11.3), respectively. For next-day volatility, the slope coefficients for RTI_L^m generally increase relative to RTI_F^m and RTI_M^m. For instance, for call options, the slope coefficient for $RTI\_L^{m=1}$ is 0.00306 (t-statistic=9.36) as compared with 0.00068 (t-statistic=1.84) for $RTI\_F^{m=1}$ and 0.00084 (t-statistic=2.9) for $RTI\_M^{m=1}$.

**** Table 10 ****

In Table 10, we investigate the relationship between intraday directional trading intensity and underlying asset returns. In general, we find that the intraday directional information measures can predict contemporaneous underlying asset returns, while end-of-day measures predict next day returns. In economic terms, the effect is positive (negative) for call (put) $DRTI\_F^m$ on contemporaneous returns and for call $DRTI\_L^m$ on next-day returns, while call (put) $DRTI\_L^m$ is negatively (positively) associated with contemporaneous returns.

Overall, the results presented in Table 9 and 10 are in line with our baseline results. Informed traders are more likely to trade during high-volume periods to hide their informational advantage and limit their price impact, in line with theories on strategic informed trading (see Easley et al., 1998). Our results also show that option trades at the end of the trading day are generally more informative than opening trades and thereby have greater impact on one-day-ahead underlying asset volatility and return.

17. Market uncertainty effects

What is the effect of market uncertainty on option trading intensity? Chordia et al. (2021a) show that that information has greater value when there is greater uncertainty. In this subsection we examine whether the predictability of our information measures varies across levels of market uncertainty. We split our sample to high and low market uncertainty days, corresponding to the top 20 percent and bottom 20 percent of the CBOE's volatility index (VIX), respectively. In Table 11 below, we report the regression results of underlying asset volatility on option trading intensity when the weekly VIX is at the bottom and top 20 percentile, respectively.

****Table 11****

For call options, $RTI_{k,t}^{m}$ remains positive and highly significant for contemporaneous and next-day underlying asset volatility. Also, the magnitude of the slope coefficient decreases when uncertainty is high. For instance, the slope coefficient of $RTI_{k,t}^{m=1}$ for contemporaneous volatility decreases from 0.01686 (t-statistic =11.58) to 0.01581 (t-statistic = 12.22) when weekly VIX increases from the bottom 20 percentile to the top 20 percentile. For put options, our information measures are positively associated with contemporaneous volatility and the magnitude of slope coefficients is generally higher when the uncertainty is high. The slope coefficient of $RTI_{k,t}^{m=1}$ for contemporaneous volatility increases from 0.01091 (t-statistic =7.8) to 0.01309 (t-statistic =10.9) when weekly VIX increases from the bottom 20 percentile to the top 20 percentile. Further, $RTI_{k,t}^{m}$ only remains highly statistically significant for next-days volatility when market uncertainty is low.

In Table 12, we report the regression results of underlying asset returns on our trading intensity measures separately for low and high levels of market uncertainty. For call options, the predictive power of our information measures is a decreasing function of market uncertainty. The magnitude of the slope coefficient of $DRTI_{k,t}^{m=1}$ for contemporaneous return decreases from 0.00952 (t-statistic =7.82) to 0.00464 (t-statistic =4.25) when weekly VIX increases from the bottom 20 percentile to the top 20 percentile. $DRTI_{k,t}^{m=1}$ is statistically insignificant for next day's return when the uncertainty is high. For put options, the magnitude of the slop coefficient of $DRTI_{k,t}^{m}$ and $DisRTI_{k,t}$ for contemporaneous return increases for increasing levels of market uncertainty.

**** Table 12 ****

Overall, our findings demonstrate that the predictive power of our information measures is generally stronger when market uncertainty is low. For index options, Chordia et al. (2021a) show that order imbalances only have predictive power at periods of high market uncertainty. Chordia et al. (2021a) argue that their results are not due to informed trading. Instead, they show that their results are due to risk protection strategies and we would expect that our results are stronger for predicting volatility rather than returns. As reported by Chordia et al. (2021a), we indeed show that, for put options, the coefficients are higher when VIX is in the top 20 percentile over the sample period. Also, during period of high uncertainty, we observe a relative increase in informed trading with respect to volatility (consistent with Chordia et al, 2021a). Our results are also consistent with prior literature that the effect of uncertainty primarily affects the volatility of options not returns (see Nofsinger and Prucyk, 2003).

18. Options trading intensity and market factors

Is the predictability of our information measures robust to other market factors? To address this question, we employ Fama-MacBeth regressions to replicate our baseline model on returns (see also Xing et al., 2010). We select five control variables to separate the predictive power of our directional information measures from other market factors[12]. We use VIX to control for daily volatility (see Tsai et al., 2015; Chordia et al., 2021a). We estimate SVOL and OVOL to capture daily stock and option trading volume, respectively. Previous literature (e.g., Easley et al., 1998; Dufour and Engle, 2000) generally shows a strongly predictive power of trading volume. Further, we estimate O/S as the options/stock trading volume ratio. Roll et al. (2010) show that the

[12] In further robustness checks, we estimate the baseline models for returns and volatility and up to 10 control variables. The results of the baseline model still hold. In order to conserve space, we report the regression results in Tables A5 and A6 in the appendix.

cross-sectional and time-series variation in the O/S ratio may be related to changes in informed trading and Johnson and So (2012) show that the O/S ratio reflects private information. Finally, we employ the volume-synchronized probability of informed trading, VPIN, to control for order flow toxicity (see Abad and Yagüe, 2012, Chakrabarty et al., 2015 and Easley et al., 2012). We present the results of the Fama-MacBeth regressions in Table 13 below.

**** Table 13 ****

For call options, $DRTI_{k,t}^{m=1}$ remains positive and highly significant after controlling for market factors. The slope coefficient of $DRTI_{k,t}^{m=1}$ decreases from 0.00686 (t-statistic =10.31) to 0.00578 (t-statistic =6.21) after including all control variables. For put options, we report similar findings except for contemporaneous return. In particular, $DRTI_{k,t}^{m=1}$ is statistically insignificant for contemporaneous return when including all control variables. In line with previous literature (Easley et al., 2012; Chordia et al., 2021a), VIX and VPIN are significantly correlated with future returns. In both cases the results of the Fama-MacBeth regressions are in line with our baseline model.[13] In addition, our information measures maintain a high significance level, which appears to be the highest alongside VPIN. This indicates that fluctuations in trading intensity in the options market is a strong indicator for the presence of private information that is subsequently translated into price changes in the underlying asset.

19. Conclusion

Prior research provides clear evidence that informed traders use the options market due to their high leverage. Also, the effect of the introduction of options markets on market completeness and efficiency depends on the type of liquidity orders. Understanding therefore the role of liquidity in options markets is key to understanding the information content of option trades. In this paper, we examine the information content of option

[13] In line with Johnson and So (2017) and Bae and Dixon (2018) in Tables A7 and A8 in the appendix we report the regression results of our baseline models of returns and volatility and up to 10 lags of underlying asset returns and volatility, respectively. The results of our baseline regressions still hold.

trades that is revealed by an option trading intensity proxy. In particular, we apply a conditional duration model (STM-ACWD) that allows us to capture the joint information content of both option trade duration and volume. Crucially, we employ the conditional duration model to identify the probability that each individual trade is initiated by different agents with progressively higher access to price unresolved information. We note that, to date, the literature is silent on the joint modelling of informed trading in options markets.

We develop a rigorous framework on the statistical properties of the STM-ACWD model and provide supportive empirical evidence of the validity of our model in capturing the time variation of information arrival in options. We report a positive and highly significant association between option trading intensity and underlying volatility and returns. We report that our information measure successfully forecast next-days volatility and stock returns. In the case of stock volatility, we show that our results are in line with the information asymmetry hypothesis in the excel volatility puzzle. For stock returns, we show that our trading intensity measure captures significant private information in the options market that does not subsequently reverse. In line with previous literature, we show that call and put options have differential information roles in predicting underlying stock returns. Finally, our findings remain robust when various situations (e.g., news announcements, moneyness effect, time-of-day effect, different levels of uncertainty, market factors) are taking into account.

Our findings contribute to our understanding of the excess volatility puzzle and the role of options contracts in the price discovery process. The joint modelling of point processes in the options market is a new research area that allows researchers to capture the multidimensionality of private information. The results are relevant for academics and practitioners who are interested in the classification of informed trades in the options market. In this paper we show that we are able to capture informed trading dynamics by the joint modelling of option trade duration and volume. However, future research should further concentrate on the relative classification of option informed trades across option moneyness as recent research has demonstrated that the distribution of informed trades across moneyness levels is not uniform (Bergsma et al., 2020).

Data availability statement

The data employed in this study is available from Thomson Reuters Tick History (via Securities Industry Research Centre of Asia-Pacific) and Algoseek. The part of data that support the finding of this study can be available from the corresponding author upon reasonable request. Restrictions apply to the availability of these data that were used under license for this research. The CBOE Volatility Index and the annoucnement data are freely downloaded from Yahoo Finance and U.S. Securities and Exchange Commission, respectively.

Data citation

[Options market data] Thomson Reuters Tick History, January 1, 2012 to June 30, 2014; Not publicly available.

[Underlying market data] Algoseek, January 1, 2012 to June 30, 2014; Not publicly available.

[Volatility Index data] CBOE Volatility Index, January 1, 2012 to June 30, 2014; https://finance.yahoo.com/

[Announcement data] SEC, January 1, 2012 to June 30, 2014; https://www.sec.gov/

Appendix

Table A1 Daily regressions of underlying asset volatility on option trading intensity with the bootstrap methods						
Panel A: Calls						
τ -Days ahead	Model 1		Model 2		Model 3	
0	0.01677***	(0.01543,0.01837)	0.01678***	(0.01543,0.01838)	0.00737***	(0.00605,0.00869)
1	0.00577***	(0.00483,0.00667)	0.00581082***	(0.00485,0.00673)	0.00332***	(0.00188,0.00466)
5	0.00153***	(0.00066,0.00240)	0.00146***	(0.00061,0.00230)	-0.00029	(-0.00147,0.00084)
10	0.00094**	(0.00008,0.00181)	0.00096**	(0.00009,0.00182)	0.00051	(-0.00079,0.00193)
Panel B: Puts						
τ -Days ahead	Model 1		Model 2		Model 3	
0	0.01237***	(0.01128,0.01357)	0.01233***	(0.01127,0.01351)	0.00332***	(0.00229,0.00426)
1	0.00424***	(0.00309,0.00574)	0.00428***	(0.00313,0.00578)	0.00005	(-0.00146,0.00123)
5	0.00118***	(0.00030,0.00208)	0.00125***	(0.00035,0.00215)	-0.00063	(-0.00168,0.00044)
10	0.00110***	(0.00010,0.00202)	0.00110***	(0.00009,0.00203)	-0.00043	(-0.00177,0.00096)

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the bootstrapping results of underlying asset volatility on options trading intensity. All bootstrapped estimates generated from 1,000 replications. Panel A and B represent the results of call and put options, respectively. They report coefficients and percentile confidence interval in parentheses. The time lag of underlying asset volatility is from day 0, 1, 5, and 10. Each model contains different independent variables ($\boldsymbol{RTI^{m=1}}$, $\boldsymbol{RTI^{m=2}}$, **and** $\boldsymbol{RTI^{m=3}}$) that refer to the relative arrival rate of trading volume with different specifications of the information dummy. All the regressions include asset and time fixed effects.

Table A2 Daily regressions of underlying asset returns on option trading intensity with the bootstrap methods

Panel A: Call options

τ -Days ahead	Model 1		Model 2		Model 3		Model 4	
0	0.00754***	(0.00651,0.00858)	0.00756***	(0.00656,0.00852)	0.00588***	(0.00444,0.00733)	0.00179***	(0.00156,0.00202)
1	0.00284***	(0.00176,0.00392)	0.00287***	(0.00182,0.00396)	0.00243***	(0.00077,0.00411)	0.00068***	(0.00041,0.00092)
5	-0.00008	(-0.00116,0.00102)	-0.00006	(-0.00116,0.00097)	-0.00028	(-0.00194,0.00125)	-0.0001	(-0.00037,0.00016)
10	0.00076	(-0.00055,0.00215)	0.00074	(-0.00056,0.00224)	0.00139	(-0.00062,0.00387)	0.00022	(-0.00012,0.00063)

Panel B: Put options

τ -Days ahead	Model 1		Model 2		Model 3		Model 4	
0	-0.00470***	(-0.00556, -0.00386)	-0.00470***	(-0.00566, -0.00383)	-0.00354***	(-0.00497, -0.00212)	-0.00116***	(-0.00137, -0.00095)
1	-0.00238***	(-0.00330, -0.00143)	-0.00244***	(-0.00341, -0.00155)	-0.00297***	(-0.00428, -0.00171)	-0.00061***	(-0.00081, -0.00040)
5	0.00048	(-0.00047,0.00144)	0.00053	(-0.00046,0.00162)	0.0009	(-0.00050,0.00222)	0.00011	(-0.00013,0.00034)
10	-0.00019	(-0.00113,0.00086)	-0.00017	(-0.00119,0.00088)	0.00013	(-0.00144,0.00180)	-0.00011	(-0.00034,0.00012)

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the bootstrapping results of underlying asset return on options trading intensity. All bootstrapped estimates generated from 1,000 replications. Panel A and B represent the results of call and put options, respectively. They report coefficients and percentile confidence interval in parentheses. The time lag of underlying asset return is from day 0, 1, 5, and 10. Each model contains different independent variables. $\boldsymbol{DRTI^{m=1}}$, $\boldsymbol{DRTI^{m=2}}$, **and** $\boldsymbol{DRTI^{m=3}}$ in Model 1, 2, and 3 refer to the directional relative arrival rate of trading volume with the different specifications of the information dummy. $\boldsymbol{DisRTI_{k,t}}$ in Model 4 refers to the inverse of the shape parameter, weighted by the direction of trade. All the regressions include asset and time fixed effects.

Table A3 Daily regressions of underlying asset volatility on option trading intensity across moneyness sub-samples

Panel A: ATM

	Call			Put		
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
0	0.00991***	0.01001***	-0.00153*	0.00714***	0.00720***	-0.00278***
	(14.69)	(14.84)	(-1.82)	(12.95)	(13.01)	(-3.90)
1	0.00138*	0.00143*	-0.00235***	0.00167**	0.00171**	-0.00333***
	(2.23)	(2.31)	(-1.99)	(2.86)	(2.96)	(-2.82)
5	-0.00170**	-0.00182**	-0.00396***	-0.00177***	-0.00169**	-0.00384***
	(-2.37)	(-2.89)	(-5.62)	(-3.20)	(-3.06)	(-4.56)
10	-0.00371***	-0.00375***	-0.00421***	-0.00131*	-0.00134*	-0.00293***
	(-4.43)	(-4.42)	(-5.58)	(-2.30)	(-2.32)	(-3.69)

Panel B: OTM

	Call			Put		
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
0	0.01761***	0.01757***	0.01618***	0.01205***	0.01200***	0.00755***
	(10.35)	(10.41)	(6.35)	(10.09)	(10.60)	(5.14)
1	0.00997***	0.01005***	0.01015***	0.00514***	0.00524***	0.00285**
	(11.08)	(11.12)	(7.02)	(3.56)	(3.59)	(2.48)
5	0.00590***	0.00587***	0.00526***	0.00286***	0.00290***	0.00145
	(7.63)	(7.63)	(5.18)	(3.00)	(3.03)	(1.49)
10	0.00693***	0.00700***	0.00721***	0.00174**	0.00181**	0.00105
	(5.93)	(5.96)	(3.55)	(2.30)	(2.38)	(1.03)

Panel C: ITM

	Call			Put		
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
0	0.01571***	0.01585***	0.01217***	0.01630***	0.01638***	0.01448***
	(15.66)	(15.65)	(9.62)	(14.10)	(14.08)	(10.57)
1	0.00467***	0.00469***	0.00621***	0.00598***	0.00599***	0.00686***
	(5.08)	(5.05)	(4.68)	(6.31)	(6.29)	(4.90)
5	0.00138	0.00141	0.00295*	0.00493***	0.00502***	0.00744***
	(1.36)	(1.40)	(2.16)	(4.39)	(4.40)	(3.01)
10	0.00214**	0.00217**	0.00420***	0.00518***	0.00517***	0.00606***
	(2.45)	(2.50)	(3.19)	(5.49)	(5.44)	(4.18)

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. The estimation regression results of underlying asset volatility on options trading intensity across moneyness level. Panel A, B, and C represent the results of ATM, OTM, and ITM options, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset volatility is from day 0, 1, 5, and 10. Each model contains different independent variables ($\boldsymbol{RTI^{m=1}}, \boldsymbol{RTI^{m=2}}$, **and** $\boldsymbol{RTI^{m=3}}$) that refer to the relative arrival rate of trading volume with different specifications of the information dummy. The adjusted R-square for these models is varying from 0.029 to 0.05. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table A4 Daily regressions of underlying asset return on option trading intensity across moneyness sub-samples								
Panel A: ATM								
	Call				**Put**			
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
0	0.00777***	0.00779***	0.00528***	0.00117***	-0.00360***	-0.00363***	-0.00176	-0.00056***
	(12.17)	(12.09)	(5.61)	(9.63)	(-5.99)	(-6.04)	(-1.96)	(-5.81)
1	0.00286***	0.00287***	0.00280**	0.00061***	-0.00193**	-0.00203***	-0.00280**	-0.00040***
	(3.95)	(3.93)	(2.69)	(4.51)	(0.43)	(0.37)	(0.08)	(0.78)
5	-0.00048	-0.00042	-0.00060	-0.00014	0.00010	0.00006	0.00106	-0.00001
	(-0.63)	(-0.57)	(-0.51)	(-1.00)	(-0.49)	(-0.37)	(-0.74)	(-0.98)
10	0.00085	0.00084	0.00188	0.00018	-0.00054	-0.00048	-0.00012	-0.00009
	(1.20)	(1.17)	(1.77)	(1.35)	(-0.31)	(-0.35)	(-0.84)	(-0.61)
Panel B: OTM								
	Call				**Put**			
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
0	0.00695***	0.00699***	0.00721***	0.00050***	-0.00602***	-0.00594***	-0.00488***	-0.00037***
	(4.64)	(4.72)	(3.91)	(8.59)	(-5.13)	(-5.26)	(-2.90)	(-6.34)
1	0.00338**	0.00351***	0.00356**	0.00014**	-0.00415***	-0.00417***	-0.00419***	-0.00021***
	(2.32)	(2.39)	(1.75)	(2.02)	(-0.46)	(-0.50)	(-1.91)	(-0.50)
5	0.00175	0.00175	0.00111	0.00004	0.00054	0.00077	0.00089	0.00004
	(1.42)	(1.43)	(0.70)	(0.68)	(0.95)	(1.05)	(0.07)	(0.58)
10	0.00242*	0.00239*	0.00326*	0.00001	0.00017	0.00013	0.00071	0.00001
	(1.19)	(1.17)	(1.04)	(0.19)	(0.57)	(0.54)	(0.23)	(1.15)
Panel C: ITM								
	Call				**Put**			
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
0	0.01070***	0.01070***	0.00804***	0.00073***	-0.00894***	-0.00894***	-0.00894***	-0.00059***
	(6.60)	(6.55)	(4.25)	(12.73)	(-5.55)	(-5.52)	(-4.18)	(-9.50)
1	0.00335**	0.00330**	0.00065	0.00022***	-0.00193	-0.0019	-0.00226	-0.00013*
	(2.23)	(2.18)	(0.34)	(3.33)	(-1.30)	(-1.30)	(-1.39)	(-1.69)
5	-0.00108	-0.00122	-0.00127	-0.0001	0.00254	0.0025	0.00113	0.00006
	(-0.84)	(-0.95)	(-0.64)	(-1.60)	(0.10)	(0.02)	(0.20)	(-0.62)
10	-0.0012	-0.00124	-0.00243	-0.00011*	0.00054	0.00047	-0.00005	-0.00004
	(-0.94)	(-0.97)	(-1.28)	(-1.91)	(0.61)	(0.79)	(-0.56)	(0.14)

Note: ***, **, and * refer to the significant level at 1%, 5%, and 10%. the estimation regression results of underlying asset return on options trading intensity across moneyness level. Panel A, B, and C represent the results of ATM, OTM, and ITM options, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset volatility is from day 0, 1, 5, and 10. Each model contains different independent variables. $\boldsymbol{DRTI^{m=1}}, \boldsymbol{DRTI^{m=2}}$, **and** $\boldsymbol{DRTI^{m=3}}$ in Model 1, 2, and 3 refer to the directional relative arrival rate of trading volume with the different specifications of the information dummy. $\boldsymbol{DisRTI_{k,t}}$ in Model 4 refers to the inverse of the shape parameter, weighted by the direction of trade. The adjusted R-square for these models is varying from 0.021 to 0.033. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table A5 Daily regressions of underlying asset volatility on option trading intensity while controlling ten variables												
Panel A: Call options												
τ -Days ahead	**0**			**1**			**5**			**10**		
Model	1	2	3	1	2	3	1	2	3	1	2	3
RTI^m	0.01437***	0.01435***	0.00251***	0.00307***	0.00308***	0.00095	-0.00041	-0.00055	-0.00170**	-0.00081	-0.00082	-0.00071
	(22.18)	(22.06)	(3.21)	(4.70)	(4.68)	(1.22)	(-0.61)	(-0.82)	(-2.13)	(-1.21)	(-1.21)	(-0.88)
VIX	0.00018***	0.00018***	0.00021***	0.00026***	0.00026***	0.00027***	0.00021***	0.00021***	0.00021***	0.00017***	0.00017***	0.00017***
	(7.01)	(6.95)	(8.10)	(10.21)	(10.20)	(10.46)	(7.85)	(7.87)	(7.89)	(6.46)	(6.46)	(6.42)
Dur	-0.00001***	-0.00001***	-0.00001***	-0.00001**	-0.00001**	-0.00001***	-0.00001**	-0.00001***	-0.00001***	-0.00001*	-0.00001*	-0.00001*
	(-6.37)	(-6.36)	(-16.79)	(-2.20)	(-2.19)	(-4.38)	(-2.52)	(-2.61)	(-3.00)	(-1.91)	(-1.91)	(-1.71)
SVOL	-0.00001	-0.00001	0.00001	-0.00001***	-0.00001***	-0.00001***	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
	(-0.04)	(-0.05)	(0.23)	(-3.69)	(-3.69)	(-3.64)	(0.27)	(0.27)	(0.27)	(0.28)	(0.28)	(0.27)
OVOL	-0.00001	-0.00001	0.00001	0.00001	0.00001	0.00001	-0.00001	-0.00001	-0.00001	0.00001	0.00001	0.00001
	(-0.46)	(-0.47)	(0.35)	(0.96)	(0.96)	(1.13)	(-0.38)	(-0.37)	(-0.38)	(0.89)	(0.89)	(0.85)
O/S	-0.00184	-0.00177	-0.00226	-0.00126	-0.00124	-0.00138	0.00024	0.00023	0.00037	-0.00182	-0.00183	-0.00176
	(-0.71)	(-0.68)	(-0.86)	(-0.48)	(-0.47)	(-0.52)	(0.09)	(0.09)	(0.14)	(-0.68)	(-0.68)	(-0.65)
VPIN	0.00474***	0.00473***	0.00549***	0.01398***	0.01398***	0.01414***	-0.00173***	-0.00172***	-0.00174***	-0.00070	-0.00070	-0.00073
	(7.88)	(7.87)	(9.02)	(23.02)	(23.02)	(23.30)	(-2.78)	(-2.77)	(-2.80)	(-1.09)	(-1.09)	(-1.16)
Lag_Dur	0.00001***	0.00001***	0.00001***	0.00001	0.00001	0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001
	(4.35)	(4.32)	(4.29)	(0.89)	(0.88)	(0.87)	(-0.81)	(-0.81)	(-0.73)	(-1.64)	(-1.63)	(-1.60)
Lag_Return	-0.00518	-0.00529	-0.00665	-0.00120	-0.00122	-0.00147	0.00353	0.00351	0.00338	-0.01592***	-0.01591***	-0.01590***
	(-0.93)	(-0.95)	(-1.18)	(-0.21)	(-0.22)	(-0.26)	(0.62)	(0.61)	(0.59)	(-2.77)	(-2.76)	(-2.76)
Lag_SVOL	-0.00001***	-0.00001***	-0.00001***	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001
	(-4.88)	(-4.87)	(-4.13)	(-1.10)	(-1.10)	(-0.96)	(-0.61)	(-0.61)	(-0.63)	(-1.02)	(-1.02)	(-1.06)
Lag_OVOL	0.00001	0.00001	0.00001	-0.00001	-0.00001	-0.00001	0.00001	0.00001	0.00001	-0.00001	-0.00001	-0.00001
	(1.44)	(1.42)	(1.58)	(-0.96)	(-0.96)	(-0.92)	(0.80)	(0.80)	(0.79)	(-1.21)	(-1.21)	(-1.22)
Adj. R2	0.089	0.089	0.065	0.072	0.072	0.071	0.034	0.034	0.034	0.032	0.032	0.032

Panel B: Put options												
τ -Days ahead	0			1			5			10		
Model	1	2	3	1	2	3	1	2	3	1	2	3
RTI^m	0.00922***	0.00915***	-0.00116	0.00135**	0.00136**	-0.00217***	-0.00058	-0.00050	-0.00179**	-0.00044	-0.00046	-0.00119
	(14.70)	(14.52)	(-1.60)	(2.14)	(2.15)	(-3.01)	(-0.90)	(-0.77)	(-2.43)	(-0.68)	(-0.71)	(-1.61)
VIX	0.00019***	0.00019***	0.00021***	0.00026***	0.00026***	0.00026***	0.00021***	0.00021***	0.00020***	0.00017***	0.00017***	0.00017***
	(7.56)	(7.53)	(8.00)	(10.18)	(10.17)	(10.19)	(7.76)	(7.76)	(7.68)	(6.32)	(6.32)	(6.27)
Dur	-0.00001***	-0.00001***	-0.00001***	-0.00001***	-0.00001***	-0.00001***	-0.00001**	-0.00001*	-0.00001**	-0.00001*	-0.00001*	-0.00001**
	(-7.31)	(-7.31)	(-16.68)	(-3.66)	(-3.63)	(-6.15)	(-1.99)	(-1.93)	(-2.44)	(-1.89)	(-1.90)	(-2.19)
SVOL	0.00001	0.00001	0.00001	-0.00001***	-0.00001***	-0.00001***	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
	(0.46)	(0.45)	(0.87)	(-2.71)	(-2.71)	(-2.61)	(0.33)	(0.32)	(0.34)	(0.37)	(0.38)	(0.38)
OVOL	-0.00003**	-0.00003**	-0.00002	-0.00004***	-0.00004***	-0.00004**	-0.00001	-0.00001	-0.00001	0.00001	0.00001	0.00001
	(-2.17)	(-2.15)	(-1.42)	(-2.70)	(-2.70)	(-2.45)	(-0.49)	(-0.49)	(-0.40)	(0.08)	(0.08)	(0.14)
O/S	0.00093	0.00080	0.00169	0.00814*	0.00812*	0.00864*	0.00074	0.00074	0.00106	-0.00194	-0.00193	-0.00173
	(0.19)	(0.16)	(0.35)	(1.67)	(1.67)	(1.77)	(0.15)	(0.15)	(0.21)	(-0.39)	(-0.38)	(-0.34)
VPIN	0.00522***	0.00522***	0.00559***	0.01416***	0.01416***	0.01423***	-0.00171***	-0.00171***	-0.00171***	-0.00066	-0.00066	-0.00067
	(8.60)	(8.60)	(9.15)	(23.30)	(23.29)	(23.43)	(-2.74)	(-2.75)	(-2.76)	(-1.04)	(-1.04)	(-1.06)
Lag_Dur	0.00001	0.00001	0.00001	-0.00001	-0.00001	-0.00001	-0.00001**	-0.00001**	-0.00001**	-0.00001***	-0.00001***	-0.00001***
	(0.47)	(0.44)	(0.37)	(-0.90)	(-0.91)	(-0.78)	(-2.14)	(-2.14)	(-2.02)	(-2.67)	(-2.66)	(-2.58)
Lag_Return	-0.00844	-0.00858	-0.00558	-0.00133	-0.00135	-0.00025	0.00443	0.00442	0.00487	-0.01476**	-0.01475**	-0.01449**
	(-1.51)	(-1.53)	(-0.99)	(-0.24)	(-0.24)	(-0.05)	(0.77)	(0.77)	(0.85)	(-2.57)	(-2.57)	(-2.52)
Lag_SVOL	-0.00001***	-0.00001***	-0.00001***	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001
	(-4.24)	(-4.24)	(-3.82)	(-0.81)	(-0.81)	(-0.74)	(-0.53)	(-0.54)	(-0.55)	(-0.96)	(-0.96)	(-0.97)
Lag_OVOL	-0.00002**	-0.00002**	-0.00002**	-0.00001	-0.00001	-0.00001	0.00001	0.00001	0.00001	-0.00002**	-0.00002**	-0.00002**
	(-2.32)	(-2.31)	(-2.17)	(-1.01)	(-1.01)	(-0.94)	(0.15)	(0.15)	(0.19)	(-2.10)	(-2.10)	(-2.07)
Adj. R2	0.074	0.074	0.063	0.072	0.072	0.072	0.034	0.034	0.034	0.033	0.033	0.033

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the estimation regression results of underlying asset volatility on option trading intensity while controlling ten variables. Panel A and B represent the results of call and put options, respectively. They report coefficients, t-statistics in parentheses, and adjusted R-square (Adj. R2). Each column shows the results based on 0, 1, 5, and 10 days ahead of underlying asset volatility. Each model contains different information measures ($RTI^{m=1}$, $RTI^{m=2}$, and $RTI^{m=3}$) as the first independent variable. Ten control variables are also included in the regression models. *VIX* refers to the daily closing price of CBOE Volatility Index. *Dur* refers to option trading duration. *SVOL* and *OVOL* refer to underlying stock trading volume and option trading volume, respectively. *O/S* refers to the options/stock trading volume ratio. *VPIN* refers to volume-synchronized probability of informed trading metric. *Lag_Dur* refers to lagged option trading duration. *Lag_Return* refers to lagged underlying asset returns. *Lag_SVOL* and *Lag_OVOL* refers to lagged underlying stock trading volume and option trading volume, respectively. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table A6 Daily regressions of underlying asset return on option trading intensity while controlling ten variables

Panel A: Call options

τ -Days ahead	0				1				5				10			
Model	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
DRTI	0.00702***	0.00703***	0.00520***	0.00167***	0.00271***	0.00272***	0.00222**	0.00064***	-0.00001	0.00001	-0.00018	-0.00008	0.00097	0.00096	0.00168*	0.00027*
	(11.38)	(11.33)	(5.85)	(11.45)	(4.34)	(4.33)	(2.48)	(4.38)	(-0.02)	(0.01)	(-0.19)	(-0.56)	(1.54)	(1.51)	(1.86)	(1.84)
VIX	-0.00060***	-0.00060***	-0.00060***	-0.00060***	-0.00002	-0.00002	-0.00002	-0.00002	0.00008**	0.00008**	0.00008**	0.00008**	0.00014***	0.00014***	0.00014***	0.00014***
	(-17.69)	(-17.69)	(-17.75)	(-17.68)	(-0.50)	(-0.51)	(-0.54)	(-0.50)	(2.45)	(2.45)	(2.44)	(2.44)	(3.91)	(3.91)	(3.93)	(3.92)
Dur	-0.00001***	-0.00001***	-0.00001***	-0.00001***	0.00001	0.00001	0.00001	0.00001	0.00001*	0.00001*	0.00001*	0.00001*	0.00001	0.00001	0.00001	0.00001
	(-9.47)	(-9.47)	(-9.89)	(-9.24)	(1.50)	(1.50)	(1.35)	(1.59)	(1.72)	(1.72)	(1.71)	(1.67)	(1.47)	(1.47)	(1.49)	(1.53)
SVOL	0.00001***	0.00001***	0.00001***	0.00001***	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
	(2.92)	(2.92)	(2.89)	(2.89)	(0.28)	(0.28)	(0.27)	(0.27)	(0.74)	(0.74)	(0.74)	(0.74)	(0.83)	(0.83)	(0.82)	(0.82)
OVOL	0.00002*	0.00002*	0.00002**	0.00002**	-0.00002**	-0.00002**	-0.00002**	-0.00002**	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001
	(1.93)	(1.90)	(1.99)	(1.97)	(-2.07)	(-2.08)	(-2.04)	(-2.05)	(-0.97)	(-0.97)	(-0.97)	(-0.96)	(-0.26)	(-0.27)	(-0.27)	(-0.26)
O/S	-0.00857**	-0.00846**	-0.00850**	-0.00865**	0.00098	0.00102	0.00100	0.00094	0.00329	0.00329	0.00329	0.00330	0.00003	0.00004	0.00005	0.00001
	(-2.47)	(-2.44)	(-2.45)	(-2.50)	(0.28)	(0.29)	(0.29)	(0.27)	(0.94)	(0.94)	(0.94)	(0.94)	(0.01)	(0.01)	(0.01)	(0.00)
VPIN	0.00061	0.00061	0.00078	0.00061	0.00860***	0.00859***	0.00866***	0.00859***	0.00007	0.00007	0.00008	0.00009	-0.00362***	-0.00362***	-0.00361***	-0.00362***
	(0.76)	(0.76)	(0.97)	(0.76)	(10.64)	(10.64)	(10.72)	(10.64)	(0.09)	(0.09)	(0.10)	(0.11)	(-4.37)	(-4.37)	(-4.37)	(-4.38)
Lag_Dur	0.00001***	0.00001***	0.00001***	0.00001***	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
	(2.89)	(2.88)	(3.13)	(2.80)	(0.33)	(0.33)	(0.42)	(0.30)	(0.22)	(0.22)	(0.23)	(0.25)	(0.35)	(0.35)	(0.36)	(0.33)
Lag_Return	0.05407***	0.05405***	0.05305***	0.05465***	-0.03321***	-0.03321***	-0.03354***	-0.03298***	0.00295	0.00295	0.00291	0.00281	0.02203***	0.02203***	0.02214***	0.02219***
	(7.32)	(7.32)	(7.16)	(7.40)	(-4.44)	(-4.44)	(-4.49)	(-4.41)	(0.39)	(0.39)	(0.39)	(0.37)	(2.94)	(2.93)	(2.95)	(2.96)
Lag_SVOL	-0.00001	-0.00001	0.00001	-0.00001	0.00001	0.00001	0.00001	0.00001	0.00001***	0.00001***	0.00001***	0.00001***	0.00001	0.00001	0.00001	0.00001
	(-0.07)	(-0.07)	(0.02)	(-0.09)	(0.77)	(0.77)	(0.80)	(0.76)	(3.08)	(3.08)	(3.08)	(3.08)	(1.40)	(1.40)	(1.40)	(1.39)
Lag_OVOL	-0.00002***	-0.00002***	-0.00002***	-0.00002***	-0.00001	-0.00001	-0.00001	-0.00001	0.00001	0.00001	0.00001	0.00001	-0.00001	-0.00001	-0.00001	-0.00001
	(-4.38)	(-4.37)	(-4.40)	(-4.42)	(-0.61)	(-0.60)	(-0.62)	(-0.62)	(0.34)	(0.34)	(0.34)	(0.34)	(-1.01)	(-1.01)	(-1.01)	(-1.01)
Adj. R2	0.055	0.055	0.050	0.055	0.032	0.032	0.031	0.032	0.028	0.028	0.028	0.028	0.029	0.029	0.030	0.030

Panel B: Put options																
τ -Days ahead	0				1				5				10			
Model	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
DRIT	-0.00445***	-0.00444***	-0.00336***	-0.00110***	-0.00236***	-0.00242***	-0.00293***	-0.00060***	0.00047	0.00052	0.00086	0.00011	-0.00037	-0.00035	-0.00001	-0.00016
	(-7.97)	(-7.91)	(-4.08)	(-8.61)	(-4.21)	(-4.29)	(-3.53)	(-4.68)	(0.84)	(0.91)	(1.03)	(0.84)	(-0.66)	(-0.62)	(-0.02)	(-1.21)
VIX	-0.00055***	-0.00055***	-0.00056***	-0.00055***	-0.00001	-0.00001	-0.00002	-0.00001	0.00007**	0.00007**	0.00007**	0.00007**	0.00014***	0.00014***	0.00014***	0.00014***
	(-16.21)	(-16.21)	(-16.37)	(-16.17)	(-0.42)	(-0.41)	(-0.48)	(-0.39)	(2.12)	(2.12)	(2.12)	(2.12)	(3.99)	(3.99)	(3.97)	(4.01)
Dur	-0.00001	-0.00001	-0.00001	-0.00001	0.00001***	0.00001***	0.00001***	0.00001***	0.00001	0.00001	0.00001	0.00001	0.00001**	0.00001**	0.00001**	0.00001**
	(-0.54)	(-0.55)	(-0.54)	(-0.57)	(2.62)	(2.61)	(2.61)	(2.60)	(0.06)	(0.06)	(0.07)	(0.06)	(2.40)	(2.40)	(2.41)	(2.40)
SVOL	0.00001**	0.00001**	0.00001**	0.00001**	-0.00001	-0.00001	-0.00001	-0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
	(2.13)	(2.12)	(2.18)	(2.14)	(-0.07)	(-0.08)	(-0.04)	(-0.07)	(0.24)	(0.24)	(0.24)	(0.24)	(0.33)	(0.33)	(0.33)	(0.33)
OVOL	0.00009***	0.00009***	0.00009***	0.00010***	-0.00001	-0.00001	-0.00001	-0.00001	0.00001	0.00001	0.00001	0.00001	0.00002	0.00002	0.00002	0.00002
	(4.73)	(4.73)	(4.61)	(4.77)	(-0.57)	(-0.57)	(-0.58)	(-0.55)	(0.11)	(0.11)	(0.10)	(0.11)	(0.81)	(0.80)	(0.78)	(0.83)
O/S	-0.02763***	-0.02762***	-0.02738***	-0.02796***	-0.00329	-0.00330	-0.00336	-0.00349	-0.00475	-0.00474	-0.00468	-0.00472	-0.00852	-0.00852	-0.00846	-0.00862
	(-4.29)	(-4.29)	(-4.24)	(-4.34)	(-0.51)	(-0.51)	(-0.52)	(-0.54)	(-0.73)	(-0.73)	(-0.72)	(-0.72)	(-1.30)	(-1.30)	(-1.29)	(-1.32)
VPIN	0.00102	0.00102	0.00101	0.00102	0.00873***	0.00873***	0.00873***	0.00873***	0.00010	0.00010	0.00010	0.00010	-0.00357***	-0.00357***	-0.00358***	-0.00357***
	(1.27)	(1.27)	(1.26)	(1.27)	(10.79)	(10.80)	(10.80)	(10.80)	(0.12)	(0.12)	(0.12)	(0.12)	(-4.31)	(-4.31)	(-4.31)	(-4.31)
Lag_Dur	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	0.00001	0.00001	0.00001	0.00001	-0.00001	-0.00001	-0.00001	-0.00001
	(-1.05)	(-1.03)	(-1.06)	(-1.03)	(-0.44)	(-0.43)	(-0.45)	(-0.43)	(1.09)	(1.08)	(1.09)	(1.08)	(-0.18)	(-0.18)	(-0.18)	(-0.18)
Lag_Return	0.05359***	0.05360***	0.05326***	0.05396***	-0.03406***	-0.03405***	-0.03403***	-0.03385***	0.00224	0.00223	0.00218	0.00221	0.02170***	0.02169***	0.02161***	0.02181***
	(7.23)	(7.23)	(7.18)	(7.28)	(-4.56)	(-4.56)	(-4.56)	(-4.53)	(0.30)	(0.30)	(0.29)	(0.30)	(2.90)	(2.89)	(2.88)	(2.91)
Lag_SVOL	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001***	0.00001***	0.00001***	0.00001***	0.00001	0.00001	0.00001	0.00001
	(0.03)	(0.04)	(0.06)	(0.05)	(0.77)	(0.77)	(0.79)	(0.78)	(2.95)	(2.95)	(2.94)	(2.95)	(1.35)	(1.35)	(1.35)	(1.35)
Lag_OVOL	-0.00001	-0.00001	-0.00001	-0.00001	0.00001	0.00001	0.00001	0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001
	(-1.30)	(-1.30)	(-1.29)	(-1.30)	(1.04)	(1.04)	(1.03)	(1.04)	(-0.17)	(-0.17)	(-0.16)	(-0.17)	(-0.51)	(-0.51)	(-0.51)	(-0.51)
Adj. R2	0.046	0.046	0.044	0.047	0.031	0.031	0.031	0.031	0.028	0.028	0.028	0.028	0.030	0.030	0.030	0.030

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the estimation regression results of underlying asset return on option trading intensity. Panel A and B represent the results of call and put options, respectively. They report coefficients, t-statistics in parentheses, and adjusted R-square (Adj. R2). Each column shows the results based on 0, 1, 5, and 10 days ahead of underlying asset return. Each model contains different information measures ($\boldsymbol{DRTI^{m=1}}$, $\boldsymbol{DRTI^{m=2}}$, $\boldsymbol{DRTI^{m=3}}$, **and** $\boldsymbol{DisRTI_{k,t}}$) as the first independent variable, *DRTI*. Ten control variables are also included in the regression models. *VIX* refers to the daily closing price of CBOE Volatility Index. *SVOL* and *OVOL* refer to underlying stock trading volume and option trading volume, respectively. *O/S* refers to the options/stock trading volume ratio. *VPIN* refers to volume-synchronized probability of informed trading metric. *Lag_Dur* refers to lagged option trading duration. *Lag_Return* refers to lagged underlying asset returns. *Lag_SVOL* and *Lag_OVOL* refers to lagged underlying stock trading volume and option trading volume, respectively. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table A7 Daily regressions of underlying asset volatility on option trading intensity

Panel A: Calls						
t days ahead	Model 1		Model 2		Model 3	
0	0.01677***	(22.86)	0.01678***	(22.94)	0.00737***	(10.13)
1	0.00577***	(11.85)	0.00581***	(11.82)	0.00332***	(4.27)
2	0.00207**	(2.51)	0.00211**	(2.51)	0.00070	(1.16)
3	0.00242***	(4.21)	0.00247***	(4.26)	0.00035	(0.36)
4	0.00067	(0.96)	0.00065	(0.95)	0.00014	(0.15)
5	0.00153***	(3.25)	0.00146***	(3.19)	-0.00029	(-0.46)
6	0.00050	(0.95)	0.00057	(1.09)	-0.00062	(-0.71)
7	0.00089	(1.13)	0.00089	(1.16)	0.00052	(0.61)
8	0.00113**	(2.45)	0.00121***	(2.62)	0.00088	(1.32)
9	0.00225***	(3.38)	0.00171***	(3.82)	0.00024	(0.38)
10	0.00094**	(2.08)	0.00096**	(2.11)	0.00051	(0.72)
Panel B: Puts						
t days ahead	Model 1		Model 2		Model 3	
0	0.01237***	(20.30)	0.01233***	(20.70)	0.00332***	(5.99)
1	0.00424***	(6.10)	0.00428***	(6.19)	0.00005	(0.07)
2	0.00194***	(3.19)	0.00197***	(3.21)	-0.00114**	(-2.05)
3	0.00215***	(3.03)	0.00209***	(3.05)	-0.00152**	(-2.24)
4	0.00084*	(1.92)	0.00091**	(2.08)	-0.00068	(-0.95)
5	0.00118**	(2.34)	0.00125**	(2.46)	-0.00063	(-1.12)
6	0.00183**	(2.48)	0.00191**	(2.57)	0.00035	(0.37)
7	0.00079	(1.09)	0.00079	(1.07)	-0.00015	(-0.18)
8	0.00173**	(2.53)	0.00176**	(2.56)	-0.00008	(-0.11)
9	0.00093**	(2.20)	0.00092**	(2.18)	-0.00093	(-1.52)
10	0.00110**	(2.17)	0.00110**	(2.15)	-0.00043	(-0.59)

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the estimation regression results of underlying asset volatility on options trading intensity. Panel A and B represent the results of call and put options, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset volatility is from day 0 to 10. Each model contains different independent variables ($\boldsymbol{RTI^{m=1}}$, $\boldsymbol{RTI^{m=2}}$, **and** $\boldsymbol{RTI^{m=3}}$) that refer to the relative arrival rate of trading volume with different specifications of the information dummy. The adjusted R-square for these models is varying from 0.028 to 0.079. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table A8 Daily regressions of underlying asset returns on option trading intensity

Panel A: Call options								
t days ahead	Model 1		Model 2		Model 3		Model 4	
0	0.00754***	(14.37)	0.00756***	(14.28)	0.00588***	(7.42)	0.00179***	(13.58)
1	0.00284***	(5.15)	0.00287***	(5.15)	0.00243***	(2.83)	0.00068***	(5.30)
2	0.00012	(0.15)	0.00017	(0.22)	0.00033	(0.37)	0.00008	(0.49)
3	-0.00060	(-0.94)	-0.00059	(-0.91)	-0.00115	(-1.47)	-0.00011	(-0.82)
4	-0.00062	(-1.18)	-0.00071	(-1.34)	-0.00056	(-0.69)	-0.00018	(-1.38)
5	-0.00008	(-0.13)	-0.00006	(-0.11)	-0.00028	(-0.33)	-0.00010	(-0.74)
6	-0.00058	(-1.03)	-0.00061	(-1.08)	-0.00052	(-0.69)	-0.00013	(-1.04)
7	0.00004	(0.06)	0.00004	(0.06)	0.00042	(0.53)	-0.00002	(-0.13)
8	0.00147***	(2.66)	0.00145***	(2.61)	0.00208**	(2.56)	0.00038***	(2.83)
9	0.00173***	(3.08)	0.00216***	(3.54)	0.00359***	(3.89)	0.00053***	(4.05)
10	0.00076	(1.11)	0.00074	(1.08)	0.00139	(1.20)	0.00022	(1.13)
Panel B: Put options								
t days ahead	Model 1		Model 2		Model 3		Model 4	
0	-0.00470***	(-9.66)	-0.00470***	(-9.72)	-0.00354***	(-4.50)	-0.00116***	(-10.27)
1	-0.00238***	(-5.15)	-0.00244***	(-5.25)	-0.00297***	(-4.40)	-0.00061***	(-5.72)
2	0.00082	(1.46)	0.00082	(1.46)	0.00108	(1.22)	0.00018	(1.53)
3	0.00026	(0.53)	0.00019	(0.38)	0.00021	(0.31)	0.00007	(0.67)
4	-0.00008	(-0.16)	-0.00012	(-0.23)	-0.00052	(-0.71)	-0.00002	(-0.17)
5	0.00048	(0.93)	0.00053	(1.00)	0.00090	(1.33)	0.00011	(0.88)
6	0.00065	(0.72)	0.00059	(0.65)	0.00067	(0.59)	0.00012	(0.61)
7	-0.00068	(-1.26)	-0.00059	(-1.10)	-0.00083	(-1.09)	-0.00019	(-1.59)
8	0.00081*	(1.69)	0.00082*	(1.71)	0.00109	(1.41)	0.00011	(1.00)
9	0.00031	(0.61)	0.00045	(0.89)	0.00124	(1.63)	0.00009	(0.80)
10	-0.00019	(-0.35)	-0.00017	(-0.31)	0.00013	(0.15)	-0.00011	(-0.93)

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the estimation regression results of underlying asset return on options trading intensity. Panel A and B represent the results of call and put options, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset return is from day 0 to 10. Each model contains different independent variables. $\boldsymbol{DRTI^{m=1}}, \boldsymbol{DRTI^{m=2}},$ **and** $\boldsymbol{DRTI^{m=3}}$ in Model 1, 2, and 3 refer to the directional relative arrival rate of trading volume with the different specifications of the information dummy. $\boldsymbol{DisRTI_{k,t}}$ in Model 4 refers to the inverse of the shape parameter, weighted by the direction of trade. The adjusted R-square for these models is varying from 0.016 to 0.03. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

References

Abad, D. & Yagüe, J. (2012). From PIN to VPIN: An introduction to order flow toxicity. *The Spanish Review of Financial Economics,* 10(2), pp. 74-83.

Andersen, T. G. & Bondarenko, O. (2014). VPIN and the flash crash. *Journal of Financial Markets,* 17, pp. 1-46.

Bae, K. H. & Dixon, P. (2018). Do investors use options and futures to trade on different types of information? Evidence from an aggregate stock index. *Journal of Futures Markets,* 38(2), pp. 175-198.

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

Biais, B. & Hillion, P. (1994). Insider and liquidity trading in stock and options markets. *The Review of Financial Studies,* 7(4), pp. 743-780.

Black, F. (1975). Fact and Fantasy in the Use of Options. *Financial Analysts Journal*, pp. 36-72.

Blasco, N., Corredor, P. & Santamaría, R. (2010). Does informed trading occur in the options market? Some revealing clues. *Accounting & Finance,* 50(3), pp. 555-579.

Boudoukh, J., Feldman, R., Kogan, S. & Richardson, M. (2019). Information, trading, and volatility: Evidence from firm-specific news. *The Review of Financial Studies,* 32(3), pp. 992-1033.

Cartea, A. & Meyer-Brandis, T. (2010). How duration between trades of underlying securities affects option prices. *Review of Finance,* 14(4), pp. 749-785.

Chakrabarty, B., Pascual, R. & Shkilko, A. (2015). Evaluating trade classification algorithms: Bulk volume classification versus the tick rule and the Lee-Ready algorithm. *Journal of Financial Markets,* 25, pp. 52-79.

Chakravarty, S., Gulen, H. & Mayhew, S. (2004). Informed trading in stock and option markets. *The Journal of Finance,* 59(3), pp. 1235-1257.

Chan, K., Chung, Y. P. & Fong, W.-M. (2002). The informational role of stock and option volume. *Review of Financial Studies,* 15(4), pp. 1049-1075.

Chan, K., Ge, L. & Lin, T. C. (2015). Informational content of options trading on acquirer announcement return. *Journal of Financial and Quantitative Analysis,* 50(5), pp. 1057-1082.

Chordia, T., Hu, J., Subrahmanyam, A. & Tong, Q. (2019). Order flow volatility and equity costs of capital. *Management Science,* 65(4), pp. 1520-1551.

Chordia, T., Kurov, A., Muravyev, D. & Subrahmanyam, A. (2021a). Index option trading activity and market returns. *Management Science,* 67(3), pp. 1758-1778.

Chordia, T., Lin, T. C. & Xiang, V. (2021b). Risk-neutral skewness, informed trading, and the cross section of stock returns. *Journal of Financial and Quantitative Analysis,* 56(5), pp. 1713-1737.

Chordia, T., Roll, R. & Subrahmanyam, A. (2011). Recent trends in trading activity and market quality. *Journal of Financial Economics,* 101(2), pp. 243-263.

Chung, K. H., Park, S. & Ryu, D. (2016). Trade duration, informed trading, and option moneyness. *International Review of Economics & Finance,* 44s, pp. 395-411.

Du, B. & Fung, S. (2018). Directional information effects of options trading: Evidence from the banking industry. *Journal of International Financial Markets, Institutions and Money,* 56, pp. 149-168.

Dufour, A. & Engle, R. F. (2000). Time and the price impact of a trade. *The Journal of Finance,* 55(6), pp. 2467-2498.

Easley, D., De Prado, M. M. L. & O'hara, M. (2014a). VPIN and the flash crash: A rejoinder. *Journal of Financial Markets,* 17, pp. 47-52.
Easley, D., De Prado, M. M. L. & O'hara, M. (2011a). The exchange of flow toxicity. *The Journal of Trading,* 6(2), pp. 8-13.
Easley, D., De Prado, M. M. L. & O'hara, M. (2011b). The microstructure of the "flash crash": flow toxicity, liquidity crashes, and the probability of informed trading. *The Journal of Portfolio Management,* 37(2), pp. 118-128.
Easley, D., Kiefer, N. M. & O'hara, M. (1997a). The information content of the trading process. *Journal of Empirical Finance,* 4(2-3), pp. 159-186.
Easley, D., Kiefer, N. M. & O'hara, M. (1997b). One day in the life of a very common stock. *The Review of Financial Studies,* 10(3), pp. 805-835.
Easley, D., Kiefer, N. M., O'hara, M. & Paperman, J. B. (1996). Liquidity, information, and infrequently traded stocks. *The Journal of Finance,* 51(4), pp. 1405-1436.
Easley, D., López De Prado, M. M. & O'hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. *The Review of Financial Studies,* 25(5), pp. 1457-1493.
Easley, D. & O'hara, M. (1992). Time and the process of security price adjustment. *The Journal of finance,* 47(2), pp. 577-605.
Easley, D., O'hara, M. & Srinivas, P. S. (1998). Option volume and stock prices: Evidence on where informed traders trade. *The Journal of Finance,* 53(2), pp. 431-465.
Easley, D., O'hara, M. & Yang, L. (2014b). Opaque trading, disclosure, and asset prices: Implications for hedge fund regulation. *The Review of Financial Studies,* 27(4), pp. 1190-1237.
Elaut, G., Frömmel, M. & Lampaert, K. (2018). Intraday momentum in FX markets: Disentangling informed trading from liquidity provision. *Journal of Financial Markets,* 37, pp. 35-51.
Engle, R. F. (2000). The econometrics of ultra‐high‐frequency data. *Econometrica,* 68(1), pp. 1-22.
Engle, R. F. & Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, pp. 1127-1162.
French, K. R. & Roll, R. (1986). Stock return variances: The arrival of information and the reaction of traders. *Journal of financial economics,* 17(1), pp. 5-26.
Gao, L., Han, Y., Li, S. Z. & Zhou, G. (2018). Market intraday momentum. *Journal of Financial Economics,* 129(2), pp. 394-414.
Gharghori, P., Maberly, E. D. & Nguyen, A. (2017). Informed trading around stock split announcements: Evidence from the option market. *Journal of financial and quantitative analysis,* 52(2), pp. 705-735.
Haldane, A. G. (2012). The race to zero. *In:* Allen, F., Aoki, M., Fitoussi, J. P., Kiyotaki, N., Gordon, R. & Stiglitz, J. E. (eds.) *The Global Macro Economy and Finance.* London: Springer.
Hu, J. (2014). Does option trading convey stock price information? *Journal of Financial Economics,* 111(3), pp. 625-645.
Hujer, R. & Vuletić, S. (2007). Econometric analysis of financial trade processes by discrete mixture duration models. *Journal of Economic Dynamics and Control,* 31(2), pp. 635-667.
Jin, W., Livnat, J. & 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. & So, E. C. (2012). The option to stock volume ratio and future returns. *Journal of Financial Economics,* 106(2), pp. 262-286.

Johnson, T. L. & So, E. C. (2017). A simple multimarket measure of information asymmetry. *Management Science,* 64(3), pp. 1055-1080.

Kacperczyk, M. & Pagnotta, E. (2019). Chasing private information. *The Review of Financial Studies,* 32(12), pp. 4997-5047.

Kalaitzoglou, I. (2020). Time Will Tell! Towards the Construction of Instantaneous Indicators of Different Agent Types. *Working Paper*.

Kalaitzoglou, I. (2021). From Calendar to Economic Time. Deciphering the Arrival of Events in Irregularly Spaced Time. *In:* Zopounidis, C., Benkraiem, R. & Kalaitzoglou, I. (eds.) *Financial Risk Management and Modeling*. Cham: Springer.

Kalaitzoglou, I. & Ibrahim, B. M. (2013). Does order flow in the European Carbon Futures Market reveal information? *Journal of Financial Markets,* 16(3), pp. 604-635.

Kalaitzoglou, I. A. & Ibrahim, B. M. (2015). Liquidity and resolution of uncertainty in the European carbon futures market. *International Review of Financial Analysis,* 37, pp. 89-102.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pp. 1315-1335.

Lee, C. & Ready, M. J. (1991). Inferring trade direction from intraday data. *The Journal of Finance,* 46(2), pp. 733-746.

Mkansi, M. & Acheampong, E. A. (2012). Research philosophy debates and classifications: students' dilemma. *Electronic Journal of Business Research Methods,* 10(2), pp. 132-140.

Ni, S. X., Pan, J. & Poteshman, A. M. (2008). Volatility information trading in the option market. *The Journal of Finance,* 63(3), pp. 1059-1091.

Nofsinger, J. R. & Prucyk, B. (2003). Option volume and volatility response to scheduled economic news releases. *Journal of Futures Markets: Futures, Options, and Other Derivative Products,* 23(4), pp. 315-345.

O'hara, M. (2003). Presidential address: Liquidity and price discovery. *The Journal of Finance,* 58(4), pp. 1335-1354.

O'hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics,* 116(2), pp. 257-270.

Pan, J. & Poteshman, A. M. (2006). The Information in Option Volume for Future Stock Prices. *Review of Financial Studies,* 19(3), pp. 871-908.

Patel, V., Putniņš, T. J., Michayluk, D. & Foley, S. (2020). Price discovery in stock and options markets. *Journal of Financial Markets,* 47, pp. 100524.

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

Ross, S. A. (1976). Options and efficiency. *The Quarterly Journal of Economics,* 90(1), pp. 75-89.

Ryu, D. & Yang, H. (2018). The directional information content of options volumes. *Journal of Futures Markets,* 38(12), pp. 1533-1548.

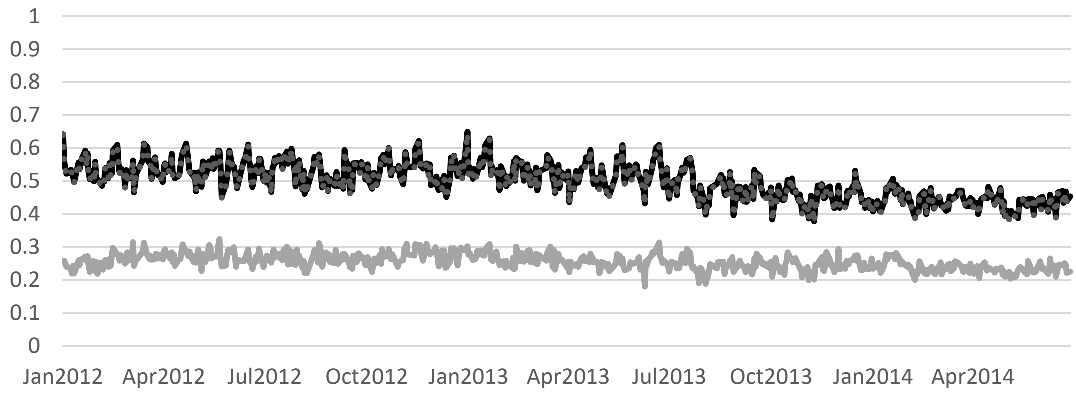
Sarwar, G. (2005). The informational role of option trading volume in equity index options markets. *Review of Quantitative Finance and Accounting,* 24(2), pp. 159-176.

Tsai, W. C., Chiu, Y. T. & Wang, Y. H. (2015). The information content of trading activity and quote changes: Evidence from VIX options. *Journal of Futures Markets,* 35(8), pp. 715-737.

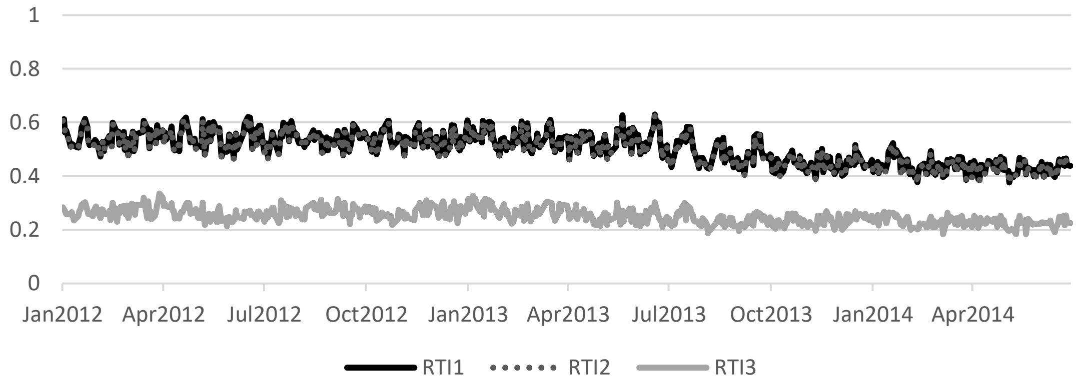
Xing, Y., Zhang, X. & Zhao, R. (2010). What does the individual option volatility smirk tell us about future equity returns? *Journal of Financial and Quantitative Analysis,* 45(3), pp. 641-662.

Figure 1: Percentage of informed trades

Panel A: Calls



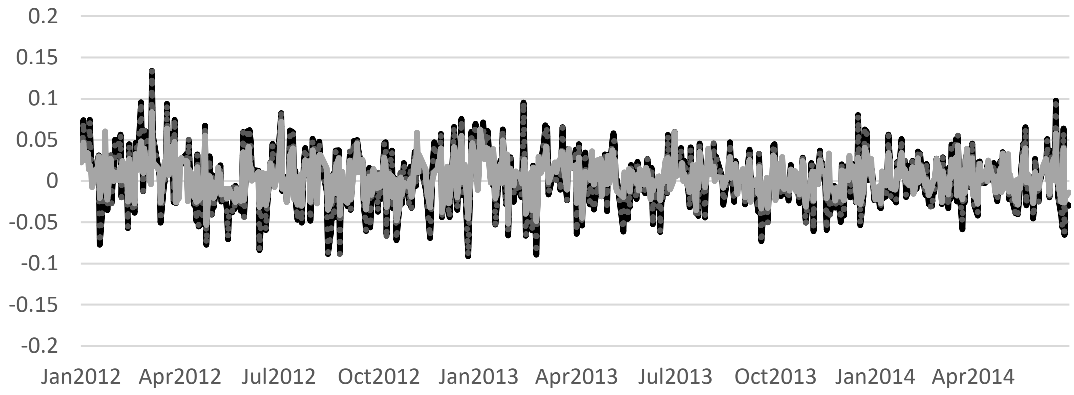
Panel B: Puts



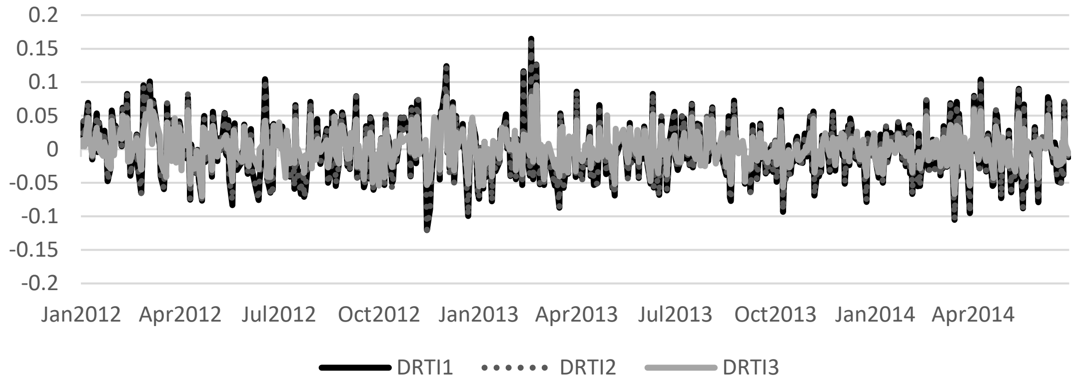
Note: the plots represent time-series of the relative arrival rate of trading volume and the information content is classified by different specifications. $RTIi$ refers to the relative arrival rate of trading volume, *i*= 1, 2, and 3 is different specifications of the dummy variable which is discussed in Section 4. The first plot represents the distribution of call options and the second plot represent the distribution of put options.

Figure 2: Distribution of directional informed dummy variables

Panel A: Calls



Panel B: Puts



Note: the plots represent time-series of the directional relative arrival rate of trading volume and the information content is classified by different specifications. $DRTIi$ refers to the directional relative arrival rate of trading volume, i = 1, 2, and 3 is different specifications of the dummy variable which is discussed in Section 4. The first plot represents the distribution of call options and the second plot represent the distribution of put options.

Table 1: Descriptive statistics										
	Calls					Puts				
		Trade duration		Volume			Trade duration		Volume	
RIC	**N**	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**N**	**Mean**	**Std Dev**	**Mean**	**Std Dev**
AXP	118,894	1,040.06	2,872.03	9.59	26.29	84,933	1,262.77	3,264.13	10.30	33.19
BA	353,005	693.19	2,346.18	8.26	43.81	242,009	884.51	2,752.61	7.32	15.03
CAT	522,326	369.13	1,593.71	8.14	22.98	441,258	412.63	1,717.09	8.34	23.68
CSCO	362,504	356.46	1,452.96	30.52	147.71	215,319	557.92	1,923.36	25.87	95.53
CVX	205,242	828.91	2,633.57	10.60	44.37	163,470	918.46	2,842.08	9.54	23.74
DD	108,012	1,157.06	3,152.20	15.71	114.59	73,376	1,516.30	3,688.02	12.38	31.89
DIS	183,033	875.99	2,580.46	11.59	36.65	109,974	1,238.78	3,208.83	13.28	34.44
GE	336,118	432.79	1,656.02	27.79	162.22	169,288	757.59	2,350.62	23.86	76.61
GS	650,816	277.94	1,286.41	6.47	16.60	394,981	416.92	1,682.74	6.04	14.19
HD	292,328	492.15	1,942.44	13.00	78.89	224,951	578.40	2,175.30	14.10	87.06
IBM	616,612	404.49	1,565.48	6.96	17.69	519,020	463.61	1,728.60	7.16	17.50
INTC	427,500	348.54	1,422.39	27.39	177.14	281,714	483.51	1,779.35	27.56	134.42
JNJ	326,924	431.87	1,927.75	14.73	91.21	241,580	543.13	2,232.08	13.40	60.78
JPM	665,879	354.73	1,570.01	16.01	74.95	455,450	474.06	1,902.92	14.89	53.95
KO	146,544	987.10	2,773.94	19.97	85.70	104,551	1,275.01	3,231.82	17.41	54.79
MCD	243,623	689.77	2,292.43	10.87	57.87	183,866	867.04	2,610.87	9.35	20.92
MMM	87,147	1,450.88	3,570.06	8.83	19.05	75,417	1,570.17	3,816.79	9.36	18.00
MRK	150,486	934.25	2,763.45	21.23	134.02	87,007	1,348.47	3,438.15	16.93	51.00
MSFT	634,897	331.82	1,404.02	25.36	122.82	409,270	475.35	1,756.61	24.84	102.40
NKE	151,714	979.11	2,865.37	9.56	29.18	117,671	1,106.11	3,129.55	10.33	32.87
PFE	194,522	706.21	2,296.28	29.68	199.24	107,695	1,080.12	2,982.84	25.01	101.43
PG	177,950	811.11	2,572.71	13.96	59.06	159,345	867.58	2,753.30	14.64	53.36
T	247,378	639.66	2,045.29	28.50	348.15	172,458	857.36	2,521.83	22.68	101.32
TRV	18,409	2,048.96	4,101.78	12.54	52.57	11,604	2,666.39	4,907.57	9.85	25.05
UNH	91,623	1,225.91	3,276.79	13.25	67.76	60,244	1,546.39	3,767.66	13.66	63.02
UTX	78,048	1,511.48	3,669.38	11.27	30.07	57,797	1,817.37	4,111.26	10.93	20.28
V	327,158	539.45	1,883.50	5.91	12.54	191,138	796.55	2,399.57	5.97	11.36
VZ	245,356	870.97	2,584.70	24.36	374.31	160,766	1,160.76	3,053.85	17.99	61.79
WMT	170,097	815.65	2,590.69	16.36	135.53	150,644	913.96	2,817.49	13.51	44.87
XOM	418,507	412.49	1,720.48	13.28	65.52	358,437	442.26	1,822.27	12.83	43.65
ALL	8,552,652	767.27	2347.08	15.72	94.95	6,025,233	976.65	2,745.64	14.31	50.27

Note: The table presents the statistic results for the options market activity from 03 Jan 2012 to 30 Jun 2014. *N* refers to the number of observations. *Trade duration* refers to the time duration between trades which is estimated in seconds. *Volume* refers to the average option volume per trade. The last line reports descriptive statistics for the whole sample.

Table 2: Estimation results of STM-ACWD															
	Calls							**Puts**							
RIC	**Omega**	**Alpha**	**Beta**	**Gamma1**	**Gamma2**	**g1**	**j1**	**Omega**	**Alpha**	**Beta**	**gamma1**	**gamma2**	**g1**	**j1**
AXP	0.094	0.100	0.900	1.032	0.222	1.332	0.520	0.111	0.106	0.894	1.004	0.215	1.252	0.531
BA	0.087	0.106	0.894	0.989	0.229	1.252	0.635	0.092	0.105	0.895	0.966	0.223	1.215	0.645
CAT	0.064	0.086	0.914	1.008	0.244	1.395	0.623	0.077	0.097	0.903	0.973	0.237	1.320	0.660
CSCO	0.072	0.083	0.917	0.981	0.233	1.280	0.645	0.079	0.085	0.915	0.967	0.224	1.212	0.646
CVX	0.082	0.084	0.916	1.036	0.226	1.281	0.506	0.080	0.087	0.913	1.000	0.221	1.250	0.585
DD	0.105	0.105	0.895	0.987	0.218	1.235	0.576	0.145	0.110	0.880	0.959	0.207	1.130	0.591
DIS	0.075	0.095	0.905	1.009	0.225	1.288	0.623	0.104	0.102	0.898	0.982	0.214	1.223	0.580
GE	0.071	0.076	0.924	1.003	0.232	1.277	0.590	0.093	0.077	0.923	1.052	0.222	1.313	0.413
GS	0.075	0.097	0.903	0.964	0.245	1.382	0.651	0.079	0.097	0.903	0.978	0.240	1.357	0.624
HD	0.078	0.074	0.926	0.891	0.217	1.138	0.739	0.096	0.086	0.914	0.834	0.210	1.064	0.811
IBM	0.070	0.095	0.905	1.006	0.240	1.358	0.624	0.083	0.107	0.893	0.974	0.235	1.288	0.658
INTC	0.065	0.071	0.929	1.035	0.242	1.408	0.524	0.080	0.079	0.921	1.032	0.233	1.343	0.500
JNJ	0.102	0.064	0.936	0.880	0.199	0.989	0.721	0.127	0.079	0.921	0.831	0.194	0.946	0.784
JPM	0.084	0.099	0.901	0.940	0.233	1.246	0.702	0.093	0.102	0.898	0.922	0.226	1.196	0.705
KO	0.085	0.085	0.915	1.057	0.223	1.387	0.462	0.099	0.089	0.911	1.059	0.215	1.315	0.436
MCD	0.092	0.102	0.898	1.019	0.229	1.282	0.556	0.079	0.085	0.915	0.995	0.224	1.300	0.566
MMM	0.101	0.080	0.912	1.015	0.215	1.298	0.507	0.112	0.095	0.905	0.982	0.209	1.158	0.541
MRK	0.119	0.106	0.894	1.001	0.219	1.233	0.526	0.191	0.112	0.868	0.972	0.207	1.177	0.517
MSFT	0.074	0.090	0.910	0.974	0.235	1.278	0.668	0.090	0.094	0.906	0.962	0.227	1.222	0.634
NKE	0.076	0.094	0.906	0.963	0.222	1.228	0.681	0.103	0.109	0.891	0.911	0.213	1.117	0.713
PFE	0.102	0.101	0.899	0.977	0.219	1.153	0.620	0.124	0.106	0.894	0.942	0.209	1.101	0.614
PG	0.102	0.104	0.896	0.965	0.220	1.170	0.630	0.129	0.125	0.875	0.873	0.210	1.048	0.760
T	0.070	0.073	0.927	1.038	0.229	1.389	0.501	0.096	0.084	0.917	1.012	0.218	1.273	0.483
TRV	0.142	0.128	0.869	1.044	0.206	1.243	0.516	0.262	0.119	0.828	1.062	0.200	1.178	0.445
UNH	0.106	0.097	0.903	0.961	0.213	1.199	0.584	0.270	0.115	0.834	0.984	0.206	1.151	0.487
UTX	0.167	0.107	0.868	0.992	0.211	1.252	0.550	0.186	0.094	0.874	0.981	0.205	1.189	0.516
V	0.069	0.098	0.902	1.011	0.238	1.352	0.632	0.083	0.098	0.902	1.019	0.228	1.300	0.566
VZ	0.084	0.089	0.911	1.022	0.222	1.289	0.550	0.109	0.097	0.903	1.022	0.213	1.203	0.496
WMT	0.087	0.090	0.910	1.021	0.224	1.244	0.550	0.096	0.097	0.903	0.966	0.218	1.183	0.626
XOM	0.092	0.098	0.902	0.989	0.231	1.228	0.607	0.114	0.131	0.870	0.898	0.224	1.156	0.764

Note: This table represent the estimation results of STM-ACWD. All results are statistically significant at 1% level.

Table 3: Average daily proportion of informed trades						
	Calls			Puts		
RIC	$RTI^{m=1}$	$RTI^{m=2}$	$RTI^{m=3}$	$RTI^{m=1}$	$RTI^{m=2}$	$RTI^{m=3}$
AXP	47.75%	47.33%	22.19%	47.65%	47.17%	21.87%
BA	47.80%	47.58%	24.83%	48.23%	47.94%	24.00%
CAT	50.66%	50.49%	26.13%	51.20%	51.03%	25.98%
CSCO	54.87%	54.67%	29.76%	53.60%	53.36%	28.43%
CVX	48.55%	48.24%	23.18%	47.70%	47.43%	23.91%
DD	45.11%	44.56%	21.09%	45.53%	44.99%	21.56%
DIS	46.06%	45.79%	22.39%	46.46%	46.11%	21.91%
GE	55.10%	54.91%	29.36%	56.03%	55.65%	27.60%
GS	52.82%	52.69%	27.22%	50.54%	50.31%	25.24%
HD	56.17%	55.99%	33.48%	57.18%	57.03%	32.49%
IBM	48.03%	47.87%	24.67%	48.02%	47.83%	23.99%
INTC	55.39%	55.18%	28.84%	55.02%	54.74%	27.85%
JNJ	68.71%	68.61%	44.93%	69.26%	69.17%	42.89%
JPM	52.73%	52.60%	29.26%	52.41%	52.26%	28.28%
KO	47.94%	47.55%	22.47%	48.96%	48.52%	22.49%
MCD	47.05%	46.76%	22.32%	47.25%	46.87%	21.90%
MMM	44.97%	44.40%	19.67%	46.37%	45.74%	20.99%
MRK	48.70%	48.37%	23.25%	48.49%	47.99%	22.84%
MSFT	53.98%	53.84%	29.13%	54.22%	54.06%	28.38%
NKE	43.23%	42.83%	21.15%	44.84%	44.59%	22.64%
PFE	50.07%	49.79%	26.58%	50.18%	49.77%	26.26%
PG	48.19%	47.87%	24.37%	49.18%	48.95%	24.88%
T	52.09%	51.83%	26.27%	52.76%	52.43%	26.65%
TRV	41.78%	40.79%	18.26%	42.64%	41.28%	17.84%
UNH	46.96%	46.46%	21.59%	48.46%	47.73%	21.11%
UTX	44.37%	43.76%	20.76%	45.83%	45.23%	21.20%
V	47.34%	47.09%	23.98%	48.12%	47.79%	23.58%
VZ	48.15%	47.80%	23.87%	49.77%	49.37%	24.86%
WMT	48.30%	47.89%	23.80%	48.89%	48.56%	23.93%
XOM	53.21%	53.07%	29.35%	53.05%	52.88%	28.72%
All	49.87%	49.55%	25.47%	50.26%	49.89%	25.14%

Note: The table presents the relative arrival rate of trading volume ($\boldsymbol{RTI_{k,t}^{m=i}}$), i = 1, 2, 3 is different specifications of the dummy variable we discussed in Section 4.

Table 4: Average daily value of the directional information measures

	Calls				Puts			
RIC	$DRTI^{m=1}$	$DRTI^{m=2}$	$DRTI^{m=3}$	$DisRTI$	$DRTI^{m=1}$	$DRTI^{m=2}$	$DRTI^{m=3}$	$DisRTI$
AXP	1.08%	1.02%	0.70%	4.05%	-0.59%	-0.60%	-0.18%	-2.49%
BA	0.43%	0.46%	0.29%	1.38%	-0.38%	-0.35%	-0.12%	-0.39%
CAT	0.14%	0.17%	0.25%	0.31%	-0.55%	-0.56%	-0.14%	-2.61%
CSCO	0.63%	0.62%	1.28%	4.16%	1.69%	1.73%	1.66%	7.32%
CVX	0.08%	0.12%	0.34%	1.03%	-1.77%	-1.76%	-1.29%	-6.62%
DD	-1.35%	-1.27%	-0.57%	-4.68%	-0.02%	-0.01%	-0.08%	0.60%
DIS	-1.07%	-1.07%	-0.54%	-5.13%	-0.57%	-0.56%	-0.37%	-1.66%
GE	0.07%	0.07%	0.44%	0.76%	-0.76%	-0.70%	-0.31%	-3.07%
GS	2.21%	2.20%	1.70%	8.10%	0.70%	0.70%	0.58%	2.18%
HD	1.05%	1.08%	1.00%	3.84%	-0.62%	-0.63%	-0.38%	-3.04%
IBM	-0.02%	0.00%	0.26%	0.62%	-0.21%	-0.21%	0.05%	0.24%
INTC	0.35%	0.38%	0.70%	2.16%	0.90%	0.92%	0.51%	3.44%
JNJ	-0.03%	-0.05%	0.59%	0.07%	1.29%	1.32%	1.14%	5.17%
JPM	0.13%	0.14%	0.41%	-0.38%	-0.28%	-0.28%	0.19%	-1.33%
KO	1.56%	1.53%	1.00%	5.33%	-1.48%	-1.49%	-0.94%	-6.73%
MCD	0.56%	0.55%	0.56%	2.02%	1.38%	1.33%	0.97%	5.94%
MMM	0.67%	0.76%	0.67%	3.22%	-0.11%	-0.05%	0.21%	-0.07%
MRK	-0.45%	-0.48%	-0.01%	-1.59%	-0.71%	-0.72%	-0.04%	-2.19%
MSFT	0.56%	0.56%	0.93%	2.54%	1.65%	1.64%	1.38%	6.28%
NKE	-1.37%	-1.36%	-0.38%	-6.36%	0.34%	0.31%	1.17%	2.14%
PFE	-0.95%	-0.89%	-0.26%	-4.80%	-0.44%	-0.43%	-0.10%	-2.75%
PG	1.87%	1.90%	1.53%	9.17%	-0.06%	-0.07%	0.89%	-1.60%
T	1.01%	1.00%	0.93%	5.05%	0.63%	0.61%	0.57%	2.89%
TRV	0.56%	0.47%	0.39%	0.44%	1.29%	0.92%	1.60%	9.22%
UNH	-0.88%	-0.83%	-0.28%	-4.63%	0.15%	0.24%	0.44%	0.09%
UTX	0.12%	0.11%	0.41%	0.92%	-0.98%	-0.97%	-0.83%	-5.33%
V	-0.91%	-0.89%	-0.29%	-3.42%	-1.91%	-1.92%	-0.93%	-8.30%
VZ	-0.28%	-0.26%	0.38%	-0.46%	-1.99%	-1.98%	-1.32%	-8.04%
WMT	1.57%	1.57%	0.87%	6.25%	0.53%	0.51%	1.02%	3.22%
XOM	2.43%	2.42%	1.83%	11.96%	-0.41%	-0.42%	0.08%	-1.37%
All	0.33%	0.33%	0.50%	1.40%	-0.11%	-0.12%	0.18%	-0.29%

Note: The table presents the average daily value of the directional information measures ($DRTI_{k,t}^{m}$ and $DisRTI_{k,t}$) as they are identified by progressively stricter classification criteria. $DRTI_{k,t}^{m=i}$ is to take the consideration of the trading direction into the relative arrival rate of trading volume, *i* = 1, 2, 3 is different specifications of the dummy variable we discussed in Section 4. $DisRTI_{k,t}$ is the inverse of the shape parameter, weighted by the direction of trade. This measure captures the magnitude of "informativeness" and its direction.

Table 5 Daily regressions of underlying asset volatility on option trading intensity

Panel A: Call options

τ -Days ahead	**Model 1**		**Model 2**		**Model 3**	
0	0.01677***	(22.86)	0.01678***	(22.94)	0.00737***	(10.13)
1	0.00577***	(11.85)	0.00581***	(11.82)	0.00332***	(4.27)
5	0.00153***	(3.25)	0.00146***	(3.19)	-0.00029	(-0.46)
10	0.00094**	(2.08)	0.00096**	(2.11)	0.00051	(0.72)

Panel B: Put options

τ -Days ahead	**Model 1**		**Model 2**		**Model 3**	
0	0.01237***	(20.30)	0.01233***	(20.70)	0.00332***	(5.99)
1	0.00424***	(6.10)	0.00428***	(6.19)	0.00005	(0.07)
5	0.00118**	(2.34)	0.00125**	(2.46)	-0.00063	(-1.12)
10	0.00110**	(2.17)	0.00110**	(2.15)	-0.00043	(-0.59)

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the estimation regression results of underlying asset volatility on options trading intensity. Panel A and B represent the results of call and put options, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset volatility is from day 0, 1, 5, and 10. Each model contains different independent variables ($\boldsymbol{RTI^{m=1}}, \boldsymbol{RTI^{m=2}}$, **and** $\boldsymbol{RTI^{m=3}}$) that refer to the relative arrival rate of trading volume with different specifications of the information dummy. The adjusted R-square for these models is varying from 0.029 to 0.079. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table 6 Daily regressions of underlying asset returns on option trading intensity

Panel A: Call options

τ -Days ahead	**Model 1**		**Model 2**		**Model 3**		**Model 4**	
0	0.00754***	(14.37)	0.00756***	(14.28)	0.00588***	(7.42)	0.00179***	(13.58)
1	0.00284***	(5.15)	0.00287***	(5.15)	0.00243***	(2.83)	0.00068***	(5.30)
5	-0.00008	(-0.13)	-0.00006	(-0.11)	-0.00028	(-0.33)	-0.00010	(-0.74)
10	0.00076	(1.11)	0.00074	(1.08)	0.00139	(1.20)	0.00022	(1.13)

Panel B: Put options

τ -Days ahead	**Model 1**		**Model 2**		**Model 3**		**Model 4**	
0	-0.00470***	(-9.66)	-0.00470***	(-9.72)	-0.00354***	(-4.50)	-0.00116***	(-10.27)
1	-0.00238***	(-5.15)	-0.00244***	(-5.25)	-0.00297***	(-4.40)	-0.00061***	(-5.72)
5	0.00048	(0.93)	0.00053	(1.00)	0.00090	(1.33)	0.00011	(0.88)
10	-0.00019	(-0.35)	-0.00017	(-0.31)	0.00013	(0.15)	-0.00011	(-0.93)

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the estimation regression results of underlying asset return on options trading intensity. Panel A and B represent the results of call and put options, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset return is from day 0, 1, 5, and 10. Each model contains different independent variables. $\boldsymbol{DRTI^{m=1}}, \boldsymbol{DRTI^{m=2}}$, **and** $\boldsymbol{DRTI^{m=3}}$ in Model 1, 2, and 3 refer to the directional relative arrival rate of trading volume with the different specifications of the information dummy. $\boldsymbol{DisRTI_{k,t}}$ in Model 4 refers to the inverse of the shape parameter, weighted by the direction of trade. The adjusted R-square for these models is varying from 0.023 to 0.03. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table 7 Two-sample t-test							
Panel A: Call options							
	Before announcement		**After announcement**				
	Obs.	Mean	Obs.	Mean	Difference in mean	t value	p value
$\boldsymbol{RTI^{m=1}}$	2962	0.512	2968	0.487	-0.024	-7.1	0
$\boldsymbol{RTI^{m=2}}$	2962	0.509	2968	0.484	-0.024	-7.15	0
$\boldsymbol{RTI^{m=3}}$	2962	0.257	2968	0.249	-0.007	-2.8	0.005
$\boldsymbol{DRTI^{m=1}}$	2962	0.002	2968	0.002	0	0.05	0.974
$\boldsymbol{DRTI^{m=2}}$	2962	0.002	2968	0.002	0	0	0.987
$\boldsymbol{DRTI^{m=3}}$	2962	0.005	2968	0.004	-0.001	-0.1	0.905
$\boldsymbol{DisRTI}$	2962	0.001	2968	0.01	0.009	0.6	0.533
Panel B: Put options							
	Before announcement		**After announcement**				
	Obs.	Mean	Obs.	Mean	Difference in mean	t value	p value
$\boldsymbol{RTI^{m=1}}$	2969	0.51	2960	0.496	-0.014	-3.85	0
$\boldsymbol{RTI^{m=2}}$	2969	0.506	2960	0.492	-0.014	-3.9	0
$\boldsymbol{RTI^{m=3}}$	2969	0.25	2960	0.249	-0.002	-0.65	0.525
$\boldsymbol{DRTI^{m=1}}$	2969	-0.002	2959	0.003	0.004	1	0.324
$\boldsymbol{DRTI^{m=2}}$	2969	-0.002	2959	0.002	0.004	0.9	0.368
$\boldsymbol{DRTI^{m=3}}$	2969	0.001	2959	0.007	0.005	2.15	0.033
$\boldsymbol{DisRTI}$	2969	-0.009	2959	0.018	0.027	1.6	0.115

Note: This table represent the results of two-sample t-test from the information measures before and after the corporate new announcements. The corporate new announcements are defined as annual and quarterly corporate releases of financial reports that appear in the U.S. Securities and Exchange Commission website. Panel A and B represent the results of call and put options, respectively. They report number of observations, mean values, difference in mean values before and after announcements, t-statistics, and p-values. $\boldsymbol{RTI^{m=1}}$, $\boldsymbol{RTI^{m=2}}$, **and** $\boldsymbol{RTI^{m=3}}$ refer to the relative arrival rate of trading volume with different specifications of the information dummy. $\boldsymbol{DRTI^{m=1}}$, $\boldsymbol{DRTI^{m=2}}$, **and** $\boldsymbol{DRTI^{m=3}}$ refer to the directional relative arrival rate of trading volume with the different specifications of the information dummy. $\boldsymbol{DisRTI_{k,t}}$ refers to the inverse of the shape parameter, weighted by the direction of trade.

Table 8 Distribution of trading intention measures across moneyness						
Panel A: Distribution of $RTI^{m=1}$						
	Calls			**Puts**		
RIC	**ITM**	**ATM**	**OTM**	**ITM**	**ATM**	**OTM**
AXP	5.56%	86.66%	7.79%	2.59%	82.50%	14.91%
BA	3.58%	84.41%	12.01%	0.96%	78.31%	20.73%
CAT	3.91%	80.79%	15.30%	2.57%	78.49%	18.94%
CSCO	7.59%	76.86%	15.55%	8.20%	76.03%	15.77%
CVX	2.49%	90.57%	6.94%	1.29%	85.30%	13.41%
DD	3.28%	88.58%	8.14%	1.44%	83.90%	14.66%
DIS	5.19%	85.46%	9.35%	2.62%	80.32%	17.07%
GE	8.67%	83.38%	7.95%	6.47%	78.61%	14.92%
GS	5.07%	78.69%	16.24%	2.30%	75.66%	22.04%
HD	5.09%	87.05%	7.86%	2.33%	81.72%	15.94%
IBM	2.08%	85.76%	12.16%	1.90%	81.37%	16.74%
INTC	8.90%	79.66%	11.44%	8.71%	76.65%	14.65%
JNJ	4.36%	91.39%	4.25%	2.29%	85.98%	11.74%
JPM	6.06%	80.82%	13.12%	4.12%	77.04%	18.84%
KO	3.70%	90.61%	5.69%	2.14%	86.44%	11.42%
MCD	2.06%	92.54%	5.41%	1.16%	88.49%	10.35%
MMM	2.87%	91.01%	6.13%	1.63%	84.70%	13.67%
MRK	3.88%	91.00%	5.12%	1.28%	84.25%	14.48%
MSFT	7.64%	80.40%	11.96%	5.99%	76.66%	17.35%
NKE	4.88%	80.83%	14.29%	4.46%	72.54%	23.00%
PFE	4.49%	88.99%	6.52%	3.29%	86.21%	10.50%
PG	3.42%	90.90%	5.67%	3.01%	76.76%	20.23%
T	5.10%	89.75%	5.15%	3.55%	86.44%	10.00%
TRV	3.39%	94.29%	2.33%	1.37%	83.71%	14.92%
UNH	4.28%	83.91%	11.81%	1.45%	79.94%	18.60%
UTX	3.12%	90.17%	6.71%	1.53%	84.81%	13.66%
V	5.18%	80.78%	14.03%	1.41%	74.95%	23.64%
VZ	4.06%	89.84%	6.10%	1.96%	86.04%	12.01%
WMT	3.48%	90.87%	5.65%	1.75%	87.45%	10.79%
XOM	2.24%	90.49%	7.26%	1.88%	81.97%	16.15%
All	4.52%	86.55%	8.93%	2.85%	81.44%	15.70%

Panel B: Distribution of $RTI^{m=2}$						
	Calls			Puts		
RIC	ITM	ATM	OTM	ITM	ATM	OTM
AXP	5.55%	86.65%	7.80%	2.59%	82.49%	14.92%
BA	3.58%	84.40%	12.02%	0.96%	78.31%	20.73%
CAT	3.91%	80.79%	15.31%	2.57%	78.49%	18.94%
CSCO	7.59%	76.85%	15.57%	8.20%	76.01%	15.79%
CVX	2.48%	90.57%	6.95%	1.28%	85.29%	13.43%
DD	3.27%	88.58%	8.15%	1.44%	83.90%	14.66%
DIS	5.19%	85.46%	9.35%	2.62%	80.31%	17.07%
GE	8.67%	83.37%	7.96%	6.48%	78.59%	14.94%
GS	5.07%	78.68%	16.24%	2.30%	75.65%	22.05%
HD	5.09%	87.04%	7.87%	2.33%	81.72%	15.95%
IBM	2.08%	85.75%	12.17%	1.90%	81.37%	16.74%
INTC	8.90%	79.65%	11.45%	8.70%	76.65%	14.65%
JNJ	4.36%	91.38%	4.25%	2.29%	85.97%	11.75%
JPM	6.06%	80.81%	13.13%	4.12%	77.03%	18.85%
KO	3.70%	90.60%	5.69%	2.14%	86.44%	11.43%
MCD	2.05%	92.54%	5.41%	1.16%	88.48%	10.36%
MMM	2.86%	91.01%	6.14%	1.63%	84.71%	13.66%
MRK	3.87%	90.99%	5.14%	1.27%	84.25%	14.48%
MSFT	7.64%	80.40%	11.97%	5.99%	76.65%	17.36%
NKE	4.86%	80.84%	14.30%	4.46%	72.53%	23.01%
PFE	4.49%	88.98%	6.53%	3.30%	86.20%	10.51%
PG	3.42%	90.90%	5.68%	3.02%	76.72%	20.27%
T	5.10%	89.75%	5.16%	3.55%	86.44%	10.01%
TRV	3.33%	94.33%	2.34%	1.39%	83.71%	14.90%
UNH	4.27%	83.89%	11.84%	1.44%	79.93%	18.62%
UTX	3.10%	90.17%	6.73%	1.52%	84.81%	13.66%
V	5.18%	80.78%	14.04%	1.41%	74.94%	23.65%
VZ	4.06%	89.83%	6.11%	1.96%	86.03%	12.01%
WMT	3.48%	90.86%	5.66%	1.75%	87.46%	10.79%
XOM	2.24%	90.48%	7.28%	1.88%	81.95%	16.17%
All	4.51%	86.54%	8.94%	2.85%	81.43%	15.71%

Panel C: Distribution of $RTI^{m=3}$						
	Calls			Puts		
RIC	ITM	ATM	OTM	ITM	ATM	OTM
AXP	5.57%	84.71%	9.72%	2.83%	79.22%	17.94%
BA	3.89%	80.96%	15.15%	1.20%	76.44%	22.36%
CAT	4.14%	77.84%	18.02%	3.21%	75.28%	21.51%
CSCO	8.93%	72.75%	18.33%	9.71%	71.36%	18.93%
CVX	3.12%	88.00%	8.87%	1.60%	79.89%	18.51%
DD	3.55%	86.28%	10.17%	1.90%	80.52%	17.58%
DIS	5.14%	83.05%	11.81%	2.76%	77.08%	20.16%
GE	10.11%	79.46%	10.43%	8.07%	73.51%	18.42%
GS	6.95%	74.39%	18.67%	2.84%	72.15%	25.02%
HD	8.66%	79.30%	12.04%	4.13%	73.54%	22.33%
IBM	2.41%	81.72%	15.88%	2.23%	79.00%	18.77%
INTC	10.21%	75.06%	14.73%	9.78%	72.17%	18.05%
JNJ	8.92%	83.38%	7.70%	3.78%	77.58%	18.64%
JPM	8.07%	75.54%	16.39%	4.56%	73.49%	21.95%
KO	3.79%	88.55%	7.66%	2.19%	83.09%	14.73%
MCD	2.20%	90.72%	7.08%	1.18%	85.78%	13.05%
MMM	2.40%	89.71%	7.88%	1.97%	81.13%	16.90%
MRK	4.04%	89.11%	6.85%	1.39%	82.52%	16.10%
MSFT	8.80%	75.98%	15.22%	7.56%	71.64%	20.80%
NKE	4.21%	79.53%	16.26%	3.69%	71.91%	24.41%
PFE	5.35%	85.67%	8.98%	2.87%	82.89%	14.24%
PG	3.77%	88.60%	7.64%	2.59%	76.01%	21.40%
T	5.72%	87.53%	6.75%	4.60%	81.36%	14.04%
TRV	3.04%	93.71%	3.24%	1.24%	81.98%	16.79%
UNH	4.37%	82.71%	12.92%	1.66%	78.94%	19.40%
UTX	3.49%	87.51%	9.00%	1.93%	83.07%	14.99%
V	5.27%	77.81%	16.92%	2.08%	72.80%	25.12%
VZ	4.45%	87.51%	8.04%	2.36%	82.58%	15.05%
WMT	3.36%	88.82%	7.82%	1.97%	84.37%	13.67%
XOM	3.16%	85.47%	11.37%	2.66%	77.45%	19.89%
All	5.24%	83.38%	11.38%	3.35%	77.96%	18.69%

Note: This table represents the distribution of trading intention measures across moneyness. Panel A, B, and C represent the distribution results of different independent variables ($RTI^{m=1}$, $RTI^{m=2}$, and $RTI^{m=3}$). Moneyness is calculated by S/K, where K refers to the strike price of the options contracts and S refers to the underlying concurrent mid-quote price. DOTM call contracts have moneyness smaller than 0.9, OTM call contracts have moneyness between 0.9 and 0.95, ATM call contracts have moneyness between 0.95 and 1.05, ITM call contracts have moneyness between 1.05 and 1.1, and DITM call contracts have moneyness over 1.1. Put contracts are based on the opposite classification.

Table 9 Daily regressions of underlying asset volatility on intraday option trading intensity									
Panel A: Call options									
	Model 1			**Model 2**			**Model 3**		
τ -Days ahead	$RTI\_F^{m=1}$	$RTI\_M^{m=1}$	$RTI\_L^{m=1}$	$RTI\_F^{m=2}$	$RTI\_M^{m=2}$	$RTI\_L^{m=2}$	$RTI\_F^{m=3}$	$RTI\_M^{m=3}$	$RTI\_L^{m=3}$
0	0.00828***	0.00330***	0.00361***	0.00832***	0.00329***	0.00361***	0.00531***	0.00228***	0.00229***
	(23.12)	(11.69)	(11.30)	(23.12)	(11.54)	(11.26)	(9.43)	(5.38)	(4.78)
1	0.00068*	0.00084***	0.00306***	0.00068*	0.00085***	0.00310***	-0.00024	0.00099**	0.00203***
	(1.84)	(2.90)	(9.36)	(1.85)	(2.91)	(9.44)	(-0.42)	(2.33)	(4.24)
5	-0.00033	0.00053*	-0.00015	-0.00039	0.00057*	-0.00017	-0.00163***	0.00040	-0.00006
	(-0.88)	(1.82)	(-0.46)	(-1.04)	(1.93)	(-0.50)	(-2.87)	(0.93)	(-0.12)
10	0.00014	0.00014	0.00048	0.00013	0.00015	0.00052	-0.00036	0.00018	0.00064
	(0.39)	(0.48)	(1.45)	(0.36)	(0.52)	(1.56)	(-0.63)	(0.41)	(1.32)
Panel B: Put options									
	Model 1			**Model 2**			**Model 3**		
τ -Days ahead	$RTI\_F^{m=1}$	$RTI\_M^{m=1}$	$RTI\_L^{m=1}$	$RTI\_F^{m=2}$	$RTI\_M^{m=2}$	$RTI\_L^{m=2}$	$RTI\_F^{m=3}$	$RTI\_M^{m=3}$	$RTI\_L^{m=3}$
0	0.00600***	0.00268***	0.00300***	0.00610***	0.00264***	0.00298***	0.00284***	0.00192***	0.00142***
	(18.79)	(10.67)	(10.77)	(19.02)	(10.44)	(10.66)	(5.68)	(5.07)	(3.38)
1	0.00085***	0.00083***	0.00186***	0.00086***	0.00085***	0.00187***	-0.00067	0.00035	0.00056
	(2.61)	(3.26)	(6.57)	(2.64)	(3.31)	(6.55)	(-1.35)	(0.93)	(1.32)
5	-0.00022	-0.00021	0.00007	-0.00023	-0.00017	0.00012	-0.00152***	-0.00084**	0.00009
	(-0.68)	(-0.81)	(0.25)	(-0.69)	(-0.67)	(0.41)	(-3.02)	(-2.19)	(0.22)
10	-0.00032	0.00038	-0.00021	-0.00029	0.00035	-0.00019	-0.00071	-0.00010	-0.00099**
	(-0.98)	(1.45)	(-0.74)	(-0.87)	(1.35)	(-0.65)	(-1.41)	(-0.25)	(-2.32)

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the estimation regression results of underlying asset volatility on options trading intensity. Panel A and B represent the results of call and put options, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset volatility is from day 0, 1, 5, and 10. Each model contains different independent variables that refer to the relative arrival rate of trading volume with different specifications of the information dummy. $RTI\_F^{m}_{k,t}$, $RTI\_M^{m}_{k,t}$, and $RTI\_L^{m}_{k,t}$ refer to the percentage of the information dummy in a 30-mins interval at the beginning, middle, and end of the trading day. The adjusted R-square for these models is varying from 0.029 to 0.065. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table 10 Daily regressions of underlying asset return on intraday option trading intensity

Panel A: Call options

	Model 1			Model 2			Model 3			Model 4		
τ -Days ahead	$DRTI\_F^{m=1}$	$DRTI\_M^{m=1}$	$DRTI\_L^{m=1}$	$DRTI\_F^{m=2}$	$DRTI\_M^{m=2}$	$RDTI\_L^{m=2}$	$DRTI\_F^{m=3}$	$DRTI\_M^{m=3}$	$RDTI\_L^{m=3}$	$DisRTI\_F$	$DisRTI\_M$	$DisRTI\_L$
0	0.00613***	-0.00041	-0.00165***	0.00621***	-0.00049	-0.00167***	0.00647***	-0.00073	-0.00252***	0.00162***	-0.00016*	-0.00049***
	(15.89)	(-1.20)	(-4.54)	(15.99)	(-1.42)	(-4.57)	(9.25)	(-1.39)	(-4.33)	(16.24)	(-1.93)	(-5.41)
1	-0.00010	0.00055	0.00074**	-0.00004	0.00052	0.00076**	-0.00003	0.00008	0.00045	-0.00003	0.00011	0.00023**
	(-0.24)	(1.59)	(2.04)	(-0.09)	(1.49)	(2.08)	(-0.04)	(0.15)	(0.77)	(-0.27)	(1.32)	(2.56)
5	0.00024	0.00013	0.00049	0.00022	0.00013	0.00042	0.00056	-0.00024	0.00067	0.00007	-0.00004	0.00008
	(0.62)	(0.38)	(1.33)	(0.56)	(0.36)	(1.14)	(0.79)	(-0.46)	(1.15)	(0.71)	(-0.48)	(0.85)
10	-0.00042	0.00009	-0.00003	-0.00042	0.00008	-0.00001	0.00032	-0.00011	-0.00043	-0.00005	0.00002	0.00002
	(-1.06)	(0.26)	(-0.07)	(-1.07)	(0.21)	(-0.03)	(0.46)	(-0.21)	(-0.73)	(-0.54)	(0.29)	(0.21)

Panel B: Put options

	Model 1			Model 2			Model 3			Model 4		
τ -Days ahead	$DRTI\_F^{m=1}$	$DRTI\_M^{m=1}$	$DRTI\_L^{m=1}$	$DRTI\_F^{m=2}$	$DRTI\_M^{m=2}$	$RDTI\_L^{m=2}$	$DRTI\_F^{m=3}$	$DRTI\_M^{m=3}$	$RDTI\_L^{m=3}$	$DisRTI\_F$	$DisRTI\_M$	$DisRTI\_L$
0	-0.00277***	0.00042	0.00089***	-0.00279***	0.00046	0.00084**	-0.00308***	0.00065	0.00119**	-0.00084***	0.00009	0.00023***
	(-8.04)	(1.32)	(2.75)	(-8.03)	(1.45)	(2.57)	(-4.94)	(1.35)	(2.30)	(-9.82)	(1.24)	(2.95)
1	-0.00034	-0.00036	-0.00111***	-0.00032	-0.00038	-0.00106***	-0.00059	-0.00062	-0.00111**	-0.00009	-0.00010	-0.00031***
	(-0.98)	(-1.14)	(-3.42)	(-0.94)	(-1.19)	(-3.26)	(-0.95)	(-1.30)	(-2.15)	(-1.03)	(-1.33)	(-3.95)
5	-0.00013	-0.00009	0.00037	-0.00013	-0.00012	0.00042	-0.00034	-0.00049	0.00034	-0.00008	0.00001	0.00009
	(-0.38)	(-0.27)	(1.14)	(-0.38)	(-0.36)	(1.28)	(-0.55)	(-1.01)	(0.65)	(-0.87)	(0.09)	(1.12)
10	-0.00078**	-0.00015	0.00036	-0.00077**	-0.00014	0.00036	-0.00079	0.00022	-0.00002	-0.00022**	-0.00002	0.00008
	(-2.26)	(-0.48)	(1.09)	(-2.21)	(-0.44)	(1.09)	(-1.25)	(0.45)	(-0.05)	(-2.52)	(-0.21)	(0.99)

Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the estimation regression results of underlying asset return on options trading intensity. Panel A and B represent the results of call and put options, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset return is from day 0, 1, 5, and 10. Each model contains different independent variables. $DRTI\_F^m, DRTI\_M^m,$ **and** $DRTI\_L^m$ in Model 1, 2, and 3 refer to the directional relative arrival rate of trading volume with the different specifications of the information dummy in a 30-mins interval at the beginning, middle, and end of the trading day. $DisRTI\_F_{k,t}$, $DisRTI\_M_{k,t}$, and $DisRTI\_L_{k,t}$ in Model 4 refers to the inverse of the shape parameter, weighted by the direction of trade, in a 30-mins interval at the beginning, middle, and end of the trading day. The adjusted R-square for these models is varying from 0.016 to 0.028. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table 11 Percentile rank regressions of underlying asset volatility on option trading intensity						
Panel A: Bottom 20 percentile of VIX						
	Call options			Put options		
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
0	0.01686***	0.01683***	0.00852***	0.01091***	0.01083***	0.00299***
	(11.58)	(11.53)	(7.69)	(7.80)	(7.71)	(3.09)
1	0.00473***	0.00488***	0.00358***	0.00182**	0.00195**	-0.00051
	(4.18)	(4.30)	(2.68)	(2.05)	(2.18)	(-0.45)
5	0.00004	0.00001	0.00039	-0.00014	-0.00000	0.00098
	(0.05)	(0.01)	(0.30)	(-0.15)	(-0.00)	(0.84)
10	-0.00111	-0.00112	-0.00013	-0.00099	-0.00097	-0.00157
	(-1.10)	(-1.10)	(-0.09)	(-0.86)	(-0.83)	(-0.99)
Panel B: Top 20 percentile of VIX						
	Call options			Put options		
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
0	0.01581***	0.01587***	0.00426***	0.01309***	0.01298***	0.00326**
	(12.22)	(12.31)	(3.01)	(10.90)	(10.84)	(2.51)
1	0.00381***	0.00376***	0.00046	0.00061	0.00055	-0.00190
	(3.30)	(3.21)	(0.30)	(0.60)	(0.55)	(-1.64)
5	-0.00019	-0.00071	-0.00214	-0.00121	-0.00109	-0.00119
	(-0.15)	(-0.60)	(-1.30)	(-0.80)	(-0.70)	(-0.82)
10	-0.00070	-0.00072	0.00209	-0.00152	-0.00157	-0.00317
	(-0.49)	(-0.52)	(0.93)	(-0.72)	(-0.73)	(-1.17)

Note: Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the estimation regression results of underlying asset volatility on options trading intensity. Panel A and B represent the results of bottom and top 20 percentile of VIX weeks, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset volatility is from day 0, 1, 5, and 10. Each model contains different independent variables ($\boldsymbol{RTI^{m=1}, RTI^{m=2}, \text{and } RTI^{m=3}}$) that refer to the relative arrival rate of trading volume with different specifications of the information dummy. The adjusted R-square for these models is varying from 0.01 to 0.148. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table 12 Percentile rank regressions of underlying asset returns on option trading intensity								
Panel A: Bottom 20 percentile of VIX								
	Call options				Put options			
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
0	0.00952***	0.00949***	0.00930***	0.00237***	-0.00535***	-0.00542***	-0.00515***	-0.00145***
	(7.82)	(7.70)	(5.65)	(8.57)	(-5.66)	(-5.72)	(-3.88)	(-7.15)
1	0.00249**	0.00252**	0.00123	0.00051**	0.00028	0.00021	0.00007	-0.00006
	(2.23)	(2.23)	(0.73)	(2.04)	(0.31)	(0.23)	(0.06)	(-0.30)
5	0.00128	0.00115	0.00023	-0.00004	0.00072	0.00084	0.00207	0.00027
	(1.16)	(1.03)	(0.15)	(-0.14)	(0.79)	(0.92)	(1.49)	(1.28)
10	0.00239*	0.00233*	0.00414**	0.00061**	0.00192*	0.00188*	0.00271*	0.00025
	(1.90)	(1.83)	(2.28)	(2.05)	(1.72)	(1.68)	(1.76)	(0.99)
Panel B: Top 20 percentile of VIX								
	Call options				Put options			
τ -Days ahead	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
0	0.00464***	0.00456***	0.00296*	0.00112***	-0.00182*	-0.00182*	0.00058	-0.00042*
	(4.25)	(4.17)	(1.93)	(4.49)	(-1.71)	(-1.69)	(0.38)	(-1.74)
1	0.00070	0.00072	0.00053	0.00010	-0.00183*	-0.00187*	-0.00197	-0.00047**
	(0.65)	(0.67)	(0.35)	(0.39)	(-1.83)	(-1.83)	(-1.35)	(-2.06)
5	-0.00012	0.00019	0.00002	0.00001	0.00013	0.00017	0.00048	-0.00001
	(-0.08)	(0.16)	(0.01)	(0.04)	(0.09)	(0.11)	(0.33)	(-0.02)
10	0.00124	0.00121	0.00210	0.00040	0.00153	0.00164	0.00130	0.00030
	(0.52)	(0.52)	(0.49)	(0.55)	(1.00)	(1.04)	(0.52)	(0.84)

Note: *** and ** refer to the significant level at 1% and 5%. This table represents the estimation regression results of underlying asset return on options trading intensity. Panel A and B represent the results of bottom and top 20 percentile of VIX weeks, respectively. They report coefficients and t-statistics in parentheses. The time lag of underlying asset return is from day 0, 1, 5, and 10. Each model contains different independent variables. $\boldsymbol{DRTI^{m=1}}, \boldsymbol{DRTI^{m=2}}$, **and** $\boldsymbol{DRTI^{m=3}}$ in Model 1, 2, and 3 refer to the directional relative arrival rate of trading volume with the different specifications of the information dummy. $\boldsymbol{DisRTI_{k,t}}$ in Model 4 refers to the inverse of the shape parameter, weighted by the direction of trade. The adjusted R-square for these models is from 0.001 to 0.106. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.

Table 13 Fama and MacBeth regression results of underlying stock return on option trading intensity

Panel A: Call options

τ -Days ahead	**0**		**1**		**5**		**10**	
Model	**1**	**2**	**1**	**2**	**1**	**2**	**1**	**2**
$\boldsymbol{DRTI^{m=1}}$	0.00686***	0.00578***	0.00161**	0.00122*	0.00010	0.00027	0.00049	0.00063
	(10.31)	(6.21)	(2.49)	(1.83)	(0.21)	(0.45)	(0.77)	(0.90)
VIX		-0.00007		-0.00024***		-0.00003		0.00006
		(-1.28)		(-4.12)		(-0.49)		(1.26)
SVOL		-0.00000		-0.00000		0.00000		-0.00000*
		(-0.61)		(-1.12)		(0.47)		(-1.69)
OVOL		0.00001		0.00000		-0.00001		0.00003*
		(0.45)		(0.17)		(-0.23)		(1.65)
O/S		0.00238		-0.00321		0.00387		-0.01400**
		(0.36)		(-0.49)		(0.46)		(-2.02)
VPIN		0.00241		0.00846***		0.00171		0.00048
		(1.54)		(3.52)		(1.04)		(0.43)
Adj. R2	0.049	0.194	0.037	0.230	0.033	0.171	0.033	0.171

Panel B: Put options

τ -Days ahead	**0**		**1**		**5**		**10**	
Model	**1**	**2**	**1**	**2**	**1**	**2**	**1**	**2**
$\boldsymbol{DRTI^{m=1}}$	-0.00269***	-0.00079	-0.00143***	-0.00187***	0.00030	0.00038	0.00034	0.00030
	(-4.19)	(-0.31)	(-2.84)	(-3.29)	(0.62)	(0.72)	(0.71)	(0.56)
VIX		0.00028		-0.00029***		0.00005		0.00005
		(0.89)		(-3.58)		(1.17)		(1.10)
SVOL		0.00002		-0.00000		-0.00000		-0.00000
		(0.97)		(-1.21)		(-1.62)		(-1.57)
OVOL		-0.00086		0.00014		0.00003		0.00003
		(-0.99)		(0.93)		(1.04)		(1.26)
O/S		-0.00592		0.00096		-0.01764		-0.01483**
		(-0.54)		(0.14)		(-1.56)		(-2.16)
VPIN		0.00233*		0.00827***		0.00120		0.00072
		(1.70)		(3.33)		(1.01)		(0.62)
Adj. R2	0.037	0.185	0.033	0.225	0.032	0.171	0.031	0.170

Note: Note: ***, **, * refer to the significant level at 1%, 5%, and 10%. This table represents the Fama-MacBath regression results of underlying asset returns on option trading intensity. Panel A and B represent the results of call and put options, respectively. They report coefficients, t-statistics in parentheses, and adjusted R-square (Adj. R2). Each column shows the results based on 0, 1, 5, and 10 days ahead of underlying asset return. Each model contains different independent variables. $\boldsymbol{DRTI^{m=1}}$ refers to the directional relative arrival rate of trading volume. *VIX* refers to the daily closing price of CBOE Volatility Index. *SVOL* and *OVOL* refer to underlying stock trading volume and option trading volume, respectively. *O/S* refers to the options/stock trading volume ratio. *VPIN* refers to volume-synchronized probability of informed trading metric. Standard errors are Newey-West autocorrelation and heteroscedasticity consistent. All the regressions include asset and time fixed effects.