Journal of Banking & Finance 37 (2013) 1148–1159

Contents lists available at SciVerse ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Public information arrival: Price discovery and liquidity in electronic limit order markets

Ryan Riordan [a,*], Andreas Storkenmaier [b], Martin Wagener [b], S. Sarah Zhang [b]

[a] University of Ontario Institute of Technology, Faculty of Business and Information Technology, 2000 Simcoe Street North, Oshawa, Ont., Canada L1H 7K4
[b] Karlsruhe Institute of Technology, Institute of Information Systems and Management, Englerstr. 14, 76131 Karlsruhe, Germany

ARTICLE INFO

Article history:
Received 4 April 2011
Accepted 5 November 2012
Available online 20 November 2012

JEL classification::
G10
G14

Keywords:
Information
Liquidity
Price discovery
News
Limit order markets

ABSTRACT

How information is translated into market prices is still an open question. This paper studies the impact of newswire messages on intraday price discovery, liquidity, and trading intensity in an electronic limit order market. We take an objective ex ante measure of the tone of a message to study the impacts of positive, negative, and neutral messages on price discovery and trading activity. As expected, we find higher adverse selection costs around the arrival of newswire messages. Negative messages are associated with higher adverse selection costs than positive or neutral messages. Liquidity increases around positive and neutral messages and decreases around negative messages. Available order book depth as well as the trading intensity increases around all news. Our results suggest that market participants possess different information gathering and processing capabilities and that negative news messages are particularly informative and induce stronger market reactions.

1. Introduction

Technology and computers have not only changed trading in financial markets, but have also revolutionized the way financial news is disseminated and analyzed. As trading technology has advanced, news providers have kept pace and deliver news to traders around the world within a fraction of a second. News providers have also started to offer newswire products with machine-learning systems that cater to algorithmic traders. However, most news is still read by professional human traders who read newswires such as Thomson Reuters, Bloomberg, or Dow Jones on a regular basis. They spend a considerable amount of time and money on these information sources and emphasize the importance of the speed and accuracy of their news. Newswire messages represent much of the overall information and real-time information traders receive. The intraday impact of newswire messages is however still not well understood. It is not entirely clear whether newswire messages actually contain new information, whether traders act in advance or after the arrival of messages, and how newswire messages impact liquidity and price dynamics in a modern electronic limit order market.

This paper studies the impact of Thomson Reuters newswire messages on the intraday price discovery, trading activity, and liquidity of stocks traded on the Toronto Stock Exchange. The Toronto Stock Exchange is well suited for such an analysis. First, it is a highly automated electronic limit order book market comparable to many international exchanges. Second, in contrast to most European markets, there are no major second language news streams. Third, the Canadian market has a very low level of fragmentation during our observation period.

In this paper, newswire messages are clustered by sentiment. The differentiation between positive, negative, and neutral news enables us to investigate potentially asymmetric reactions to newswire messages based on their tone. Liquidity increases around positive and neutral messages and decreases around negative messages. Trading intensity increases around all types of newswire messages. In general, we find higher adverse selection costs around newswire messages. Negative messages are associated with significantly higher adverse selection costs than positive messages.

Traditional financial theory does not differentiate between positive and negative public information. However, psychological studies from the field of impression formation show that humans react stronger to bad news than to good news (cf. Soroka, 2006). Overall, our results suggest that participants' possess different information

* Corresponding author. Tel.: +1 905 721 8668x5397; fax: +1 905 721 3178.
E-mail addresses: ryan.riordan@uoit.ca (R. Riordan), andreas.storkenmaier@kit.edu (A. Storkenmaier), martin.wagener@kit.edu (M. Wagener), sarah.zhang@kit.edu (S. Sarah Zhang).

0378-4266/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbankfin.2012.11.008

gathering and information processing capabilities and that these participants react differently to good and bad news.

The remainder of the paper is structured as follows. Section 2 presents related literature. Section 3 gives an overview on the institutional structure of the Toronto Stock Exchange. Section 4 provides details on newswire messages, trade and order book data, and the sample selection. Section 5 introduces the research design and methodology. Section 6 provides the results and interpretation and Section 7 concludes.

2. Related work

The existing public information literature studies different types of public information, from unstructured media content to scheduled earnings announcements. The Thomson Reuters newswire messages are in-between these extremes. Ranaldo (2008) analyzes the intraday market dynamics of firm specific unstructured news at the Paris Bourse. The six months of news data is based on the Reuters alert system without ex ante measures news sentiment (i.e. positive, negative, or neutral). He finds a marginally significant increase in liquidity and slightly higher adverse selection costs around news arrivals. Groß-Klußmann and Hautsch (2011) use a news data set similar to ours in a study of high-frequency returns and profitability. They find that "high-frequency trading activity indeed significantly reacts to intraday company-specific news items".

Krinsky and Lee (1996) analyze the impact of scheduled earnings announcements at the NYSE and AMEX. They find that the adverse selection component of the spread increases around earnings announcements. The authors attribute this effect to temporary information advantages of informed investors and to faster news processing capabilities of public information processors. In a study of macroeconomic announcements on US government bond trading, Green (2004) shows higher adverse selection costs around macroeconomic news releases. US government bond trading is organized as a dealer market which might yield different results than a public limit order market. Berry and Howe (1994) also analyze the intraday impact of public information arrivals. Their proxy for public information is the number of news releases by Thomson Reuters' news service per unit of time. Their results suggest a "positive, moderate relationship between public information and trading volume, and no relationship with price volatility" (Berry and Howe, 1994). Algorithmic traders in the foreign exchange market monitor macroeconomic news and reduce their supply of liquidity directly after economic news arrivals to being adversely selected (Chaboud et al., 2009).

Kim and Verrecchia (1991) formulate a theoretical model that explains higher adverse selection costs prior to an scheduled announcement such as earnings announcements. Traders acquire costly private information to trade in advance of a public announcement. Kim and Verrecchia (1994) introduce the notion that different traders have varying capabilities to interpret earnings announcements. This might lead to an increase in adverse selection costs after earnings announcements due to higher information asymmetry. However, trading volume might still increase despite a decrease in liquidity around earnings announcements due to informed traders' willingness to pay the spread. Another theoretical model developed by Harris and Raviv (1993) attributes effects around the announcement of public information to speculative trading. Traders disagree as a result of asymmetric private information or different information processing capabilities which leads to a surge in market activity.

One challenge in the analysis of public information is the transformation of ambiguous news and media content to variables that can be studied in econometric models. Several papers analyze ambiguous content and study its impact on financial markets. Newspaper content is one of the most frequently studied types of media content. Niederhoffer (1971) provides one of the earliest papers that analyzes media content. He shows that world events are followed by larger price changes than under normal market conditions. Tetlock (2007) shows that high pessimism in the WSJ column "Abreast of the Market" is followed by lower market prices and thereafter by a reversal to fundamentals.

In a study of Internet stock message board postings, Antweiler and Frank (2004) find that an increase in postings correlates with an increase in volatility. News media also affects individual buyers' perception of and their attention towards specific stocks (Barber and Odean, 2008). Individual buyers are more prone to buy stocks which have drawn their attention through media outlets. All these studies have in common that they quantify ambiguous media content or otherwise derive quantitative information from qualitative linguistic messages.

For our analyses, we use objectively quantified news messages from Thomson Reuters. First, we study how newswire messages affect price discovery, trading activity, and liquidity. Theory and empirical literature suggests an increase in trading activity around news arrivals. Empirical evidence for liquidity and adverse selection is mixed. Financial theory however suggests an increase in adverse selection costs and a reduction of liquidity. Second, we investigate how trading activity, liquidity, and adverse selection interact around news messages.

3. Institutional details

The Toronto Stock Exchange (TSX) is Canada's most important equity exchange and is operated by the TMX Group.[1] The TSX is North America's third largest equity exchange by trading volume after Nasdaq and the New York Stock Exchange.[2] We use prices on the TSX to calculate the S&P/TSX 60 index, Canada's blue chip stock market index maintained by Standard and Poors.

The TSX operates an entirely electronic market with a centralized public limit order book. The market features basic limit and market orders. The TSX market model is based on price and time priority. Iceberg orders that display only a portion of their total size are available for a minimum of 500 shares. They sacrifice time priority on the non-displayed portion of the order. Liquidity is provided by public limit orders displayed in the order book.

4. Data and sample selection

Our news data consists of Thomson Reuters NewsScope Content and is tagged using the Thomson Reuters NewsScope Sentiment Engine (RNSE).[3] The RNSE real-time data stream is disseminated to approximately 370,000 Reuters screens worldwide. According to Thomson Reuters, they "deliver over 500,000 alerts and over two million unique stories a year".[4] The RNSE data provide three pieces of automatically generated information: *Sentiment*, *Relevance*, and Novelty.[5] *Sentiment* measures the stock specific tone of a news item

[1] Alternative trading systems do not play an important role during our observation period. The TSX's market share by trading volume was still 94.2% in January 2009, directly after the end of our observation period, and close to 100% one year earlier. (*Source:* Financial Times, 20 November 2009, "Toronto's trading platforms draw regulatory scrutiny".)

[2] World Federation of Exchanges, 2008, http://www.world-exchanges.org/statistics.

[3] We thank Thomson Reuters for providing access to Thomson Reuters NewsScope Sentiment Engine archive data.

[4] Thomson Reuters, http://www.online.reuters.com/productinfo/newsscoperealtime.

[5] See further: Reuters Newsscope Sentiment Engine: Guide to sample data and system overview and Reuters Newsscope Sentiment Engine: Output image and file format.

and is either positive, negative, or neutral. The sentiment measure can take the values of −1 for negative sentiment, 0 for neutral sentiment, and +1 for positive sentiment. *Relevance* measures how important a message is to a specific firm. *Relevance* is defined on the interval [0, 1] and increases in the relevance of the news message for the specific stock. Novelty measures whether news with the same content has been previously released and allows us to remove stale news. Messages are tagged by stock. For example, a news message that is positive for 'Research in Motion' might be negative for 'Rogers Communications' and may be more relevant for 'Research in Motion' than 'Rogers Communications'. Table 1 reports one sample RNSE newswire message for 'Research in Motion'.

The data are cleaned as follows. We delete news that link to a news message with similar content during the last 24 h. This ensures that news are relatively novel. We further delete messages that arrive outside of the continuous trading period. Overall, we have 6625 novel intraday news messages over the sample period. The important measures for our analyses are *Sentiment* and *Relevance*.

Fig. 1 plots the number of news over the observation period and Fig. 2 plots the distribution of news over weekdays.

We retrieve trade and quote (TAQ) data and order book data from the Thomson Reuters DataScope Tick History archive through SIRCA.[6] Our sample period ranges from January 1, 2005 to December 31, 2008 for S&P/TSX 60 constituents. Trades and quotes are timestamped to the millisecond. All prices are reported in Canadian dollars. We remove the first and last five minutes of a trading day to avoid biases associated with the opening and closing processes. We also delete crossing trades and on-close orders.

Sample firms in the S&P/TSX 60 index are required to have at least 20 distinct newswire messages per year and at least 10 newswire messages per sentiment category over the years 2005–2008. The final sample consists of 33 highly liquid, actively traded S&P/TSX 60 constituents. Table 2 provides an overview of the sample and descriptive statistics for each firm in the sample.

Table 1
Sample news. Table 1 shows one novel intraday RNSE news message for the firm Research in Motion (RIM.TO).

Sample RNSE news item	
timestamp	24 OCT 2007 16:30:02.064
bcast_ref	RIM.TO
stock_ric	RIM.TO
item_id	2007-10-24_16.30.01.nN24487523.T1.8da5a8b6
relevance	0.150756
sentiment	1
sent_pos	0.559651
sent_neut	0.358283
sent_neg	0.0820664
lnkd_cnt1	0
lnkd_cnt2	0
lnkd_cnt3	0
lnkd_cnt4	0
lnkd_cnt5	0
lnkd_id1	.
...	
lnkd_idpv1	.
...	
item_type	ARTICLE
item_genre	NOT DEFINED
bcast_text	RIM rolls out Facebook software for BlackBerry
dsply_name	2
pnac	nN24487523
story_type	S
cross_ref	.
proc_date	24-OCT-2007
take_time	16:30:01
story_date	24-OCT-2007
story_time	16:30:01
named_item	.
take_seqno	1
attribtn	RTRS
prod_code	E U CAN G PSC RNP DNP PGE PCO PCU EMK
topic_code	BUS CA US DE INV TEL WWW SFWR HDWR ENT LEI TEEQ TECH COMS ELC CEEU EUROPE WEU LEN RTRS
co_ids	RIMM.O RIM.TO DT.N
lang_ind	EN

5. Research design

5.1. Price discovery

There are several methodologies to decompose stock price movements around information events into various components. The most common models are the MRR model introduced by Madhavan et al. (1997) and the model by Huang and Stoll (1997). More recent developments employ a variance decomposition into information-driven and noise-induced volatility. This is applied for macroeconomic news using a state-space model as in Hautsch et al. (2011) and rather concentrates on volatility-based price discovery in contrast to our analyses.

We modify the MRR model by Madhavan et al. (1997) and extend it in the spirit of Green (2004) by including variables for thirty minute intervals around a news event.[7] The MRR model is a common market microstructure model to approximate the adverse selection component of the bid-ask spread, the information revealed through trades. The adverse selection component measures the portion of the bid-ask spread that is required as compensation for liquidity suppliers' risk of losing to informed traders. Hence, the adverse selection component of the bid-ask spread can be interpreted as private information that is impounded into prices through trading. The MRR also captures an inventory and order processing cost component of the bid-ask spread.[8]

We prefer the Madhavan et al. (1997) to the Huang and Stoll (1997) model. The Huang and Stoll (1997) model offers an appealing decomposition of order processing costs into processing and inventory costs. Like the MRR model, it relies on the estimation of the autocorrelation of order flow on a market maker market. The downside is that modern limit order markets, in contrast to specialist or market maker markets, exhibit positive autocorrelation at high frequencies, which may cause the estimation of negative inventory costs. To rectify this, data can be aggregated to lower frequencies at which order flow is negatively autocorrelated, as in market maker markets, but this requires the econometrician to select an appropriate level of aggregation without undue loss of information or statistical power. Given that inventory costs are typically attributed to market makers and it is unclear what role they play in modern limit order markets, the use of the MRR model appears to be the best choice.

We specify the MRR model as follows. Let x_t be the directed order flow, 1 for a market buy order and −1 for a market sell order, at time t. p_t denotes the transaction price. t is a single observation in the trade process. The Madhavan et al. (1997) model is then:

$$p_t - p_{t-1} = (\phi + \theta)x_t - (\phi + \rho\theta)x_{t-1} + \epsilon_t, \tag{1}$$

[6] We thank SIRCA (Securities Industry Research Centre of Asia Pacific) for providing access to the Thomson Reuters DataScope Tick History archive, http://www.sirca.org.au/.

[7] We also break the pre and post periods up into 15 min intervals and arrive at similar results.

[8] We use the common Lee and Ready (1991) algorithm to sign trades with contemporaneous quotes. For the TSX, a limit order market without executions inside the spread, this algorithm signs trades without ambiguity if trade and quote timestamps match.

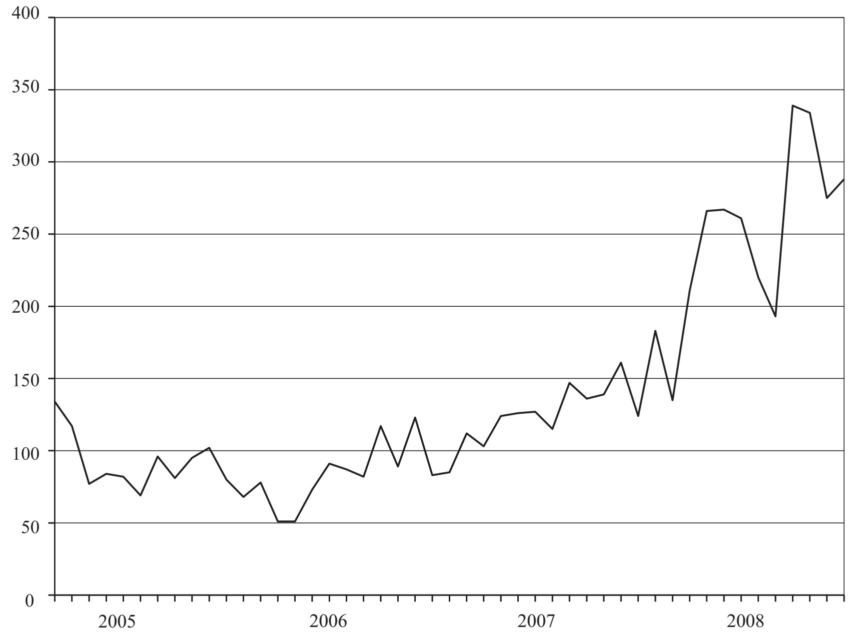

Fig. 1. Novel Intraday News Per Year and Month 2005–2008. The figure shows the number of novel intraday news messages per year and month.

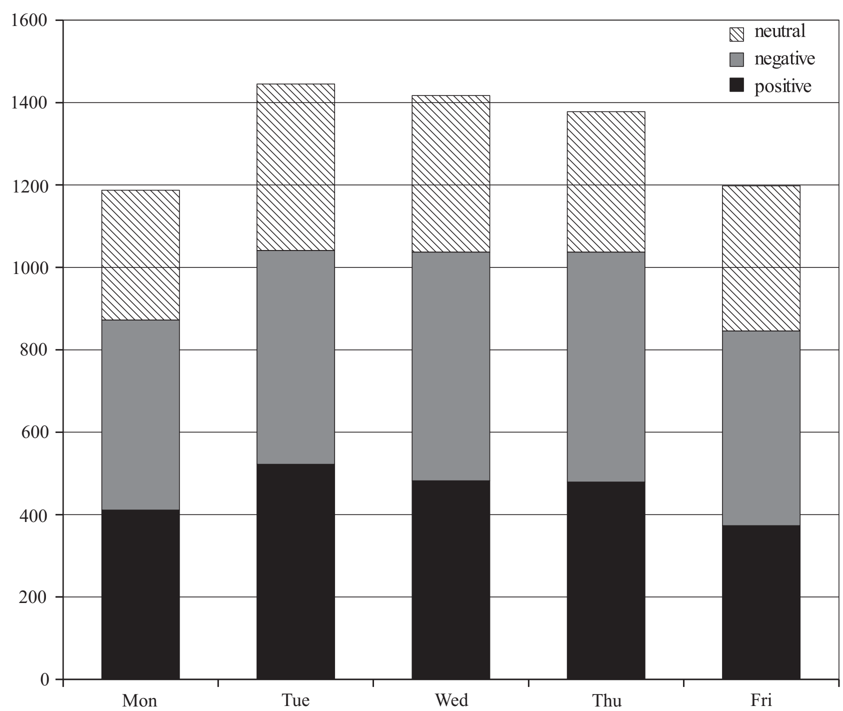

Fig. 2. Novel Intraday News Per Weekday 2005–2008. The figure shows the number of novel intraday news messages per day of the week.

where ρ is the first-order autocorrelation, θ is the asymmetric information component, and ϕ captures inventory and order processing costs.[9] The assumption in the model is deviations from expected order flow are information driven. Expected order flow ρx_{t-1} is based on order flow autocorrelation ρ. θ measures "the degree of information asymmetry or the so-called permanent impact of the order flow innovation" (Madhavan et al., 1997). Beliefs about asset values might change through new public information without trading and through information impounded by the order flow whereas the change in belief is positively correlated with order flow innovation. Inventory and order processing costs represent the transitory effect of order flow on prices. The inventory and order processing cost components are independent of trade direction. As such, ϕ is independent of the trade process' autocor-

[9] The original estimation of the model also includes λ in its moment conditions, the probability of an inside the spread execution which is identified through a trade direction of zero. This component is not estimated for the TSX as this market does not feature inside the spread execution. The TSX features a completely electronic order book without inside the spread executions.

Table 2
Descriptive statistics for sample companies. The sample is based on stocks continuously listed in the S&P/TSX 60 index between 2005 and 2008. Thirty-three stocks qualify for the sample after filtering based on newswire data. Table 2 reports descriptive statistics for the number of news, news sentiment, market value, and economic sector. News measures are derived from Thomson Reuters RNSE data wheres market value and economic sector are based on Compustat data. The average per company sentiment is denoted 'Sent'. The overall number of news (#) as well as the number of news differentiated by sentiment are reported. 'MVal' stands for the average market value in Million Candian dollars. The table is sorted by the descending number of news per company in the analysis.

Company name	#News	#Pos	#Neg	#Neut	Sent	MVal	Sector
Barrick Gold	518	156	169	193	−0.0251	30,944	Materials
Research in Motion	366	112	181	71	−0.1896	32,982	IT
Royal Bank of Canada	361	125	134	102	−0.0249	62,805	Financials
EnCana	302	102	129	71	−0.0894	45,014	Energy
Toronto-Dominion Bank	290	103	135	52	−0.1103	45,919	Financials
Nortel Networks	287	86	116	85	−0.1045	8922	IT
Bank of Nova Scotia	270	119	83	68	0.1333	45,957	Financials
Goldcorp	252	68	67	117	0.0040	21,011	Materials
Canadian Imperial Bank	247	59	135	53	−0.3077	27,133	Financials
BCE	239	92	100	47	−0.0335	25,807	TelCo
Petro-Canada	215	74	75	66	−0.0047	21,602	Energy
Bank of Montreal	205	70	84	51	−0.0683	29,202	Financials
Suncor Energy	205	68	89	48	−0.1024	36,976	Energy
Cameco	201	87	41	73	0.2289	21,713	Energy
Potash Corp. of Sask.	193	69	89	35	−0.1036	24,745	Materials
Canad. Natural Res.	179	59	83	37	−0.1341	32,472	Energy
Canad. National Railway	178	52	94	32	−0.2360	23,564	Industrials
Bombardier	175	69	58	48	0.0629	7030	Industrials
Teck Resources	175	70	67	38	0.0171	12,575	Materials
Imperial Oil	162	57	54	51	0.0185	40,966	Energy
Enbridge	158	51	69	38	−0.1139	14,086	Energy
Telus	145	52	68	25	−0.1103	15,428	TelCo
TransCanada	143	87	39	17	0.3357	20,011	Energy
Agrium	139	53	46	40	0.0504	6510	Materials
Kinross Gold	134	30	39	65	−0.0672	8689	Materials
Nexen	133	37	52	44	−0.1128	14,857	Energy
Talisman Energy	131	44	36	51	0.0611	18,680	Energy
Manulife Financial	118	43	44	31	−0.0085	52,332	Financials
National Bank of Canada	118	40	52	26	−0.1017	8876	Financials
Magna International	114	40	50	24	−0.0877	8233	Consumer Discr.
Rogers Communications	96	38	38	20	0.0000	22,898	TelCo
Yamana Gold	91	29	13	49	0.1758	5573	Materials
Agnico-Eagle Mines	87	26	37	24	−0.1264	6385	Materials
Mean	201	69	78	54	−0.0356	23,967	
Standard deviation	93	31	41	34	0.1278	14,965	
Median	178	68	68	48	−0.0672	21,620	
Minimum	87	26	13	17	−0.3077	5573	
Maximum	518	156	181	193	0.3357	62,805	
Sum	6625	2267	2566	1792		790,914	

relation ρ. The model assumes a fixed order size and does not consider volume.[10]

To analyze information around news, we introduce three dummy variables I_1 and I_2 for 30 min intervals around a Thomson Reuters news message and I_3 for all other trading periods:

$$I_{n,t} = \begin{cases} 1 & \text{n} = 1: \text{ if } t \text{ is 30 min before a news}, \\ & \text{n} = 2: \text{ if } t \text{ is 30 min after a news}, \\ 0 & \text{else}. \end{cases} \tag{2}$$

We also create dummy variables for news depending on the sentiment of the news. One dummy variable represents positive news, one negative news, and one neutral news. Periods without news are assigned no sentiment dummy variable.

$$S_{m,t} = \begin{cases} 1 & \text{if sentiment of news is positive } (m=1), \\ & \text{negative } (m=2), \text{ or neutral } (m=3), \\ 0 & \text{else}. \end{cases}$$

We estimate the following extended model with dummy variables for different time intervals and different news types:

$$\begin{aligned} p_t - p_{t-1} = \sum_{n=1}^{2}\sum_{m=1}^{3}\big[&(\phi_{n,m}+\theta_{n,m})x_t I_{n,t} S_{m,t} \\ &- (\phi_{n,m}+\rho_{n,m}\theta_{n,m})x_{t-1} I_{n,t-1} S_{m,t-1}\big] \\ &+ (\phi_3+\theta_3)x_t I_{3,t} - (\phi_3+\rho_3\theta_3)x_{i-1} I_{3,t-1} + \epsilon_t. \end{aligned} \tag{3}$$

In contrast to Madhavan et al. (1997), we use relative price changes in basis points to estimate Eq. (3) excluding overnight returns. The results are robust to using absolute or relative price changes. We estimate the model using generalized methods of moments (GMM) with Newey and West (1987) standard errors.[11] Let r_t denote the relative price change, which is calculated as r_t = 10,000 × ln (p_t/p_{t-1}), then the model can be estimated as:

$$\begin{aligned} u_t = r_t + \sum_{n=1}^{2}\sum_{m=1}^{3}[&-(\phi_{n,m}+\theta_{n,m})x_t I_{n,t} S_{m,t} + (\phi_{n,m} \\ &+ \rho_{n,m}\theta_{n,m})x_{t-1} I_{n,t-1} S_{m,t-1}] - (\phi_3+\theta_3)x_t I_{5,t} + (\phi_3 \\ &+ \rho_3\theta_3)x_{t-1} I_{3,t-1} - \sum_{td=1}^{12}\tau_{td} T_{td} - \sum_{wd=1}^{4}\omega_{wd} W_{wd} \end{aligned} \tag{4}$$

[10] Madhavan et al. (1997) find that trade direction is more reflective of information than signed volume.

[11] We apply the Newey–West estimator with five lags. Using more lags does not change the significance of our results.

with half hour dummy variables T_{td} for the time of day and dummy variables for the day of the week W_{wd}.[12] Let α be a constant and

$$v_t = x_t - x_{t-1}\left(\sum_{n=1}^{2}\sum_{m=1}^{3} I_{n,t}S_{m,t}\rho_{n,m} + \sum_{m=1}^{3} I_{3,t}\rho_{3,m}\right). \quad (5)$$

Then the parameter vector:

$$\beta = (\theta_{1,1},\theta_{1,2},\theta_{1,3},\theta_{2,1},\theta_{2,2},\theta_{2,3},\theta_3,\phi_{1,1},\phi_{1,2},\phi_{1,3},\phi_{2,1},\phi_{2,2},\phi_{2,3},\phi_3,\rho_{1,1},\rho_{1,2},\rho_{1,3},\rho_{2,1},\rho_{2,2},\rho_{2,3},\rho_3)$$

is exactly identified using OLS with the following additional moment conditions:

$$E\begin{pmatrix} v_t I_{1,t-1}S_{1,t}x_{t-1} \\ v_t I_{1,t-1}S_{2,t}x_{t-1} \\ v_t I_{1,t-1}S_{3,t}x_{t-1} \\ v_t I_{2,t-1}S_{1,t}x_{t-1} \\ v_t I_{2,t-1}S_{2,t}x_{t-1} \\ v_t I_{2,t-1}S_{3,t}x_{t-1} \\ v_t I_{3,t-1}x_{t-1} \end{pmatrix} = 0. \quad (6)$$

The moment conditions in Eq. (6) capture the autocorrelation in the trade direction. The estimation of Eq. (4) gives results for asymmetric information θ, the inventory and order processing cost or cost of supplying liquidity ϕ, and the autocorrelation of order flow ρ for each t.

To assess the model's statistical significance, we use the likelihood ratio tests as in Green (2004). These tests compare the GMM criterion function of the unrestricted model with restricted models. The tests reject the null at all conventional levels and are suppressed for brevity. To consider robustness, we also compare model implied spreads with actual quoted spreads from the data. Since model implied spreads are solely based on the order flow, they do not necessarily need to be exactly the same as data based quoted spreads. However, they should be roughly similar in their order of magnitude. The medians of the differences of individual model coefficients of the 33 stock specific models are compared with a Wilcoxon signed rank test.

5.2. Trading intensity, liquidity, and volatility

To measure trading intensity, we transform the trade process into a process with one observation per minute and calculate the number of trades per minute, number of shares traded per minute, and traded dollar volume per minute. For estimation purposes, we use the natural logarithms of the number of shares traded per minute and traded dollar volume per minute. For the news dummy variable definition, we resort to the MRR information model definition of dummy variables and use exactly the same. The no-news variable does not need to be included, it is the basis of comparison and coefficients capture the difference to no-news periods. We include time dummy variables γ_{qy} or each quarter for a specific year into all regressions to account for market trends. In our data, we have 16 year quarter combinations. We also include firm dummy variables F_x with $x \in \{1, 2, \ldots, 33\}$ where x denotes a single firm. In the regression, one firm serves as the base category which results in 32 firm dummy variables. Additionally, the equation includes half hour dummy variables T_{td} for the time of day and dummy variables for the day of the week W_{wd}. Let l denote the minutes in the data and $M_{x,l}$ denotes the respective trading intensity measure on a minute and per firm basis then the following model is used to assess trading intensity around news messages:

$$M_{x,l} = a + \sum_{n=1}^{2}\sum_{m=1}^{3}\psi_{n,m}I_{n,x,l}S_{m,x,l} + \sum_{x=1}^{32}\iota_x F_x + \sum_{qy=1}^{15}\zeta_{qy}\gamma_{qy} + \sum_{td=1}^{12}\tau_{td}T_{td} + \sum_{wd=1}^{4}\omega_{wd}W_{wd} + e_{x,l}. \quad (7)$$

To estimate the pooled regression model in Eq. (7), we again use the Newey–West procedure based on five lags.

Quote based, ex ante observable, liquidity is measured based on three different measures: quoted half spread, the volume at the best bid and ask, and the volume at three order book depth levels. All three liquidity measures are based on a quote-to-quote process, which is then aggregated to minute averages for estimation purposes. Let a_t denote the best ask and b_t the best bid at time t, then quoted half spreads qs_t (as in Bessembinder and Kaufman, 1997) are calculated as follows in basis points:

$$qs_t = \left(\frac{a_t - b_t}{(a_t + b_t)/2}\right)\Big/ 2 \times 10{,}000. \quad (8)$$

Quoted spreads qs_t are aggregated to per firm and minute average quoted spreads $qs_{x,l}$.

To further analyze liquidity, we consider Canadian dollar volume at the best bid and ask (Depth0). Let again be a_t the best ask, b_t the best bid, an_t the number of shares available at the ask, and bn_t the number of shares available at the bid then the Canadian dollar volume at the best bid and ask eV_t is calculated as:

$$ev_t = bn_t \times b_t + an_t \times a_t. \quad (9)$$

Depth0 eV_t is aggregated to per firm and minute averages Depth0 $eV_{x,l}$. For estimation purposes we use the natural logarithm of Canadian dollar volume at the best bid and ask ($lev_{x,l} = \ln eV_{x,l}$). eV_t only provides information about volume directly at the spread. Order book data allows to analyze available volume deeper into the order book and liquidity which is used and needed for larger trades. We thus also consider depth at three levels (Depth3). Let $a_{t,dl}$ be the ask at time t on depth level dl, $b_{t,dl}$ denotes the bid on depth level dl, $an_{t,dl}$ is the Canadian dollar volume on a certain depth level at the ask, and $bn_{t,dl}$ denotes the volume at depth level dl on the bid. Then, the depth measure for three depth levels d_t is calculated as:

$$d_t = \sum_{dl=1}^{3} bn_{t,dl} \times b_{t,dl} + \sum_{dl=1}^{3} an_{t,dl} \times a_{t,dl}. \quad (10)$$

Again, Depth3 d_t is aggregated to per firm and minute average Depth3 $d_{x,l}$. As for volume at the best bid and ask, we use the natural logarithm of Depth3 for the estimation ($ld_{x,l} = \ln d_{x,l}$).

The effective spread, a trade process based liquidity measure, is the spread paid when an incoming market order trades against a limit order in the order book. The effective spread also captures institutional features of a market such as hidden liquidity through e.g. iceberg orders and market depth. Let p_t be the execution price and D_t the trade direction then the effective spread es_t is defined in basis points as:

$$es_t = D_t \times \frac{p_t - (a_t + b_t)/2}{(a_t + b_t)/2} \times 10{,}000. \quad (11)$$

The regression model from Eq. (7) is also used to analyze liquidity.

To estimate realized volatility also known as realized variance (cf. Hansen and Lunde, 2005), we construct one minute midpoint to midpoint returns from quote data. To calculate the square of returns for the realized volatility, we use logarithmic midpoint returns. Let mp_l denote a 1 min midpoint then the one minute realized volatility is defined as:

$$rv_l = \left(ln\frac{m_l}{m_{l-1}}\right)^2 \times 1{,}000{,}000. \quad (12)$$

[12] Excluding the dummy variables from the analysis does not significantly change the results.

Table 3
Descriptive statistics market measures. Table 3 provides descriptive statistics for the market measures over all companies in the sample. Descriptives are shown overall and for different news periods and no-news periods. 'QSpread' denotes the average quoted spread per minute whereas 'QSpreadT' denotes the average quoted spread per minute at trades. 'ESpread' denotes the average effective spread per minute. 'Depth0' is the average Canadian dollar volume at the best bid and ask and 'Depth3' is the dollar volume three levels into the order book. '#TradesMin' is the average number of trades per minute, '#SharesMin' the average numbers of shares traded per minute, and 'VolumeMin' the average Canadian dollar volume traded per minute. 'RV' represents realized volatility based on minute-to-minute midpoint returns. Spread measures are in basis points (bps). Measures are calculated for the years 2005–2008 over the whole sample.

	QSpread (in bps)	QSpreadT (in bps)	ESpread (in bps)	Depth0 (in C$1,000)	Depth3 (in C$1,000)	#TradesMin	#SharesMin	VolumeMin (in C$1,000)	RV
Overall									
Mean	4.7422	3.7639	3.8118	132	416	9.70	5746	234	0.0120
StdDev	4.7786	4.3977	4.4004	116	404	11.97	17,197	352	0.1103
No news									
Mean	4.7480	3.7681	3.8159	132	415	9.53	5667	230	0.0118
StdDev	4.7808	4.3926	4.3960	116	401	11.73	17,025	345	0.1104
Before news (positive)									
Mean	4.4162	3.5294	3.5746	142	458	13.55	8188	343	0.0171
StdDev	4.0826	3.8690	3.8552	143	523	15.99	23,887	498	0.1338
After news (positive)									
Mean	4.3673	3.4886	3.5362	143	461	13.16	7906	335	0.0112
StdDev	3.9593	3.7472	3.7556	135	488	15.73	21,781	516	0.0541
Before news (negative)									
Mean	4.7265	3.7557	3.8097	121	399	16.93	8343	357	0.0228
StdDev	5.5519	5.4426	5.4032	115	465	19.16	21,407	490	0.1412
After news (negative)									
Mean	4.6377	3.7116	3.7622	126	411	15.97	8090	341	0.0163
StdDev	5.5925	5.5496	5.5166	122	522	17.89	21,126	484	0.0914
Before news (neutral)									
Mean	4.5722	3.6538	3.7091	145	476	15.57	8744	330	0.0176
StdDev	4.1409	3.9847	3.9928	127	537	17.76	22,770	441	0.1190
After news (neutral)									
Mean	4.4957	3.6117	3.6653	148	476	16.08	9503	373	0.0152
StdDev	3.7754	3.5745	3.5820	134	516	18.90	21,084	487	0.0791

The original realized volatility measure is multiplied by 1,000,000 to enhance readability of the numbers.[13] Scaling realized volatility by 1,000,000 does not change its statistical properties. The same regressions as above are used to analyze realized volatility around newswire messages for news with different sentiments. All measures are winsorized at 0.1% and 99.9% to account for potential extreme values through technical data recording errors.

In order to test for robustness of our results and to correct for cross/auto-correlations of variables, we apply a vector-autoregressive (VAR) model. We follow the VAR model structure of Groß-Klußmann and Hautsch (2011) who use a VAR model which includes the 6 market activity variables, specifically *Money value*, *Volatility*, *Absolute trade imbalance*, *Depth*, *Spread*, and *Average trade size*. We include one variable in each of our four analyzed categories in our VAR model: one variable for trading intensity (traded dollar volume, number of trades, or number of shares traded per minute), one spread measure (qs_t or es_t), one depth measure (eV_t or d_t), and the realized volatility variable (rv_t). The base model uses traded dollar volume, quoted spread, Depth0, and realized volatility. We choose a lag length of 5 minutes which has a sufficiently high value of the Bayes Information Criterion and leaves enough observations for the analysis. As in Groß-Klußmann and Hautsch (2011), all variables are standardized by the mean for the respective stock day. Spread variables are further divided by 100 and depth variables are divided by 10,000 for exposition.

The following model equations are estimated separately using OLS:

$$\boldsymbol{y}_t = \boldsymbol{c} + \sum_{i=1}^{5}\boldsymbol{\Gamma}_i\boldsymbol{y}_{t-i} + \sum_{n=1}^{4}\boldsymbol{\Xi}_{n,pos}\boldsymbol{I}_{n,pos,t} + \sum_{n=1}^{4}\boldsymbol{\Xi}_{n,neg}\boldsymbol{I}_{n,neg,t} + \sum_{n=1}^{4}\boldsymbol{\Xi}_{n,neut}\boldsymbol{I}_{n,neut,t} + \boldsymbol{\epsilon}_t, \epsilon \sim N(0,\Omega), \quad (13)$$

where $\boldsymbol{\Gamma}_i$ denote a (4×4) coefficient matrix and $\boldsymbol{\Xi}_{n,pos}$, $\boldsymbol{\Xi}_{n,neg}$, and $\boldsymbol{\Xi}_{n,neut}$ denote (4×5) coefficient matrices. $\boldsymbol{I}_{1,neut,t}$ is the (4×1) vector for the interval $[-30,-15)$ before a neutral news event and similarly defined for positive and negative news events. $\boldsymbol{I}_{2,neut,t}$, $\boldsymbol{I}_{3,neut,t}$, and $\boldsymbol{I}_{4,neut,t}$ are analogously defined for the interval $[-15,0)$, $(0,15]$, and $(15,30]$ before and after a news event respectively. An extended version of the VAR model is in Appendix A.

6. Results and interpretation

The descriptive statistics for each firm in our sample are in Table 2. The average firm has 201 distinct news items and a marginally negative sentiment of −0.04. The average sentiment is in line with studies that report a bias towards negative media coverage (Soroka, 2006). The average market capitalization of a firm over the years 2005–2008 is approximately C$24bn. Market capitalization ranges from C$5.5bn to C$63bn. With roughly 63% of total market capitalization, our sample covers a large part of the Canadian market capitalization. Table 2 also shows that the sample represents a broad cross-section of industries.

In Table 3, we present summary statistics for periods without news, before, and after news for positive, negative, and neutral news separately. We present results for each "setting" for quoted spreads over all quote changes and for quoted spreads only at trade-time, effective spreads, depth at the best (Depth0), depth at three levels into the order book (Depth3), trading intensity

[13] Usually, one minute returns are quite small in magnitude which would lead to very small numbers in the result tables.

Table 4
Information estimations around news. Table 4 provides the results of the MRR model for no-news and news periods. Results comprise of the adverse selection components θ, order processing costs ϕ, and trade autocorrelations ρ. The terms positive, negative, and neutral relate to the RNSE news sentiment. 'Before' and 'after' describe thirty minute intervals before and thirty minute intervals after a news is disseminated over Thomson Reuters' newswire systems. The MRR model is estimated on a per company basis for the years 2005–2008. By-company estimation results in Panel A consist of the medians and means of GMM estimations for each single company in the sample. Robust median t-statistics can be found below estimates in parantheses. Panel B provides differences between different intervals and no-news periods and between pre and post-news periods. The medians of the differences ΔEst are compared with Wilcoxcon Signed Rank tests.

		No news (nn)	Positive		Negative		Neutral	
			Before	After	Before	After	Before	After
Panel A: By-company information estimations – MRR Model								
Adv. selection	Median Est.	1.4741	1.7630	1.5632	1.5574	1.6086	1.6194	1.3884
θ	Mean Est.	1.4282	1.6808	1.6478	1.8309	1.9807	1.8085	1.7461
	Median t-stat	(270.50)	(25.71)	(26.20)	(28.88)	(28.09)	(23.81)	(20.65)
Order proc.	Median Est.	1.1374	1.0011	0.8889	1.0131	0.8777	0.7810	1.0036
ϕ	Mean Est.	1.5922	1.3123	1.2335	1.5121	1.1915	1.1322	1.2462
	Median t-stat	(198.81)	(17.37)	(13.37)	(16.72)	(13.19)	(11.95)	(16.81)
Autocorr.	Median Est.	0.3500	0.3948	0.3996	0.4176	0.4050	0.4270	0.3998
ρ	Mean Est.	0.3585	0.4326	0.4443	0.4229	0.4401	0.4646	0.4231
	Median t-stat	(565.91)	(46.68)	(54.75)	(60.01)	(66.81)	(50.65)	(42.57)

		Positive		Negative		Neutral		Positive	Negative	Neutral
		Before – nn	After – nn	Before – nn	After – nn	Before – nn	After – nn	Before – after	Before – after	Before – after
Panel B: By-company information estimation differences – MRR model										
Adv. selection	Median ΔEst	0.1524[b]	0.0861[b]	0.2000[a]	0.2140[a]	0.1965[b]	0.0368	0.0649	0.0178	0.0749
θ	p-value (Wilcoxon T.)	0.0015	0.0098	<.0001	<.0001	0.0043	0.3655	0.5815	0.6812	0.1813
Order proc.	Median ΔEst	−0.0606[c]	−0.1524[a]	0.0298	−0.1215[c]	−0.2142[a]	−0.0669[c]	−0.0081	0.1155	−0.1015
ϕ	p-value (Wilcoxon T.)	0.0116	0.0007	0.4888	0.0445	<.0001	0.0262	0.5576	0.1754	0.0980
Autocorr.	Median ΔEst	0.0432[b]	0.0594[a]	0.0312[a]	0.0440[a]	0.0690[a]	0.0289[c]	−0.0204	0.0023	0.0217
ρ	p-value (Wilcoxon T.)	0.0015	0.0004	0.0002	0.0006	<.0001	0.0192	0.4778	0.7071	0.2947

[a] Significance at the 0.1% level.
[b] Significance at the 1% level.
[c] Significance at the 5% level.

(number of trades per minute), numbers of shares traded per minute, volume per minute, and share price volatility (realized volatility).

The overall average quoted spread is 4.74 basis points (bps), the average quoted spread at trade-time is 3.76 bps, and the average effective spread is 3.81 bps. Spreads at TSX are very narrow and are evidence of a liquid market. Quoted spreads are significantly larger than effective spreads which indicates that market monitoring occurs and market participants trade when it is inexpensive to do so. Depth3 (avg. C$416k) is approximately three times higher than Depth0 (avg. C$132k). Trading intensity measures show that stocks in our sample are actively traded with an average of ten trades per minute and firm. Descriptive statistics provide some evidence that liquidity increases around positive and neutral news and do not change for negative news. In line with existing literature, we find that trading activity increases around news announcements (cf. Green, 2004; Berry and Howe, 1994).

6.1. Price discovery

To understand information processing around news announcements, we present the results of the extended MRR model. Theory (cf. Kim and Verrecchia, 1994) suggests that informed trading after information events should increase. Depending on the type of an event, scheduled versus unscheduled, and the instrument traded, informed trading may also increase before an event (cf. Kim and Verrecchia, 1991). We aim to understand both of these components by studying pre and post-news periods and comparing these to periods without news.

Panel A of Table 4 provides the estimates for the adverse selection component of the spread, the order processing cost component, and the trade autocorrelations.[14] The adverse selection component of the spread is 1.47 bps during no-news periods and increases for all settings around news announcements except after neutral news. This result makes sense as after the release of neutral news there is no information to be impounded into prices. Consistent with intuition and in contrast to Green (2004) in a dealer market, we find a positive order processing cost component of spreads for all news settings. Order processing costs fall as adverse selection costs increase. Adverse selection is the larger of the two spread components. The median t-statistics for all MRR coefficients are significantly different than zero at all conventional levels. The model likelihood ratio chi-squared test statistics are significantly different than zero at all conventional levels for all dependent variables, they are thus not reported for brevity.

Panel B of Table 4 provides information on the difference between no-news periods and news periods before and after arrival for positive, negative, and neutral news. As we would expect, adverse selection costs are higher after positive and negative news than in periods without news. The only period around news messages in which adverse selection is not statistically different than for no-news periods is after neutral news. The differences for posi-

[14] In general, the MRR estimation is consistent with quoted and effective spreads calculated from the trade and quote data. Since spread components are estimated based on the trade point process in the MRR model, they do not necessarily need to exactly compare to quoted spreads. However, they should be comparable in magnitude.

Table 5

Pooled regression of liquidity and trading intensity around news. Table 5 provides results for liquidity and trading intensity measures around news in contrast to no-news periods over the years 2005–2008. The terms positive, negative, and neutral relate to the RNSE news sentiment. The terms 'before' and 'after' describe thirty minute intervals before and thirty minute intervals after a news is disseminated over Thomson Reuters' newswire systems. Overall GMM estimations for liquidity measures are reported in Panel A and for trading intensity in Panel B. All estimations are calculated with firm, year/quarter, day of the week, and time of day dummy variables. Robust *t*-statistics can be found below estimates in parantheses. Quoted spreads and effective spreads are measured in basis points and 'lnDepth0' represents the natural logarithm of the available volume at the best bid and ask in Canadian dollars. 'lnDepth3' is the natural logarithm of available volume in Canadian dollars at the top three order book levels. Liquidity measures are aggregated to minute averages prior to estimation. The number of shares traded per minute and the traded dollar volume per minute are transformed through the natural logarithm for the regressions. We omit estimates for control variables variables.

		Positive		Negative		Neutral		Positive	Negative	Neutral
		Before	After	Before	After	Before	After	Before − after	Before − after	Before − after
Panel A: Liquidity estimations										
Quoted spread	Estimate	−0.0526[c]	−0.0257	0.0794[c]	0.0835[c]	−0.0841[b]	−0.0463	−0.0269	−0.0041	−0.0379
	t-stat	(−2.09)	(−1.08)	(2.22)	(2.31)	(−3.10)	(−1.82)			
lnDepth0	Estimate	0.0404[a]	0.0378[a]	−0.0117[b]	−0.0037	0.0743[a]	0.0577[a]	0.0026	−0.0080	0.0167
	t-stat	(8.98)	(8.39)	(−2.95)	(−0.92)	(14.52)	(10.59)			
lnDepth3	Estimate	0.0497[a]	0.0497[a]	−0.0025	0.0067	0.0886[a]	0.0688[a]	0.0000	−0.0092	0.0198
	t-stat	(11.41)	(11.44)	(−0.65)	(1.71)	(17.21)	(12.69)			
Effective spread	Estimate	−0.0323	−0.0140	0.0811[c]	0.0947[c]	−0.0487[b]	−0.0208	−0.0183	−0.0136	−0.0278
	t-stat	(−1.39)	(−0.63)	(2.28)	(2.59)	(−1.87)	(−0.92)			
Panel B: Trading intensity estimations										
#Trades per min.	Estimate	1.8201[a]	1.6631[a]	2.8623[a]	2.2139[a]	2.2551[a]	3.5241[a]	0.1570	0.6484	−1.2690
	t-stat	(18.09)	(16.62)	(25.48)	(20.90)	(18.56)	(24.99)			
ln #Shares per min.	Estimate	0.1890[a]	0.1809[a]	0.2015[a]	0.2956[a]	0.1555[a]	0.2769[a]	0.0081	−0.0042	−0.1214
	t-stat	(22.84)	(21.90)	(26.43)	(26.80)	(17.39)	(28.15)			
ln volume per min.	Estimate	0.2053[a]	0.1983[a]	0.1860[a]	0.1885[a]	0.1614[a]	0.2787[a]	0.0070	−0.0025	−0.1172
	t-stat	(24.49)	(23.72)	(23.73)	(24.13)	(17.77)	(27.64)			
Panel C: Realized volatility estimations										
Realized volatility	Estimate	0.0016[a]	0.0014[a]	0.0031[a]	0.0015[a]	0.0018[a]	0.0032[a]	0.0003	0.0017	−0.0014
	t-stat	(6.21)	(5.99)	(9.50)	(4.78)	(5.04)	(8.65)			

[a] Significance at the 0.1% level.
[b] Significance at the 1% level.
[c] Significance at the 5% level.

Table 6

VAR estimation of liquidity and trading intensity around news. This table provides results for liquidity, trading intensity, and volatility measures around news in contrast to no-news periods over the years 2005–2008. The terms positive, negative, and neutral relate to the RNSE news sentiment. The time intervals describe minute intervals before and after a news is disseminated over Thomson Reuters' news wire systems. The coefficients of the estimated VAR model in Eq. (13) for liquidity measures are reported in Panel A, for trading intensity in Panel B, and for volatility in Panel C. *T*-statistics can be found below estimates in parantheses. Quoted spreads and effective spreads are measured in 100 basis points and 'lnDepth0' represents the natural logarithm of the available volume at the best bid and ask in Canadian dollars divided by 10,000. 'lnDepth3' is the natural logarithm of available volume in Canadian dollars at the top three order book levels divided by 10,000. Liquidity measures are aggregated to minute averages prior to estimation. The number of shares traded per minute and the traded dollar volume per minute are transformed through the natural logarithm for the regressions. All variables are divided by the average of the corresponding underlying trade minute of each stock.

	Positive				Negative				Neutral			
	Before news		After news		Before news		After news		Before news		After news	
	[−30; −15)	[−15; 0)	(0; 15]	(15; 30]	[−30; −15)	[−15; 0)	(0; 15]	(15; 30]	[−30; −15)	[−15; 0)	(0; 15]	(15; 30]
Panel A: Liquidity estimations												
Qspread	−0.20	0.21	0.12	−0.02	0.83[a]	0.41	0.44[c]	0.73[b]	−0.59[c]	−0.33	−0.76[b]	−0.20
	(−0.82)	(0.90)	(0.50)	(−0.07)	(3.56)	(1.85)	(1.97)	(3.19)	(−2.15)	(−1.24)	(−2.80)	(−0.63)
lnDepth0	6.93[b]	7.09[b]	7.34[b]	6.19[b]	5.65[c]	9.06[a]	6.89[b]	8.57[a]	7.14[b]	7.00[b]	9.52[a]	6.09[c]
	(2.95)	(3.14)	(3.27)	(2.69)	(2.57)	(4.28)	(3.27)	(3.96)	(2.73)	(2.76)	(3.73)	(2.05)
lnDepth3	3.14[c]	3.58[b]	2.38[c]	2.66[c]	2.98[c]	4.27[a]	3.90[a]	2.88[c]	3.09[c]	2.69[c]	3.68[b]	2.81
	(2.54)	(3.00)	(2.01)	(2.19)	(2.57)	(3.82)	(3.51)	(2.52)	(2.23)	(2.01)	(2.73)	(1.79)
Espread	0.21	0.49	0.68	0.53	2.01[a]	1.33[a]	1.70[a]	2.16[a]	−0.22	−0.21	−0.72	0.25
	(0.55)	(1.33)	(1.84)	(1.40)	(5.58)	(3.82)	(4.92)	(6.08)	(−0.50)	(−0.51)	(−1.72)	(0.51)
Panel B: Trading intensity estimations												
# Trades	0.05[a]	0.05[a]	0.05[a]	0.04[a]	0.10[a]	0.10[a]	0.09[a]	0.10[a]	0.08[a]	0.08[a]	0.08[a]	0.05[a]
	(9.37)	(9.74)	(10.07)	(8.55)	(20.98)	(20.83)	(20.75)	(21.09)	(13.69)	(14.99)	(13.79)	(8.48)
lnShares	0.15[a]	0.18[a]	0.17[a]	0.13[a]	0.23[a]	0.24[a]	0.25[a]	0.24[a]	0.17[a]	0.18[a]	0.22[a]	0.16[a]
	(12.25)	(14.63)	(14.07)	(10.47)	(19.57)	(21.34)	(22.46)	(20.68)	(12.03)	(13.44)	(16.35)	(10.07)
lnVolume	0.18[a]	0.20[a]	0.19[a]	0.15[a]	0.20[a]	0.21[a]	0.22[a]	0.20[a]	0.17[a]	0.19[a]	0.24[a]	0.17[a]
	(14.32)	(16.48)	(16.33)	(12.53)	(17.24)	(18.35)	(19.54)	(17.82)	(12.40)	(14.31)	(17.88)	(10.80)
Panel C: Realized volatility estimations												
Rvola	0.09[a]	0.11[a]	0.09[a]	0.13[a]	0.34[a]	0.31[a]	0.27[a]	0.25[a]	0.18[a]	0.25[a]	0.27[a]	0.20[a]
	(4.72)	(5.92)	(4.77)	(6.45)	(18.15)	(17.32)	(15.34)	(13.53)	(7.93)	(11.48)	(12.38)	(8.03)

[a] Significance at the 0.1% level.
[b] Significance at the 1% level.
[c] Significance at the 5% level.

Table 7
VAR estimation of liquidity and trading intensity around relevant news. Table 7 provides results for liquidity, trading intensity, and volatility measures around relevant (i.e. relevance equal or higher than 0.5) news in contrast to no-news periods over the years 2005–2008. The terms positive, negative, and neutral relate to the RNSE news sentiment. The time intervals describe minute intervals before and after a news is disseminated over Thomson Reuters' news wire systems. The coefficients of the estimated VAR model in Eq. (13) for liquidity measures are reported in Panel A, for trading intensity in Panel B, and for volatility in Panel C. *T*-statistics can be found below estimates in parantheses. Quoted spreads and effective spreads are measured in 100 basis points and 'lnDepth0' represents the natural logarithm of the available volume at the best bid and ask in Canadian dollars divided by 10,000. 'lnDepth3' is the natural logarithm of available volume in Canadian dollars at the top three order book levels divided by 10,000. Liquidity measures are aggregated to minute averages prior to estimation. The number of shares traded per minute and the traded dollar volume per minute are transformed through the natural logarithm for the regressions. All variables are divided by the average of the corresponding underlying trade minute of each stock.

	Positive				Negative				Neutral			
	Before		After		Before		After		Before		After	
	[−30; −15)	[−15; 0)	(0; 15]	(15; 30]	[−30; −15)	[−15; 0)	(0; 15]	(15; 30]	[−30; −15)	[−15; 0)	(0; 15]	(15; 30]
Panel A: Liquidity estimations												
Qspread	−0.83[c]	−0.50	−0.89[c]	−0.65	0.81[c]	0.47	0.78	0.86[c]	−0.95[b]	−0.71[c]	−1.47[a]	−0.51
	(−2.28)	(−1.38)	(−2.16)	(−1.84)	(2.23)	(1.32)	(1.90)	(2.41)	(−2.75)	(−2.03)	(−3.62)	(−1.16)
lnDepth0	7.76[c]	10.99[b]	3.99	6.84[c]	7.15[c]	10.61[b]	6.87	8.08[c]	9.27[b]	9.36[b]	13.18[a]	4.97
	(2.26)	(3.24)	(1.03)	(2.04)	(2.09)	(3.12)	(1.77)	(2.40)	(2.85)	(2.85)	(3.44)	(1.20)
lnDepth3	2.44	5.33[b]	2.28	2.14	4.39[c]	3.95[c]	5.98[b]	2.09	3.59[c]	4.21[c]	4.39[c]	2.33
	(1.35)	(2.97)	(1.11)	(1.21)	(2.43)	(2.20)	(2.92)	(1.18)	(2.09)	(2.43)	(2.17)	(1.07)
Espread	−0.84	−0.67	−1.14	−0.61	2.09[a]	1.54[b]	1.91[b]	2.72[a]	−0.78	−0.73	−1.77[b]	0.30
	(−1.49)	(−1.19)	(−1.79)	(−1.10)	(3.71)	(2.75)	(3.00)	(4.92)	(−1.46)	(−1.34)	(−2.81)	(0.45)
Panel B: Trading intensity estimations												
# Trades	0.05[a]	0.05[a]	0.04[a]	0.04[a]	0.09[a]	0.10[a]	0.10[a]	0.10[a]	0.09[a]	0.10[a]	0.06[a]	0.06[a]
	(6.36)	(7.38)	(4.95)	(6.02)	(11.61)	(14.02)	(11.84)	(13.39)	(12.31)	(14.24)	(7.26)	(7.09)
lnShares	0.19[a]	0.23[a]	0.17[a]	0.18[a]	0.23[a]	0.28[a]	0.31[a]	0.27[a]	0.18[a]	0.19[a]	0.25[a]	0.20[a]
	(10.37)	(12.88)	(8.41)	(10.10)	(12.75)	(15.57)	(15.18)	(15.06)	(10.49)	(10.89)	(12.26)	(9.18)
lnVolume	0.22[a]	0.26[a]	0.21[a]	0.21[a]	0.19[a]	0.22[a]	0.25[a]	0.21[a]	0.18[a]	0.20[a]	0.26[a]	0.21[a]
	(11.95)	(14.37)	(10.17)	(12.04)	(10.55)	(12.48)	(12.09)	(11.95)	(10.46)	(11.31)	(13.04)	(9.53)
Panel C: Realized volatility estimations												
Rvola	0.09[b]	0.10[a]	0.03	0.07[c]	0.36[a]	0.38[a]	0.38[a]	0.33[a]	0.21[a]	0.27[a]	0.19[a]	0.19[a]
	(2.98)	(3.34)	(0.76)	(2.49)	(12.21)	(12.99)	(11.52)	(11.38)	(7.48)	(9.60)	(5.77)	(5.52)

[a] Significance at the 0.1% level.
[b] Significance at the 1% level.
[c] Significance at the 5% level.

tive news of 0.15 bps and 0.09 bps are significant at the 1% level. The values for negative news at 0.20 bps and 0.21 bps are higher than the values of positive news and highly significant at the 0.1% level. Neutral news messages show a significant increase in adverse selection pre-news arrival and no statistically significant effect after arrival in comparison to no-news periods. Our results are not driven by the financial crisis in 2008 and 2009. We obtain similar results for a control period from 2003 to 2006.

Interesting is the highly significant increase in adverse selection around negative news in comparison to positive news. As described above, the theoretical models of Kim and Verrecchia (1991), Kim and Verrecchia (1994) predict higher adverse selection costs around information events in general. Market participants put different levels of effort into information gathering and thus posses different levels of private information. These different levels raise the level of information asymmetry in a market pre-news which induces higher adverse selection costs. Pre-announcement reaction might be driven by insider trading (information leakage) or it could be more innocuous such as news announcements before the Reuters' release by a competitor or other information sources, e.g. rumors. Krinsky and Lee (1996) provide empirical evidence for Kim and Verrecchia (1991), Kim and Verrecchia (1994) and find higher adverse selection costs around announcements comparable to our results for positive and negative news. After neutral news, adverse selection costs are comparable to no-news periods. One possible explanation in the light of existing models is that information gatherers cannot agree pre-news whether information is positive or negative which induces higher adverse selection costs.

Recent empirical evidence is also suggestive of asymmetric reactions to positive and negative information. Tetlock (2007) only finds significant market reactions to bad news in a Wall Street Journal column. Akhtar et al. (2010) study the effect that the monthly release of the Australian consumer sentiment has on the Australian stock market. They find a significant impact of negative information on prices and no impact on prices for positive information. They attribute this result to the 'negativity effect' found in psychology literature (cf. Soroka, 2006). Stock markets react stronger to monetary policy decisions that are bad for stock markets than to those that are good for stock markets on an intraday level (Chuli et al., 2010).

6.2. Liquidity, trading intensity, and volatility

To understand liquidity and to further understand the price dynamics around news arrivals, it is important to bear in mind that liquidity, information, and trading intensity are inherently related. Tables 5 and 6 provides regression results on liquidity, trading intensity, and realized volatility. Results for liquidity are more mixed than the information results in the previous section. However, compared with each other, all liquidity measures provide consistent results. We find that before and after positive news liquidity increases. For negative news, liquidity generally falls, more so before than after news. Theory would suggest a consistent reduction in liquidity over all news types which is at odds with our empirical findings (cf. Kim and Verrecchia, 1994). Consistent with existing literature (Berry and Howe, 1994), trading intensity increases around all different types of news.

All coefficients in Table 5 represent the difference of the respective period to no-news periods. Panel A shows that liquidity increases significantly around positive and around neutral news. Liquidity increases if spreads tighten and depth increases. The quoted spreads decrease 0.05 bps before positive news and decrease 0.03 bps after positive news however not statistically significant after positive news. The increase in liquidity for neutral news is almost double the increase for positive news. The quoted spread increases, corresponding to a decrease in liquidity, before negative news by 0.08 bps and also 0.08 bps after negative news. Results for effective spreads are similar to those for quoted spreads. Results for Depth0 and Depth3 point to a positive effect of news messages on depth, although results for negative news are insignificant and thus inconclusive. Comparing negative and positive messages, the former seem to have a stronger influence on spreads while the latter more strongly affect available depth.

Panel B of Table 5 provides results for trading intensity. Trading intensity increases around all types of news. It increases slightly more for negative than for positive news if measured in the number of trades per minute and it increases stronger for positive news if measured in Canadian dollar volume. However, trading intensity increases even more around neutral news compared to both positive and negative news.[15]

Since theory (cf. Kim and Verrecchia, 1991; Kim and Verrecchia, 1994) predicts a reduction of liquidity around information events, our results for positive and neutral news may at first seem contradictory. However, the types of news that we analyze are different to the mostly studied macroeconomic announcements or earnings announcements. On average, newswire messages have a lower impact than earnings or macroeconomic announcements and also their implications are on average much lower than those of major world events. Trading intensity increases around news announcements which reflects changes in expectations of individual investors who trade to adjust their holdings to reflect their new expectations.

The findings on realized volatility are overall slightly higher around information arrivals than during no news periods. Panel C reports the coefficients for realized volatility around news arrivals.[16] The results for realized volatility pre and post-news arrival are also consistent with the MRR information results. Around news, adverse selection costs are higher than during no-news periods for all three different news sentiments (see Table 4, Panel B) which indicates higher private information flow around news. French and Roll (1986) find that a major determinant of return volatility is trading of informed market participants, i.e. private information flow revealed to the market through trades. We find this pattern in our data with higher realized volatility around news.

In order to test for the robustness of the results of the univariate model, we conduct a multivariate VAR analysis for the whole sample and for a subsample consisting of news items with relevance measures greater equal 0.5. The VAR model in Eq. (13) gives a more robust estimate of the news impact on the observed variables by accounting for interrelation of the lagged observation variables. We also distinguish between the four 15 min intervals before and after news arrivals.

The VAR results confirm the results from the pooled regression: While spread measures decrease for positive and neutral news arrivals, i.e. liquidity increases, spread measures increase for negative news. A difference is presented in the results for depth, since all depth measures increase for all sentiment groups. We conclude from this difference that news in fact have an asymmetric impact on spread measures, but that depth consistently increases around news arrivals. Any possible negative effects on depth are rather driven by the effect on spreads than by the news events.

Results for trading intensity and realized volatility are consistent with our univariate regression results. Results for negative news messages are even higher than in the univariate analysis, which strengthens the negativity bias as found in Akhtar et al. (2010).

We conduct the same analysis only for news items with relevance measures over 0.5. The results are presented in Table 7. The results are qualitatively similar and do not yield different results to the whole sample.

Connecting our spread results to the MRR results in Table 4, we can see that the VAR results support the MRR results for adverse selection. The adverse selection component decreases for positive and neutral news and increases for positive news. This suggests that the spread results hold even after accounting for order processing costs. Furthermore, it highlights the differences between positive, negative, and neutral news. The reactions are asymmetric as suggested by previous literature.

7. Conclusion

In this paper, we analyze the impact of Thomson Reuters newswire messages on intraday price discovery, liquidity, and trading intensity at the Toronto Stock Exchange. In contrast to existing literature, we are able to cluster news based on message content. We split the news data into groups of news messages with positive, negative, and neutral sentiment which gives us the opportunity to study asymmetric reactions to news messages. News are not sorted based on ex post return measures but on ex ante message content based measures. We estimate the adverse selection component of the spread with an extension of the Madhavan et al. (1997) model.

Our results provide evidence of asymmetric reactions to news. In general, we find higher adverse selection costs around news messages, which can be explained through information gathering pre-news arrival and differential information processing capabilities of market participants post-news arrival. On the sentiment level, negative news messages induce significantly higher adverse selection costs than positive news messages. Liquidity, as proxied by spread measures, increases around positive and neutral messages whereas it decreases around negative messages. Depth and trading intensity consistently increases around all types of news messages. Using a VAR model, we further control for interaction of the variables and find similarly strong results. A possible explanation for the difference between the impact of news messages with different sentiments, the asymmetric reaction, could be ambiguity aversion. Ambiguity averse traders react asymmetrically to ambiguous information such as news messages. If our market is composed of a proportion of ambiguity averse traders, this provides a possible explanation for our results. The main contribution of our study is that we show that traders react asymmetrically to positive, negative, and neutral intraday news arrivals. This result is in line with psychological studies from the field of impression formation showing that humans react stronger to negative information. However, we find that newswire messages with both negative and positive content generally have a significant impact on trading in an electronic limit order market.

Acknowledgements

The authors thank Ike Mathur (the managing editor) and an anonymous referee for valuable comments and suggestions. Financial support from Stuttgart Stock Exchange and from the IME Graduate School at Karlsruhe Institute of Technology (KIT) funded by the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

[15] All LR tests are highly significant which suggests that all models are better specified than the restricted models from the LR tests.

[16] The LR test χ^2 statistic for the realized volatility estimation is 34 which is highly significant at the 0.01% level.

Appendix A. Extended VAR specification

$$
\begin{aligned}
V_t &= c_1 + \sum_{i=1}^{5}\gamma_{1,1,i}V_{t-i} + \sum_{i=1}^{5}\gamma_{1,2,i}qs_{t-i} + \sum_{i=1}^{5}\gamma_{1,3,i}ev_{t-i} + \sum_{i=1}^{5}\gamma_{1,4,i}rv_{t-i} \\
&\quad + \sum_{n=1}^{4}\xi_{1,n,pos}I_{n,pos,t} + \sum_{n=1}^{4}\xi_{1,n,neg}I_{n,neg,t} + \sum_{n=1}^{4}\xi_{1,n,neut}I_{n,neut,t} + \epsilon_{1,t} \\
qs_t &= c_2 + \sum_{i=1}^{5}\gamma_{2,1,i}V_{t-i} + \sum_{i=1}^{5}\gamma_{2,2,i}qs_{t-i} + \sum_{i=1}^{5}\gamma_{2,3,i}ev_{t-i} + \sum_{i=1}^{5}\gamma_{2,4,i}rv_{t-i} \\
&\quad + \sum_{n=1}^{4}\xi_{2,n,pos}I_{n,pos,t} + \sum_{n=1}^{4}\xi_{2,n,neg}I_{n,neg,t} + \sum_{n=1}^{4}\xi_{2,n,neut}I_{n,neut,t} + \epsilon_{2,t} \\
ev_t &= c_3 + \sum_{i=1}^{5}\gamma_{3,1,i}V_{t-i} + \sum_{i=1}^{5}\gamma_{3,2,i}qs_{t-i} + \sum_{i=1}^{5}\gamma_{3,3,i}ev_{t-i} + \sum_{i=1}^{5}\gamma_{3,4,i}rv_{t-i} \\
&\quad + \sum_{n=1}^{4}\xi_{3,n,pos}I_{n,pos,t} + \sum_{n=1}^{4}\xi_{3,n,neg}I_{n,neg,t} + \sum_{n=1}^{4}\xi_{3,n,neut}I_{n,neut,t} + \epsilon_{3,t} \\
rv_t &= c_4 + \sum_{i=1}^{5}\gamma_{4,1,i}V_{t-i} + \sum_{i=1}^{5}\gamma_{4,2,i}qs_{t-i} + \sum_{i=1}^{5}\gamma_{4,3,i}ev_{t-i} + \sum_{i=1}^{5}\gamma_{4,4,i}rv_{t-i} \\
&\quad + \sum_{n=1}^{4}\xi_{4,n,pos}I_{n,pos,t} + \sum_{n=1}^{4}\xi_{4,n,neg}I_{n,neg,t} + \sum_{n=1}^{4}\xi_{4,n,neut}I_{n,neut,t} + \epsilon_{4,t}
\end{aligned}
\tag{A.1}
$$

with

$$
I_{n,t} = \begin{cases} 1 & \text{if } t \text{ is within } [-30,-15) \text{ (for } n=1\text{)}, \\ & [-15,0] \text{ (for } n=2\text{)}, \\ & (0,15] \text{ (for } n=3\text{)}, \\ & (15,30] \text{ (for } n=4\text{)} \quad \text{minutes around a news}, \\ 0 & \text{else}. \end{cases}
\tag{A.2}
$$

References

Akhtar, S., Faff, R., Oliver, B., Subrahmanyam, A., 2010. The power of bad: the negativity bias in Australian consumer sentiment announcements on stock returns. Journal of Banking and Finance 35 (5), 1239–1249.

Antweiler, W., Frank, M.Z., 2004. Is all that talk just noise? The information content of internet stock message boards. The Journal of Finance 59 (3), 1139–1167.

Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. The Review of Financial Studies 21 (2), 785–818.

Berry, T.D., Howe, K.M., 1994. Public information arrival. The Journal of Finance 49 (4), 1331–1346.

Bessembinder, H., Kaufman, H.M., 1997. A cross-exchange comparison of execution costs and information flow for NYSE-listed stocks. Journal of Financial Economics 46 (3), 293–319.

Chaboud, A., Chiquoine, B., Hjalmarsson, E., Vega, C., 2009. Rise of the machines: algorithmic trading in the foreign exchange market. FRB International Finance Discussion Paper No. 980.

Chuli, H., Martens, M., van Dijk, D., 2010. Asymmetric effects of federal funds target rate changes on S&P100 stock returns, volatilities and correlations. Journal of Banking and Finance 34 (4), 834–839.

French, K.R., Roll, R., 1986. Stock return variances: the arrival of information and the reaction of traders. Journal of Financial Economics 17 (1), 5–26.

Green, T.C., 2004. Economic news and the impact of trading on bond prices. The Journal of Finance 59 (3), 1201–1233.

Groß-Klußmann, A., Hautsch, N., 2011. When machines read the news: using automated text analytics to quantify high frequency news-impled market reactions. Journal of Empirical Finance 18 (2), 321–340.

Hansen, P., Lunde, A., 2005. A realized variance for the whole day based on intermittent high-frequency data. Journal of Financial Econometrics 3 (4), 525–554.

Harris, M., Raviv, A., 1993. Differences of opinion make a horse race. The Review of Financial Studies 6 (3), 473–506.

Hautsch, N., Hess, D., Veredas, D., 2011. The impact of macroeconomic news on quote adjustments, noise, and informational volatility. Journal of Banking and Finance 35 (10), 2733–2746.

Huang, R.D., Stoll, H.R., 1997. The components of the bid-ask spread: a general approach. The Review of Financial Studies 10 (4), 995–1034.

Kim, O., Verrecchia, R.E., 1991. Market reactions to anticipated announcements. Journal of Financial Economics 30 (2), 273–309.

Kim, O., Verrecchia, R.E., 1994. Market liquidity and volume around earnings announcements. Journal of Accounting and Economics 17 (1–2), 41–67.

Krinsky, I., Lee, J., 1996. Earnings announcements and the components of the bid-ask spread. The Journal of Finance 51 (4), 1523–1535.

Lee, C., Ready, M., 1991. Inferring trade direction from intraday data. The Journal of Finance 46 (2), 733–746.

Madhavan, A., Richardson, M., Roomans, M., 1997. Why do security prices change? A transaction-level analysis of NYSE stocks. The Review of Financial Studies 10 (4), 1035–1064.

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

Niederhoffer, V., 1971. The analysis of world events and stock prices. The Journal of Business 44 (2), 193–219.

Ranaldo, A., 2008. Intraday market dynamics around public information arrivals. In: Gregoriou, G.N., Lhabitant, F.-S. (Eds.), Stock Market Liquidity: Implications for Market Microstructure and Asset Pricing. John Wiley and Sons, Hoboken (NJ).

Soroka, S.N., 2006. Good news and bad news: asymmetric responses to economic information. The Journal of Politics 68 (2), 372–385.

Tetlock, P.C., 2007. Giving content to investor sentiment: the role of media in the stock market. The Journal of Finance 62 (3), 1139–1167.