Does News, Order Flow, or Illiquidity drive jumps in stock returns? In the day or in the night?*

Yoontae Jeon[†] Thomas H. McCurdy[‡] Stephen Szaura[§]

Current Version: May 2023

Abstract

We investigate how firm-level news, stock illiquidity, and order imbalances are reflected in stock return jumps and idiosyncratic jump risk. We analyze these relationships for the entire day as well as for the daytime and overnight trading periods. Our results show that information flows and trading frictions are significantly related to non-parametric measures of jump intensity and jump-size distributions and reveal variations over the trading day and across individual firms. Our analyses could enrich the economic content of models for stock return dynamics which typically have treated the sources of jumps as latent, and also help identify jumps due to information arrival as opposed to liquidity or strategic trading based on private information.

Keywords: Jumps; News flow; Illiquidity; Overnight Returns
JEL Classification: G10, G14, G19

*We thank seminar participants at BI Norwegian School of Business, Norwegian School of Economics, Oxford Man Institute of Quantitative Finance, and Toronto Metropolitan University. We thank Zhi Da, Darya Yuferova and Xiaofei Zhao for helpful comments. We acknowledge financial support from the Social Sciences and Humanities Research Council Insight Grant number 435-2020-1333.

[†]DeGroote School of Business, McMaster University, yoontae.jeon@mcmaster.ca
[‡]Rotman School of Management, University of Toronto, tom.mccurdy@rotman.utoronto.ca
[§]BI Norwegian Business School, stephen.w.szaura@bi.no

1 Introduction

One of the most fundamental questions in asset pricing theory seeks to find driving forces of changes in asset prices. In particular, understanding the source of large price movements, typically labelled as "jumps", carries significant importance for investment and risk management decisions. On one hand, information flows have been thought to be the main drivers of price movements in a market where traders update their beliefs of fair asset price based on all available information to them. On the other hand, changes in the asset trading environment due to asset order flow or trading costs can also cause large price movements, especially over shorter periods of time.

Many news announcements are made after the traditional market close. However, with the advent of new market trading platforms, trading activities taking place in the overnight trading environment are documented to be dramatically different in terms of the liquidity and order flow. As sources of stock return jumps, the daytime versus overnight trading environment and news arrival remains relatively under explored.

We have compiled a comprehensive dataset, which includes the frequency and content of firm-specific intraday news articles, stock trading cost measures, and stock order-flow measures, to investigate how those variables affect return jumps for publicly-traded individual stocks. We also focus on comparing overnight versus daytime periods to highlight differences in the sources and features of jumps during those alternative trading periods which exhibit significant differences in the variables of interest.

Trading frictions, information flows, and daytime versus overnight (traditionally referred to as 'market' versus 'non-market') trading periods have all been the focus of a large and growing number of research studies.

Chordia and Subrahmanyam (2004) found that contemporaneous order imbalances are positively related to returns (their empirical analysis was entirely based on open to close returns excluding the overnight trading period in order to avoid the bid-ask bounce). Amihud and Mendelson (1986) found that the expected stock return is positively associated with a higher bid-ask spread. Hence, it is possible that stock return jumps are positively associated with a higher relative effective spread (RES).

Furthermore, since the US stock markets became fully electronic after 2004, they are largely a fragmented electronic network with quoting and trading occurring on several exchanges and trading platforms (e.g., O'Hara and Ye, 2011; Menkveld and Yueshen, 2018). This fragmented nature can lead to cases of mismatch between the supply and demand of market participants and cause large price changes. For example, during the well-known *Flash Crash* on May 6, 2010 between 14 : 00 and 15 : 00, there were large and temporary selling pressures associated with stock market indexed securities and many individual stocks experienced very large drops in return (Kirilenko et al., 2017). Weller (2019) formalizes this intuition and proposes a framework linking the information content embedded in the bid-ask spread to the probability of future jumps in asset prices.

Christoffersen et al. (2021) find that market illiquidity shocks, not market-order flow, induce stock market index jumps. Khan and Riordan (2020) show that intra-day liquidity fragmentation across different exchanges predicts intra-day price jumps in SPX 100 stocks. Using intra-day instrumental variable estimation, they show that jumps are predictable; the liquidity suppliers' information is reflected in the liquidity fragmentation. The ask (bid) side liquidity fragmentation increases the probability of positive (negative) jumps. However, these papers do not consider news announcements or the overnight trading period in their empirical analysis,

This literature raises the following questions concerning the relationship between stock return jumps and illiquidity and/or order-flow.

Question 1: How is stock return jump probability related to illiquidity (RES) and order imbalance shocks (OIB)?

Question 2: Given a stock return jump, how are expected returns (jump sizes) related to the RES and the OIB?

Macroeconomic and firm-specific news announcement information have also been shown to be important determinants of asset return jumps.[1] Using textual analyses, Baker et al. (2021b) and Gurkaynak, Kisacikoglu, and Wright (2020) document that news source clarity and non-headline text are important in helping to explain jumps in asset prices. Using a large sample of stocks and firm-specific news, Jeon, McCurdy, and Zhao (2021) report that jumps are related to the frequency, tone, and uncertainty associated with news announcements, as well as the news type and source. For a panel of international stock return indices, Bongaerts et al. (2022) find more support for the impact of macroeconomic news than market-wide illiquidity jumps or order flow.

Many news announcements are made after traditional market-close time. Substantial differences between overnight returns and daytime returns were noted in Cliff, Cooper, and Gulen (2008) who found that overnight returns are positive, large and significantly different from zero whereas open to close day returns are sometimes negative but largely close to zero (also see Berkman et al., 2012).

However, due to advances of electronic trading platforms developing across the globe, market participants can now trade through alternative venues at nearly any time of the day. Barclay and Hendershott (2004) note that despite the fact that the overnight period has reduced trading activity with much higher trading costs than during the trading day and greater adverse selection costs, liquidity provision remains competitive.[2]

News is an important source of idiosyncratic risk in overnight returns (Boudoukh et al. (2019)) and earn significant abnormal returns (see Wong and Yang (2019) and Cui and Gozluklu (2021) amongst others).[3] Therefore, it would be natural to expect that jumps in asset prices during the regular trading period versus those occurring during after-hours trading to be significantly different, for example, with respect to frequency, size, and source. However,

[1] For example, see Lee and Mykland (2008), Lee (2012); as well as Jiang, Lo, and Verdelhan (2011) for U.S. Treasury bond return jumps.

[2] see Barclay and Hendershott (2003) for information about overnight price discovery.

[3] As documented in Bollerslev, Li, and Todorov (2016) and Hendershott, Livdan, and Rosch (2020), firms' CAPM Beta is positive (negative) in the overnight (daytime) period. Jiang and Zhu (2017) find that trading strategies based on lagged overnight stock return jumps earn significant returns over the following months which the authors show is consistent with the under-reaction of investors.

most papers that study what drives stock return jumps omit the overnight trading period which could be important in order to separately analyze the news impact and illiquidity trading costs during the daytime versus the overnight period.[4]

Our paper is also related to recent studies exploring sources of large stock price movements and associated risk premiums. Hong and Wang (2000) show theoretically that there should be higher returns over trading periods instead of non-trading periods with more volatile returns over trading periods. Bollerslev, Li, and Xue (2018) investigate the relationship between trading intensity (stock volume) and spot volatility around public news announcements which has implications for the way in which financial markets process information. Brogaard et al. (2018) note that high-frequency traders tend to be absent around large price movements. Pelger (2020) generates high-frequency intraday factors that capture intraday risk premiums that reverse in overnight returns. Kapadia and Zekhnini (2019) and Bégin, Dorion, and Gauthier (2020) highlight the importance of idiosyncratic stock return jump risk, however, they do not characterize what economic mechanisms drive this risk.[5] Our results also add to a large and growing literature on understanding how news impacts stock prices.[6]

Those examples of prior research findings raise the following questions concerning the relationship between stock return jumps and information flows during daytime versus nighttime trading periods.

Question 3: What is the relationship between the frequency of jumps and volatility during the overnight versus daytime trading periods?

Question 4: Is the sensitivity of jump probability to the frequency of news articles different

[4]To the best of our knowledge our paper is the first to analyze what drives stock return jumps in the overnight period.

[5]Herskovic et al. (2016) highlight economic sources of idiosyncratic volatility risk.

[6]See for instance: Cutler, Poterba, and Summers (1989), Berry and Howe (1994), Mitchell and Mulherin (1994), Chan (2003), Antweiler and Frank (2004), Tetlock (2007b), Neuhierl, Scherbina, and Schiusene (2008), Tetlock, Saar-Tsechansky, and Macskassy (2008), Tetlock (2007a), Tetlock (2011), Savor (2012), Baker, Bloom, and Davis (2016), Manela and Moreira (2017), Zhao (2017), Baker et al. (2021a), Ke and Kelly (2022), Bybee et al. (2021), and Baker et al. (2021b). Additionally see Gentzkow, Kelly, and Taddy (2019), Gurkaynak, Kisacikoglu, and Wright (2020), Kelly, Manela, and Moreira (2021), and Bybee et al. (2022), Fedyk and Hodson (2021), Fedyk (2022), and van Binsbergen et al. (2022) for a literature on the application of text data content to stock returns.

in the overnight versus daytime trading periods?

Question 5: Is the sensitivity of jump probability to the sentiment of news articles different in the overnight versus daytime trading periods?

Question 6: Given the risks associated with jumps, are returns different in the overnight versus daytime trading periods?

We contribute to these growing lines of research by investigating how firm-specific intraday news flow and content, stock trading cost measures, and stock order-flow measures affect return jumps for publicly-traded individual stocks. We focus on comparing overnight versus day-time periods to highlight differences in the sources and features of jumps during those alternative trading periods. Our approach incorporates variables that are are compiled from time-stamped intraday data, designed to capture public firm-specific information flows, stock level illiquidity, and order-flow measures. Using non-parametric analyses we do not impose any prior structure on stock returns.[7] In the following paragraphs, we briefly summarize our results, organized to address the six questions listed above.

Jump probabilities are related positively to illiquidity and order imbalance shocks. Baseline results for all firms for the entire day suggest that a one standard deviation increase in stock illiquidity (stock order imbalances) increases the odds of a jump by 48% (22%). Given a negative (positive) jump, expected returns are negatively (positively) associated with illiquidity and order imbalance shocks.

Our baseline results for all firms for the entire day reveal that a one standard deviation increase in news count increases the odds of a jump by 37%. When comparing stock return jumps during the day to those in overnight trading periods, we find lower volatility and a lower frequency of jumps in the overnight versus daytime periods. However, the sensitivity of stock return jump probability to the number of after-hours news articles is larger than the sensitivity to the number of news articles during the daytime trading period. For example,

[7] In contrast, Brogaard et al. (2022) estimate a structural return variance decomposition model to distinguish different sources of return variance, and find that 31% of return variance comes from noise, or stock illiquidity, while 37% of return variance can be attributed to public firm-specific information.

a one standard deviation increase in number of overnight (daytime) news articles increases the odds of a jump by 41% (14%).

The jump-size mean is statistically significantly related to the microstructural measures as well as the news content. Including all jump returns, there is a positive coefficient on news counts, both microstructural variables remain positive and significant, and the R^2 increases to 10.05% when those illiquidity and order imbalance variables are included. Conditional on negative jumps, the increase in R^2 is from 7.67% to 14.18%.

The average article news sentiment is higher for larger firms than for the average firm in the cross section; and the average night news article sentiment is higher than the average day news article sentiment for all firms in the cross section. Baseline results for all firms for the entire day reveal that a one standard deviation increase in the average article sentiment increases the odds of a jump by 11%. Jumps are positively associated with average daytime news sentiment but negatively associated with average overnight news sentiment for the sample of all firms. But return jumps for the top 500 firms are positively associated with both average daytime and overnight news sentiment.

Our results show how daytime and overnight news flow, stock illiquidity, and order flow are reflected in stock return jumps and idiosyncratic jump risk. We find a higher alpha for both positive and negative jumps associated with the higher number of daytime news articles belonging to the Top 10% news category, suggesting that the higher number of news articles during the day leads to substantially higher idiosyncratic returns, particularly for larger firms.

Finally, we perform several robustness tests, sub-sample analysis associated with the electronic market era subset (January 2004 to December 2020), removal of earnings announcement days, removal of financial crisis years, and find that the results do not change significantly.

The remainder of this paper is organized as follows: our data, variable construction, and stock return jump identification methodology are presented in Section 2, and the main results are presented in Section 3. Section 4 provides several robustness tests and Section 5 concludes.

2 Methodology and Data

2.1 Data and Variables

Our empirical analyses require merging daily data from three different data sources: daily stock data, daily frequency and contents contained in news articles, and daily stock trade level illiquidity and order flow data. We obtain daily stock data from CRSP on WRDS for US common stocks by keeping observations with share codes 10 or 11. Opening (9:30 EST) and closing (16:00 EST) stock prices are first available in the CRSP daily files beginning January 1993.

Daily stock illiquidity and order-flow data have historically been obtained and aggregated from the intra-day trades and quotes (TAQ) files on WRDS. Recently WRDS created a daily data set of various commonly used stock illiquidity and order flow measures called WRDS Intra-Day Indicators. We use WRDS MTAQ Intra-Day Indicators based on the MTAQ monthly intra-day trades and quotes files from January 1, 1993 to December 31, 2014, then WRDS DTAQ Intra-Day Indicators based on the DTAQ monthly intra-day trades and quotes files from January 1, 2015 to December 31, 2020.

In our empirical analyses, our stock illiquidity trade-based measure is defined by the relative effective spread (RES_t) measure as defined in Goyenko, Holden, and Trzcinka (2009).

$$RES_t = \frac{\sum_k^{N_t} VOL_k \times \left[\frac{2|P_k - M_k|}{M_k}\right]}{\sum_k^{N_t} VOL_k} \tag{2.1}$$

where M_k is midpoint bid, ask for transaction k, P_k is stock trade price for transaction k, N_t is the number of stock trades on day t, VOL_k is dollar trade size for transaction k.[8]

[8] In WRDS MTAQ Intra-Day Indicators for the sample period 2000 to 2014 we use the variable

Daily stock order flow data is computed as stock net order imbalances (OIB_t) which are defined as dollar buy volume minus dollar sell volume scaled by total buy sell dollar volume (as per Chordia and Subrahmanyam (2004)).

$$OIB_t = \frac{\sum_k^{N_t} Buy_k \times VOL_k - \sum_k^{N_t} Sell_k \times VOL_k}{\sum_k^{N_t} Buy_k \times VOL_k + \sum_k^{N_t} Sell_k \times VOL_k} \tag{2.2}$$

where Buy_k ($Sell_k$) is 1 and 0 if trade k is buy (sell) order.[9]

Stock news article data are obtained from the RavenPack database. We focus on how novel (innovative or surprising) news is related to stock market jumps. The RavenPack news dataset provides a variable that measures how *novel* a news article is by comparing the content of the news article with previous news article about the same company. The highest novelty score is 100. In our analysis we only retain news articles that have a novelty score of 100 in order to focus on news that is most likely to be a surprise. In addition to the number of novel news, we also measure the tone of these news articles using the proprietary sentiment measure that RavenPack provides. The sentiment measure from RavenPack ranges from 0 to 100; we subtract 50 so that it ranges from –50 to 50 with a negative value of the re-centred measure representing negative sentiment and a positive value representing positive sentiment.[10]

Num News is the total number of daily news announcements for a particular firm (timestamps). Using the sentiment score provided by RavenPack for each article, Avg. Sentiment is calculated as the equal-weighted average sentiment of all daily news announcements for a particular firm.

Since we are provided with the timestamps of each news article, we compute the total number of news articles whose time stamps fall between the stock market opening

ESpreadPctVWi which is stock daily dollar-weighted relative effective spreads. In WRDS DTAQ Intra-Day Indicators for the sample period 2015 to 2020 we use the variable EffectiveSpreadPercentDW which is stock daily dollar-weighted relative effective spreads.

[9] In WRDS MTAQ Intra-Day Indicators we compute the order imbalance as the difference between BUYDOLLARLR1 and SELLDOLLARLR1 over their sum.

[10] Note that RavenPack does not provide the news text, so we are not able to generate an uncertain words measure.

time (09:30 EST) and the stock market closing time (16:00 EST) and term this variable Num News_{Day} with corresponding average sentiment of those news articles represented as Avg. Sentiment_{Day}.

Similarly we compute the total number of news articles whose time stamps fall between the stock market closing time (16:00 EST) and the stock market opening time (09:30 EST) the following trading day variable as Num News_{Night} with corresponding average sentiment of those news articles represented as Avg. Sentiment_{Night}.

Summary statistics of total daily news counts, total day and night news counts, average sentiment for all daily news as well as for day and night news, relative effective spreads and order imbalances for all stocks (largest 500 firms by market capitalization) are presented in Panel A (B) of Table 1. The data sample period is from January 2000 to the end of 2020. The average number of total daily news articles is higher for larger firms, 1.21 for largest 500 firms compared to 0.37 for the average firm in the cross section. The average article news sentiment is higher for larger firms, 3.44 for the largest 500 firms compared to 1.3 for the average firm in the cross section.

INSERT TABLE 1 HERE

The average number of night news articles is higher than the number of day news articles for all firms in the cross section, 0.23 versus 0.12, as well as for the largest 500 firms, 0.54 versus 0.43. Similarly, the average night news article sentiment is higher than the average day news article sentiment for all firms in the cross section, 0.24 versus 0.11, as well as for the largest 500 firms, 0.46 versus 0.33.

Unsurprisingly, stock illiquidity is higher for the average firm in the cross section, 0.02 with a standard deviation of 0.13, than for the largest 500 firms in the sample, 0.01 with a standard deviation of 0.06. Stock order imbalance is higher for the average firm in the cross section, -0.02 with a standard deviation of 0.37, whereas for the largest 500 firms in the sample the average stock order imbalance is positive 0.01 with a standard deviation of 0.18.

In our subsequent empirical analyses, introduced in equation 3.1 in Section 3.1, we standardized all news related, stock illiquidity, and stock order imbalance variables, in order to have the same mean and standard deviation across firms. We also removed observations of stock illiquidity and stock-order imbalance that are larger than two, i.e. observations of stock illiquidity and stock order imbalance that are beyond two standard deviations of the mean, in order to mitigate the impact of outliers in our analyses.

2.2 Jump Detection

Realized jumps in stock returns are identified using the well known and widely used non-parametric approach (Lee and Mykland (2008)).[11] Jumps are identified in the returns of three trading periods: (1) the entire daily trading period which begins at 16:00 EST and goes until 16:00 EST of the next trading day, (2) the night trading period which begins at 16:00 EST and goes until 09:30 EST the following trading day, and (3) the daytime period which begins at 09:30 EST the following day and ends on that same next calendar day at 16:00 EST.

To identify jumps, we use the distribution from Lee and Mykland (2008), and use J to denote using the distribution of maximums and the number subscript to J refers to the percentile of the corresponding jump identification criteria. If there is additionally a D (N) subscript this indicates a day (night) trading period and no subscript indicates the entire daily trading period is used.

For example, J_{99} corresponds to the 99th percentile of the distribution of maximums as in Lemma 1 in Lee and Mykland (2008) using the entire daily trading period returns, whereas $J_{D,99}$ ($J_{N,99}$) corresponds to the 99th percentile of the distribution of maximums as in Lemma 1 in Lee and Mykland (2008) using the daytime (overnight) period returns. Effectively, we identify each return observation as jump if the absolute value of the return is above 5.1024, 4.4881 times the time-varying daily spot volatility, respectively. As a result, the thresholds

[11]Note that we use the Gilder, Shackleton, and Taylor (2014) modified test statistic, also used in Jeon, McCurdy, and Zhao (2021).

to identify jumps according to these criteria are also time-varying. Correspondingly, $J_{D,99}$ ($J_{N,99}$) identifies a day (night) trading period jump if the absolute value of the day (night) trading period return is above 5.28 times the time-varying daily spot volatility, respectively.

Daily log stock returns, r_{t_i}, are defined as being the log of the ratio between the closing stock prices $r_{t_i} = \log\left(S_C(t_i)/S_C(t_{i-1})\right)$ which leads to the natural decomposition $r_{t_i} = r_{N,t_i} + r_{D,t_i}$ where the night and day returns are defined as $r_{D,t_i} = \log\left(S_C(t_i)/S_O(t_i)\right)$ and $r_{N,t_i} = \log\left(S_O(t_i)/S_C(t_{i-1})\right)$ respectively where $S_O(t_i)$ is the opening stock price ($9:30AM$ EST) and $S_C(t_i)$ is the closing stock price ($16:00PM$ EST). The Lee and Mykland (2008) for day and overnight jump identification are

$$\begin{aligned}\mathcal{L}_D(t_i) &= \frac{r_D(t_i)}{\widehat{\sigma_D}(t_i)} = \frac{\log\left(S_C(t_i)/S_O(t_i)\right)}{\widehat{\sigma_D}(t_i)}\\ \mathcal{L}_N(t_i) &= \frac{r_N(t_i)}{\widehat{\sigma_N}(t_i)} = \frac{\log\left(S_O(t_i)/S_C(t_{i-1})\right)}{\widehat{\sigma_N}(t_i)}\end{aligned} \tag{2.3}$$

where

$$\begin{aligned}\widehat{\sigma_D}^2(t_i) &= \frac{\sum_{j=i-K+2}^{i-1} |\log\left(S_O(t_i)/S_C(t_{i-1})\right| \cdot |\log\left(S_C(t_{i-1})/S_O(t_{i-1})\right|}{K-2}\\ \widehat{\sigma_N}^2(t_i) &= \frac{\sum_{j=i-K+2}^{i-1} |\log\left(S_C(t_{i-1})/S_O(t_{i-1})\right| \cdot |\log\left(S_O(t_{i-1})/S_C(t_{i-2})\right|}{K-2}\end{aligned} \tag{2.4}$$

where a jump occurs during the daytime period at t_i if $|\mathcal{L}_D(t_i)| > 5.28$, i.e.

$$J_{D,99} = \begin{cases} 1 & \text{if } |\mathcal{L}_D(t_i)| > 5.28 \\ 0 & \text{if } |\mathcal{L}_D(t_i)| < 5.28 \end{cases}$$

and a jump occurs during the overnight period at t_i if $|\mathcal{L}_N(t_i)| > 5.28$, i.e.

$$J_{N,99} = \begin{cases} 1 & \text{if } |\mathcal{L}_N(t_i)| > 5.28 \\ 0 & \text{if } |\mathcal{L}_N(t_i)| < 5.28 \end{cases}$$

Table 2 presents summary statistics associated with the daily as well as day (night) trading period realized jumps. Each indicator variable takes a value of 1 if there is a jump identified using the specific criteria for stock i on daytime period t, and is 0 otherwise.

INSERT TABLE 2 HERE

The total number of days where J_{99} equals 1 for all firms is 298,110 (out of around 23.9 million days with non-missing returns), indicating that there is on average one jump every 80 (=24/0.30) days.[12] The frequency of daily jumps varies significantly inversely with the size of the firm; the largest 500 firms is one out of every 112 days ($= 20/2,242$) which shows that the frequency of daily jumps is far lower for larger firms than the average firm in the cross section.

The total number of days where the daytime (overnight) period jumps $J_{D,99}$ ($J_{N,99}$) equals 1 for all firms is 384,523 (95,224) out of around 15.76 million days with non-missing returns, indicating that there is on average one jump every 41 (165) day (night) trading periods. As with the frequency daily jumps, the frequency of day and night trading period return jumps varies significantly inversely with the size of the firm. The frequency of day (night) trading period jumps varies significantly inversely with the size of the firm as for the largest 500 firms is one out of every 60 (153) days. Our summary statistics results of $J_{D,99}$ and $J_{N,99}$ for all firms in Panel A of Table 2 shows that the frequency of daytime period jumps is higher than the frequency of night trading period jumps for the average firm in the cross section. Similarly, Panel B of Table 2 displays the same conclusion for the largest 500

[12] J_{95} equals 1 for all firms is 432,890 (out of around 23.9 million days with non-missing returns), indicating that there is on average one jump every 55 (=24/0.43) days.

firms. In unreported results (available upon request) we find that the same finding holds when excluding earnings announcement days, during the electronic market era subsample of $[2004, 2020]$, when excluding the financial crisis 2008 and 2009 years as well as the COVID 2020 year.

3 Main Results

3.1 Jump Probability Results

We investigate links between realized jumps and news flows, stock illiquidity, and stock order flow. We start by using logistic regressions to examine how the probability of jumps is related to the news flows (measured by the news count and news tone) versus individual stock level microstructural frictions (measured by stock illiquidity and order imbalances).

$$\begin{aligned} \text{logit}(p_{it}) = b_0 &+ b_1 \cdot \text{Num News} + b_2 \cdot |\text{Avg. Sentiment}| + b_3 \cdot |ret_{i,t-1}| + b_4 \cdot RES \\ &+ b_5 \cdot OIB + \epsilon_{i,t} \end{aligned} \tag{3.1}$$

In our logistic regressions, the explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances. By doing so, we primarily rely on the time series variation in the explanatory variables to explain the probability of jumps. The sample period is from January 2000 to February 2020. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity; the t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses.

INSERT TABLE 3 HERE

Table 3 reports the results where Panel A reports the coefficient estimates for the whole sample and Panel B reports the standardized odds ratio associated with the corresponding variable. As a baseline for comparison, we begin with the logistic regression specification of Jeon, McCurdy, and Zhao (2021) which only uses news flows variables. In column 1 we find that probability of a jump is statistically significantly related to the news count and average article sentiment as was found in Jeon, McCurdy, and Zhao (2021). Column 2 replicates the same logistic regression specification as in column 1 for non-missing values of stock illiquidity and order imbalance measures in order to compare the incremental additional impact of stock illiquidity and order imbalance measures in columns (3)–(5). This will be the main baseline for comparison of the incremental addition of stock illiquidity and order imbalances. The standardized odds ratio associated with the news count is 1.37 and for average sentiment is 1.1. The baseline results suggest that one standard deviation increase in news count increases the odds of a jump by 37% (and 11% percentage for average article sentiment) with an R^2 5.41%.

We expect the coefficients b4 and b5 for both RES and OIB to be positive and statistically significant. Column 3 adds stock illiquidity (RES) to the specification in column 2 and order imbalance measure (OIB) in column 4 and column 5 adds both together. The probability of a jump is statistically significantly related to the stock illiquidity and stock order imbalance. The standardized odds ratio associated with the stock illiquidity (stock order imbalances) is 1.48 (1.22), suggesting that one standard deviation increase in stock illiquidity (stock order imbalances) increases the odds of a jump by 48% (22%). The R^2 increases from 5.41% to 5.88% when both variables are added together.

Table A.1 reports the results using the Factiva data from Jeon, McCurdy, and Zhao (2021) where Panel A reports the coefficient estimates for the whole sample and Panel B reports the standardized odds ratio associated with the corresponding variable. As a baseline for comparison we begin with the logistic regression specification of Jeon, McCurdy, and Zhao (2021) with news flows versus individual stock level microstructural frictions. In column 1 we find that probability of a jump is statistically significantly related to the news count, absolute value of news tone, and percentage of uncertain words as was found in Jeon, McCurdy, and Zhao (2021). The standardized odds ratio associated with the news count is 1.22, absolute value of news tone is 1.02, and percentage of uncertain words is 1.01. The baseline results suggest that one standard deviation increase in news count increases the odds of a jump by 22% (and 2% and 1% for absolute value of news tone and percentage of uncertain words respectively) with an R^2 1.47%. Column 2 replicates the same logistic regression specification as in column 1 for non-missing values of stock illiquidity and order imbalance measures in order to compare the incremental additional impact of stock illiquidity and order imbalance measures. We expect the coefficients b5 and b6 for both RES and OIB to be positive and statistically significant. Column 3 adds stock illiquidity (RES) to the specification in column 2 and order imbalance measure (OIB) in column 5 and column 6 adds both together. The probability of a jump is statistically significantly related to the stock illiquidity and stock order imbalance. The standardized odds ratio associated with the stock illiquidity (stock order imbalances) is 1.38 (1.31), suggesting that one standard deviation increase in stock illiquidity (stock order imbalances) increases the odds of a jump by 38% (31%). The R^2 increases from 2.48% to 3.41% when both variables are added together

The patterns are, in general, similar when analyzing the realized jumps of the subset of 500 largest firms in columns 2–5 of Table A.2. However, for the 500 large firms, the number of news articles is more important than the stock illiquidity and order imbalances in explaining jumps (as indicated by a higher odds ratio), whereas for the average firm in the cross section, stock illiquidity is of most important.

The results in Tables 3, A.1, and A.2 show that stock return jump probability is positively related with stock illiquidity (RES) and order imbalance shocks (OIB) both for all stocks and largest 500 firms. Table 3 and A.1 show that for the average firm in the cross section there is higher probability of jump due to stock illiquidity than news (both for the Factiva and Ravepack datasets). Whereas although stock illiquidity and order imbalances are important for large firm stock return jumps, news counts are associated with the most likely jump source as indicated in Table A.2 with the higher standardized odds ratio for news count over stock illiquidity or order flow. Economically, our results show that the average firm jump risk is more likely to be from stock illiquidity shocks, however, larger firm stock return jumps are more sensitive to news source information.

3.2 Jump Size Results

Next, we analyze how the jump-size distribution is affected by news flows, stock illiquidity, and stock order flow. For the jump-size mean, we run the following regression:

$$\begin{aligned} E[r_{i,t}|J_{99}=1] = b_0 + b_1 \cdot \text{Num News} + b_2 \cdot |\text{Avg. Sentiment}| + b_3 \cdot ret_{i,t-1} \\ + b_4 \cdot RES + b_5 \cdot OIB + \epsilon_{i,t} \end{aligned} \tag{3.2}$$

INSERT TABLE 4 HERE

Tables 4 reports the results for all firms using the Ravenpack news data set. Panel A reports the results for all jumps (regardless of the sign of jumps); Panels B and C split the analyses for positive versus negative jumps respectively. In Panel A, the jump-size mean is statistically significantly related to the news content: positively related to news counts and the news tone as was found in Jeon, McCurdy, and Zhao (2021). Column 2 replicates the same regression specification as in column 1 for non-missing values of stock illiquidity and order imbalance measures in order to compare the incremental impact of stock illiquidity and order imbalance measures. In Panel A, the jump-size mean is statistically significantly related to the stock microstructural content: positively related to stock illiquidity and stock order imbalances. When we put all the jump returns in the same regressions, in Panel A, there is a positive coefficient on news counts, RES and OIB remains positive and significant.

The R^2 increases from 0.59% to 10.05% when both RES and OIB variables are added together in Panel A. In Panel B, positive jump returns are significantly and positively related to the stock illiquidity and order imbalance: higher illiquidity is associated with higher trading costs which is associated with higher positive returns on jump days as is higher net buying pressure (as higher buying pressure forces market makers to increase the price). The R^2 increases from 1.35% to 3.68% when both RES and OIB variables are added together in Panel B. In Panel C, we find that given a negative jump, expected returns are negatively associated with RES and OIB. The R^2 increases from 7.77% to 14.818% when both RES and OIB variables are added together in Panel C.

Table A.3 reports the results for all firms using the Factiva news data set. Panel A reports the results for all jumps (regardless of the sign of jumps) and Panels B and C we split the analyses for positive versus negative jumps respectively. In Panel A, the jump-size mean is statistically significantly related to the news content: positively related to news counts, the news tone and the percentage of uncertain words as was found in Jeon, McCurdy, and Zhao (2021). Column 2 replicates the same regression specification as in column 1 for non-missing values of stock illiquidity and order imbalance measures in order to compare the incremental additional impact of stock illiquidity and order imbalance measures. In Panel A, the jump-size mean is statistically significantly related to the stock microstructural content: negatively related to stock illiquidity and positively related to stock order imbalances. When we put all the jump returns in the same regressions, in Panel A, there is a positive coefficient on news counts, RES changes sign and is no longer significant, however, OIB remains positive and significant. The R^2 increases from 0.92% to 13.81% when both RES and OIB variables are added together in Panel A. In Panel B, positive jump returns are significantly and positively related to the stock illiquidity and order imbalance (although not significant for OIB): higher illiquidity is associated with higher trading costs which is associated with higher positive returns on jump days. The R^2 increases from 1.77% to 3.15% when both RES and OIB variables are added together in Panel B. In Panel C, negative jump returns are significantly and negatively related both to the stock illiquidity and order imbalance: lower illiquidity is associated with lower trading costs which is associated with higher negative returns on jump days. The R^2 increases from 8.68% to 14.15% when both RES and OIB variables are added together in Panel C.

The patterns are, in general, similar when analyzing the realized jumps of the subset of 500 largest firms in columns 2–5 of Table A.4 to that of all firms in Table 4. However, for the 500 large firms the OIB is no longer significant when including the RES in either Panel B or C, News count and stock illiquidity have the largest coefficients of magnitude (and significance) for both positive and negative jumps.

The evidences in Tables 4, A.3, and A.4 support the idea that given a negative (positive) stock return jump, expected returns are negatively (positively) associated with stock illiquidity and order imbalances.

3.3 Separating Day and Night News, Illiquidity, and Order Flow Jump Probability and Size Results

Motivated by recent advances in electronic trading systems to allow for after-hours market trading, we re-investigate our findings in sections 3.1 and 3.2 in order to quantify the impact on the likelihood sensitivity of jumps to news announcements, stock illiquidity, order imbalances during the daytime period and during the overnight trading periods.

$$\begin{aligned}\text{logit}(p_{it}) = b_0 &+ b_1 \cdot \text{Num News}_{Night} + b_2 \cdot |\text{Avg. Sentiment}_{Night}| \\ &+ b_3 \cdot \text{Num News}_{Day} + b_4 \cdot |\text{Avg. Sentiment}_{Day}| \\ &+ b_5 \cdot |ret_{i,t-1}| + b_6 \cdot RES + b_7 \cdot OIB + \epsilon_{i,t} \qquad (3.3)\end{aligned}$$

INSERT TABLE 5 HERE

We estimate the logistic regression in equation 3.3 with the news count and average sentiment separated during the overnight and daytime periods for all firms and report the results in Tables 5. The probability of a jump is positively statistically significantly related to the the number of night news articles, number of day news articles, and average sentiment of day news articles (as well as positively related to stock illiquidity and stock order imbalance as in section 3.1). The probability of a jump is negatively statistically significantly related to average sentiment of night news articles. The standardized odds ratio associated with the number of overnight (daytime) news articles is 1.41 (1.14), suggesting that a one standard deviation increase in number of overnight (daytime) news articles increases the odds of a jump by 41% (14%). The results indicate sensitivity of stock return jump probability to the number of after-hours news articles is larger than the sensitivity to the number of daytime news articles.

The patterns are, in general, similar when analyzing the realized jumps of the subset of 500 largest firms in columns 2–5 of Table A.5. However, for the 500 large firms, the number of night news articles and average sentiment are more important than the stock illiquidity and order imbalances in explaining jumps, whereas for the average firm in the cross section, stock illiquidity is the most important. Table 5 shows that all stock return jumps are positively associated with average daytime news sentiment but negatively associated with average overnight news sentiment, however, for top 500 (A.5) firms stock return jumps are positively associated with both average daily and night news sentiment.

Next we re-estimate the stock return regressions in equation 3.4 in Section 3.2 in Tables 4 for all firms with the news count and average sentiment separated during the overnight and daytime periods and report the results in Tables 6.

$$\begin{aligned}E[r_{i,t}|J_{99}=1] = b_0 &+ b_1 \cdot \text{Num News}_{Night} + b_2 \cdot |\text{Avg. Sentiment}_{Night}| \\ &+ b_3 \cdot \text{Num News}_{Day} + b_4 \cdot |\text{Avg. Sentiment}_{Day}| \\ &+ b_5 \cdot |ret_{i,t-1}| + b_6 \cdot RES + b_7 \cdot OIB + \epsilon_{i,t} \qquad (3.4)\end{aligned}$$

INSERT TABLE 6 HERE

Table 6 shows that the sensitivity of stock return jump probability to the number of overnight news articles is larger than the sensitivity due to the number of daytime news articles. Moreover, although illiquidity is more likely to drive a stock return jump than news for the average firm, for larger firms there is a higher likelihood that the stock return jump is driven by overnight information shocks.[13]

Additionally we re-estimate the logistic regression in equation 3.3 by taking into account overnight stock microstructural effects such as overnight stock quoted spreads and overnight trade volume (which proxies for overnight stock order imbalance). The results are reported in Tables 7 for all firms.[14]

INSERT TABLE 7 HERE

3.4 Day and Night Jump Probability and Jump Size Results

When we separated the daily stock return into day and night trading period returns in Section 3.3 there was a clear differential impact of the number of news articles in the night trading period versus the number in the daytime period in explaining the likelihood of a daily stock return jump. Motivated by those findings, we use the day and night trading period jumps based on the jump identification in Section 2.2.

In the spirit of understanding the impact of news and stock microstructural effects separated into day and night trading period returns, we separate the logistic regression model specification of equation 3.1 into day and night trading period estimation periods in equations 3.5 and 3.6.[15]

[13] Table A.6 reports the results for the 500 largest firms. The results are qualitatively the same.

[14] Table A.7 reports the results for the 500 largest forms. We display the results separately from Table 5, A.5 as the sample period of $[2000, 2014]$ where WRDS MTAQ has stock quoted spreads and after hours trading volume is available.

[15] For stock illiquidity in the overnight trading period we use relative quoted spreads (RQS_{Night}) and for stock demand pressure we use overnight stock trade volume ($\log(VOL_{Night})$).

$$\text{logit}(p_{D,it}) = b_0 + b_1 \cdot \text{Num News}_{Day} + b_2 \cdot |\text{Avg. Sentiment}_{Day}| \\ + b_3 \cdot |ret_{D,i,t-1}| + b_4 \cdot RES + b_5 \cdot OIB + \epsilon_{D,i,t} \quad (3.5)$$

and

$$\text{logit}(p_{N,it}) = b_0 + b_1 \cdot \text{Num News}_{Night} + b_2 \cdot |\text{Avg. Sentiment}_{Night}| \\ + b_3 \cdot |ret_{N,i,t-1}| + b_4 \cdot RQS_{Night} + b_5 \cdot \log(VOL_{Night}) + \epsilon_{N,i,t} \quad (3.6)$$

The dependent variable ($\{J_{D,99}\}$) is the same for each regression where $\{J_{D,99}\}$ identifies a jump period if the absolute value of daytime period return is above $\{5.28\}$ times the time-varying period spot volatility. The same applies for the night trading period return ($\{J_{N,99}\}$). Results of logistic regression estimation for equations 3.5 and 3.6 for day and night trading periods are presented in Tables 8 (A.8) and 9 (A.9) for all stocks (largest 500 stocks).

INSERT TABLES 8 and 9 HERE

3.5 Interaction of News and illiquidity Jump Probability and Jump Size Results

Our analysis so far has largely considered the effects of news count, news sentiment, stock illiquidity, and stock order flow separately and has not looked at the interaction of any of the effects in tandem. Table 10 shows the fully specified probit model framework from Section 3.1 augmented (separately in each column) with interactions of news count with stock illiquidity and then stock order imbalance, as well as correspondingly average news sentiment with stock illiquidity and then stock order imbalance.

INSERT TABLE 10 HERE

For all firms in the cross section in Table 10, we find negative statistically significant interactions of news count with stock illiquidity (with odds ratio of 0.93), news count with stock order imbalance (with odds ratio of 0.99), and average news sentiment with stock illiquidity and then stock order imbalance (with odds ratio of 0.94). In general the magnitude of the decrease in the odds ratio of each of the three interaction variables is very small relative to the increase in odds ratio of the stock illiquidity, order imbalance, news count, or average news sentiment. For the largest 500 firms in the cross section in Table A.10, we only find a small positive statistically significant interaction of average news sentiment with stock order imbalance (with odds ratio of 1.02).

Taken together, our findings suggest that the interaction between news and stock microstructural effects has a small economical impact on stock return jumps.

3.6 Abnormal Returns around Jumps by News Count, Illiquidity and Order Flow

What drives the idiosyncratic jump risk? Kapadia and Zekhnini (2019) and Bégin, Dorion, and Gauthier (2020) highlight the importance of idiosyncratic stock return jump risk, however, they do not characterize what economic mechanisms drive this idiosyncratic stock return jump risk.

Following Kapadia and Zekhnini (2019), we compute abnormal returns around the stock return jump days (identified using the $\{J_{99}\}$ test statistic) across our different variables of interest (total news count, night news count, etc.) in order to compare the jump size when associated with within firm variables of interest. Each firm-day is classified as top 10%, 10%–25%, 25%–50%, or bottom 50% using within-firm variable of interest observations. For example, a firm-day is classified as Top 10% if the number of news articles arriving on that day for a particular firm falls in the top 10 percentile of the news count distribution in our sample for that firm.[16] In our framework, abnormal returns are computed by adjusting for

[16] Our results in Section 3.2, when we present the estimates of equation 3.2, should not be confused with

the Fama and French (1993) and Cahart (1997) factor model as per Kapadia and Zekhnini (2019). We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day.

Tables 11–15 are sorted by total news count, night news count, day news count, stock illiquidity and order imbalances respectively, across all firms. Additionally in each table the lower panel recomputes the results excluding $[-3, +3]$ days around earnings announcement days.

INSERT TABLES 11–15 HERE

For the cross section of all stocks, when comparing the size of the stock return jump alpha of largest (Top 10%) across each of the variable sorted categories for positive (negative) jumps we find an alpha of 18.13% (-14.03%) for positive (negative) jumps when sorting on the stock illiquidity, 17.23% (-14.24%) for positive (negative) jumps when sorting on the total number of news, and 12.70% (-8.28%) for positive (negative) jumps when sorting on the stock order imbalances. A higher positive alpha is noted for stock illiquidity and a larger negative alpha is noted for total number of news. However, when decomposing the total news count in the day and night trading periods and sorting on the number of day news we see a 17.62% (-15.00%) alpha for positive (negative) jumps and an alpha of 15.75% (-13.97%) for positive (negative) jumps when sorting on the number of night news. When excluding $[-3, +3]$ days around earnings announcement days, we find similar sizes of alpha of 18.68% (-13.78%) for positive (negative) jumps when sorting on the stock illiquidity, 12.83% (-7.61%) for positive (negative) jumps when sorting on the stock order imbalances, and 20.87% (-15.98%) for positive (negative) jumps when sorting on the total number of news. Decomposing the total news count in the day and night trading periods and sorting yields 19.67% (-15.95%) for positive (negative) jumps when sorting on the number of day news, 19.05% (-15.97%) for positive (negative) jumps when sorting on the number of night news. In general we find the

estimation of idiosyncratic stock return jump risk as the results in Section 3.2 are not idiosyncratic jump risk estimates since they are not adjusted for systematic risks.

largest alphas for both positive and negative jumps when sorting on the total number of news articles or the number of day news.

Correspondingly, Tables A.11–A.15 are sorted by total news count, night news count, day news count, stock illiquidity and order imbalances respectively, across the largest 500 firms. Additionally in each table the lower panel recomputes the results excluding $[-3, +3]$ days around earnings announcement days.

For the sample of the largest 500 stocks, when comparing the size of the stock return jump alpha of largest (Top 10%) across each of the variable sorted categories for positive (negative) jumps, we find an alpha of 9.43% (-9.84%) for positive (negative) jumps when sorting on the number of day news, 8.49% (-8.83%) for positive (negative) jumps when sorting on the number of night news, 8.98% (-9.17%) for positive (negative) jumps when sorting on the stock illiquidity, 8.64% (-9.01%) for positive (negative) jumps when sorting on the total number of news, and 6.58% (-5.35%) for positive (negative) jumps when sorting on the stock order imbalances. Similarly when excluding $[-3, +3]$ days around earnings announcement days we also find the largest alphas for both positive and negative jumps when sorting on number of day news.

Our results also show a higher alpha for both positive and negative jumps associated with a higher number of day news articles belonging to the Top 10% news category, suggesting that higher number of news articles during the day lead to substantially higher idiosyncratic returns, particularly for larger firms.

4 Robustness Section

4.1 Removal of Earnings Announcement Days

We follow Huang, Tan, and Wermers (2019) and remove $[-3, +3]$ days around earnings announcement days. Table 16 re-estimated the logistic regression results of Tables 3 excluding $[-3, +3]$ days around earnings announcement days from 2000 to 2020 for all stocks and largest 500 firms respectively.

INSERT TABLES 16 and 17 HERE

Table 17 re-estimated the logistic regression results of Tables 5, with night and day news count and average sentiment separated, excluding $[-3, +3]$ days around earnings announcement days from 2000 to 2020 for all stocks.[17]

The results in Tables 16 and 17 serve as a robustness test to support the finding that stock return jump probability is positively related with stock illiquidity (RES) and order imbalance shocks (OIB) both for all stocks and largest 500 firms. Although stock illiquidity and order imbalances are important for large firm stock return jumps, news counts are associated with the most likely jump source (higher standardized odds ratio for news count over stock illiquidity or order flow). As well it also supports the finding that the sensitivity of stock return jump probability to the number of overnight news articles is larger than the sensitivity to the number of daytime news articles, and although illiquidity is more likely to drive a stock return jump than news for the average firm, for larger firms there is a higher likelihood that the stock return jump is driven by overnight information shocks.

4.2 Electronic Market Era Subset Analysis: January 2004 to December 2020

As an additional robustness test, we investigate the impact on our conclusions during the post electronic market trading era beyond January 2004. By 2004 most U.S. stock exchanges had become fully electronic and more fragmented, allowing for more market participants that do not need to be physically present at an exchange to trade. Given the dramatic changes in market structure, as well as changes in liquidity, order flow and how information is processed, we believe additional analysis during this subset is warranted. Tables A.18 and A.19 re-estimated the logistic regression results of tables 3 from 2004 to 2020 for all stocks

[17] Tables A.16 and A.17 report the results for the largest 500 firms.

and largest 500 firms respectively.

Tables A.20 and A.21 re-estimated the regression equation 3.2 of Tables 4 from 2004 to 2020 for all stocks and largest 500 firms respectively.

Probit tables for both all stocks and largest 500 firms show that the sensitivity of stock return jump probability to illiquidity has increased from the $[2000, 2020]$ sample in the both the period $[2004, 2020]$ and $[2004, 2020]$ with crises removed. This shows that stock return jumps have become more sensitive to illiquidity in a more liquid market (as the post 2004 trading period is when the stock markets were largely fully electronic). Despite the fact that sensitivity of stock return jump probability associated with illiquidity has increased, we can see from both the positive and negative jump expected return tables, for both sample periods, that the expected return given either a positive or negative jump has decreased from the $[2000, 2020]$ sample in both the periods $[2004, 2020]$ and $[2004, 2020]$ with crises removed. This shows that stock return jump size has decreased despite the fact that the sensitivity has increased.[18]

5 Conclusion

Stock prices exhibit large, discrete changes, typically labelled as "jumps". In this paper, we investigate potential sources of jumps in individual stock returns, focusing on differences during the daytime versus overnight trading periods. We compile time-stamped intraday data on the frequency and content of information flows, as well as illiquidity and order-imbalance measures. We then investigate the relationship between those variables and jumps in stock returns.

In particular, we analyze the impact of firm-level (novel) news frequency and content, stock illiquidity, and order imbalances on individual stock return jumps. We analyze these

[18] In additional unreported results we have probits and jump size results robust to the 2004 to 2022 sub-sample with 2008, 2009, 2020 removed (crisis years) as well as the results robust to removing stocks with price less than 5 dollars (that is, our results are not sensitive to small stocks) which are not presented in the paper but are available upon request.

relationships for the entire day as well as for the daytime and overnight trading periods. Our results show that the information flow and trading frictions are significantly related to non-parametric measures of jump intensity and jump-size distributions and explain an important fraction of variations in the jumps across individual firm stock returns. The differential effects associated with the daytime versus the overnight periods are informative.

We also provide results from a comprehensive range of robustness analyses (for example, sub-sample analyses and removal of earnings announcements). These tests reveal that the effects of news flows, stock illiquidity, and order imbalance on stock return jumps, and their differential effects over time and across companies, are robust.

Our analyses of the relationships between intraday news flows, as well as trading friction measures, with daily stock return jumps for a large panel of companies could enrich the economic content of models for stock return dynamics which typically have treated the sources of jumps as latent. Explicitly incorporating news processes and liquidity measures in models of stock return jumps can potentially help identify jumps due to information arrival as opposed to jumps due to other reasons, such as liquidity or strategic trading based on private information.

References

Amihud, Y., and H. Mendelson. 1986. Asset Pricing and the Bid-Ask Spread. *Journal of Financial Economics* 17:223–49.

Antweiler, W., and M. Z. Frank. 2004. Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance* 59:1259–94.

Baker, S. R., N. Bloom, and J. Davis. 2016. Measuring economic policy uncertainty. *Quarterly Journal of Economics* 131:1593–636.

Baker, S. R., N. Bloom, J. Davis, and K. Kost. 2021a. Public news and stock market volatility. *NBER Working Paper* 1:1–76.

Baker, S. R., N. Bloom, J. Davis, and M. C. Sammon. 2021b. What Triggers Stock Market Jumps? *NBER Working Paper* 1:1–76.

Barclay, M. J., and T. Hendershott. 2003. Price Discovery and Trading After Hours. *Review of Financial Studies* 16:1041–73.

———. 2004. Liquidity Externalities and Adverse Selection: Evidence from Trading after Hours. *Journal of Finance* 59:681–710.

Bégin, J.-F., C. Dorion, and G. Gauthier. 2020. Idiosyncratic Jump Risk Matters: Evidence from Equity Returns and Options. *Review of Financial Studies* 33:155–211.

Berkman, H., P. D. Koch, L. Tuttle, and Y. J. Zhang. 2012. Paying Attention: Overnight Returns and the Hidden Cost of Buying at the Open. *Journal of Financial and Quantitative Analysis* 47:715–44.

Berry, T. D., and K. M. Howe. 1994. Public Information Arrival. *Journal of Finance* 49:1331–46.

Bollerslev, T., J. Li, and Y. Xue. 2018. Volume, Volatility, and Public News Announcements. *Review of Economic Studies* 85:2005–41.

Bollerslev, T., S. Z. Li, and V. Todorov. 2016. Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns. *Journal of Financial Economics* 120:464–90.

Bongaerts, D., R. Roll, R. Rosch, M. Van Dijck, and D. Yuferova. 2022. How Do Shocks Arise and Spread Across Stock Markets? A Microstructure Perspective. *Management Science* 68:3071–89.

Boudoukh, J., R. Feldman, S. Kogan, and M. Richardson. 2019. Information, Trading, and Volatility: Evidence from Firm-Specific News. *Review of Financial Studies* 32:992–1033.

Brogaard, J., A. Carrion, T. Moyaert, R. Riordan, A. Shkilko, and K. Sokolov. 2018. High frequency trading and extreme price movements. *Journal of Financial Economics* 128:253–65.

Brogaard, J., T. H. Nguyen, T. J. Putnins, and E. Wu. 2022. What Moves Stock Prices? The Roles of News, Noise, and Information. *Review of Financial Studies* 35:4341–86.

Bybee, L., B. T. Kelly, A. Manela, and Y. Su. 2022. Narrative Asset Pricing: Interpretable Systematic Risk Factors from News Text. *Yale University Working Paper* 1:1–59.

Bybee, L., B. T. Kelly, A. Manela, and D. Xiu. 2021. Business News and Business Cycles. *Yale University Working Paper* 1:1–65.

Cahart, M. 1997. On Persistence in Mutual Fund Performance. *Journal of Finance* 52:57–82.

Chan, W. S. 2003. Stock price reaction to news and no-news: drift and reversal after headlines. *Journal of Financial Economics* 70:223–60.

Chordia, T., and A. Subrahmanyam. 2004. Order Imbalance and Individual Stock Returns: Theory and Evidence. *Journal of Financial Economics* 72:485–518.

Christoffersen, P., B. Feunou, Y. Jeon, and C. Orthanalai. 2021. Time-Varying Crash Risk: The Role of Stock Market Liquidity. *Review of Finance* 25:1261–98.

Cliff, M., M. J. Cooper, and H. Gulen. 2008. Return Differences between Trading and Non-trading Hours: Like Night and Day. *Working Paper* 1:1–48.

Cui, B., and A. E. Gozluklu. 2021. News and Trading After Hours. *SSRN Working Paper* 1:1–70.

Cutler, D. M., J. M. Poterba, and L. H. Summers. 1989. What moves stock prices? *Journal of Portfolio Management* 15:4–12.

Fama, E. F., and K. R. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33:3–56.

Fedyk, A. 2022. Front Page News: The Effect of News Positioning on Financial Markets. *Journal of Finance (FORTHCOMING)* 0:1–66.

Fedyk, A., and J. Hodson. 2021. When Can the Market Identify Old News? *Hass Berkely Working Paper* 1:1–55.

Gentzkow, M., B. T. Kelly, and M. Taddy. 2019. Text As Data. *Journal of Economic Literature* 57:535–74.

Gilder, D., M. B. Shackleton, and S. J. Taylor. 2014. Cojumps in stock prices: Empirical Evidence. *Journal of Banking and Finance* 40:443–58.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka. 2009. Do liquidity measures measure liquidity? *Journal of Financial Economics* 92:153–81.

Gurkaynak, R. S., B. Kisacikoglu, and J. H. Wright. 2020. Missing events in event studies: identifying the effects of partially measured news surprises. *American Economic Review* 110:3871–912.

Hendershott, T., D. Livdan, and D. Rosch. 2020. Asset pricing: A tale of night and day. *Journal of Financial Economics* 138:635–62.

Herskovic, B., B. Kelly, H. Lustig, and S. Van Nieuwerburgh. 2016. Stock returns after major price shocks: The impact of information. *Journal of Financial Economics* 119:249–83.

Hong, H., and J. Wang. 2000. Trading Returns Under Periodic Market Closures. *Journal of Finance* 60:297–354.

Huang, A. G., H. Tan, and R. Wermers. 2019. Institutional trading around corporate news: evidence from textual analysis. *Review of Financial Studies* 33:4627–75.

Jeon, Y., T. H. McCurdy, and X. Zhao. 2021. News as sources of jumps in stock returns: Evidence from 26 million news articles. *Journal of Financial Economics* 0:0–.

Jiang, G. J., I. Lo, and A. Verdelhan. 2011. Information Shocks, Liquidity Shocks, Jumps, and Price Discovery: Evidence from the U.S. Treasury Market. *Journal of Financial and Quantitative Analysis* 46:527–51.

Jiang, G. J., and K. X. Zhu. 2017. Information Shocks and Short-Term Market Underreaction. *Journal of Financial Economics* 124:43–64.

Kapadia, N., and M. Zekhnini. 2019. Do idiosyncratic jumps matter? *Journal of Financial Economics* 131:666–92.

Ke, T., and D. Kelly, B. T. Xiu. 2022. Predicting Returns with Text Data. *Yale University Working Paper* 1:1–66.

Kelly, B. T., A. Manela, and A. Moreira. 2021. Text Selection. *Journal of Business Economics and Statistics* 39:859–79.

Khan, S. A., and R. Riordan. 2020. Intraday Jump Dynamics: What Predicts Price Jumps? *Working Paper* 1:1–48.

Kirilenko, A., A. S. Kyle, S. Samadi, and T. Tuzun. 2017. The Flash Crash: High-Frequency Trading in an Electronic Market. *Journal of Finance* 72:967–98.

Lee, S. S. 2012. Jumps and Information Flow in Financial Markets. *Review of Financial Studies* 25:439–79.

Lee, S. S., and P. A. Mykland. 2008. Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. *Review of Financial Studies* 21:2535–63.

Loughran, T., and B. McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance* 66:35–65.

Manela, A., and A. Moreira. 2017. News implied volatility and disaster concerns. *Journal of Financial Economics* 123:137–62.

Menkveld, A. J., and B. Z. Yueshen. 2018. The flash crash: A cautionary tale about highly fragmented markets. *Management Science* 65:449–70.

Mitchell, M. L., and J. H. Mulherin. 1994. The Impact of Public Information on the Stock Market. *Journal of Finance* 49:923–50.

Neuhierl, A., A. Scherbina, and B. Schiusene. 2008. Market reaction to corporate press releases. *Journal of Financial and Quantitative Analysis* 48:1207–40.

O'Hara, M., and M. Ye. 2011. Is market fragmentation harming market quality? *Journal of Financial Economics* 100:459–74.

Pelger, M. 2020. Understanding Systematic Risk: A High-Frequency Approach. *Journal of Finance* 75:2179–220.

Savor, P. G. 2012. Stock returns after major price shocks: The impact of information. *Journal of Financial Economics* 106:635–59.

Tetlock, P. C. 2007a. Does public financial news resolve asymmetric information? *Review of Financial Studies* 23:1437–67.

———. 2007b. Giving content to investor sentiment: the role of media in the stock market. *Journal of Finance* 62:1139–68.

———. 2011. All the news that's fit to reprint: do investors react to stale information? *Review of Financial Studies* 24:1481–512.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy. 2008. More than words: quantifying language to measure firms' fundamentals. *Journal of Finance* 63:1437–67.

van Binsbergen, J. H., S. Bryzgalova, M. Mukhopadhyay, and V. Sharma. 2022. (Almost) 200 Years of News-Based Economic Sentiment. *London Business School Working Paper* 1:1–48.

Weller, B. M. 2019. Measuring Tail Risks at High Frequency. *Review of Financial Studies* 32:3571–616.

Wong, L., and J. W. Yang. 2019. The timing of information arrival and overnight returns. *Working Paper* 1:1–35.

Zhao, X. 2017. Does information intensity matter for stock returns? Evidence from Form 8-K filings. *Management Science* 63:1382–404.

Table 1 Summary Statistics

	All Stocks		
Variable Name	N obs	Mean	Std. Dev.
Num News Night	18,052,448	0.23	1.13
Num News Day	18,052,448	0.12	0.56
Avg. Sentiment Night	18,052,448	0.24	1.36
Avg. Sentiment Day	18,052,448	0.11	0.52
Relative Effective Spreads (RES)	21,721,722	0.02	0.13
Order Imbalances (OIB)	21,700,054	−0.02	0.37
Num News	23,865,089	0.37	1.6
Avg. Sentiment	23,865,089	1.3	4.23
	Top 500 Firms		
Variable Name	N obs	Mean	Std. Dev.
Num News Night	2,094,233	0.54	2.08
Num News Day	2,094,233	0.43	1.11
Avg. Sentiment Night	2,094,233	0.46	1.76
Avg. Sentiment Day	2,094,233	0.33	0.9
Relative Effective Spreads (RES)	2,151,016	0.01	0.06
Order Imbalances (OIB)	2,151,010	0.03	0.18
Num News	2,242,279	1.21	3.54
Avg. Sentiment	2,242,279	3.44	6.26

This table reports the summary statistics for main variables used in the analysis. We report number of news and average sentiment score from the RavenPack database for all stocks and top 500 firms. Stock illiquidity (RES) and order imbalances (OIB) are computed using the MTAQ and DTAQ database as outlined in Section 2.1.

Table 2 Jump Summary Statistics

	All Stocks			
Variable Name	Sum obs	N obs	Mean	Std. Dev.
J_{99}	$298,110$	$23,865,089$	0.01	0.11
J_{95}	$432,890$	$23,865,089$	0.02	0.13
$J_{D,99}$	$384,523$	$15,762,893$	0.02	0.15
$J_{N,99}$	$95,224$	$15,756,908$	0.01	0.08
	Top 500			
Variable Name	Sum obs	N obs	Mean	Std. Dev.
J_{99}	$20,103$	$2,242,279$	0.01	0.09
J_{95}	$30,366$	$2,242,279$	0.01	0.12
$J_{D,99}$	$34,567$	$2,087,123$	0.02	0.13
$J_{N,99}$	$13,616$	$2,086,688$	0.01	0.08

We report the summary statistics of jump indicators for all stocks and top 500 firms. For each stock trading day $\{J_{99}\}$ identifies a jump day if the absolute value of the daily stock return is above $\{5.1024\}$ times the time-varying daily spot volatility. For each stock $J_{D,99}$ ($J_{N,99}$) identifies a jump trading period if the absolute value of day (night) return is above $\{5.28\}$ times the time-varying day (night) spot volatility.

Table 3 Effects of news measures and stock illiquidity on probability of daily jumps

Panel A: Probit Coefficient Estimates					
Variable	(1)	(2)	(3)	(4)	(5)
Num News	0.3	0.31	0.31	0.31	0.32
	(140.44)	(143.06)	(143.52)	(142.95)	(143.43)
Avg. Sentiment	0.08	0.09	0.09	0.09	0.09
	(26.8)	(33.35)	(33.74)	(33.12)	(33.51)
$\|ret_{t-1}\|$	0.1	0.09	0.08	0.09	0.08
	(70.3)	(57.33)	(47.43)	(57.15)	(47.06)
RES_s			0.39		0.4
			(67.03)		(67.33)
OIB_s				0.2	0.2
				(50.24)	(51.18)
R^2	4.95	5.41	5.66	5.62	5.88
N obs	23, 808, 488	20, 601, 951	20, 601, 951	20, 601, 951	20, 601, 951

Panel B: Odds Ratio					
Variable	(1)	(2)	(3)	(4)	(5)
Num News	1.36	1.37	1.37	1.37	1.37
Avg. Sentiment	1.08	1.1	1.1	1.1	1.1
$\|ret_{t-1}\|$	1.11	1.1	1.08	1.1	1.08
RES_s			1.48		1.49
OIB_s				1.22	1.22

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table 4 Effects of news measures and stock illiquidity on daily jump size

	Panel A: All Jumps				
Variable	(1)	(2)	(3)	(4)	(5)
Num News	$-3.7e-3$	$-4.2e-3$	$-3.7e-3$	$-2.8e-3$	$-2.7e-3$
	(-9.57)	(-11.62)	(-9.97)	(-7.78)	(-7.47)
Avg. Sentiment	0.01	0.01	0.01	0.01	0.01
	(22.38)	(26.9)	(27.62)	(26.12)	(26.28)
ret_{t-1}	-0.09	-0.05	0.03	0.01	0.01
	(-4.09)	(-2.29)	(1.59)	(0.41)	(0.53)
RES_s			$4.5e-3$		0.01
			(3.47)		(6.22)
OIB_s				0.07	0.07
				(106.1)	(105.95)
R^2	0.54	0.59	0.68	10.02	10.05
N obs	297,172	275,051	260,155	260,155	260,155

	Panel B: Positive Jumps				
Variable	(1)	(2)	(3)	(4)	(5)
Num News	0.01	0.01	0.01	0.01	0.01
	(9.86)	(9.58)	(11.66)	(10.47)	(11.75)
Avg. Sentiment	0.01	0.01	0.01	0.01	0.01
	(9.75)	(12.33)	(13.5)	(12.79)	(13.63)
ret_{t-1}	-0.09	-0.06	0.1	0.07	0.1
	(-2.26)	(-1.58)	(4.52)	(3.12)	(4.48)
RES_s			0.06		0.06
			(33.02)		(32.82)
OIB_s				0.01	$3.7e-3$
				(9.04)	(5.04)
R^2	1.28	1.35	3.67	1.95	3.69
N obs	169,976	158,840	150,160	150,160	150,160

	Panel C: Negative Jumps				
Variable	(1)	(2)	(3)	(4)	(5)
Num News	-0.01	-0.01	-0.01	-0.01	-0.01
	(-42.94)	(-43)	(-47.25)	(-40.17)	(-43.36)
Avg. Sentiment	$-4.6e-3$	-0.01	-0.01	$-5e-3$	-0.01
	(-15.13)	(-18.22)	(-18.12)	(-17.56)	(-17.85)
ret_{t-1}	-0.08	-0.06	-0.06	-0.06	-0.06
	(-8.63)	(-5.88)	(-6.44)	(-5.79)	(-6.06)
RES_s			-0.04		-0.05
			(-47.35)		(-50.87)
OIB_s				-0.01	-0.01
				(-21.8)	(-31.73)
R^2	7.67	7.77	12.99	8.81	14.18
N obs	127,196	116,211	109,995	109,995	109,995

This table reports results from regressions of daily jump sizes, conditional on the jump indicator being 1, on daily news measures as well as stock illiquidity and order imbalances from TAQ for the all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panels B and C report the results for positive jump sizes and negative jump sizes, respectively. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table 5 Effects of day and night news measures and stock illiquidity on probability of daily jumps

Panel A: Probit Coefficient Estimates						
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	0.36		0.34	0.34	0.34	0.34
	(165.23)		(165.95)	(166.34)	(166.17)	(166.55)
Num News Day$_s$		0.16	0.13	0.14	0.13	0.14
		(100.28)	(92.77)	(93.3)	(92.32)	(92.86)
Avg. Sentiment Night$_s$	-0.01		-0.02	-0.01	-0.02	-0.01
	(-5.94)		(-7.41)	(-6.65)	(-7.48)	(-6.71)
Avg. Sentiment Day$_s$		0.1	0.07	0.07	0.07	0.07
		(78.14)	(60.01)	(60.94)	(59.64)	(60.57)
$\|ret_{t-1}\|$	0.1	0.1	0.09	0.08	0.09	0.08
	(56.44)	(58.05)	(53.42)	(42.97)	(53.14)	(42.45)
RES_s				0.4		0.4
				(61.23)		(61.3)
OIB_s					0.25	0.25
					(64.99)	(65.75)
R^2	6.72	2.18	7.91	8.17	8.24	8.5
N obs	16,622,936	16,614,922	16,609,504	16,609,504	16,609,504	16,609,504

Panel B: Odds Ratio						
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	1.43		1.4	1.4	1.4	1.41
Num News Day$_s$		1.17	1.14	1.15	1.14	1.14
Avg. Sentiment Night$_s$	0.99		0.98	0.99	0.98	0.99
Avg. Sentiment Day$_s$		1.11	1.08	1.08	1.08	1.08
$\|ret_{t-1}\|$	1.1	1.1	1.1	1.08	1.1	1.08
RES_s				1.49		1.49
OIB_s					1.28	1.28

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table 6 Effects of day and night news measures and stock illiquidity on daily jump size.

Panel A: All Jumps						
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night_s	$-4.8e-3$		-0.01	-0.01	$-4.1e-3$	$-4e-3$
	(-24.21)		(-23.41)	(-23.03)	(-19.84)	(-19.31)
Num News Day_s		$0.1e-3$	$0.9e-3$	$0.9e-3$	$0.9e-3$	$0.9e-3$
		(0.25)	(1.51)	(1.52)	(1.55)	(1.56)
Avg. Sentiment Night_s	0.01		0.01	0.01	0.01	0.01
	(11.24)		(10.95)	(10.97)	(11.85)	(11.89)
Avg. Sentiment Day_s		$0.2e-3$	$1e-3$	$1e-3$	$1e-3$	$1.1e-3$
		(0.65)	(2.55)	(2.59)	(2.83)	(2.93)
ret_{t-1}	0.02	0.02	0.02	0.02	0.01	0.01
	(1.06)	(1.09)	(1.05)	(1.08)	(0.32)	(0.4)
RES_s				$2.4e-3$		0.01
				(1.7)		(4.1)
OIB_s					0.08	0.08
					(98.18)	(98.09)
R^2	0.46	0.01	0.5	0.5	9.3	9.32
N obs	218,914	218,914	218,914	218,914	218,914	218,914

Panel B: Positive Jumps						
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night_s	$1.2e-3$		$0e-3$	$0.8e-3$	$0.2e-3$	$0.9e-3$
	(5.27)		(-0.14)	(2.94)	(0.89)	(3.4)
Num News Day_s		0.01	0.01	0.01	0.01	0.01
		(10.58)	(10.17)	(10.49)	(10.24)	(10.52)
Avg. Sentiment Night_s	0.01		0.01	0.01	0.01	0.01
	(12.63)		(11.48)	(12.35)	(11.53)	(12.37)
Avg. Sentiment Day_s		$0.6e-3$	$0.3e-3$	$0.6e-3$	$0.4e-3$	$0.7e-3$
		(1.37)	(0.69)	(1.4)	(0.79)	(1.44)
ret_{t-1}	0.08	0.08	0.08	0.1	0.08	0.1
	(3.25)	(3.12)	(3.12)	(4.37)	(3.09)	(4.35)
RES_s				0.06		0.06
				(29.92)		(29.81)
OIB_s					0.01	$2.5e-3$
					(7.38)	(3.28)
R^2	0.44	1.47	1.65	3.35	1.69	3.36
N obs	128,979	128,979	128,979	128,979	128,979	128,979

Panel C: Negative Jumps						
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night_s	-0.01		$-3.9e-3$	$-4.6e-3$	$-3.6e-3$	$-4.2e-3$
	(-37.54)		(-27.55)	(-32.99)	(-25.8)	(-29.94)
Num News Day_s		-0.01	-0.01	-0.01	-0.01	-0.01
		(-30.2)	(-26.1)	(-26.63)	(-25.72)	(-25.91)
Avg. Sentiment Night_s	$-4.5e-3$		$-4e-3$	$-3.8e-3$	$-4e-3$	$-3.7e-3$
	(-10.67)		(-10.01)	(-9.62)	(-9.99)	(-9.56)
Avg. Sentiment Day_s		$-2.3e-3$	$-1.5e-3$	$-1.7e-3$	$-1.4e-3$	$-1.7e-3$
		(-9.24)	(-5.77)	(-7.24)	(-5.57)	(-6.9)
ret_{t-1}	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06
	(-5.84)	(-6.37)	(-6.17)	(-6.39)	(-6.08)	(-6.22)
RES_s				-0.05		-0.05
				(-45.55)		(-47.8)
OIB_s					-0.01	-0.01
					(-10.91)	(-21.4)
R^2	5.11	6.31	8.84	14.07	8.98	14.62
N obs	89,935	89,935	89,935	89,935	89,935	89,935

This table reports results from regressions of daily jump sizes, conditional on the jump indicator being 1, on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panels B and C report the results for positive jump sizes and negative jump sizes, respectively. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table 7 Effects of news measures (day and night) and stock illiquidity (day and night) on probability of daily jumps (All Firms)

	Panel A: Probit Coefficient Estimates			
Variable	(1)	(2)	(3)	(4)
Num News Night$_s$	0.34	0.34	0.36	0.36
	(163.85)	(163.84)	(172.87)	(173.27)
Num News Day$_s$	0.14	0.14	0.15	0.15
	(90.71)	(90.71)	(89.36)	(89.07)
Avg. Sentiment Night$_s$	$2.2e-3$	$2.2e-3$	0.01	0.02
	(0.96)	(0.96)	(6.37)	(7.51)
Avg. Sentiment Day$_s$	0.08	0.08	0.08	0.08
	(59.47)	(59.47)	(58.06)	(58.03)
$\|ret_{t-1}\|$	0.08	0.08	0.08	0.06
	(39.38)	(39.42)	(30.48)	(23.5)
$RQS_{A,s}$		$-4.4e-3$		-0.01
		(-1.49)		(-4.03)
log(DollTradingVolume$_{A,s}$)			0.1	0.14
			(22.22)	(30.76)
RES_s				0.4
				(46.53)
OIB_s				0.27
				(50.21)
R^2	8.24	8.24	10.3	10.85
N obs	12, 397, 112	12, 397, 112	10, 187, 859	10, 187, 859

	Panel B: Odds Ratio			
Variable	(1)	(2)	(3)	(4)
Num News Night$_s$	1.4	1.4	1.43	1.43
Num News Day$_s$	1.15	1.15	1.16	1.16
Avg. Sentiment Night$_s$	1	1	1.01	1.02
Avg. Sentiment Day$_s$	1.08	1.08	1.08	1.08
$\|ret_{t-1}\|$	1.09	1.09	1.08	1.06
$RQS_{A,s}$		1		0.99
log(DollTradingVolume$_{A,s}$)			1.11	1.15
RES_s				1.5
OIB_s				1.31

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for the largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. Panels B and C report the results for positive jump sizes and negative jump sizes, respectively. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table 8 Effects of day news measures and day stock illiquidity on probability of daytime period jumps (All Firms)

Panel A: Probit Coefficient Estimates				
Variable	(1)	(2)	(3)	(4)
Num News Day_s	0.11	0.11	0.11	0.11
	(82.72)	(83.56)	(82.56)	(83.4)
Avg. Sentiment Day_s	0.06	0.06	0.06	0.06
	(51.06)	(51.94)	(50.92)	(51.8)
$\lvert ret_{D,t-1} \rvert$	0.09	0.08	0.09	0.08
	(62.76)	(53.16)	(62.46)	(52.75)
RES_s		0.3		0.3
		(44.02)		(44.89)
OIB_s			0.12	0.12
			(42.45)	(43.58)
R^2	0.67	0.82	0.76	0.91
N obs	14,938,325	14,938,325	14,938,325	14,938,325

Panel B: Odds Ratio				
Variable	(1)	(2)	(3)	(4)
Num News Day_s	1.11	1.11	1.11	1.11
Avg. Sentiment Day_s	1.06	1.06	1.06	1.06
$\lvert ret_{D,t-1} \rvert$	1.1	1.08	1.09	1.08
RES_s		1.35		1.35
OIB_s			1.13	1.13

This table reports results from pooled logistic regressions of daytime period jump indicators, defined using Lee and Mykland (2008), on day news measures as well as day stock illiquidity and order imbalances from TAQ for all firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the number of news articles reported on the Ravenpack database within the trading day, the absolute value of news tone of within the trading day news, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The jump variable ($\{J_{D,99}\}$) is the same for each regression where $\{J_{D,99}\}$ identifies a jump in the daytime period if the absolute value of the daytime period return is above $\{5.28\}$ times the time-varying daytime period spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table 9 Effects of night news measures and night stock illiquidity on probability of night trading period jumps (All Firms)

	Panel A: Probit Coefficient Estimates				
Variable	(1)	(2)	(3)	(4)	(5)
Num News Night$_s$	0.48	0.48	0.48	0.48	
	(190.06)	(190.07)	(189.98)	(189.99)	
Avg. Sentiment Night$_s$	0.09	0.09	0.09	0.09	
	(36.56)	(36.63)	(36.32)	(36.36)	
$\|ret_{N,t-1}\|$	-0.18	-0.18	-0.17	-0.17	
	(-20.42)	(-20.02)	(-20.05)	(-19.86)	
$RQS_{A,s}$		-0.06		-0.04	
		(-10.25)		(-5.93)	
log(DollTradingVolume$_{A,s}$)			0.3	0.3	
			(47.29)	(46.56)	
R^2	19.58	19.61	20.05	20.06	
N obs	$9,835,365$	$9,835,365$	$9,835,365$	$9,835,365$	

	Panel B: Odds Ratio				
Variable	(1)	(2)	(3)	(4)	(5)
Num News Night$_s$	1.62	1.62	1.62	1.62	
Avg. Sentiment Night$_s$	1.09	1.09	1.09	1.09	
$\|ret_{N,t-1}\|$	0.83	0.84	0.84	0.84	
$RQS_{A,s}$		0.94		0.96	
log(DollTradingVolume$_{A,s}$)			1.35	1.35	

This table reports results from pooled logistic regressions of night trading period jump indicators, defined using Lee and Mykland (2008), on night news measures as well as night stock illiquidity and log night trading period dollar volume from TAQ for all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of night news articles reported on the Ravenpack database in the night trading period, the absolute value of news tone of the news articles reported in the night trading period, the absolute value of the previous day's night return, the stock relative quoted spreads in the night trading period, and the log night trading period dollar volume from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2014. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{N,99}\}$) is the same for each regression where $\{J_{N,99}\}$ identifies a night trading period return jump if the absolute value of the night trading period return is above $\{5.28\}$ times the time-varying night trading period return spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table 10 Probability of Daily Jumps with Interactions Jumps (All Firms)

	Panel A: Probit Coefficient Estimates			
Variable	(1)	(2)	(3)	(4)
Num News	0.3	0.32	0.31	0.32
	(135.89)	(143.66)	(143.29)	(143.56)
Avg. Sentiment	0.09	0.09	0.08	0.09
	(33.76)	(33.44)	(29.81)	(32.4)
$\|ret_{t-1}\|$	0.08	0.08	0.08	0.08
	(46.85)	(46.92)	(47.1)	(46.96)
RES_s	0.45	0.4	0.42	0.4
	75.88	(67.27)	71.42	(67.3)
OIB_s	0.2	0.2	0.2	0.2
	51.26	49.74	(51.19)	(48.65)
Num News x RES_s	-0.08			
	-25.98			
Num News x OIB_s		-0.01		
		-7.61		
Avg. Sentiment x RES_s			-0.06	
			(-16.22)	
Avg. Sentiment x OIB_s				$2.9e-3$
				(1.41)
R^2	5.88	5.94	5.9	5.88
N obs	$20,601,951$	$20,601,951$	$20,601,951$	$20,601,951$

	Panel B: Odds Ratio			
Variable	(1)	(2)	(3)	(4)
Num News	1.35	1.37	1.37	1.37
Avg. Sentiment	1.1	1.1	1.09	1.1
$\|ret_{t-1}\|$	1.08	1.08	1.08	1.08
RES_s	1.57	1.49	1.53	1.49
OIB_s	1.22	1.23	1.22	1.22
Num News x RES_s	0.93			
Num News x OIB_s		0.99		
Avg. Sentiment x RES_s			0.94	
Avg. Sentiment x OIB_s				1

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for the largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ, as well as all interacted pairs of variables. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.28\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table 11 Abnormal Returns around Jumps by News Count (All Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.6e-3$	0.03	$0e-3$	65,958	$-0.5e-3$	0.17	$0.4e-3$	36,434	$-0.6e-3$	-0.14	$-0.4e-3$	29,501
	(-15.49)	(32.89)	(0.93)		(-9.92)	(139.4)	(6.06)		(-12.34)	(-220.00)	(-5.75)	
10% - 25%	$-0.6e-3$	0.03	$0.1e-3$	17,220	$-0.6e-3$	0.13	$0.3e-3$	10,014	$-0.6e-3$	-0.1	$-0.2e-3$	7,203
	(-8.07)	(24.98)	(1.26)		(-5.66)	(82.31)	(3)		(-6.03)	(-92.42)	(-1.23)	
25% - 50%	$-0.4e-3$	0.01	$0.2e-3$	4,845	$-0.6e-3$	0.08	$0.1e-3$	2,585	$-0.2e-3$	-0.06	$0.2e-3$	2,260
	(-5.36)	(9.72)	(1.6)		(-5.11)	(47.64)	(0.77)		(-2.1)	(-46.68)	(1.47)	
Bottom 50%	$-1.3e-3$	0.05	$0e-3$	130,663	$-1.4e-3$	0.14	$-0.6e-3$	79,798	$-1.1e-3$	-0.09	$1e-3$	50,805
	(-36.02)	(103.71)	(0.91)		(-28.73)	(239.9)	(-9.99)		(-21.84)	(-260.00)	(12.99)	
News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.9e-3$	0.05	$-0.2e-3$	32,643	$-0.6e-3$	0.21	$-0.1e-3$	18,929	$-1.2e-3$	-0.16	$-0.5e-3$	13,639
	(-13.28)	(32.22)	(-2.45)		(-7.2)	(95.84)	(-0.53)		(-12.21)	(-140.00)	(-3.15)	
10% - 25%	$-0.7e-3$	0.04	$0.1e-3$	11,923	$-0.7e-3$	0.12	$0.3e-3$	7,002	$-0.7e-3$	-0.09	$-0.3e-3$	4,921
	(-7.54)	(22.77)	(0.54)		(-5.46)	(62.85)	(2.12)		(-5.35)	(-68.16)	(-1.31)	
25% - 50%	$-0.5e-3$	0.01	$0.2e-3$	3,420	$-0.6e-3$	0.07	$0.1e-3$	1,763	$-0.3e-3$	-0.06	$0.2e-3$	1,654
	(-4.44)	(6.76)	(1.37)		(-3.46)	(33.78)	(0.75)		(-2.83)	(-36.4)	(1.16)	
Bottom 50%	$-1.3e-3$	0.05	$0e-3$	118,294	$-1.4e-3$	0.14	$-0.6e-3$	72,992	$-1.1e-3$	-0.09	$1.1e-3$	45,633
	(-34.5)	(102.16)	(0.44)		(-27.76)	(226.72)	(-10.6)		(-20.58)	(-240.00)	(12.88)	

This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm number of Ravenpack news articles on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3, +3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table 12 Abnormal Returns around Jumps by Night News Count (All Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
Night News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.5e-3$	0.02	$0.2e-3$	63,874	$-0.5e-3$	0.16	$0.7e-3$	34,655	$-0.6e-3$	-0.14	$-0.5e-3$	29,219
	(-16.47)	(25.01)	(3.83)		(-10.76)	(150.7)	(11.82)		(-12.78)	(-230.00)	(-6.83)	
10% - 25%	$-0.8e-3$	0.03	$0.4e-3$	6,324	$-0.7e-3$	0.11	$0.6e-3$	3,681	$-0.9e-3$	-0.09	$0.2e-3$	2,643
	(-7.2)	(14.05)	(2.79)		(-4.2)	(51.46)	(3.78)		(-7.15)	(-49.01)	(0.67)	
25% - 50%	$-0.7e-3$	0.02	$-0.4e-3$	488	$-0.6e-3$	0.08	$0.5e-3$	275	$-0.7e-3$	-0.06	$-1.5e-3$	213
	(-2.16)	(3.3)	(-0.61)		(-1.78)	(11.36)	(1.08)		(-1.31)	(-10.28)	(-1.18)	
Bottom 50%	$-1.2e-3$	0.05	$0e-3$	147,941	$-1.3e-3$	0.14	$-0.6e-3$	90,247	$-1e-3$	-0.09	$0.9e-3$	57,694
	(-35.97)	(105.3)	(-0.35)		(-28.68)	(232.64)	(-10.81)		(-21.8)	(-260.00)	(12.39)	
Night News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.9e-3$	0.04	$0e-3$	27,713	$-0.7e-3$	0.19	$0.4e-3$	15,522	$-1.2e-3$	-0.16	$-0.5e-3$	12,200
	(-14.62)	(22.9)	(0.08)		(-7.84)	(91.59)	(3.76)		(-13.49)	(-130.00)	(-3.97)	
10% - 25%	$-0.8e-3$	0.03	$0.2e-3$	4606	$-0.7e-3$	0.11	$0.5e-3$	2,700	$-1e-3$	-0.09	$-0.1e-3$	1,907
	(-6.24)	(12.31)	(1.28)		(-3.53)	(42.5)	(2.59)		(-6.53)	(-41.39)	(-0.13)	
25% - 50%	$-0.3e-3$	0.02	$-0.1e-3$	388	$-0.2e-3$	0.09	$1e-3$	220	$-0.3e-3$	-0.06	$-1.5e-3$	168
	(-0.9)	(3.28)	(-0.12)		(-0.6)	(9.75)	(1.69)		(-0.68)	(-9.55)	(-0.94)	
Bottom 50%	$-1.2e-3$	0.05	$0e-3$	133,587	$-1.3e-3$	0.14	$-0.7e-3$	82,136	$-1.1e-3$	-0.09	$1e-3$	51,630
	(-34.41)	(103.5)	(-0.76)		(-27.71)	(220.23)	(-11.5)		(-20.47)	(-240.00)	(12.28)	

This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm number of Ravenpack news articles on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3,+3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table 13 Abnormal Returns around Jumps by Day News Count (All Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
Day News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.6e-3$	0.03	$0e-3$	33,402	$-0.4e-3$	0.18	$0.2e-3$	18,299	$-0.9e-3$	-0.15	$-0.3e-3$	15,086
	(-11.62)	(21.57)	(-0.01)		(-4.52)	(107.54)	(2.22)		(-14.1)	(-150.00)	(-2.73)	
10% - 25%	$-0.3e-3$	0.01	$0.2e-3$	8,112	$-0.5e-3$	0.09	$0.3e-3$	4,218	$-0.1e-3$	-0.08	$0e-3$	3,897
	(-4.91)	(6.01)	(2.57)		(-5.39)	(74.9)	(3.62)		(-1.08)	(-74.08)	(0.16)	
25% - 50%	$-0.4e-3$	$1.3e-3$	$-0.1e-3$	1,157	$-0.3e-3$	0.06	$-0.1e-3$	581	$-0.5e-3$	-0.05	$0e-3$	576
	(-3.22)	(0.54)	(-0.44)		(-1.91)	(23.95)	(-0.41)		(-2.62)	(-23.58)	(-0.2)	
Bottom 50%	$-1.1e-3$	0.05	$0.1e-3$	175,953	$-1.2e-3$	0.14	$-0.3e-3$	105,743	$-0.9e-3$	-0.1	$0.6e-3$	70,223
	(-38.14)	(102.41)	(1.5)		(-30.84)	(255.97)	(-6.31)		(-22.46)	(-300.00)	(9.55)	
Day News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.8e-3$	0.05	$-0.2e-3$	19,804	$-0.3e-3$	0.2	$0e-3$	11,405	$-1.5e-3$	-0.16	$-0.4e-3$	8,399
	(-10)	(23.1)	(-1.7)		(-2.66)	(81.63)	(-0.31)		(-14.45)	(-110.00)	(-2.23)	
10% - 25%	$-0.4e-3$	0.01	$-0.1e-3$	4,181	$-0.6e-3$	0.08	$-0.1e-3$	2,130	$-0.2e-3$	-0.07	$0e-3$	2,051
	(-4.96)	(5.31)	(-0.51)		(-4.51)	(45.06)	(-0.79)		(-2.23)	(-45.34)	(0.01)	
25% - 50%	$-0.4e-3$	$3.7e-3$	$0.1e-3$	801	$-0.2e-3$	0.05	$0e-3$	406	$-0.6e-3$	-0.05	$0.1e-3$	395
	(-2.61)	(1.33)	(0.33)		(-1)	(17.58)	(0.01)		(-2.69)	(-17.36)	(0.45)	
Bottom 50%	$-1.2e-3$	0.05	$0e-3$	141,510	$-1.4e-3$	0.14	$-0.5e-3$	86,658	$-1.1e-3$	-0.09	$0.8e-3$	54,872
	(-36.21)	(100.81)	(0.07)		(-29.4)	(223.57)	(-9.4)		(-21.15)	(-250.00)	(10.86)	

This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm number of Ravenpack news articles on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3, +3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table 14 Abnormal Returns around Jumps by Stock Illiquidity (All Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
Stock Illiquidity Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-2.7e-3$	0.04	$0.9e-3$	29,718	$-3.2e-3$	0.18	$-0.4e-3$	16,942	$-2.1e-3$	-0.14	$2.6e-3$	12,776
	(-28.95)	(31.08)	(6.61)		(-22.95)	(111.46)	(-2.16)		(-17.85)	(-130.00)	(11.26)	
10% - 25%	$-1.6e-3$	0.04	$0.3e-3$	40,599	$-1.9e-3$	0.16	$-0.1e-3$	23,879	$-1.2e-3$	-0.12	$0.8e-3$	16,807
	(-24.47)	(42.67)	(3.04)		(-21.52)	(128)	(-1.05)		(-12.15)	(-150.00)	(5.91)	
25% - 50%	$-0.9e-3$	0.04	$0.1e-3$	63,167	$-0.9e-3$	0.14	$0e-3$	37,417	$-0.8e-3$	-0.1	$0.2e-3$	25,750
	(-22.94)	(53.52)	(1.23)		(-17.39)	(159.26)	(-0.18)		(-15.19)	(-190.00)	(2.18)	
Bottom 50%	$-0.1e-3$	0.04	$-0.4e-3$	85,158	$0e-3$	0.13	$-0.3e-3$	50,618	$-0.3e-3$	-0.09	$-0.4e-3$	34,484
	(-4.06)	(67.78)	(-9.17)		(-0.04)	(161.96)	(-6.19)		(-7)	(-210.00)	(-7.19)	
Stock Illiquidity Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-3e-3$	0.05	$0.8e-3$	25,117	$-3.5e-3$	0.19	$-0.8e-3$	14,477	$-2.3e-3$	-0.14	$3e-3$	10,692
	(-28.16)	(30.8)	(5.48)		(-22.47)	(101.24)	(-4.16)		(-17.21)	(-120.00)	(11.34)	
10% - 25%	$-1.8e-3$	0.05	$0.1e-3$	31,649	$-2.2e-3$	0.16	$-0.4e-3$	18,876	$-1.3e-3$	-0.11	$1e-3$	12,773
	(-22.05)	(42.13)	(1.36)		(-19.5)	(110.68)	(-3.22)		(-10.77)	(-130.00)	(5.85)	
25% - 50%	$-1.1e-3$	0.05	$0e-3$	45,484	$-1.2e-3$	0.14	$-0.2e-3$	27,609	$-1.1e-3$	-0.1	$0.3e-3$	17,911
	(-22.14)	(52.53)	(-0.37)		(-16.32)	(129.53)	(-2.78)		(-15.48)	(-150.00)	(2.86)	
Bottom 50%	$-0.1e-3$	0.05	$-0.4e-3$	64,046	$0e-3$	0.13	$-0.5e-3$	39,739	$-0.4e-3$	-0.08	$-0.3e-3$	24,341
	(-3.45)	(70.07)	(-8.95)		(0.47)	(135.42)	(-7.8)		(-6.69)	(-160.00)	(-4.44)	

This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm number of Ravenpack news articles on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3, +3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table 15 Abnormal Returns around Jumps by Stock Order Imbalances (All Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
Stock Order Imbalances Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-1.5e-3$	0.11	$-0.1e-3$	22,271	$-1.7e-3$	0.13	$-0.1e-3$	20,168	$-0.3e-3$	-0.08	$0e-3$	2,103
	(-20.5)	(114.99)	(-1.07)		(-20.55)	(139.83)	(-1.11)		(-1.96)	(-50.77)	(0.01)	
10% - 25%	$-1.1e-3$	0.1	$-0.3e-3$	41,528	$-1.3e-3$	0.14	$-0.3e-3$	34,477	$-0.5e-3$	-0.11	$-0.2e-3$	7,051
	(-19.42)	(114.48)	(-3.44)		(-18.86)	(162.51)	(-3.28)		(-4.73)	(-100.00)	(-1.04)	
25% - 50%	$-0.6e-3$	0.07	$-0.2e-3$	64,837	$-0.5e-3$	0.17	$-0.4e-3$	43,718	$-0.6e-3$	-0.12	$0.3e-3$	21,119
	(-11.53)	(75.69)	(-3.14)		(-8.52)	(147)	(-5.44)		(-8.41)	(-170.00)	(3.24)	
Bottom 50%	$-1.1e-3$	-0.02	$0.4e-3$	89,970	$-1.1e-3$	0.12	$0.1e-3$	30,495	$-1e-3$	-0.1	$0.5e-3$	59,496
	(-30.89)	(-47.44)	(8.19)		(-18.65)	(154.98)	(2.3)		(-24.62)	(-270.00)	(7.99)	
Stock Order Imbalances Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-1.7e-3$	0.11	$-0.3e-3$	19,285	$-1.8e-3$	0.13	$-0.3e-3$	17,772	$-0.6e-3$	-0.08	$0.1e-3$	1,513
	(-19.79)	(111.6)	(-2.33)		(-19.65)	(129.77)	(-2.45)		(-2.7)	(-37.7)	(0.27)	
10% - 25%	$-1.3e-3$	0.12	$-0.4e-3$	33,481	$-1.4e-3$	0.15	$-0.5e-3$	29,218	$-0.7e-3$	-0.1	$0e-3$	4,263
	(-18.22)	(116.13)	(-4.5)		(-17.5)	(147.86)	(-4.8)		(-5.09)	(-67.07)	(0.12)	
25% - 50%	$-0.6e-3$	0.09	$-0.4e-3$	46,772	$-0.6e-3$	0.18	$-0.7e-3$	33,446	$-0.8e-3$	-0.12	$0.6e-3$	13,326
	(-9.83)	(74.61)	(-4.79)		(-7.33)	(123.05)	(-7.75)		(-7.08)	(-120.00)	(4.18)	
Bottom 50%	$-1.3e-3$	-0.03	$0.5e-3$	66,758	$-1.6e-3$	0.13	$-0.1e-3$	20,143	$-1.2e-3$	-0.1	$0.8e-3$	46,615
	(-30.09)	(-48.16)	(7.5)		(-18.47)	(114.35)	(-1.36)		(-23.78)	(-230.00)	(8.89)	

This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm number of Ravenpack news articles on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3, +3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table 16 Probability of Daily Jumps: Removing Earnings Window (All Firms)

	Panel A: Probit Coefficient Estimates			
Variable	(1)	(2)	(3)	(4)
Num News	0.3	0.3	0.3	0.3
	(106.19)	(107.14)	(106.17)	(107.15)
Avg. Sentiment	0.14	0.15	0.14	0.14
	(63.3)	(64.19)	(62.61)	(63.48)
$\|ret_{t-1}\|$	0.11	0.09	0.11	0.09
	(58.56)	(47.68)	(58.39)	(47.18)
RES_s		0.48		0.48
		(69.73)		(69.78)
OIB_s			0.26	0.26
			(62.44)	(63.52)
R^2	3.38	3.77	3.76	4.17
N obs	$15,563,087$	$15,563,087$	$15,563,087$	$15,563,087$

	Panel B: Odds Ratio			
Variable	(1)	(2)	(3)	(4)
Num News	1.34	1.35	1.34	1.35
Avg. Sentiment	1.15	1.16	1.15	1.16
$\|ret_{t-1}\|$	1.12	1.1	1.12	1.1
RES_s		1.61		1.62
OIB_s			1.3	1.3

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ, as well as all interacted pairs of variables. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity. Reported results excluding $[-3, +3]$ days around earnings announcement days.

Table 17 Effects of day and night news measures and stock illiquidity on probability of daily jumps: Removing Earnings Window

	Panel A: Probit Coefficient Estimates					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	0.32		0.3	0.3	0.3	0.3
	(114.07)		(112.75)	(112.86)	(112.64)	(112.74)
Num News Day$_s$		0.15	0.13	0.13	0.13	0.13
		(85.97)	(78.87)	(79.41)	(78.49)	(79.05)
Avg. Sentiment Night$_s$	−0.01		−0.01	−0.01	−0.01	−0.01
	(−3.44)		(−4.93)	(−3.85)	(−4.94)	(−3.87)
Avg. Sentiment Day$_s$		0.09	0.07	0.08	0.07	0.08
		(61.3)	(52.18)	(53.14)	(51.74)	(52.7)
$\lvert ret_{t-1} \rvert$	0.12	0.12	0.12	0.1	0.12	0.1
	(67.06)	(68.65)	(63.92)	(52.32)	(63.77)	(51.79)
RES_s				0.47		0.48
				(67.9)		(67.94)
OIB_s					0.27	0.26
					(62.47)	(63.49)
R^2	2.51	1.57	3.5	3.89	3.9	4.29
N obs	15,560,887	15,553,386	15,548,077	15,548,077	15,548,077	15,548,077

	Panel B: Odds Ratio					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	1.38		1.35	1.35	1.35	1.35
Num News Day$_s$		1.16	1.14	1.14	1.14	1.14
Avg. Sentiment Night$_s$	0.99		0.99	0.99	0.99	0.99
Avg. Sentiment Day$_s$		1.09	1.08	1.08	1.08	1.08
$\lvert ret_{t-1} \rvert$	1.13	1.13	1.12	1.1	1.12	1.1
RES_s				1.6		1.61
OIB_s					1.31	1.3

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity. Reported results excluding $[-3, +3]$ days around earnings announcement days.

Table 18 Effects of day and night news measures and stock illiquidity on probability of daily jumps

	Panel A: Probit Coefficient Estimates					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	0.37		0.35	0.35	0.35	0.35
	(163.08)		(163.06)	(163.54)	(163.43)	(163.9)
Num News Day$_s$		0.16	0.13	0.13	0.13	0.13
		(92.9)	(82.75)	(82.89)	(82.35)	(82.48)
Avg. Sentiment Night$_s$	-0.01		-0.02	-0.02	-0.02	-0.02
	(-6.19)		(-7.52)	(-7.28)	(-7.62)	(-7.37)
Avg. Sentiment Day$_s$		0.11	0.08	0.08	0.08	0.08
		(73.39)	(55.94)	(56.65)	(55.73)	(56.46)
$\lvert ret_{t-1}\rvert$	0.1	0.1	0.1	0.08	0.1	0.08
	(48.5)	(51.57)	(45.77)	(39.46)	(45.98)	(39.5)
RES_s				0.41		0.42
				(46.43)		(47.59)
OIB_s					0.24	0.24
					(54.91)	(55.99)
R^2	8.12	2.36	9.34	9.54	9.62	9.82
N obs	13,137,073	13,134,578	13,131,389	13,131,389	13,131,389	13,131,389

	Panel B: Odds Ratio					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	1.45		1.42	1.42	1.43	1.43
Num News Day$_s$		1.17	1.14	1.14	1.14	1.14
Avg. Sentiment Night$_s$	0.99		0.98	0.98	0.98	0.98
Avg. Sentiment Day$_s$		1.11	1.08	1.08	1.08	1.08
$\lvert ret_{t-1}\rvert$	1.1	1.1	1.1	1.09	1.1	1.09
RES_s				1.51		1.52
OIB_s					1.27	1.27

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2004 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

A Appendix

The appendix contains additional results that support the main findings of the paper.

Table A.1 Effects of news measures and stock illiquidity on probability of daily jumps (Factiva)

	Panel A: Probit Coefficient Estimates				
Variable	(1)	(2)	(3)	(4)	(5)
$NewsCount$	0.21	0.24	0.24	0.24	0.25
	(108.61)	(110.66)	(109.27)	(108.94)	(109.06)
$NewsTone$	0.01	0.02	0.02	0.02	0.02
	(7.66)	(10.82)	(9.76)	(9.6)	(9.76)
$UncWords$	0.01	0.01	0.01	0.01	0.01
	(6.12)	(5.93)	(5.64)	(5.72)	(5.66)
$\|ret_{t-1}\|$	0.07	0.08	0.05	0.07	0.05
	(38.09)	(38.42)	(22.64)	(30.01)	(21.28)
RES_s			0.33		0.34
			(51.25)		(53.06)
OIB_s				0.27	0.27
				(67.9)	(70.26)
R^2	1.47	2.48	2.85	3.08	3.41
N obs	$17,397,790$	$11,471,715$	$10,595,021$	$10,595,021$	$10,595,021$

	Panel B: Odds Ratio				
Variable	(1)	(2)	(3)	(4)	(5)
$NewsCount$	1.23	1.27	1.28	1.27	1.28
$NewsTone$	1.01	1.02	1.02	1.02	1.02
$UncWords$	1.01	1.01	1.01	1.01	1.01
$\|ret_{t-1}\|$	1.08	1.08	1.05	1.07	1.05
RES_s			1.38		1.41
OIB_s				1.31	1.32

Notes: This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Factiva database each day, the absolute value of news tone, the percentage of uncertain words, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed from the percentage of positive and negative words using the list in Loughran and McDonald (2011). The sample period is from January 1980 to July 2012 corresponding to the Factiva database as outlined in Jeon, McCurdy, and Zhao (2021) but shrinks to January 1993 to July 2012 when TAQ data became available. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.2 Effects of news measures and stock illiquidity on probability of daily jumps (Largest 500 firms)

	Panel A: Probit Coefficient Estimates				
Variable	(1)	(2)	(3)	(4)	(5)
Num News	0.45	0.45	0.45	0.45	0.45
	(96.95)	(95.13)	(95.38)	(95.18)	(95.39)
Avg. Sentiment	0.2	0.21	0.21	0.21	0.21
	(28.59)	(28.25)	(28.33)	(28.3)	(28.36)
$\lvert ret_{t-1} \rvert$	-0.04	-0.05	-0.05	-0.05	-0.05
	(-4.38)	(-4.87)	(-5.04)	(-4.87)	(-5.01)
RES_s			0.07		0.05
			(2.16)		(1.63)
OIB_s				0.08	0.08
				(6.91)	(6.73)
R^2	1.29	1.33	1.33	1.33	1.33
N obs	$2,242,279$	$2,069,113$	$2,069,113$	$2,069,113$	$2,069,113$

	Panel B: Odds Ratio				
Variable	(1)	(2)	(3)	(4)	(5)
Num News	1.57	1.57	1.57	1.57	1.57
Avg. Sentiment	1.23	1.23	1.23	1.23	1.23
$\lvert ret_{t-1} \rvert$	0.96	0.96	0.96	0.96	0.96
RES_s			1.05		1.05
OIB_s				1.08	1.08

Notes: This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for the largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.3 Effects of news measures and stock illiquidity on daily jump size. (Factiva)

Panel A: All Jumps					
Variable	(1)	(2)	(3)	(4)	(5)
$NewsCount$	$3.5e-3$	$2.2e-3$	$2.5e-3$	$3.6e-3$	$3.6e-3$
	(11.69)	(6.83)	(7.74)	(11.05)	(11.14)
$NewsTone$	0.01	0.01	0.01	0.01	0.01
	(22.42)	(20.31)	(18.67)	(17.01)	(17.01)
$UncWords$	$4.5e-3$	$2.8e-3$	$2.7e-3$	$2.3e-3$	$2.3e-3$
	(8.39)	(3.9)	(3.77)	(3.42)	(3.43)
ret_{t-1}	-0.21	-0.06	-0.01	-0.01	-0.01
	(-11.18)	(-2.53)	(-0.46)	(-0.59)	(-0.57)
RES_s			-0.01		$1.3e-3$
			(-5.46)		(1.12)
OIB_s				0.09	0.09
				(103.75)	(103.35)
R^2	1.28	0.92	0.88	13.81	13.81
N obs	310,113	171,302	157,169	157,169	157,169

Panel B: Positive Jumps					
Variable	(1)	(2)	(3)	(4)	(5)
$NewsCount$	0.01	0.01	0.01	0.01	0.01
	(29.39)	(22.98)	(24.04)	(22.36)	(23.95)
$NewsTone$	$1.5e-3$	$0.9e-3$	$1.5e-3$	$1.1e-3$	$1.5e-3$
	(2.3)	(1.16)	(1.79)	(1.37)	(1.79)
$UncWords$	$3.2e-3$	$4.7e-3$	$4.8e-3$	$4.6e-3$	$4.8e-3$
	(5.13)	(5.37)	(5.37)	(5.18)	(5.37)
ret_{t-1}	-0.26	-0.06	0.03	0.01	0.03
	(-9.94)	(-1.73)	(1.15)	(0.18)	(1.15)
RES_s			0.04		0.04
			(26.13)		(26.12)
OIB_s				$1.9e-3$	$0.6e-3$
				(2.3)	(0.76)
R^2	2.83	1.77	3.15	1.85	3.15
N obs	182,239	105,159	97,016	97,016	97,016

Panel C: Negative Jumps					
Variable	(1)	(2)	(3)	(4)	(5)
$NewsCount$	-0.01	-0.01	-0.01	-0.01	-0.01
	(-41.42)	(-33.5)	(-33.52)	(-29.88)	(-31.38)
$\|NewsTone\|$	$4.3e-3$	0.01	$4.8e-3$	0.01	$4.8e-3$
	(11.57)	(13.56)	(11.65)	(12.82)	(11.79)
$UncWords$	$0e-3$	$-1.4e-3$	$-1.5e-3$	$-1.1e-3$	$-1.3e-3$
	(-0.06)	(-2.94)	(-3.73)	(-2.58)	(-3.15)
ret_{t-1}	-0.05	-0.02	-0.02	-0.02	-0.02
	(-4.34)	(-1.24)	(-1.15)	(-1.17)	(-1.03)
RES_s			-0.04		-0.04
			(-39.02)		(-41.73)
OIB_s				-0.01	-0.02
				(-22.43)	(-28.33)
R^2	7.79	8.68	13.02	9.28	14.54
N obs	127,874	66,143	60,153	60,153	60,153

This table reports results from regressions of daily jump sizes, conditional on the jump indicator being 1, on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms in each month in the sample. The explanatory variables are the total number of news articles reported on the Factiva database each day, news tone, the percentage of uncertain words, the previous day's return the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. Each of the variables is standardized to have the same mean and standard deviation across firms. The news tone measure is constructed from the percentage of positive and negative words using the list in Loughran and McDonald (2011). The news tone measure is constructed from the percentage of positive and negative words using the list in Loughran and McDonald (2011). The sample period is from January 1980 to July 2012 corresponding to the Factiva database as outlined in Jeon, McCurdy, and Zhao (2021) but shrinks to January 1993 to July 2012 when TAQ data became available. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panels B and C report the results for positive jump sizes and negative jump sizes, respectively. $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.4 Effects of news measures and stock illiquidity on daily jump size. (Largest 500 firms)

Panel A: All Jumps					
Variable	(1)	(2)	(3)	(4)	(5)
Num News	$-1.2e-3$	$-1.3e-3$	$-1.1e-3$	$-1e-3$	$-1e-3$
	(-3.81)	(-4.07)	(-3.42)	(-3.07)	(-3.17)
Avg. Sentiment	0.01	0.01	0.01	0.01	0.01
	(16.65)	(16.39)	(16.27)	(16.06)	(16.01)
ret_{t-1}	0.19	0.19	0.18	0.15	0.14
	(2.28)	(2.29)	(2.17)	(1.75)	(1.71)
RES_s			$-1.3e-3$		-0.01
			(-0.35)		(-1.92)
OIB_s				0.03	0.03
				(18.54)	(18.51)
R^2	1.89	1.92	1.97	5.53	5.57
N obs	20,101	19,238	18,666	18,666	18,666

Panel B: Positive Jumps					
Variable	(1)	(2)	(3)	(4)	(5)
Num News	$4e-3$	$4e-3$	$4.5e-3$	$4.3e-3$	$4.5e-3$
	(14.05)	(13.78)	(15.71)	(14.74)	(15.63)
Avg. Sentiment	$3e-3$	$3e-3$	$3.3e-3$	$3.2e-3$	$3.3e-3$
	(5.98)	(5.91)	(6.54)	(6.28)	(6.69)
ret_{t-1}	-0.29	-0.29	-0.29	-0.29	-0.29
	(-3.46)	(-3.35)	(-3.53)	(-3.39)	(-3.53)
RES_s			0.05		0.05
			(9.4)		(9.23)
OIB_s				$4.6e-3$	$1.2e-3$
				(3.92)	(1.03)
R^2	7.27	7.21	7.77	11.54	11.56
N obs	9,859	9,499	9,219	9,219	9,219

Panel C: Negative Jumps					
Variable	(1)	(2)	(3)	(4)	(5)
Num News	-0.01	-0.01	-0.01	-0.01	-0.01
	(-20.22)	(-19.45)	(-20)	(-18.77)	(-19.42)
Avg. Sentiment	$-3.5e-3$	$-3.6e-3$	$-3.5e-3$	$-3.5e-3$	$-3.5e-3$
	(-6.26)	(-6.16)	(-6.19)	(-6.05)	(-6.19)
ret_{t-1}	0.49	0.49	0.47	0.49	0.47
	(7.61)	(7.47)	(6.83)	(7.2)	(6.82)
RES_s			-0.04		-0.04
			(-11.58)		(-11.6)
OIB_s				$0.7e-3$	$0.6e-3$
				(0.57)	(0.51)
R^2	13.17	12.99	16.41	13.28	16.41
N obs	10,242	9,739	9,447	9,447	9,447

This table reports results from regressions of daily jump sizes, conditional on the jump indicator being 1, on daily news measures as well as stock illiquidity and order imbalances from TAQ for the largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panels B and C report the results for positive jump sizes and negative jump sizes, respectively. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.5 Effects of day and night news measures and stock illiquidity on probability of daily jumps (500 Largest Firms)

	Panel A: Probit Coefficient Estimates					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	0.43		0.37	0.37	0.37	0.37
	(95.71)		(79.72)	(79.75)	(79.86)	(79.85)
Num News Day$_s$		0.29	0.19	0.19	0.19	0.19
		(56.82)	(38.22)	(37.92)	(37.96)	(37.95)
Avg. Sentiment Night$_s$	0.05		0.04	0.04	0.04	0.04
	(10.79)		(6.81)	(6.85)	(6.85)	(6.87)
Avg. Sentiment Day$_s$		0.17	0.11	0.11	0.11	0.11
		(49.84)	(27.05)	(26.72)	(26.69)	(26.69)
$\|ret_{t-1}\|$	-0.02	-0.02	-0.03	-0.03	-0.03	-0.03
	(-1.72)	(-2.03)	(-3.01)	(-3.03)	(-2.93)	(-2.99)
RES_s				0.04		0.02
				(1.14)		(0.67)
OIB_s					0.07	0.07
					(6)	(5.9)
R^2	1.27	0.61	1.47	1.47	1.47	1.47
N obs	2,013,555	2,013,335	2,013,335	2,013,335	1,998,044	1,998,041

	Panel B: Odds Ratio					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	1.54		1.45	1.45	1.45	1.45
Num News Day$_s$		1.33	1.21	1.21	1.21	1.21
Avg. Sentiment Night$_s$	1.05		1.04	1.04	1.04	1.04
Avg. Sentiment Day$_s$		1.19	1.11	1.11	1.11	1.11
$\|ret_{t-1}\|$	0.98	0.98	0.97	0.97	0.97	0.97
RES_s				1.04		1.02
OIB_s					1.08	1.07

Notes: This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.6 Effects of day and night news measures and stock illiquidity on daily jump size. (500 Largest Firms)

	Panel A: All Jumps					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	$-0.7e-3$		$-0.7e-3$	$-0.7e-3$	$-0.7e-3$	$-0.7e-3$
	(-2.7)		(-2.45)	(-2.6)	(-2.28)	(-2.46)
Num News Day$_s$		$-1.7e-3$	$-1.5e-3$	$-1.5e-3$	$-1.6e-3$	$-1.5e-3$
		(-2.87)	(-2.55)	(-2.51)	(-2.61)	(-2.55)
Avg. Sentiment Night$_s$	$3e-3$		$2.9e-3$	$2.9e-3$	$2.9e-3$	$3e-3$
	(2.89)		(2.76)	(2.83)	(2.99)	(3.09)
Avg. Sentiment Day$_s$		$1.8e-3$	$1.9e-3$	$1.9e-3$	$1.8e-3$	$1.8e-3$
		(4.8)	(4.93)	(4.84)	(4.8)	(4.69)
ret_{t-1}	0.21	0.2	0.2	0.2	0.16	0.16
	(2.42)	(2.32)	(2.37)	(2.36)	(1.93)	(1.91)
RES_s				$-3e-3$		$-3.9e-3$
				(-1.56)		(-2.12)
OIB_s					0.03	0.03
					(21.45)	(21.6)
R^2	0.41	0.52	0.69	0.69	5.18	5.3
N obs	18,822	18,822	18,822	18,822	18,822	18,822

	Panel B: Positive Jumps					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	$2.3e-3$		$1e-3$	$1.2e-3$	$1.2e-3$	$1.3e-3$
	(8.81)		(3.89)	(4.34)	(4.46)	(4.62)
Num News Day$_s$		0.01	$4.9e-3$	$4.9e-3$	$5e-3$	$4.9e-3$
		(14.26)	(12.36)	(12.45)	(12.43)	(12.49)
Avg. Sentiment Night$_s$	0.01		$4e-3$	$4e-3$	$4.1e-3$	$4e-3$
	(5.39)		(4.25)	(4.24)	(4.3)	(4.27)
Avg. Sentiment Day$_s$		$1e-3$	$0.5e-3$	$0.5e-3$	$0.5e-3$	$0.6e-3$
		(3.52)	(1.8)	(2)	(1.87)	(2.03)
ret_{t-1}	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25
	(-2.83)	(-2.92)	(-2.91)	(-2.92)	(-2.95)	(-2.94)
RES_s				0.01		0.01
				(4.41)		(4.35)
OIB_s					$4.6e-3$	$2.5e-3$
					(4.45)	(2.26)
R^2	4.54	7.74	8.58	10.56	8.84	10.64
N obs	9,289	9,289	9,289	9,289	9,289	9,289

	Panel C: Negative Jumps					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	$-3.5e-3$		$-1.7e-3$	$-1.8e-3$	$-1.8e-3$	$-1.9e-3$
	(-14.82)		(-6.56)	(-7.15)	(-6.89)	(-7.36)
Num News Day$_s$		-0.01	-0.01	-0.01	-0.01	-0.01
		(-14.95)	(-12.72)	(-12.62)	(-12.75)	(-12.62)
Avg. Sentiment Night$_s$	$-4.5e-3$		$-3.2e-3$	$-2.9e-3$	$-3.2e-3$	$-2.9e-3$
	(-4.17)		(-3.12)	(-2.91)	(-3.12)	(-2.91)
Avg. Sentiment Day$_s$		$-0.4e-3$	$0.4e-3$	$0.2e-3$	$0.3e-3$	$0.2e-3$
		(-0.92)	(0.91)	(0.53)	(0.85)	(0.49)
ret_{t-1}	0.47	0.44	0.46	0.45	0.45	0.45
	(6.83)	(6.58)	(6.82)	(6.74)	(6.79)	(6.72)
RES_s				-0.01		-0.01
				(-4.59)		(-4.57)
OIB_s					$2.5e-3$	$1.6e-3$
					(2.2)	(1.48)
R^2	9.35	14.51	15.74	17.9	15.81	17.93
N obs	9,533	9,533	9,533	9,533	9,533	9,533

This table reports results from regressions of daily jump sizes, conditional on the jump indicator being 1, on daily news measures as well as stock illiquidity and order imbalances from TAQ for the largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panels B and C report the results for positive jump sizes and negative jump sizes, respectively. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.7 Effects of news measures (day and night) and stock illiquidity (day and night) on probability of daily jumps (Top 500 Firms)

	Panel A: Probit Coefficient Estimates			
Variable	(1)	(2)	(3)	(4)
Num News Night$_s$	0.37	0.37	0.37	0.37
	(68.26)	(68.11)	(67.4)	(67.57)
Num News Day$_s$	0.21	0.21	0.2	0.2
	(36.43)	(36.38)	(34.92)	(35.21)
Avg. Sentiment Night$_s$	0.07	0.07	0.07	0.07
	(13.56)	(13.6)	(12.18)	(12.5)
Avg. Sentiment Day$_s$	0.11	0.11	0.11	0.11
	(25.13)	(25.03)	(24.25)	(24.02)
$\|ret_{t-1}\|$	-0.03	-0.03	-0.03	-0.04
	(-2.14)	(-2.2)	(-2.36)	(-2.95)
$RQS_{A,s}$		0.01		0.02
		(0.67)		(1.84)
log(DollTradingVolume$_{A,s}$)			0.35	0.4
			(18.52)	(21.78)
RES_s				0.25
				(7.26)
OIB_s				0.22
				(12.49)
R^2	1.45	1.47	1.49	1.51
N obs	1,470,103	1,468,070	1,468,362	1,467,678

	Panel B: Odds Ratio			
Variable	(1)	(2)	(3)	(4)
Num News Night$_s$	1.45	1.45	1.44	1.44
Num News Day$_s$	1.23	1.23	1.22	1.23
Avg. Sentiment Night$_s$	1.08	1.08	1.07	1.07
Avg. Sentiment Day$_s$	1.12	1.12	1.12	1.12
$\|ret_{t-1}\|$	0.97	0.97	0.97	0.96
$RQS_{A,s}$		1.01		1.02
log(DollTradingVolume$_{A,s}$)			1.41	1.49
RES_s				1.28
OIB_s				1.25

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for the largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. Panels B and C report the results for positive jump sizes and negative jump sizes, respectively. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.8 Effects of day news measures and day stock illiquidity on probability of daytime period jumps (Top 500 Firms)

	Panel A: Probit Coefficient Estimates			
Variable	(1)	(2)	(3)	(4)
Num News Day_s	0.16	0.16	0.16	0.16
	(35.84)	(36.53)	(36.04)	(36.67)
Avg. Sentiment Day_s	0.09	0.09	0.09	0.09
	(28.84)	(28.96)	(28.9)	(29.01)
$\lvert ret_{D,t-1}\rvert$	0.04	0.03	0.04	0.03
	(5.76)	(4.54)	(5.68)	(4.56)
RES_s		0.21		0.19
		(8.34)		(7.59)
OIB_s			0.1	0.09
			(9.3)	(8.67)
R^2	0	0	0	0
N obs	$1,990,305$	$1,990,305$	$1,990,305$	$1,990,305$

	Panel B: Odds Ratio			
Variable	(1)	(2)	(3)	(4)
Num News Day_s	1.17	1.17	1.17	1.17
Avg. Sentiment Day_s	1.1	1.1	1.1	1.1
$\lvert ret_{D,t-1}\rvert$	1.04	1.03	1.04	1.03
RES_s		1.23		1.21
OIB_s			1.1	1.09

Notes: This table reports results from pooled logistic regressions of daytime period jump indicators, defined using Lee and Mykland (2008), on day news measures as well as day stock illiquidity and order imbalances from TAQ for the largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the number of news articles reported on the Ravenpack database within the trading day, the absolute value of news tone of within the trading day news, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The jump variable ($\{J_{D,99}\}$) is the same for each regression where $\{J_{D,99}\}$ identifies a jump in the daytime period if the absolute value of the daytime period return is above $\{5.28\}$ times the time-varying daytime period spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.9 Effects of night news measures and night stock illiquidity on probability of night trading period jumps (Largest 500 Firms)

	Panel A: Probit Coefficient Estimates				
Variable	(1)	(2)	(3)	(4)	(5)
Num News Night$_s$	0.5	0.51	0.51	0.5	0.5
	(94.6)	(85.1)	(85.12)	(83.87)	(83.85)
Avg. Sentiment Night$_s$	0.09	0.14	0.14	0.14	0.14
	(23.21)	(35.51)	(35.5)	(34.63)	(34.63)
$\lvert ret_{t-1} \rvert$	-0.05	-0.23	-0.23	-0.24	-0.24
	(-3.4)	(-9.42)	(-9.42)	(-9.67)	(-9.67)
$RQS_{A,s}$			$-4.5e-3$		$2.7e-3$
			(-0.36)		(0.22)
log(DollTradingVolume$_{A,s}$)				0.31	0.31
				(13.88)	(13.9)
R^2	20.53	21.85	21.85	22.11	22.11
N obs	1,990,126	1,460,089	1,460,089	1,460,089	1,460,089

	Panel B: Odds Ratio				
Variable	(1)	(2)	(3)	(4)	(5)
Num News Night$_s$	1.65	1.67	1.67	1.66	1.66
Avg. Sentiment Night$_s$	1.1	1.15	1.15	1.15	1.15
$\lvert ret_{N,t-1} \rvert$	0.95	0.79	0.79	0.79	0.79
$RQS_{A,s}$			1		1
log(DollTradingVolume$_{A,s}$)				1.36	1.36

Notes: This table reports results from pooled logistic regressions of night trading period jump indicators, defined using Lee and Mykland (2008), on night news measures as well as night stock illiquidity and log night trading period dollar volume from TAQ for the largest 500 firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of night news articles reported on the Ravenpack database in the night trading period, the absolute value of news tone of the news articles reported in the night trading period, the absolute value of the previous day's night return, the stock relative quoted spreads in the night trading period, and the log night trading period dollar volume from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2014. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{N,99}\}$) is the same for each regression where $\{J_{N,99}\}$ identifies a night trading period return jump if the absolute value of the night trading period return is above $\{5.28\}$ times the time-varying night trading period return spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.10 Probability of Daily Jumps with Interactions Jumps (Top 500 Firms)

	Panel A: Probit Coefficient Estimates			
Variable	(1)	(2)	(3)	(4)
Num News	0.45	0.45	0.45	0.45
	(80.04)	(95.23)	(95.12)	(95.25)
Avg. Sentiment	0.21	0.21	0.21	0.21
	(28.31)	(26.29)	(28.23)	(28.02)
$\lvert ret_{t-1} \rvert$	-0.05	-0.05	-0.05	-0.05
	(-4.73)	(-4.73)	(-4.72)	(-4.72)
RES_s	0.06	0.03	0.04	0.04
	1.81	(1.12)	1.24	(1.23)
OIB_s	0.08	0.08	0.07	0.07
	6.78	6.78	(5.81)	(5.62)
Num News x RES_s	-0.02			
	-1.14			
Num News x OIB_s		0.01		
		0.44		
Avg. Sentiment x RES_s			0.01	
			(1.52)	
Avg. Sentiment x OIB_s				0.02
				(2.44)
R^2	1.37	1.37	1.37	1.37
N obs	$1,998,041$	$1,998,041$	$1,998,041$	$1,998,041$

	Panel B: Odds Ratio			
Variable	(1)	(2)	(3)	(4)
Num News	1.57	1.58	1.58	1.58
Avg. Sentiment	1.23	1.24	1.23	1.23
$\lvert ret_{t-1} \rvert$	0.95	0.95	0.95	0.95
RES_s	1.06	1.04	1.04	1.04
OIB_s	1.08	1.08	1.07	1.07
Num News x RES_s	0.98			
Num News x OIB_s		1.01		
Avg. Sentiment x RES_s			1.01	
Avg. Sentiment x OIB_s				1.02

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for the largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ, as well as all interacted pairs of variables. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.11 Abnormal Returns around Jumps by News Count (500 Largest Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.3e-3$	$-1.3e-3$	$0e-3$	$8,661$	$-0.3e-3$	0.09	$0.2e-3$	$4,358$	$-0.4e-3$	-0.09	$-0.1e-3$	$4,303$
	(-6.97)	(-1.07)	(0.28)		(-3.87)	(86.04)	(2.77)		(-5.91)	(-83.81)	(-0.68)	
10% - 25%	$-0.5e-3$	$5e-3$	$0e-3$	$1,860$	$-0.6e-3$	0.07	$-0.3e-3$	928	$-0.4e-3$	-0.06	$0.3e-3$	931
	(-4.31)	(2.65)	(-0.25)		(-3.5)	(36.62)	(-2)		(-2.54)	(-33.86)	(1.37)	
25% - 50%	$-0.3e-3$	$4.4e-3$	$0.1e-3$	$2,295$	$-0.5e-3$	0.05	$-0.1e-3$	$1,123$	$-0.2e-3$	-0.04	$0.3e-3$	$1,172$
	(-4.23)	(3.32)	(1.03)		(-3.88)	(38.37)	(-0.55)		(-2.02)	(-36.8)	(1.85)	
Bottom 50%	$-0.3e-3$	$4.8e-3$	$0.1e-3$	$5,444$	$-0.5e-3$	0.05	$0e-3$	$2,601$	$-0.1e-3$	-0.04	$0.1e-3$	$2,843$
	(-5.12)	(6.14)	(0.73)		(-5.64)	(58.8)	(0.25)		(-1.4)	(-51.87)	(0.79)	
News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.7e-3$	$-3.5e-3$	$0e-3$	$3,471$	$-0.3e-3$	0.09	$0.1e-3$	$1,713$	$-1.1e-3$	-0.1	$-0.2e-3$	1757
	(-6.86)	(-1.57)	(-0.07)		(-2.29)	(44.87)	(0.98)		(-7.12)	(-46.19)	(-0.35)	
10% - 25%	$-0.5e-3$	$4.5e-3$	$-0.2e-3$	1386	$-0.5e-3$	0.06	$-0.5e-3$	675	$-0.4e-3$	-0.05	$0e-3$	712
	(-3.72)	(2.22)	(-1.49)		(-2.83)	(29.77)	(-2.25)		(-2.42)	(-28.39)	(-0.08)	
25% - 50%	$-0.4e-3$	$3.1e-3$	$0e-3$	1923	$-0.5e-3$	0.05	$-0.2e-3$	910	$-0.3e-3$	-0.04	$0.2e-3$	$1,013$
	(-4.81)	(2.23)	(-0.13)		(-3.87)	(32.33)	(-1.35)		(-2.88)	(-34.14)	(1.04)	
Bottom 50%	$-0.2e-3$	$4.7e-3$	$0e-3$	$4,922$	$-0.4e-3$	0.05	$0e-3$	$2,352$	$0e-3$	-0.03	$0.1e-3$	$2,572$
	(-3.67)	(5.78)	(0.47)		(-4.62)	(54.8)	(0.11)		(-0.31)	(-48.21)	(0.56)	

Notes: This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm number of Ravenpack news articles on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3, +3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table A.12 Abnormal Returns around Jumps by Night News Count (500 Largest Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
Night News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.3e-3$	$-1.4e-3$	$0.2e-3$	7,803	$-0.3e-3$	0.08	$0.2e-3$	3,916	$-0.3e-3$	-0.09	$0.1e-3$	3,887
	(-6)	(-1.12)	(2.15)		(-4)	(82.58)	(3.81)		(-4.47)	(-81.54)	(0.81)	
10% - 25%	$-0.6e-3$	$3.9e-3$	$0e-3$	1,473	$-0.5e-3$	0.06	$-0.1e-3$	754	$-0.6e-3$	-0.05	$0.1e-3$	719
	(-4.73)	(1.86)	(0.03)		(-2.86)	(26.52)	(-0.72)		(-3.96)	(-24.99)	(0.59)	
25% - 50%	$-0.5e-3$	0.01	$-1.9e-3$	317	$-0.3e-3$	0.06	$0.1e-3$	169	$-0.7e-3$	-0.05	$-4.3e-3$	148
	(-1.71)	(0.99)	(-1.26)		(-0.79)	(11.31)	(0.43)		(-1.57)	(-7.7)	(-1.32)	
Bottom 50%	$-0.3e-3$	$4.3e-3$	$0e-3$	8,666	$-0.5e-3$	0.06	$-0.1e-3$	4,168	$-0.2e-3$	-0.04	$0.1e-3$	4,495
	(-7.15)	(5.77)	(-0.29)		(-6.69)	(74)	(-1.09)		(-3.27)	(-63.36)	(0.56)	
Night News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.7e-3$	-0.01	$0.3e-3$	2,758	$-0.3e-3$	0.09	$0.3e-3$	1,313	$-0.9e-3$	-0.1	$0.4e-3$	1,445
	(-6.04)	(-2.58)	(1.48)		(-2.34)	(38.31)	(2.02)		(-5.91)	(-42.1)	(0.94)	
10% - 25%	$-0.5e-3$	0.01	$-0.1e-3$	1244	$-0.4e-3$	0.06	$-0.3e-3$	649	$-0.7e-3$	-0.05	$0.2e-3$	596
	(-4.18)	(2.31)	(-0.42)		(-2.13)	(24.98)	(-1.65)		(-3.92)	(-22.46)	(0.56)	
25% - 50%	$-0.4e-3$	$3.7e-3$	$-2.3e-3$	256	$-0.1e-3$	0.06	$0.1e-3$	135	$-0.7e-3$	-0.06	-0.01	122
	(-1.15)	(0.59)	(-1.24)		(-0.33)	(9.16)	(0.35)		(-1.23)	(-6.67)	(-1.29)	
Bottom 50%	$-0.3e-3$	$4.5e-3$	$-0.1e-3$	7,453	$-0.5e-3$	0.05	$-0.1e-3$	3,558	$-0.2e-3$	-0.04	$0e-3$	3,895
	(-6.24)	(5.72)	(-0.93)		(-5.99)	(65.15)	(-1.28)		(-2.68)	(-56.97)	(-0.13)	

Notes: This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm number of Ravenpack news articles that are recorded as having been reported after the market closing on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3, +3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table A.13 Abnormal Returns around Jumps by Day News Count (Largest 500 Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
Day News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.5e-3$	$-4.5e-3$	$0.1e-3$	$4,957$	$-0.4e-3$	0.09	$0.1e-3$	$2,417$	$-0.6e-3$	-0.1	$0.1e-3$	$2,540$
	(-7.35)	(-2.55)	(0.72)		(-4.13)	(62.71)	(1.16)		(-6.14)	(-62.91)	(0.43)	
10% - 25%	$-0.2e-3$	$2.6e-3$	$0.2e-3$	$2,921$	$-0.3e-3$	0.07	$0.2e-3$	$1,431$	$-0.1e-3$	-0.06	$0.2e-3$	$1,490$
	(-2.79)	(1.6)	(2.09)		(-2.61)	(44.71)	(1.43)		(-1.18)	(-47.43)	(1.53)	
25% - 50%	$-0.4e-3$	$0.6e-3$	$0.1e-3$	$1,171$	$-0.3e-3$	0.05	$0.1e-3$	583	$-0.4e-3$	-0.05	$0.2e-3$	588
	(-3.24)	(0.3)	(0.89)		(-1.78)	(26.14)	(0.47)		(-2.8)	(-23.75)	(0.8)	
Bottom 50%	$-0.3e-3$	0.01	$-0.1e-3$	$9,210$	$-0.4e-3$	0.06	$0e-3$	$4,577$	$-0.1e-3$	-0.05	$-0.1e-3$	$4,628$
	(-6.42)	(7.31)	(-1.17)		(-6.57)	(85.17)	(-0.23)		(-2.39)	(-70.57)	(-1.37)	
Day News Count Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.8e-3$	$-4.3e-3$	$0e-3$	$2,475$	$-0.4e-3$	0.1	$-0.1e-3$	$1,210$	$-1.3e-3$	-0.1	$0.2e-3$	$1,265$
	(-6.86)	(-1.56)	(0.14)		(-2.45)	(37.54)	(-0.43)		(-6.88)	(-38.4)	(0.25)	
10% - 25%	$-0.3e-3$	$2.5e-3$	$0e-3$	$1,639$	$-0.4e-3$	0.07	$-0.1e-3$	762	$-0.2e-3$	-0.05	$0.2e-3$	877
	(-2.25)	(1.23)	(0.26)		(-1.91)	(29.25)	(-0.52)		(-1.19)	(-31.52)	(0.79)	
25% - 50%	$-0.4e-3$	$3.2e-3$	$0.4e-3$	818	$-0.3e-3$	0.05	$0.2e-3$	411	$-0.5e-3$	-0.05	$0.5e-3$	407
	(-2.74)	(1.24)	(1.59)		(-1.34)	(19.35)	(0.65)		(-2.56)	(-17.32)	(1.5)	
Bottom 50%	$-0.3e-3$	$4.1e-3$	$-0.1e-3$	$6,781$	$-0.5e-3$	0.05	$-0.1e-3$	$3,272$	$-0.2e-3$	-0.04	$-0.2e-3$	$3,506$
	(-5.97)	(5.26)	(-1.51)		(-5.98)	(64.84)	(-0.57)		(-2.36)	(-56.63)	(-1.52)	

Notes: This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm number of Ravenpack news articles on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3,+3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table A.14 Abnormal Returns around Jumps by Stock Illiquidity (500 Largest Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
Stock Illiquidity Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.6e-3$	-0.01	$0.3e-3$	1843	$-0.8e-3$	0.09	$0.2e-3$	863	$-0.4e-3$	-0.09	$0.3e-3$	980
	(-4)	(-2.37)	(1.5)		(-3.42)	(34.03)	(0.82)		(-2.18)	(-33.66)	(1.27)	
10% - 25%	$-0.5e-3$	$2.3e-3$	$0.3e-3$	3,926	$-0.5e-3$	0.08	$0.2e-3$	1,965	$-0.6e-3$	-0.07	$0.4e-3$	1,961
	(-6.44)	(1.38)	(2.3)		(-4.07)	(53.07)	(1.91)		(-5.04)	(-45.53)	(1.61)	
25% - 50%	$-0.3e-3$	$4.5e-3$	$0e-3$	5,591	$-0.4e-3$	0.07	$0e-3$	2924	$-0.3e-3$	-0.07	$0e-3$	2,667
	(-6.05)	(3.82)	(0.01)		(-5.12)	(73.63)	(-0.4)		(-3.47)	(-58.26)	(0.19)	
Bottom 50%	$-0.2e-3$	$1.8e-3$	$-0.1e-3$	6,896	$-0.2e-3$	0.06	$0e-3$	3,252	$-0.1e-3$	-0.05	$-0.3e-3$	3,641
	(-4.21)	(2.14)	(-1.89)		(-3.82)	(61.18)	(-0.19)		(-2.13)	(-68.13)	(-2)	
Stock Illiquidity Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.6e-3$	-0.01	$0.3e-3$	1,308	$-0.8e-3$	0.09	$0.2e-3$	603	$-0.5e-3$	-0.09	$0.4e-3$	705
	(-3.58)	(-2)	(1.18)		(-2.68)	(27.48)	(0.58)		(-2.37)	(-26.28)	(1.05)	
10% - 25%	$-0.6e-3$	$-0.7e-3$	$0.3e-3$	2,459	$-0.5e-3$	0.07	$0.2e-3$	1,163	$-0.7e-3$	-0.07	$0.5e-3$	1,296
	(-4.89)	(-0.33)	(1.56)		(-2.73)	(33.83)	(0.98)		(-4.16)	(-31.3)	(1.27)	
25% - 50%	$-0.5e-3$	$4e-3$	$-0.1e-3$	3,370	$-0.4e-3$	0.06	$-0.2e-3$	1,734	$-0.5e-3$	-0.06	$0e-3$	1,636
	(-5.99)	(2.72)	(-0.74)		(-4.03)	(48.31)	(-1.77)		(-4.42)	(-38.42)	(-0.08)	
Bottom 50%	$-0.2e-3$	$4.5e-3$	$-0.2e-3$	4,576	$-0.3e-3$	0.05	$-0.1e-3$	2,155	$-0.2e-3$	-0.04	$-0.3e-3$	2,421
	(-4.62)	(4.66)	(-1.98)		(-3.86)	(42.08)	(-1.04)		(-2.63)	(-49.29)	(-1.71)	

Notes: This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm stock illiquidity (relative effective spread) on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3, +3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table A.15 Abnormal Returns around Jumps by Stock Order Flow (500 Largest Firms)

Full Sample	All Jumps				Positive Jumps				Negative Jumps			
Stock Order Flow Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$-0.2e-3$	0.04	$0.2e-3$	1,542	$-0.3e-3$	0.07	$0.2e-3$	1,184	$0.3e-3$	-0.05	$0.2e-3$	358
	(-1.3)	(20.27)	(1.48)		(-2.07)	(40.9)	(1.33)		(1.71)	(-20.82)	(0.66)	
10% - 25%	$-0.3e-3$	0.01	$-0.1e-3$	2,565	$-0.6e-3$	0.06	$-0.2e-3$	1,590	$0.2e-3$	-0.07	$0e-3$	975
	(-3.97)	(6.26)	(-0.99)		(-5.58)	(45.34)	(-1.39)		(1.41)	(-39.71)	(0.31)	
25% - 50%	$-0.4e-3$	$-0.1e-3$	$0e-3$	5,272	$-0.5e-3$	0.07	$0.1e-3$	2,672	$-0.3e-3$	-0.07	$-0.1e-3$	2,600
	(-6.81)	(-0.06)	(0.03)		(-5.56)	(63.15)	(0.83)		(-3.97)	(-64.22)	(-0.67)	
Bottom 50%	$-0.3e-3$	-0.01	$0.1e-3$	8875	$-0.2e-3$	0.07	$0.1e-3$	3,561	$-0.4e-3$	-0.06	$0.1e-3$	5,316
	(-7.18)	(-5.86)	(0.71)		(-3.11)	(68.78)	(1.24)		(-6.57)	(-66.25)	(0.36)	
Stock Order Flow Percentile	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N	$t-21:t-1$	t	$t+1:t+21$	N
Top 10%	$0e-3$	0.04	$0e-3$	1,172	$-0.1e-3$	0.06	$0e-3$	910	$0.5e-3$	-0.04	$0.2e-3$	262
	(0.11)	(19.78)	(0.18)		(-0.62)	(36.93)	(-0.12)		(2.01)	(-15.58)	(0.71)	
10% - 25%	$-0.4e-3$	0.02	$0.1e-3$	1,689	$-0.7e-3$	0.06	$-0.2e-3$	1,135	$0.1e-3$	-0.06	$0.6e-3$	554
	(-3.67)	(10.81)	(0.54)		(-4.44)	(33.73)	(-0.93)		(0.53)	(-23.81)	(2.66)	
25% - 50%	$-0.5e-3$	0.01	$-0.1e-3$	3,124	$-0.6e-3$	0.06	$0e-3$	1,688	$-0.5e-3$	-0.06	$-0.2e-3$	1,433
	(-6.14)	(3.97)	(-1.01)		(4.67)	(42.06)	(-0.31)		(-3.98)	(-40.36)	(-1.06)	
Bottom 50%	$-0.5e-3$	-0.01	$0e-3$	5,723	$-0.3e-3$	0.07	$0e-3$	1,916	$-0.5e-3$	-0.05	$0e-3$	3,807
	(-7.29)	(-11.02)	(-0.08)		(-3.12)	(40.47)	(0.05)		(-6.64)	(-49.03)	(-0.1)	

Notes: This table reports Fama-French-Cahart 4 factor abnormal returns (alphas) around the jump days identified using the $\{J_{99}\}$ test statistic. We report the average aggregate alphas for 21 days prior to and after the jump day, and average alpha on the jump day. We split into 4 different news categories based on the within-firm stock order imbalance on the jump day. The top panel reports results using entire sample and the bottom panel reports results excluding $[-3, +3]$ days around earnings announcement days. The sample period is from January 2000 to December 2020. t-stats are reported in the parentheses.

Table A.16 Probability of Daily Jumps (Top 500 Largest Companies)

	Panel A: Probit Coefficient Estimates			
Variable	(1)	(2)	(3)	(4)
Num News	0.44	0.45	0.44	0.45
	(48.9)	(49.05)	(48.95)	(49.08)
Avg. Sentiment	0.19	0.19	0.19	0.19
	(21.61)	(21.76)	(21.66)	(21.79)
$\|ret_{t-1}\|$	0.01	0.01	0.01	0.01
	(1.02)	(0.66)	(1.02)	(0.7)
RES_s		0.13		0.12
		(3.85)		(3.42)
OIB_s			0.07	0.07
			(5.34)	(5.02)
R^2	0.45	0.45	0.45	0.45
N obs	$1,853,956$	$1,853,956$	$1,853,956$	$1,853,956$

	Panel B: Odds Ratio			
Variable	(1)	(2)	(3)	(4)
Num News	1.56	1.56	1.56	1.56
Avg. Sentiment	1.21	1.21	1.21	1.21
$\|ret_{t-1}\|$	1.01	1.01	1.01	1.01
RES_s		1.14		1.13
OIB_s			1.08	1.07

This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ, as well as all interacted pairs of variables. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2000 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity. Reported results excluding $[-3, +3]$ days around earnings announcement days.

Table A.17 Effects of day and night news measures and stock illiquidity on probability of daily jumps (Top 500 Largest Companies)

Panel A: Probit Coefficient Estimates						
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	0.42		0.35	0.35	0.35	0.35
	(47.3)		(39.51)	(39.53)	(39.53)	(39.55)
Num News Day$_s$		0.25	0.2	0.2	0.2	0.2
		(42.53)	(35.02)	(35.13)	(35.07)	(35.16)
Avg. Sentiment Night$_s$	0.04		0.03	0.03	0.03	0.03
	(7.04)		(4.09)	(4.18)	(4.13)	(4.21)
Avg. Sentiment Day$_s$		0.15	0.12	0.12	0.12	0.12
		(36.09)	(27.48)	(27.51)	(27.42)	(27.45)
$\lvert ret_{t-1}\rvert$	0.04	0.04	0.02	0.02	0.02	0.02
	(3.76)	(3.84)	(2.25)	(1.94)	(2.26)	(1.98)
RES_s				0.11		0.1
				(3.14)		(2.74)
OIB_s					0.07	0.07
					(4.98)	(4.72)
R^2	0.31	0.28	0.48	0.48	0.48	0.48
N obs	$1,853,956$	$1,853,956$	$1,853,956$	$1,853,956$	$1,853,956$	$1,853,956$

Panel B: Odds Ratio						
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	1.53		1.42	1.42	1.42	1.42
Num News Day$_s$		1.29	1.23	1.23	1.23	1.23
Avg. Sentiment Night$_s$	1.04		1.03	1.03	1.03	1.03
Avg. Sentiment Day$_s$		1.16	1.13	1.13	1.13	1.13
$\lvert ret_{t-1}\rvert$	1.04	1.04	1.02	1.02	1.02	1.02
RES_s				1.12		1.1
OIB_s					1.07	1.07

Notes: This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2004 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity. Reported results excluding $[-3, +3]$ days around earnings announcement days.

Table A.18 Effects of news measures and stock illiquidity on probability of daily jumps

	Panel A: Probit Coefficient Estimates			
Variable	(1)	(2)	(3)	(4)
Num News	0.33	0.33	0.33	0.33
	(139.91)	(140.11)	(139.85)	(140.05)
Avg. Sentiment	0.08	0.08	0.08	0.08
	(24.96)	(24.9)	(24.78)	(24.72)
$\|ret_{t-1}\|$	0.09	0.08	0.09	0.08
	(48.39)	(42.2)	(48.61)	(42.27)
RES_s		0.41		0.42
		(51.39)		(52.62)
OIB_s			0.19	0.19
			(40.62)	(41.8)
R^2	6.22	6.42	6.4	6.6
N obs	16,349,414	16,349,414	16,349,414	16,349,414

	Panel B: Odds Ratio			
Variable	(1)	(2)	(3)	(4)
Num News	1.39	1.39	1.39	1.39
Avg. Sentiment	1.09	1.09	1.08	1.08
$\|ret_{t-1}\|$	1.1	1.09	1.1	1.09
RES_s		1.51		1.52
OIB_s			1.2	1.2

Notes: This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2004 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.19 Effects of news measures and stock illiquidity on probability of daily jumps (Largest 500 Firms)

	Panel A: Probit Coefficient Estimates			
Variable	(1)	(2)	(3)	(4)
Num News	0.45	0.45	0.45	0.45
	(92.01)	(92.06)	(91.98)	(92.02)
Avg. Sentiment	0.21	0.21	0.21	0.21
	(25.78)	(25.7)	(25.78)	(25.7)
$\|ret_{t-1}\|$	-0.08	-0.09	-0.08	-0.08
	(-7.51)	(-7.78)	(-7.44)	(-7.71)
RES_s		0.34		0.33
		(5.21)		(5.12)
OIB_s			0.07	0.07
			(5.99)	(5.84)
R^2	1.47	1.47	1.47	1.47
N obs	$1,689,651$	$1,689,651$	$1,689,651$	$1,689,651$

	Panel B: Odds Ratio			
Variable	(1)	(2)	(3)	(4)
Num News	1.57	1.57	1.57	1.57
Avg. Sentiment	1.23	1.23	1.23	1.23
$\|ret_{t-1}\|$	0.92	0.92	0.92	0.92
RES_s		1.4		1.39
OIB_s			1.08	1.08

Notes: This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2004 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.20 Effects of news measures and stock illiquidity on daily jump size.

	Panel A: All Jumps			
Variable	(1)	(2)	(3)	(4)
Num News	$-2.1e-3$	$-2.1e-3$	$-1.6e-3$	$-1.5e-3$
	(-5.28)	(-5.18)	(-3.88)	(-3.7)
Avg. Sentiment	0.01	0.01	0.01	0.01
	(23.42)	(23.48)	(22.52)	(22.63)
$\|ret_{t-1}\|$	0.03	0.03	0.01	0.01
	(1.45)	(1.48)	(0.78)	(0.84)
RES_s		$3.9e-3$		0.01
		(2.27)		(4.35)
OIB_s			0.07	0.07
			(84.79)	(84.76)
R^2	0.67	0.68	7.84	7.86
N obs	$201,858$	$201,858$	$201,858$	$201,858$

	Panel B: Positive Jumps			
Variable	(1)	(2)	(3)	(4)
Num News	0.01	0.01	0.01	0.01
	(10.7)	(11.55)	(10.89)	(11.57)
Avg. Sentiment	0.01	0.01	0.01	0.01
	(10.47)	(11.19)	(10.67)	(11.33)
$\|ret_{t-1}\|$	0.15	0.17	0.15	0.17
	(6.42)	(7.47)	(6.39)	(7.44)
RES_s		0.05		0.05
		(21.59)		(21.39)
OIB_s			0.01	$4.1e-3$
			(7.26)	(4.75)
R^2	2.67	3.78	2.72	3.8
N obs	$114,601$	$114,601$	$114,601$	$114,601$

	Panel C: Negative Jumps			
Variable	(1)	(2)	(3)	(4)
Num News	-0.01	-0.01	-0.01	-0.01
	(-39.91)	(-42.42)	(-36.57)	(-38.47)
Avg. Sentiment	-0.01	-0.01	-0.01	-0.01
	(-16.89)	(-17.02)	(-16.77)	(-16.94)
$\|ret_{t-1}\|$	-0.09	-0.08	-0.08	-0.08
	(-9.93)	(-10.08)	(-9.59)	(-9.62)
RES_s		-0.04		-0.05
		(-32.29)		(-35.41)
OIB_s			-0.01	-0.02
			(-24.21)	(-31.37)
R^2	9.23	12.37	10.05	13.73
N obs	$87,257$	$87,257$	$87,257$	$87,257$

This table reports results from regressions of daily jump sizes, conditional on the jump indicator being 1, on daily news measures as well as stock illiquidity and order imbalances from TAQ for the all firms in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2004 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panels B and C report the results for positive jump sizes and negative jump sizes, respectively. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.21 Effects of news measures and stock illiquidity on daily jump size. (500 Largest Firms)

	Panel A: All Jumps			
Variable	(1)	(2)	(3)	(4)
Num News	$-0.2e-3$	$-0.2e-3$	$-0.3e-3$	$-0.2e-3$
	(-0.65)	(-0.63)	(-0.81)	(-0.79)
Avg. Sentiment	0.01	0.01	0.01	0.01
	(17.08)	(17.07)	(16.79)	(16.78)
$\lvert ret_{t-1} \rvert$	0.03	0.03	0.02	0.02
	(0.28)	(0.3)	(0.2)	(0.22)
RES_s		0.01		0.01
		(1.51)		(1.42)
OIB_s			0.01	0.01
			(7.57)	(7.53)
R^2	2.24	2.28	2.68	2.71
N obs	$16,047$	$16,047$	$16,047$	$16,047$

	Panel B: Positive Jumps			
Variable	(1)	(2)	(3)	(4)
Num News	$4.8e-3$	$4.8e-3$	$4.6e-3$	$4.7e-3$
	(16.06)	(16.32)	(15.62)	(15.87)
Avg. Sentiment	$4.1e-3$	$4.1e-3$	$4.1e-3$	$4e-3$
	(8.5)	(8.5)	(8.35)	(8.35)
$\lvert ret_{t-1} \rvert$	-0.4	-0.39	-0.4	-0.39
	(-3.87)	(-4.03)	(-3.84)	(-4)
RES_s		0.05		0.05
		(4.32)		(4.33)
OIB_s			-0.01	-0.01
			(-4.95)	(-5.2)
R^2	12.41	14.39	12.66	14.66
N obs	$7,884$	$7,884$	$7,884$	$7,884$

	Panel C: Negative Jumps			
Variable	(1)	(2)	(3)	(4)
Num News	$-4.9e-3$	$-4.9e-3$	$-4.7e-3$	$-4.7e-3$
	(-18.82)	(-18.81)	(-17.77)	(-17.78)
Avg. Sentiment	$-4.3e-3$	$-4.3e-3$	$-4.1e-3$	$-4.2e-3$
	(-7.36)	(-7.46)	(-7.08)	(-7.18)
$\lvert ret_{t-1} \rvert$	0.45	0.45	0.46	0.45
	(4.74)	(4.7)	(4.79)	(4.75)
RES_s		-0.04		-0.04
		(-4.81)		(-4.73)
OIB_s			-0.01	-0.01
			(-7.54)	(-7.21)
R^2	14.52	15.87	15.07	16.37
N obs	$8,163$	$8,163$	$8,163$	$8,163$

This table reports results from regressions of daily jump sizes, conditional on the jump indicator being 1, on daily news measures as well as stock illiquidity and order imbalances from TAQ for the largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2004 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panels B and C report the results for positive jump sizes and negative jump sizes, respectively. The jump variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.

Table A.22 Effects of day and night news measures and stock illiquidity on probability of daily jumps (500 Largest Firms)

Panel A: Probit Coefficient Estimates						
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	0.43		0.37	0.37	0.37	0.37
	(92.69)		(77.43)	(77.57)	(77.5)	(77.61)
Num News Day$_s$		0.27	0.18	0.18	0.18	0.18
		(53.38)	(34.68)	(34.33)	(34.39)	(34.36)
Avg. Sentiment Night$_s$	0.04		0.03	0.03	0.03	0.03
	(7.36)		(4.07)	(4.14)	(4.08)	(4.13)
Avg. Sentiment Day$_s$		0.19	0.11	0.11	0.11	0.11
		(46.6)	(26.57)	(26.4)	(26.34)	(26.38)
$\|ret_{t-1}\|$	-0.04	-0.04	-0.06	-0.06	-0.05	-0.06
	(-3.58)	(-3.44)	(-5)	(-5.21)	(-4.86)	(-5.13)
RES_s				0.38		0.37
				(5.84)		(5.75)
OIB_s					0.07	0.07
					(5.37)	(5.22)
R^2	1.42	0.65	1.63	1.63	1.63	1.63
N obs	$1,643,421$	$1,643,421$	$1,643,421$	$1,630,635$	$1,630,632$	$1,630,632$

Panel B: Odds Ratio						
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Num News Night$_s$	1.54		1.45	1.45	1.45	1.45
Num News Day$_s$		1.31	1.19	1.19	1.19	1.19
Avg. Sentiment Night$_s$	1.04		1.03	1.03	1.03	1.03
Avg. Sentiment Day$_s$		1.21	1.12	1.12	1.12	1.12
$\|ret_{t-1}\|$	0.96	0.97	0.95	0.94	0.95	0.94
RES_s				1.46		1.45
OIB_s					1.07	1.07

Notes: This table reports results from pooled logistic regressions of daily jump indicators, defined using Lee and Mykland (2008), on daily news measures as well as stock illiquidity and order imbalances from TAQ for largest 500 firms each month in the sample. The explanatory variables, which are standardized to have the same mean and standard deviation across firms, are the total number of news articles reported on the Ravenpack database each day, the absolute value of news tone, the absolute value of the previous day's return, the stock relative effective spreads, and order imbalances from TAQ. The news tone measure is constructed by Ravenpack using a proprietary framework resulting in a numerical score. The sample period is from January 2004 to February 2020. The t-statistics, computed using standard errors clustered at individual firm levels, are reported in parentheses. Panel B reports the odds ratios associated with each variable in brackets. The dependent variable ($\{J_{99}\}$) is the same for each regression where $\{J_{99}\}$ identifies a jump day if the absolute value of daily return is above $\{5.1024\}$ times the time-varying daily spot volatility. All regression specifications include a constant term that is not reported for brevity.