No. 2009/31

Quantifying High-Frequency Market Reactions to Real-Time News Sentiment Announcements

Axel Groß-Klußmann and Nikolaus Hautsch

Center for Financial Studies
Goethe-Universität Frankfurt ■ House of Finance
Grüneburgplatz 1 ■ 60323 Frankfurt ■ Deutschland

Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077
http://www.ifk-cfs.de ■ E-Mail: info@ifk-cfs.de

Center for Financial Studies

The *Center for Financial Studies* is a nonprofit research organization, supported by an association of more than 120 banks, insurance companies, industrial corporations and public institutions. Established in 1968 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia.

The CFS Working Paper Series presents the result of scientific research on selected topics in the field of money, banking and finance. The authors were either participants in the Center´s Research Fellow Program or members of one of the Center´s Research Projects.

If you would like to know more about the *Center for Financial Studies*, please let us know of your interest.

Jan Krahnen

Prof. Dr. Jan Pieter Krahnen

Center for Financial Studies
Goethe-Universität ▪ House of Finance
Grüneburgplatz 1 ▪ 60323 Frankfurt am Main ▪ Deutschland

Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077
http://www.ifk-cfs.de ▪ E-Mail: info@ifk-cfs.de

CFS Working Paper No. 2009/31

Quantifying High-Frequency Market Reactions to Real-Time News Sentiment Announcements*

Axel Groß-Klußmann[1] and Nikolaus Hautsch[2]

This Version: December 2009

Abstract:
We examine intra-day market reactions to news in stock-specific sentiment disclosures. Using pre-processed data from an automated news analytics tool based on linguistic pattern recognition we extract information on the relevance as well as the direction of company-specific news. Information-implied reactions in returns, volatility as well as liquidity demand and supply are quantified by a high-frequency VAR model using 20 second intervals. Analyzing a cross-section of stocks traded at the London Stock Exchange (LSE), we find market-wide robust news-dependent responses in volatility and trading volume. However, this is only true if news items are classified as highly relevant. Liquidity supply reacts less distinctly due to a stronger influence of idiosyncratic noise. Furthermore, evidence for abnormal highfrequency returns after news in sentiments is shown.

JEL-Classifications: G14, C32

Keywords: Firm-specific News, News Sentiment, High-frequency Data, Volatility, Liquidity, Abnormal Returns.

* For helpful comments and discussions we thank Boris Drovetsky, Lada Kyj, Roel Oomen, and the participants of workshops at Humboldt-Universit¨at zu Berlin and at the Quantitative Products Laboratory. This research is supported by the Deutsche Bank AG via the Quantitative Products Laboratory and the Deutsche Forschungsgemeinschaft (DFG) via the Collaborative Research Center 649 ”Economic Risk”.

1 Quantitative Products Laboratory (QPL), Berlin and Institute for Statistics and Econometrics, Humboldt-Universit¨at zu Berlin. Email: axel.gross-klussmann@db.com. Address: Alexander Str. 5, D-10178 Berlin, Germany.

2 Institute for Statistics and Econometrics and Center for Applied Statistics and Economics (CASE), Humboldt-Universit¨at zu Berlin as well as Quantitative Products Laboratory (QPL), Berlin, and Center for Financial Studies (CFS), Frankfurt. Email: nikolaus.hautsch@wiwi.hu-berlin.de. Address: Spandauer Str. 1, D-10178 Berlin, Germany.

1 Introduction

Trading on financial markets is ultimately driven by news. However, news are difficult to observe and to identify. Due to the enormous amount of information continuously released by modern electronic communication media it is virtually impossible to process all information associated with a certain financial asset. In particular, it is problematic to distinguish between relevant and less relevant news and to interpret them accordingly. Because of these difficulties nearly all empirical studies examine the impact of news by solely focusing on specific news events, such as scheduled macroeconomic announcements, political interventions, or certain firm-specific news such as earnings announcements which are in most cases easily identifiable.

This paper addresses the challenge of linking a virtually continuous asset-specific news flow to high-frequency market activity. The news flow consists of messages from an automated news sentiment engine. This engine performs a linguistic pre-processing of stock-specific public news. It transforms the news content into items indicating news' relevance and the author's sentiment of the underlying story. Exploiting this source we link stock-specific news arrivals to high-frequency returns, volatility, trading intensity, trade sizes, spreads and market depth.

The question of how news is incorporated into asset prices is analyzed by a wide range of studies. The predominant part of this literature focuses on macroeconomic news and company-specific earnings announcements. Starting with Beaver (1968), numerous studies have quantified the link between abnormal volatility and trading volume induced by the disclosure of earnings information, see, e.g., Malatesta and Thompson (1985), Landsman and Maydew (2002) or Graham et al. (2006). Comparable results are found for macroeconomic news, see, e.g., Ederington and Lee (1993), DeGennaro and Shrieves (1997), Fleming and Remolona (1999), Hautsch and Hess (2002) and Andersen et al. (2003). However, only very few studies try to link asset prices and trading activities to an intraday flow of information. This is due to the fact that high-frequency news items are difficult to record and are typically considered to be too noisy due to the interference with other sources of information. As a consequence, Berry and Howe (1994) and Mitchell and Mulherin (1994) construct aggregated news measures and document a positive relationship between the amount of news and market activity.

Kalev et al. (2004) use the number of public news items as an explanatory variable in a GARCH specification to test the influence of the news arrival rate on stock market volatility. Ranaldo (2008) is the only study examining the impact of single firm-specific news items on intra-day trading processes. Still, a major problem of his analysis is the vast amount of virtually non-informative news. As a result, the estimated news impact is comparably low, particularly, if earnings announcements are discarded. These results indicate that the distinction between relevant and irrelevant news as well as the filtering of noise is very crucial.

To our best knowledge, the present study is the first one exploiting data from an automated news engine. We use the Reuters NewsScope Sentiment Engine which classifies firm-specific news according to positive and negative author sentiments and provides an indicator for news' relevance. Each sentiment and relevance measure is derived from a linguistic pattern analysis of the respective news story. Supposing that the news engine captures a major part of intradaily news arrivals in a pre-filtered and structured way, it opens up a new direction to examine the effects of a continuous news flow on intraday trading. Using this data we aim to answer the following research questions: (i) Can we identify significant reactions in returns, volatility and liquidity induced by the arrival of a news item? (ii) Does the magnitude of the reactions depend on the indicated relevance and sign of news? (iii) Are the results robust across different stocks or are they overlaid by stock-specific noise? (iv) Are news in sentiments anticipated or known by the market *prior* to publication? (v) Is there a different reaction to sentiments on days of earnings announcements?

Using 20 second aggregates of transaction data from 35 liquid stocks traded at the London Stock Exchange (LSE), we study news' impact on abnormal returns, squared returns, cumulated trading volume, spreads and market depth. Particularly the behavior of liquidity supply and demand around news announcements is still widely unexplored. To our knowledge only Fleming and Remolona (1999) provide a systematic analysis of trading intensities, volumes and spreads around scheduled (macroeconomic) news releases. While many studies analyze news effects based on fixed windows around the event dates, we model the complete underlying trading process. To avoid spurious regression results due to neglected dynamics and cross-dependencies between the variables, we employ a high-frequency Vector Autoregressive (VAR) model which is augmented by news-specific explanatory variables and explicitly accounts for the naturally high proportion of zero variables arising from non-trading in a 20-second interval.

A major finding of our analysis is that high-frequency trading activity significantly reacts to news items which are identified as relevant. Conversely, for less relevant news no significant responses can be quantified. In this sense, the sentiment relevance indicator carries information that is obviously taken into account by the market. Most distinct news effects are shown for volatility and trading volume which react strongly and fast. While volume and volatility reactions are widely stable across the market and are robust with respect to dynamics and cross-dependencies, for bid-ask spreads and market depth less distinct news effects are shown. For these variables, we observe stronger market-wide variations and generally weaker responses to news as soon as multivariate trading dynamics are taken into account. This finding is attributed to a higher impact of idiosyncratic noise and a stronger dependence on general market dynamics and thus spillovers from other trading variables. Moreover, we find evidence for significant abnormal returns after the arrival of relevant news items. This is particularly true on days of company earnings announcements. Finally, there are significant above-average market activities *before* the publication of an information item indicating the existence of other sources of news and an overall clustering thereof.

The remainder of the paper is organized as follows. In the next section, we describe the underlying data set and present descriptive statistics. Section 3 reports evidence for unconditional news impacts without explicitly controlling for time series dynamics in the processes. In Section 4, the econometric framework and corresponding results based on a high-frequency VAR model are given. Section 5 concludes.

2 Data

To facilitate the processing of new information, several news vendors offer software environments capturing particular characteristics of information in real time. These tools electronically analyze available information using linguistic pattern recognition algorithms. Words, word patterns, the novelty of a news item, its type and other characteristics are translated into indicators of the relevance as well as of the tone of the item.

We use pre-processed news data from a news-analytics tool of the Reuters company, the Reuters NewsScope Sentiment Engine. The data contain all 2007 news headlines as observed on traders' screens. Each news item provides a sentiment and relevance indicator. These indicators are produced based on pattern recognition algorithms. The sentiment attributes of the news are coded +1, 0 and -1 for a positive, neutral and negative tone of the underlying story, respectively. Relevance is indicated by a number in the $[0, 1]$ interval. News arrival is recorded based on time stamps up to a micro-second precision.

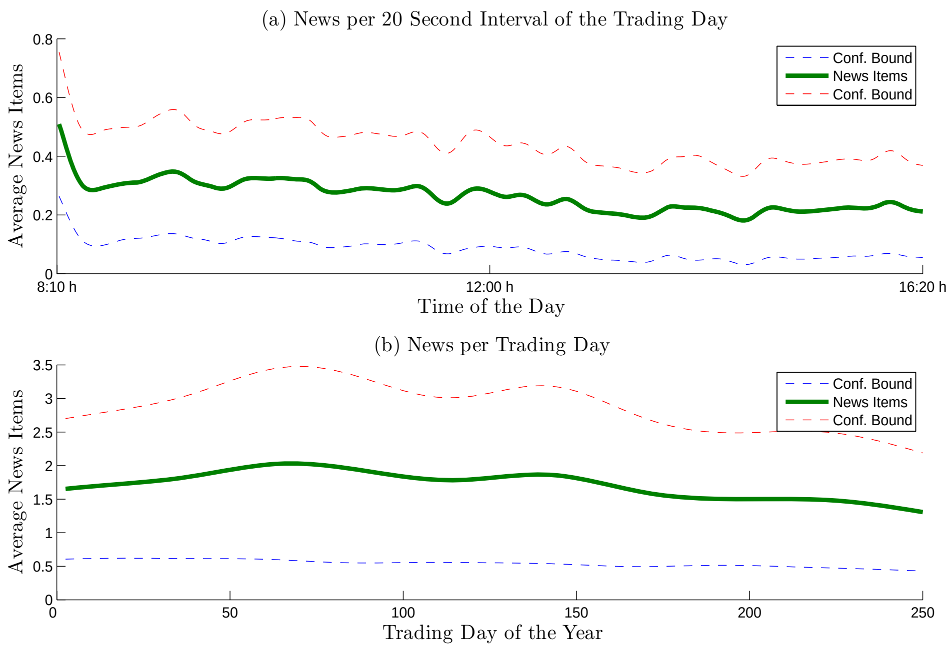

Figure 1: Distribution of news over a day and over the year. Smoothed via kernel regression.

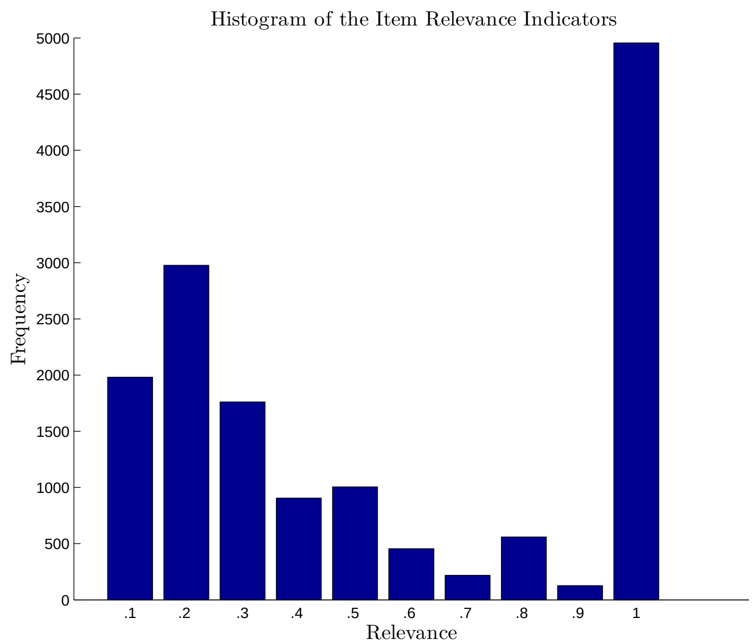

Figure 2: Distribution of the Relevance Indicator.

Table 1: Descriptive Statistics

Name / RIC	Money Value	Price Change	Average Spread	# of Trades	# of News	# of Rel. News	# of Rel. Positive	# of Rel. Negative	News on Days of EA
AAL	166	0,24	2,11	7300	631	267	93	114	9
AV	60,7	-0,18	0,67	3587	248	119	57	40	6
AZN	136	-0,21	1,51	5325	379	216	78	102	32
BATS	61,8	0,38	1,41	3455	119	55	32	15	8
BA	42,7	-0,42	0,51	3407	560	292	130	129	14
BARC	218	-0,31	0,57	7608	1303	699	357	230	14
BG	74,1	0,65	0,75	3977	197	91	44	30	10
BP	303	0,08	0,52	6075	1280	722	249	339	47
BT	72,3	-0,09	0,30	3519	241	125	79	33	24
CBRY	50,9	0,14	0,72	3054	197	112	45	53	5
DGE	62,6	0,07	0,92	3488	117	49	31	12	8
EMG	51,0	0,08	0,70	3304	88	53	22	17	2
GSK	177	-0,04	1,09	5399	567	209	63	87	18
HBOS	117	-0,36	0,91	5171	398	175	55	75	11
HSBA	305	-0,09	0,54	6178	1100	554	200	135	20
IMT	45,0	0,32	2,08	3087	202	109	74	30	1
LLOY	98,5	-0,17	0,55	4382	254	124	44	51	20
MKS	55,9	-0,22	0,71	3059	177	85	42	31	6
NG	50,1	0,13	0,69	3096	152	103	39	44	6

Table 1: Descriptive Statistics (Cont'd)

Name / RIC	Money Value	Price Change	Spread	# of Trades	# of News	# of Rel. News	# of Rel. Positive	# of Rel. Negative	News on Days of EA
NXT	42,9	-0,11	2,01	3229	117	82	44	32	9
PRU	77,3	0,01	0,70	3985	191	111	65	22	11
BG	27,8	0,14	0,69	2503	175	137	76	43	17
BLT	210	0,66	0,68	7464	187	121	41	61	9
FP	24,7	-0,29	0,46	2212	165	104	59	31	4
III	24,6	-0,07	1,37	2158	165	65	35	13	4
ITV	18,8	-0,18	0,18	1763	159	95	27	57	5
RBS	207	-0,78	0,73	7692	975	406	198	140	10
RIO	228	0,97	2,29	7847	495	169	95	44	4
SAB	41,2	0,22	1,49	2754	144	89	50	20	6
SL	11,9	-0,15	0,66	1471	131	86	46	29	6
STAN	86,3	0,24	1,45	4329	371	171	104	35	15
TSCO	88,7	0,18	0,33	4217	303	158	77	58	10
ULVR	65,6	0,34	1,40	3005	159	76	40	22	15
VOD	235	0,33	0,15	6284	948	421	249	109	16
XTA	155	0,38	2,49	6508	484	201	97	76	9

Note: RIC denotes the Reuters Identifier Code. Money Value (traded) is computed as the trade size times the respective price (turnover total in 2007 in million). Price Change is the % price change from 01/03/07 to 12/31/07. Spread and Nr. of trades are averages per trading day. The News column refers to the number of news items per firm in 2007 without overnight news and identical updates. Relevant news items are classified to be the ones with a relevance indicator greater than .5. Rel. Positive and Rel. Negative give the numbers of relevant positive and negative items, respectively. The last column gives the number of news on days of earnings announcements (EA).

We select 40 stocks from the FTSE 100 Index which are most active in terms of the number of published news items. As we require data availability for 230 trading days, the sample is ultimately cut down to 35 stocks. The fact that the selected stocks are also very actively traded (see Table 1) allows us to study market dynamics based on a high frequency.

The underlying transaction data is aggregated to 20 second intervals. We consider this aggregation level to be a good compromise between exploiting a maximum of information on the one hand and making the analysis still computationally tractable (given a year of data). To reduce the impact of market opening and closing effects, we discard the first ten and last ten minutes of a trading day. Intraday returns, volatility and liquidity are captured by the following variables computed over 20 second intervals:

(i) cumulated trade size,

(ii) average trade size, defined as the cumulated trade size divided by the corresponding number of trades per interval,

(iii) bid-ask spread evaluated at the endpoint of each interval,

(iv) mid-quote returns over each interval,

(v) money value traded, defined as trade sizes in the intervals weighted by the corresponding mid-quotes,

(vi) depth, defined as as the volume pending at the best bid and ask level, evaluated at the endpoint of each interval,

(vii) volatility, defined as the sum of squared mid-quote transaction returns over each interval.

All volatility and liquidity variables exhibit pronounced intraday trading patterns. Figure 3 shows the widely documented daily U-shape pattern for cumulated trade sizes. As shown in the Appendix, similar shapes are also revealed for the other variables. To capture theses patterns, we standardize all processes by the yearly average of the corresponding underlying 20 seconds interval, i.e.,

$$x_{jd} = \frac{x_{jd}}{1/n \sum_{d=1}^{n} x_{jd}},$$

where j denotes the specific interval of the trading day d and x represents the corresponding variable.

Under the assumption that updates of a news story do not carry much extra information compared to the initial one, we only employ the first message from a sequence of news updates. Subsequent updates with identical headlines as the initial one are deleted from the sample. In addition, we only focus on the news flow within a trading day and do not exploit overnight news. Incorporating the latter would considerably increase the complexity of the study.

After pre-filtering, the number of news range from a minimum of 117 to a maximum of 1303 disclosures per stock in 2007 (see Table 1). We observe that news tend to cluster in the first half of a day. Figure 1 a) shows the average number of news per 5-minute interval during a trading day. It turns out that the news intensity peaks at the beginning of the trading period but is relatively stable during the rest of the day. Figure 1 b) gives the average number of news items per day through the year 2007. Similarly to the intra-day shape there is no pronounced yearly pattern.

We distinguish between different types of news. First, we separate between scheduled and non-scheduled news by identifying days on which company-specific earnings estimates are released. Second, we distinguish between relevant and less relevant news. Since we expect the reported relevance indicator to be a relatively noisy measure, we classify news items with an indicator value above or at (below) 0.6 as relevant (irrelevant) (see Figure 2).

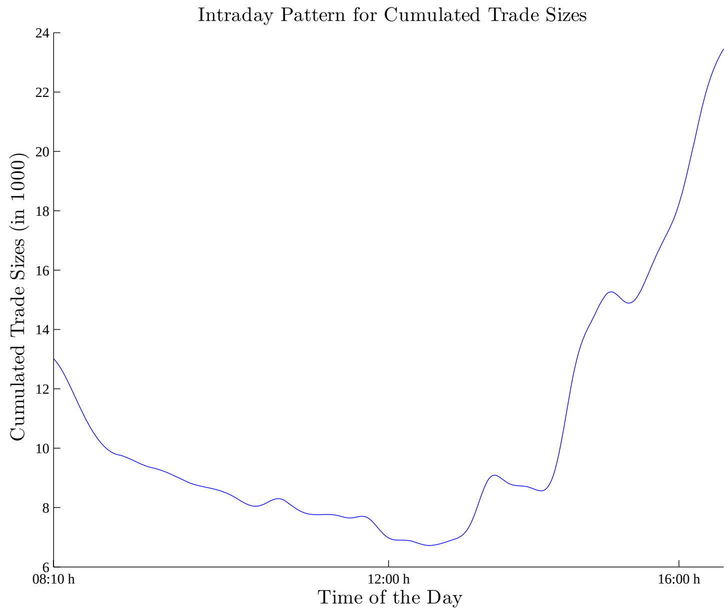

Figure 3: Intraday seasonality pattern of the cumulated trading volume. Smoothed via kernel regression.

3 Unconditional News Impacts

3.1 Impact on Volatility and Liquidity

In this section, we study the unconditional impact of the news flow without explicitly controlling for market dynamics and cross-dependencies between the variables. Such an analysis already provides important insights and serves as a basis for the econometric modelling in Section 4. Here, we analyze 400 20-second intervals around news arrivals capturing 100 intervals before each disclosure and 300 thereafter.

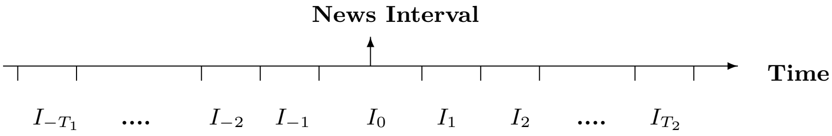

Figure 4: Intervals around News Arrival

Figure 4 illustrates the timing of the intervals. I_0 denotes the specific 20-second interval around the news item, whereas T_1 and T_2 are the numbers of intervals before and after the news period, respectively. For each stock, we compute the average market reaction and corresponding standard errors over all event windows. For sake of brevity, we refrain from showing results for individual stocks but report pooled averages over the cross-section of stocks. Correspondingly, by denoting the market reaction of variable X to news item i during interval I_j as X_{iI_j}, the pooled average across all news events and all stocks is computed as $\overline{X}_{I_j} = 1/n \sum_{i=1}^{n} X_{iI_j}$, where n is the total number of news for all stocks. Given that the stocks have quite similar empirical characteristics (see Table 1), this proceeding allows us to highlight the results common to all stocks. Assuming (approximative) normally distributed reactions, the 95% confidence intervals of $\overline{X}_{I_j}$ are computed as two times the standard errors of $\overline{X}_{I_j}$. Since these standard errors reflect variations across all event windows as well as across the market they capture overall news responses and statistical confidence thereof. Two robustness checks underscore the validity of the inference. First, the confidence intervals closely match those obtained from a parametric bootstrap. Second, to account for the fact that stocks with a high number of news naturally have a stronger weight in $\overline{X}_{I_j}$, we perform a robustness check using a group-means estimator instead of a pooled average. The results are qualitatively identical.[1]

Figures 5 to 10 show the money value traded, realized volatility, spreads, market depth, average trade sizes and cumulated trade sizes around news arrivals of differ-

[1] See in the Appendix for more details on the computation of standard errors.

ent types. Note that by construction of the seasonality adjustment the mean of each series equals one. We differentiate between relevant news on days with earnings announcements (henceforth EA), relevant news on days without earnings announcements (henceforth noEA) and less relevant news which virtually always occur on noEA days.

The following findings can be summarized: First, during the analyzed time window each of the variables is significantly above its mean. For instance, money value traded is on a level of more than 50% above its mean. For most variables, above-average activities start already more than thirty minutes before the item arrival. This finding is a strong hint for market participants having different and more timely sources of information and for news itself being clustered.

Second, though prior information seems to be present, *relevant* news items still induce significant reactions at the event time. In contrast, less relevant information does not cause any distinct market response. Hence, we find convincing evidence for the fact that market participants seem to distinguish between important and less important news and thus extract information from the sentiment ticker.

Third, we observe significant responses in volatility, bid-ask spreads and money value traded. As shown by Figure 7, spreads are significantly increased indicating that liquidity providers tend to post less competitive quotes and protect themselves against possible informational disadvantage and adverse selection. Interestingly, such behavior is not accompanied by changes in the corresponding market depth which remains relatively stable and widely unaffected by news arrivals. On the other hand, liquidity demand, as measured by the money value traded and cumulated trade sizes, significantly peaks around the event time. Interestingly, this reaction is predominantly induced by faster trading but not by higher trade sizes (see Figures 9 and 10). Moreover, we observe strong reactions in high-frequency volatility and trading volumes. Both are obviously closely related. Overall, trading activity remains on an above-average level for at least 60 minutes after news arrival.

Fourth, we observe a stronger news response on EA days than on noEA days. This might be due to the fact that news on EA days convey more information or markets are simply more sensitive.

In order to test for the existence of possible asymmetric market reactions in dependence of the sign of news, we define a sentiment indicator to have a distinct direction (positive or negative) whenever the probability p measuring the assessment's confidence exceeds 0.7. This allows us to filter out noisy and unreliable information. Figure 11 (a) shows the volatility reaction to positive and negative news items on days without earnings announcements. Figure 11 (b) depicts the volatility response to news items in-

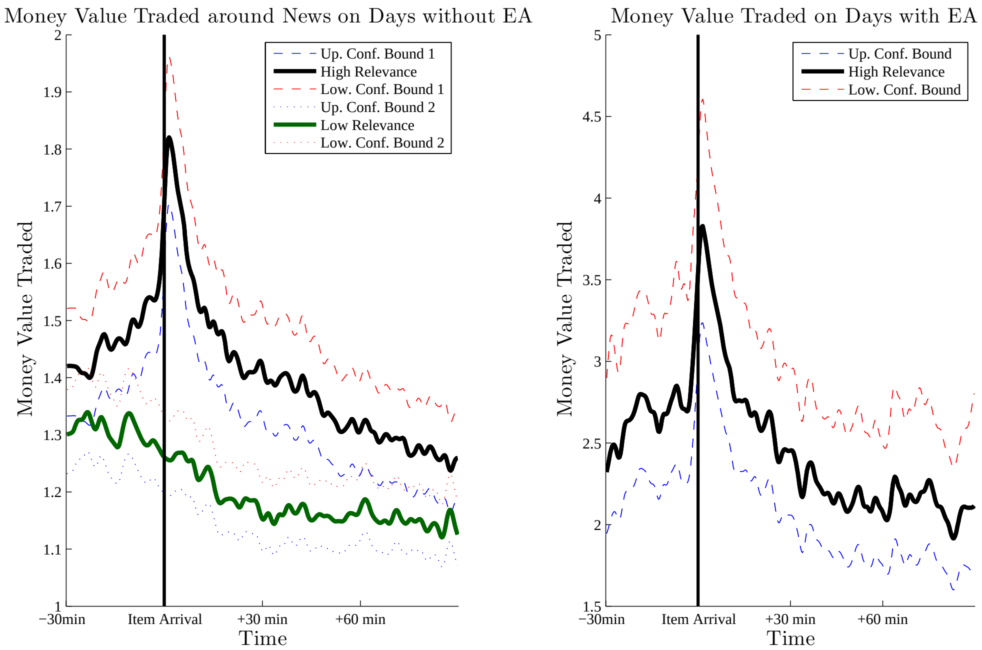

Figure 5: Money Value around News Arrivals. Smoothed via kernel regression.

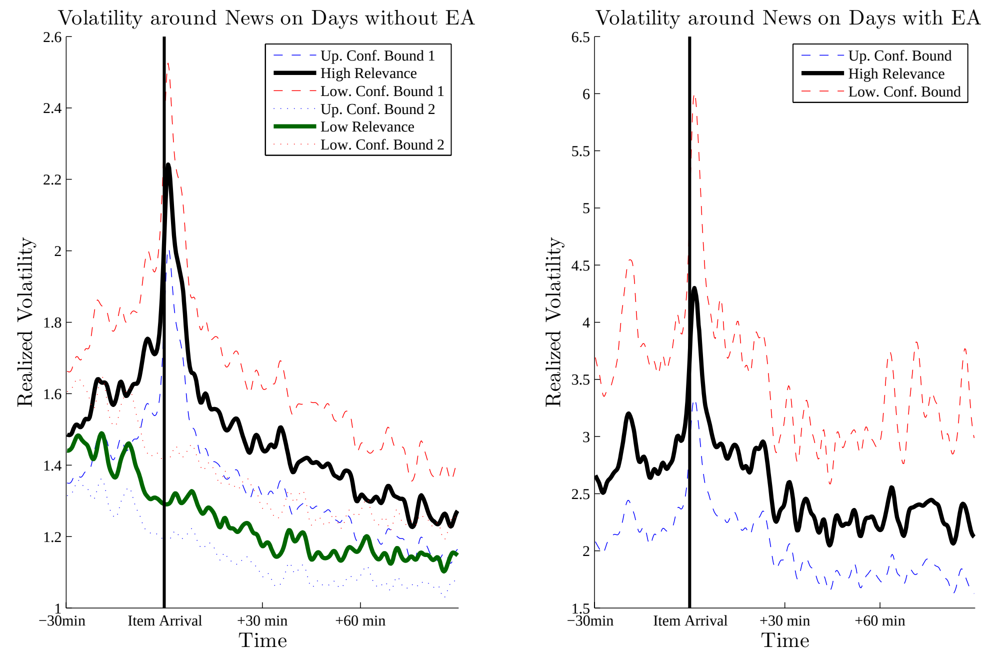

Figure 6: Realized Volatility around News Arrivals. Smoothed via kernel regression.

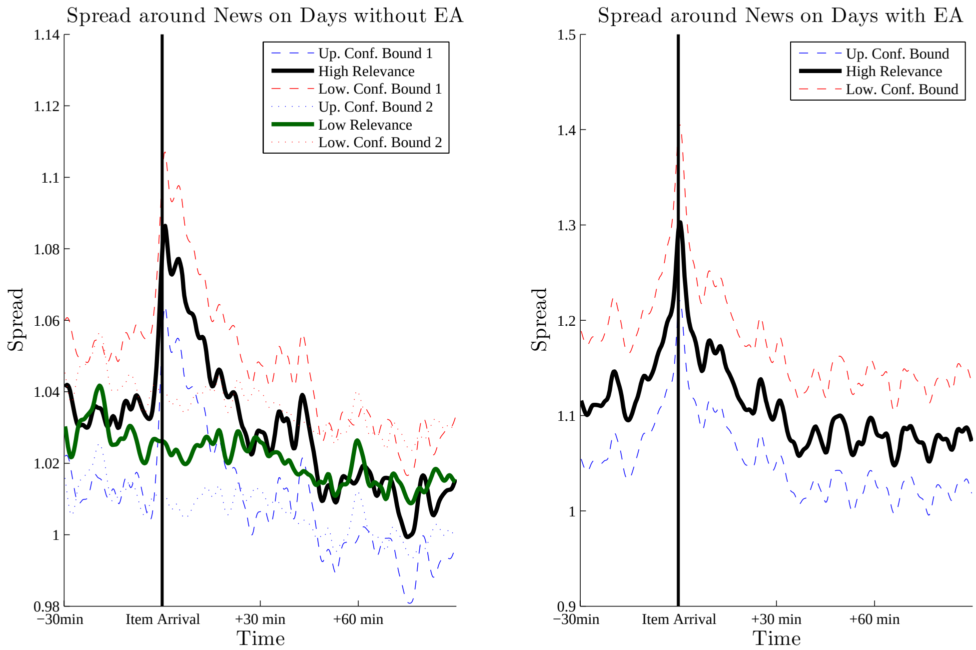

Figure 7: Spread around News Arrivals. Smoothed via kernel regression.

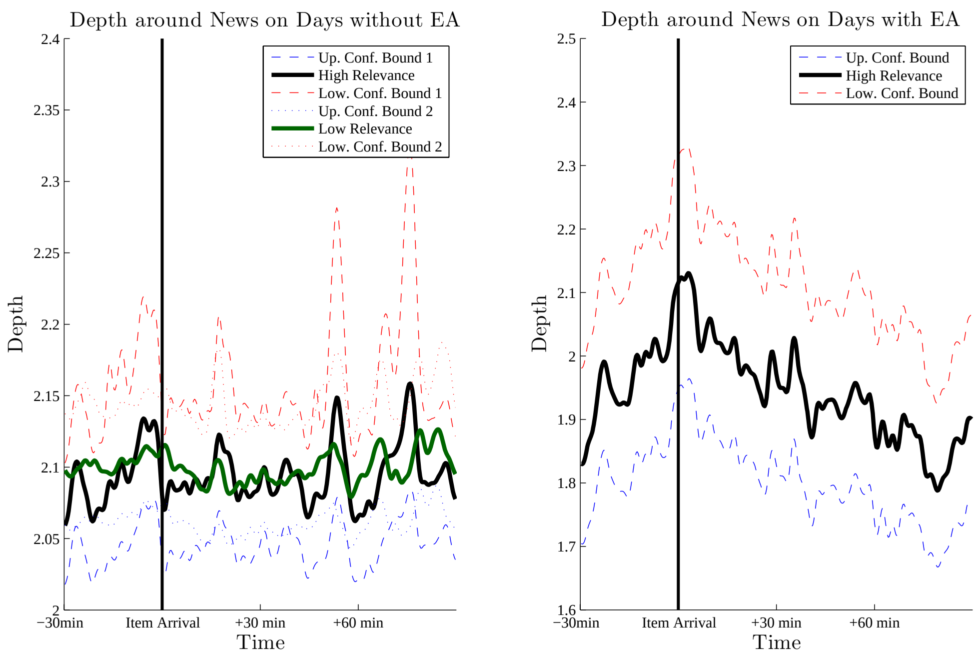

Figure 8: Cumulated Ask and Bid Depth around News Arrivals. Smoothed via kernel regression.

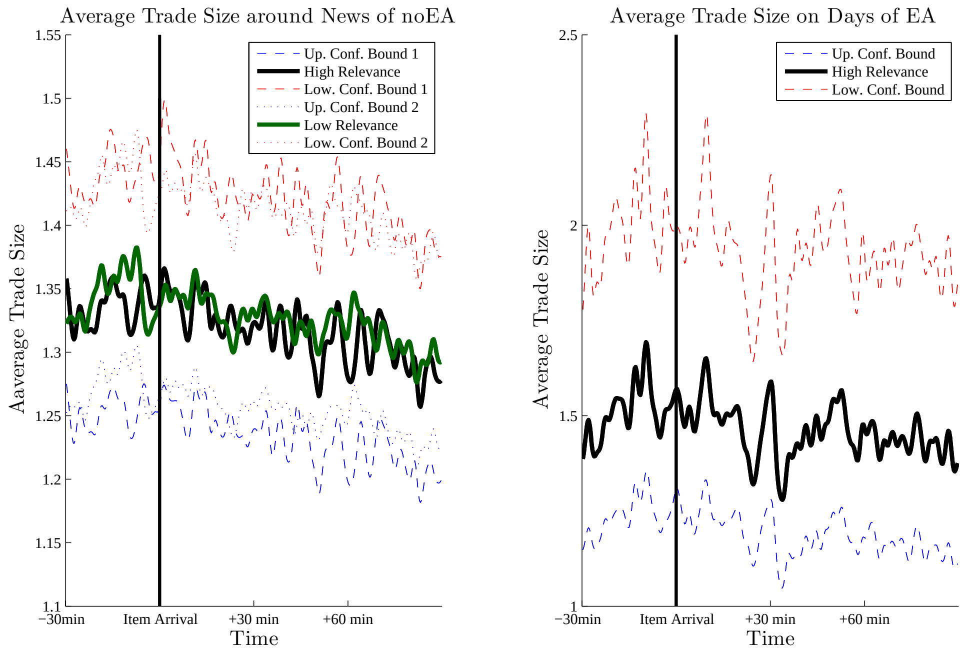

Figure 9: Average Trade Size around News Arrivals. Smoothed via kernel regression.

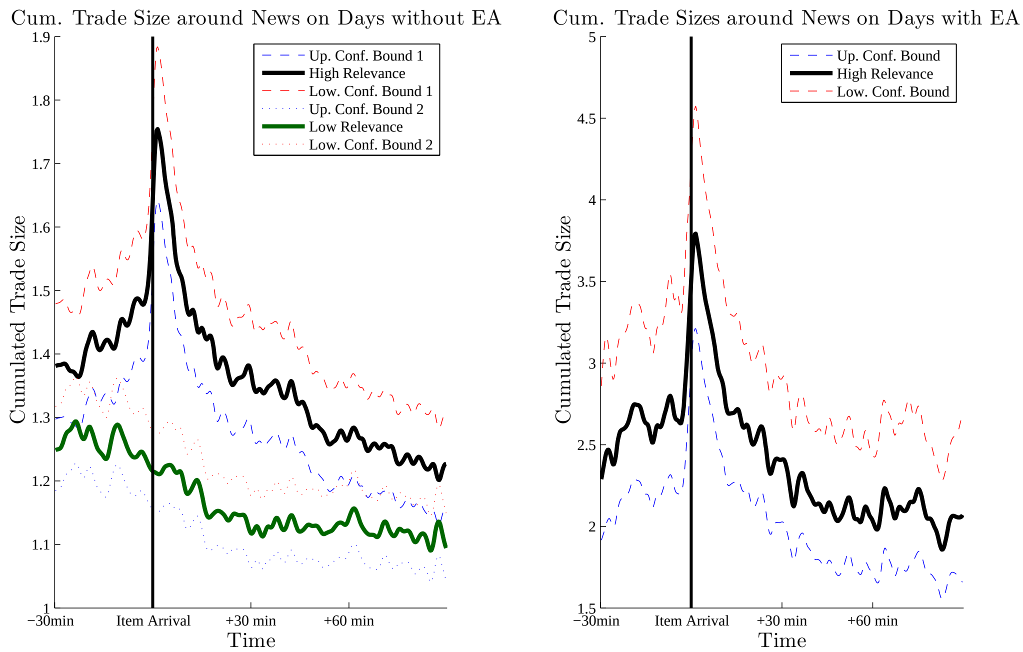

Figure 10: Cumulated Trade Size around News Arrivals. Smoothed via kernel regression.

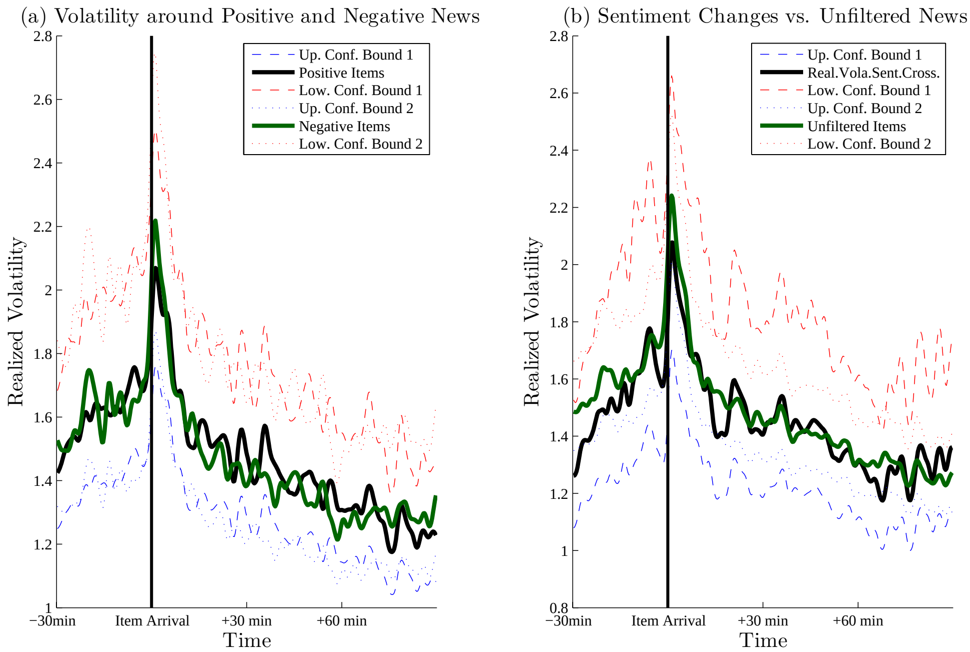

Figure 11: Volatility Reaction to News Filtered Based on Sign and Sign Changes. Smoothed via kernel regression.

dicating *changes* of sentiments. Here, we select news items only if their sign is contrary to that of a sequence of at least three previous news items with identical signs. The underlying idea is that a negative (positive) news disclosure might have a stronger impact when the recent market sentiment has been positive (negative). As depicted by both figures, we observe virtually no evidence for market reactions in volatility depending on the sign of news. This is in contrast to corresponding results based on macroeconomic announcements as reported, e.g., by Hautsch and Hess (2002) and might be explained by the existence of too much idiosyncratic noise in company-specific news. Similar findings are also obtained for the other variables.

3.2 Return Behavior

To test for abnormal returns we employ the event study framework as outlined in Campbell et al. (1997). As a model for 'normal' returns we assume the market model

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim (0, \sigma_i^2), \tag{1}$$

where t denotes the underlying (20 second) intervals, R_{mt} is the market return, computed as the return of the FTSE 100 index, and R_{it} is the return for stock i. Model

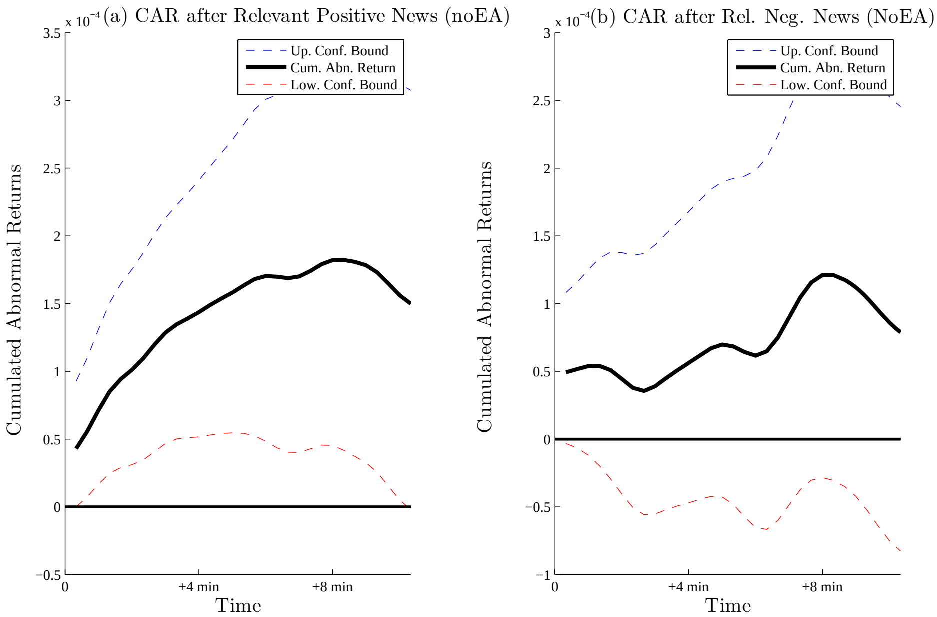

Figure 12: Cumulated Abnormal Returns after Positive and Negative News (High Relevance on NoEA Days). Smoothed via kernel regression.

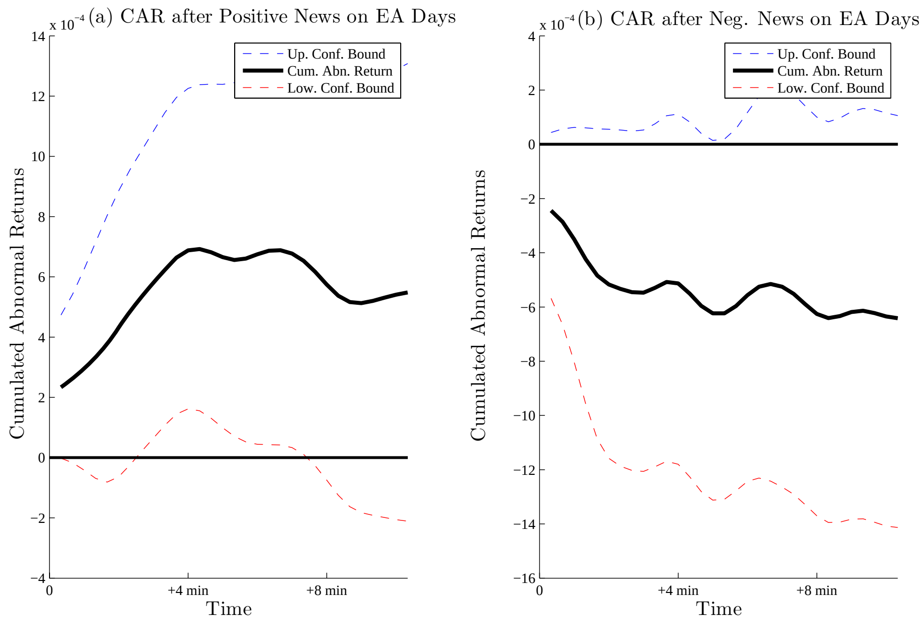

Figure 13: Cumulated Abnormal Returns after Positive and Negative News (News on EA Days). Smoothed via kernel regression.

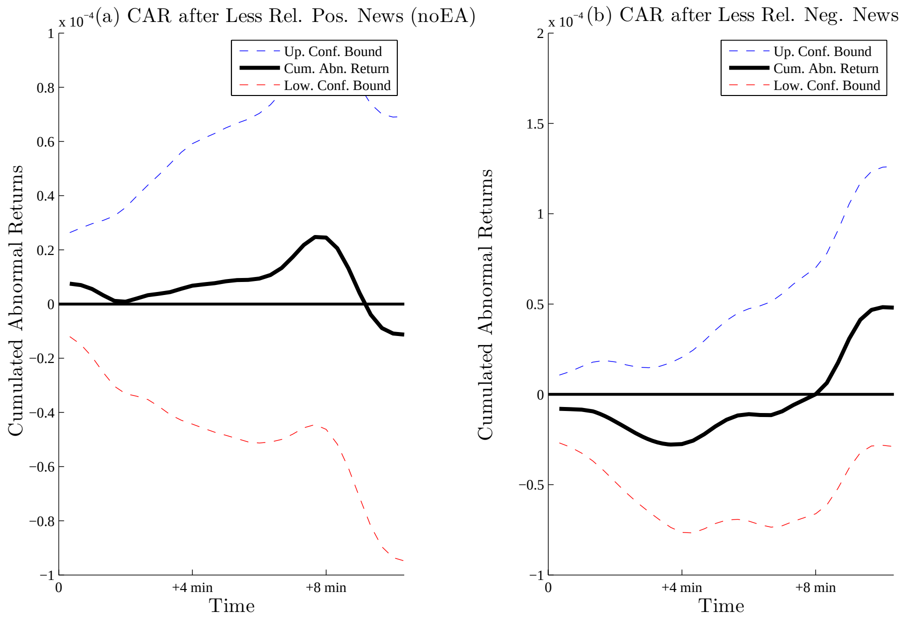

Figure 14: Cumulated Abnormal Returns after Positive and Negative News (Low Relevance News on EA Days). Smoothed via kernel regression.

(1) is estimated based on the complete 20-second return time series *without* including the event windows. Using the resulting parameter estimates, we compute the abnormal returns $\widehat{AR}_{it} := R_{it} - \widehat{\alpha}_i - \widehat{\beta}_i R_{mt}$ during the event windows. Let $\widehat{\mathbf{AR}}_i^k$ denote the $((T_2+1)\times 1)$ vector of abnormal returns for event k of stock i computed between time points I_0 and I_{T_2} in Figure 4. Let γ_j be a $(j \times 1)$ vector consisting of j ones, $1 \leq j \leq T_2+1$. Then, we define the cumulated abnormal return for interval j *after* the event time as

$$\widehat{CAR}_{ij}^k := \gamma_j' \widehat{\mathbf{AR}}_i^k. \tag{2}$$

Averaging $\widehat{CAR}_{ij}^k$ yields

$$\overline{\widehat{CAR}}_j = \frac{1}{n}\left(\sum_i \sum_k \widehat{CAR}_{ij}^k\right), \tag{3}$$

where n is the total number of events over all stocks. Assuming (asymptotic) normality, 95% confidence intervals are computed as two times the standard deviation of the estimates $\widehat{CAR}_j$.

Figure 12 shows the averaged cumulated abnormal returns (ACAR) $\overline{\widehat{CAR}}$ employing the relevant noEA news set. We observe significantly positive cumulated abnormal returns after positive news arrivals. In case of negative relevant news arrivals ACARs are surprisingly still positive, but less significant and lower in magnitude. A more distinct pattern is observed for the EA news set (Figure 13). Here, price movements are significant and in line with news' direction. This finding indicates the specific information content of news related to earnings announcements compared to other news items. Not surprisingly, less relevant news (noEA) do not induce significant abnormal returns (see Figure 14). The overall stronger reactions after positive news might be explained by the fact that during 2007 stock markets have been generally bearish making positive news items more striking than negative news.

4 Market Dynamics around News

4.1 Econometric Methodology

The unconditional analysis of the previous section provides strong indications for information-driven market reactions to news disclosures. However, as shown by Figures 15 to 17 (for a representative sample of stocks), we observe significant autocorrelations as well as cross-correlations in volatility and trading activity (see in the Appendix for the cross- and autocorrelations of the other variables). In order to avoid spurious results, these interdependencies have to be explicitly taken into account. Therefore, we suggest a four-dimensional model for the realized variance, the money value traded, the bid-ask spread and market depth. Money value traded is highly correlated with cumulated and average trade sizes and thus sufficiently captures the overall trading intensity. Moreover, as high-frequency volatility and liquidity are only weakly related to (signed) returns, we refrain from including the latter in the model. Accordingly, the vector of endogenous variables is

$$\mathbf{y}_t = \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \end{pmatrix} := \begin{pmatrix} \textit{money value traded} \\ \textit{realized volatility} \\ \textit{bid} - \textit{ask spread} \\ \textit{market depth} \end{pmatrix}. \tag{4}$$

The fact that even for liquid stocks there is not necessarily a transaction in every 20 second interval induces a non-trivial fraction of zero observations for money value traded and realized volatility (see Figure 18). To account for these effects, we suggest explicitly differentiating between the cases of trading, $y_{1t} > 0$, and no trading, $y_{1t} = 0$,

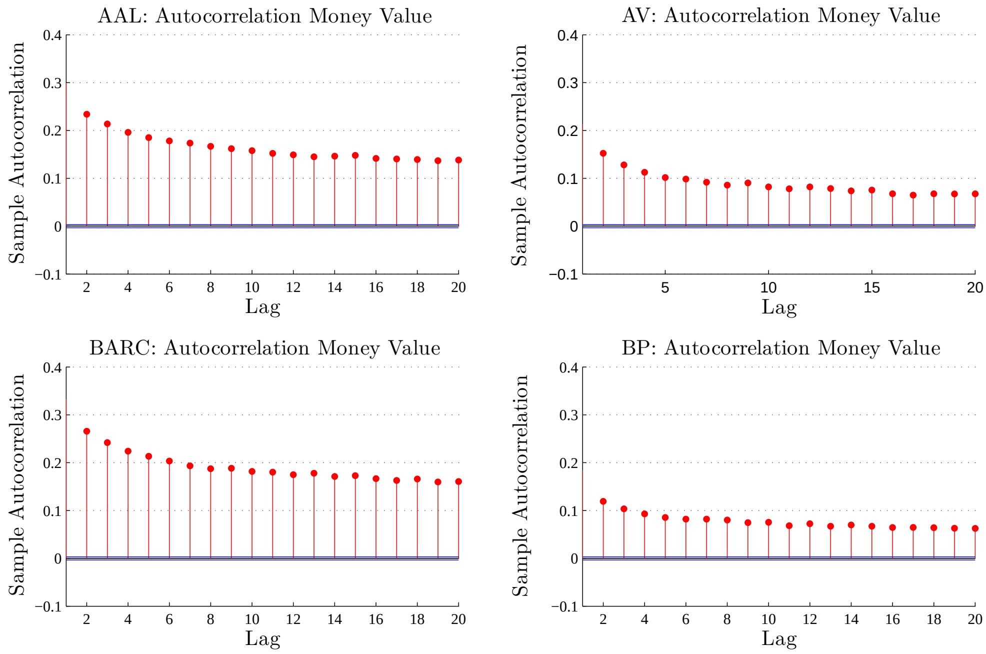

Figure 15: Autocorrelation Plots for Money Value Traded

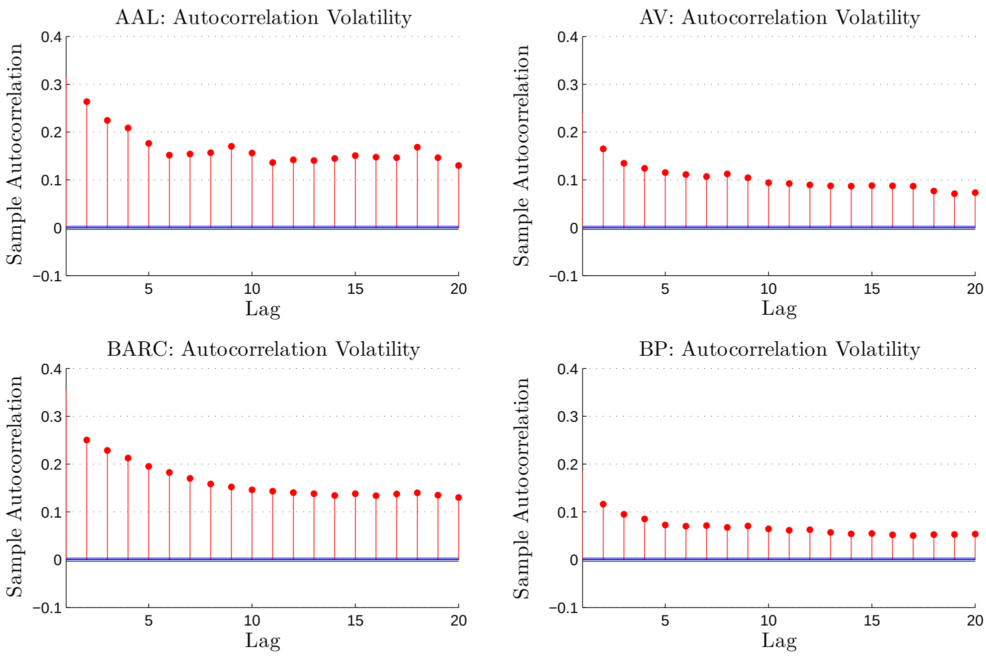

Figure 16: Autocorrelation Plots for Volatility

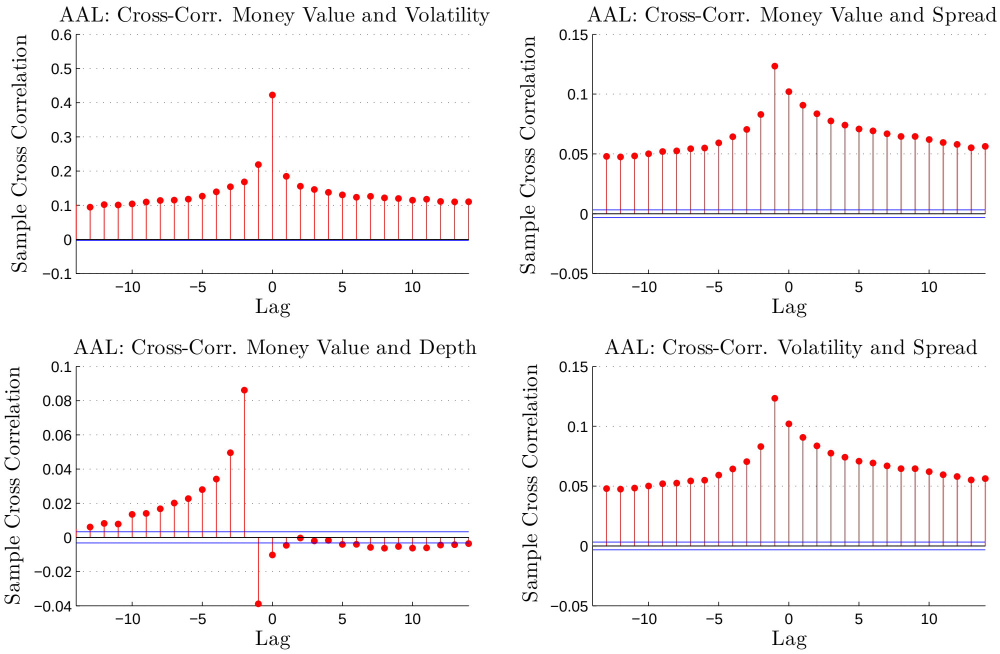

Figure 17: Cross-Correlations for the AAL stock

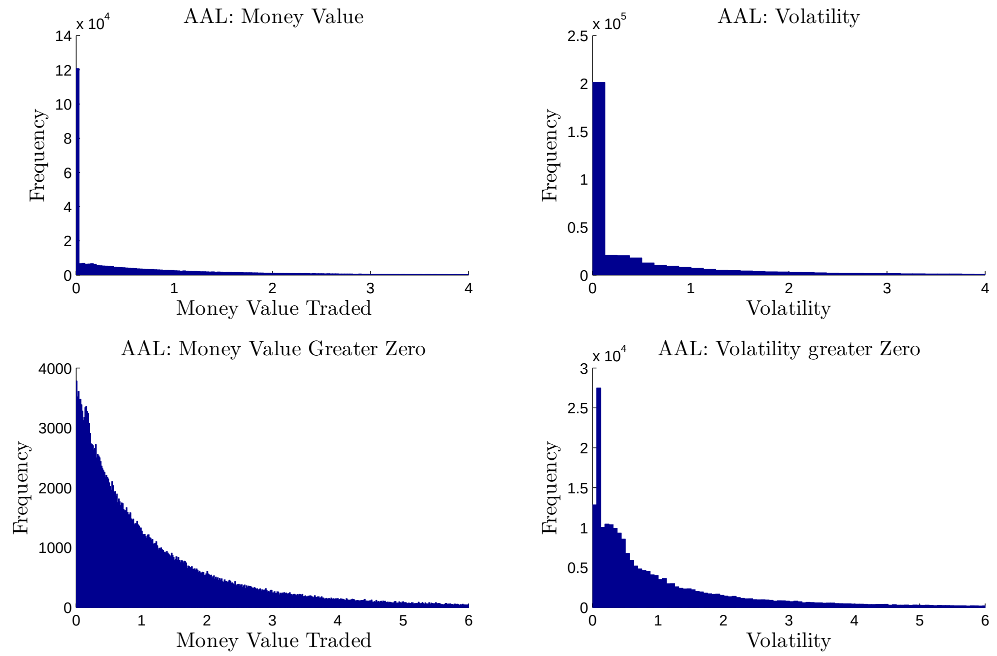

Figure 18: Histograms for Money Value and Volatility for AAL (upper two: unconditional, lower two: $(y_t|y_t > 0)$)

in interval t. Correspondingly, the log likelihood function is given by

$$\begin{aligned}\ln \mathcal{L}(\mathbf{y}; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3) = &\sum_{t=1}^{T} \{\ln f(\mathbf{y}_t | y_{1t} > 0; \boldsymbol{\theta}_1) + \ln \mathrm{P}(y_{1t} > 0; \boldsymbol{\theta}_2)\} \cdot \mathbb{1}(y_{1t} > 0) \\ &+ \sum_{t=1}^{T} \{\ln \mathrm{P}(y_{1t} = 0; \boldsymbol{\theta}_2) + \ln f(\mathbf{y}_t | y_{1t} = 0; \boldsymbol{\theta}_3)\} \cdot \mathbb{1}(y_{1t} = 0),\end{aligned}$$

where $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$ denote corresponding parameter sets.

As long as the parameter sets $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$ are disjoint, the likelihood components can be maximized separately. Since $f(\mathbf{y}_t | y_{1t} = 0; \boldsymbol{\theta}_3)$ is not in the core of our interest, we leave it unspecified. To parameterize $f(\mathbf{y}_t | y_{1t} > 0; \boldsymbol{\theta}_1)$, we suggest a VAR specification given by

$$\mathbf{y}_t | y_{1t} > 0 = \mathbf{c} + \sum_{i=1}^{p} (\boldsymbol{\Gamma}_i \mathbf{y}_{t-i} + \boldsymbol{\Psi}_i Z_{t-i}) + \boldsymbol{\Xi} \cdot \mathbf{D}_t^x + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Omega}), \tag{5}$$

where $\boldsymbol{\Gamma}_i$ and $\boldsymbol{\Xi}$ denote (4×4) and $(4 \times (p_1 + p_2 + 1))$ coefficient matrices.[2] Lags of the dummy $Z_t := \mathbb{1}_{(y_{1t}=0)}$ capture previous periods of nontrading with corresponding (4×1) coefficient vectors $\boldsymbol{\Psi}_i$. In order to capture the time-dependent impact of news we define appropriate dummy variables

$d_t^r = 1$ in case of relevant (noEA) news in t and zero otherwise,

$d_t^l = 1$ in case of less relevant (noEA) news in t and zero otherwise,

$d_t^{ea} = 1$ in case of EA news in t and zero otherwise.

Then, $\mathbf{D}_t^x := (d_{t+p_1}^x ....d_{t-p_2}^x)'$ with $x \in \{r, l, ea\}$ is a vector of time dummies indicating the different types of news and covering p_1 intervals before and p_2 intervals after news arrival. Model (5) can be consistently (though not necessarily efficiently) estimated equation by equation using ordinary least squares.

The conditional probabilities for the occurrence of zero observations (i.e., no trading) in period t, $\mathrm{P}(y_{1t} = 0; \boldsymbol{\theta}_2)$, are parameterized in terms of a probit specification for the money value equation. Let $\mathbf{x}_t$ contain all right-hand side variables of equation (5), i.e., $\mathbf{x}_t' := [1 \ \mathbf{y}_{t-1}' ... \mathbf{y}_{t-p}' \ Z_{t-1}' ... Z_{t-p}' \ \mathbf{D}_t'^x]$. Assuming a normally distributed

[2] Alternatively, one could use a multivariate multiplicative error model (MEM) as proposed by Manganelli (2005). However, since a MEM can be re-written in terms of a V(ARMA) model both frameworks are ultimately not very different.

latent process $y^*_{1t} \sim N(\mathbf{x}_t'\boldsymbol{\theta}_2, 1)$ underlying the trading "decision", we have

$$P(y^*_{1t} > 0) = \Phi(\mathbf{x}_t'\boldsymbol{\theta}_2), \qquad \text{if } y^*_{1t} > 0 \;\Leftrightarrow\; y_{1t} > 0, \tag{6}$$

$$P(y^*_{1t} \leq 0) = 1 - \Phi(\mathbf{x}_t'\boldsymbol{\theta}_2), \qquad \text{if } y^*_{1t} \leq 0 \;\Leftrightarrow\; y_{1t} = 0, \tag{7}$$

for the binary decision $y_{1t} > 0$ vs. $y_{1t} = 0$. The probit model is straightforwardly estimated by maximum likelihood.

The model is applied to each stock in our sample. In order to obtain equal lag structures in all equations which eases cross-sectional comparisons and the computation of cross-sectional averages, we choose a universal lag length of 10 for all stocks. This lag length is sufficiently close to the individually optimal lag length according to the Bayes Information Criterion and does not restrict the validity of the results discussed below. In the following we show the cross-sectional averages of point estimates and corresponding standard errors.

4.2 Estimation Results

In order to keep the model computationally tractable and parsimonious, the three types of news dummies $\mathbf{D}_t^r$, $\mathbf{D}_t^l$ and $\mathbf{D}_t^{ea}$ are included separately. Since the VAR dynamics in the individual specifications are very similar, we concentrate on the estimates of the model including the noEA dummy set associated with high relevance ($\mathbf{D}_t^r$). Depending on the number of underlying trading days, the individual time series for the 35 stocks in the sample contain up to 369,000 observations. Table 2 reports the corresponding averaged estimates. For sake of brevity, we do not show coefficients for lags of the dependent variables greater than two. Likewise, coefficient estimates for the dummies Z_t are not reported.[3] News dummies cover 40 seconds before the disclosure and 100 seconds thereafter.

Analyzing the dynamics of volatility and liquidity, we can summarize the following findings: First, all variables reveal significantly positive own dynamics. This is strongly expected given the underlying autocorrelations reported above. Second, we observe a significantly positive relationship between money value traded and volatility. Hence, volatility and trading activity are closely dependent not only on a daily level as suggested by Clark (1973) and Tauchen and Pitts (1983), among others, but obviously also on a high-frequency level (see, e.g., Hautsch (2008)). Third, bid-ask spreads are higher in periods of high liquidity demand and volatility but are lower in periods of high liquidity supply (represented by the depth). Similarly, depth is lower if recent

[3]These results are available upon request from the authors.

Table 2: Average VAR Results: Dynamics ($\mathbf{y}|y_{1t} > 0$)

Model

	Variable	VAR ($\mathbf{y}_t|y_{1t} > 0$)				Probit Model
		Money Value	Volatility	Spread	Depth	Money Value $\mathbb{1}(y_{1t} > 0)$
	c	-1,749***	3,125***	0,342***	0,422***	0,731***
		(0,307)	(0,507)	(0,047)	(0,026)	(0,297)
Money Value	mv_{t-1}	0,183***	0,125***	0,001***	-0,001***	0,082***
		(0,003)	(0,005)	(0,000)	(0,001)	(0,007)
	mv_{t-2}	0,094***	0,037***	0,001***	0,001	0,035***
		(0,002)	(0,003)	(0,000)	(0,000)	(0,004)
Real. Vola.	rv_{t-1}	0,013***	0,168***	0,000	0,000	0,016***
		(0,001)	(0,008)	(0,000)	(0,000)	(0,003)
	rv_{t-2}	0,008***	0,082***	0,000	0,000*	0,008***
		(0,001)	(0,004)	(0,000)	(0,000)	(0,000)
Spread	spr_{t-1}	1,724***	2,078***	0,533***	-0,104***	0,762***
		(0,323)	(0,402)	(0,014)	(0,032)	(0,200)
	spr_{t-2}	-1,742***	0,124***	0,103***	0,043**	-0,985***
		(0,337)	(0,057)	(0,004)	(0,037)	(0,090)
Depth	$dpth_{t-1}$	-2,002***	-1,855***	-0,117***	0,480***	-1,088***
		(0,283)	(0,251)	(0,019)	(0,017)	(0,259)
	$dpth_{t-2}$	2,594***	-0,413	-0,006	0,088***	0,463***
		(0,314)	(0,081)	(0,002)	(0,007)	(0,060)

Table 2: Average VAR Results: NoEA News Dummies (high relevance) (Cont'd)

Model

	Variable	$\mathbf{D}_t^r$ for VAR ($\mathbf{y}_t \| y_{1t} > 0$)				Probit Model
		Money Value	Volatility	Spread	Depth	Money Value $\mathbb{1}(y_{1t} > 0)$
Dummy Leads	d_{t+2}	0,525	0,156	0,007	-0,029	0,012
		(0,203)	(0,240)	(0,027)	(0,025)	(0,010)
	d_{t+1}	0,328	0,108	0,029	-0,034	0,045
		(0,128)	(0,211)	(0,016)	(0,018)	(0,000)
Item Dummy	d_t	0,564*	0,940	0,091	0,017	0,079
		(0,135)	(0,304)	(0,027)	(0,015)	(0,030)
Dummy Lags	d_{t-1}	1,173***	1,326***	0,019	-0,022	0,076
		(0,271)	(0,333)	(0,015)	(0,020)	(0,003)
	d_{t-2}	0,999**	1,241**	0,029	-0,010	0,077
		(0,292)	(0,390)	(0,015)	(0,024)	(0,010)
	d_{t-3}	0,730	0,913	0,062	0,033	0,004
		(0,205)	(0,310)	(0,024)	(0,018)	(0,003)
	d_{t-4}	0,444	0,616	-0,006	0,003	0,061
		(0,101)	(0,319)	(0,020)	(0,017)	(0,014)
	d_{t-5}	0,646	0,739	0,043	0,015	0,355
		(0,170)	(0,273)	(0,018)	(0,018)	(0,106)

Note: The first four columns show OLS estimation results of system (5) with relevant noEA news dummies $\mathbf{D}_t^r$. The last column shows the ML estimation results of the corresponding probit model (6) with the same set of news dummies. Reported coefficients are averages of the estimates for each individual stock. Significance is reported based on average t-statistics. (Cross-sectional) standard errors of the averaged coefficients are given in parentheses below. $(***)$ denotes significance of the average coefficient estimates at the 1 % level, $(**)$ at the 5 % level, and $(*)$ at the 10 % level.

Table 3: Average VAR Results for EA News Dummies and NoEA Low Relevance News Dummies ($\mathbf{y}_t|y_{1t} > 0$)

Model

	Variable	$\mathbf{D}_t^l$ for VAR ($\mathbf{y}_t\|y_{1t} > 0$)				$\mathbf{D}_t^{ea}$ for VAR ($\mathbf{y}_t\|y_{1t} > 0$)			
		Money Value	Volatility	Spread	Depth	Money Value	Volatility	Spread	Depth
Dummy Leads	d_{t+2}	0,224	0,293	0,015	0,010	1,077	0,222	0,076	0,024
		(0,081)	(0,139)	(0,009)	(0,013)	(0,436)	(0,400)	(0,119)	(0,073)
	d_{t+1}	0,031	0,031	-0,005	-0,016	2,056**	0,282	0,207	0,081
		(0,071)	(0,105)	(0,010)	(0,014)	(0,563)	(0,476)	(0,172)	(0,091)
Item Dummy	d_t	0,151	0,013	0,005	0,017	1,280	0,500	0,059	-0,050
		(0,071)	(0,089)	(0,012)	(0,018)	(0,431)	(0,592)	(0,091)	(0,055)
Dummy Lags	d_{t-1}	0,227	0,372	-0,003	0,010	1,069	0,520	-0,070	0,149
		(0,065)	(0,119)	(0,013)	(0,015)	(0,394)	(0,647)	(0,060)	(0,094)
	d_{t-2}	0,186	0,156	-0,015	-0,006	1,800*	1,260	0,101	0,011
		(0,093)	(0,125)	(0,010)	(0,009)	(0,446)	(0,521)	(0,103)	(0,057)
	d_{t-3}	0,018	-0,098	0,005	0,009	1,030	0,591	-0,024	0,081
		(0,061)	(0,091)	(0,011)	(0,010)	(0,438)	(0,674)	(0,056)	(0,092)
	d_{t-4}	0,316	0,074	-0,014	0,017	1,151	0,684	0,039	-0,001
		(0,122)	(0,102)	(0,008)	(0,010)	(0,455)	(0,816)	(0,065)	(0,064)
	d_{t-5}	0,174	0,210	0,005	-0,002	1,286	0,610	-0,117	-0,026
		(0,099)	(0,085)	(0,011)	(0,009)	(0,356)	(0,349)	(0,077)	(0,051)

*Note:*The columns show OLS estimates for the news dummies of two VAR estimations using the model (5). The left-hand columns show coefficient estimates for the news dummies $\mathbf{D}_t^l$, whereas the right-hand columns refers to $\mathbf{D}_t^{ea}$. Reported coefficients are averages of the estimates for each individual stock. Significance is reported based on average t-statistics. (Cross-sectional) standard errors of the averaged coefficients are given in parentheses below. $(***)$ denotes significance of the average coefficient estimates at the 1 % level, $(**)$ at the 5 % level, and $(*)$ at the 10 % level.

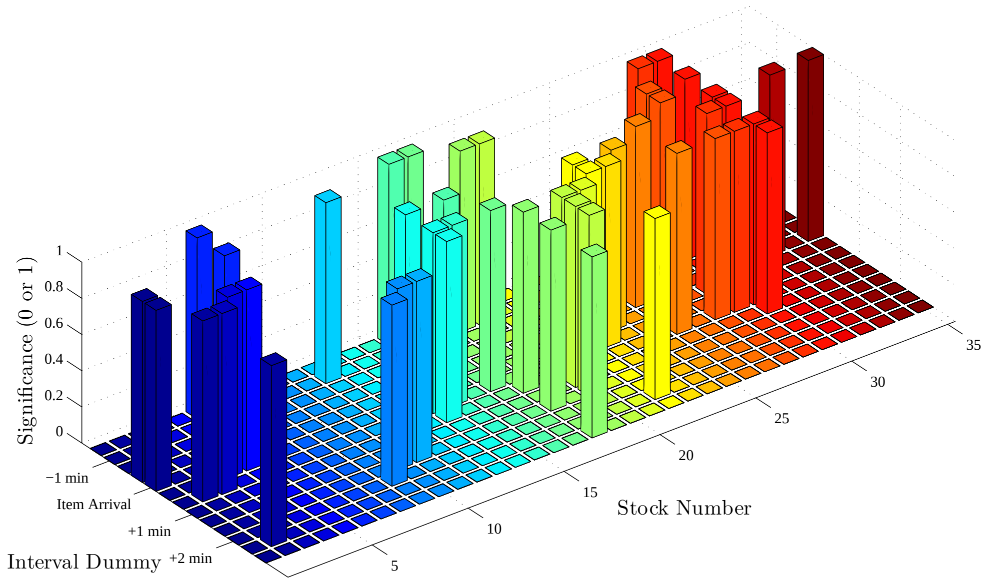

Figure 19: Proportions of significant news dummies in the spread equation (5 % level) based on relevant noEA news. All coefficient signs are positive.

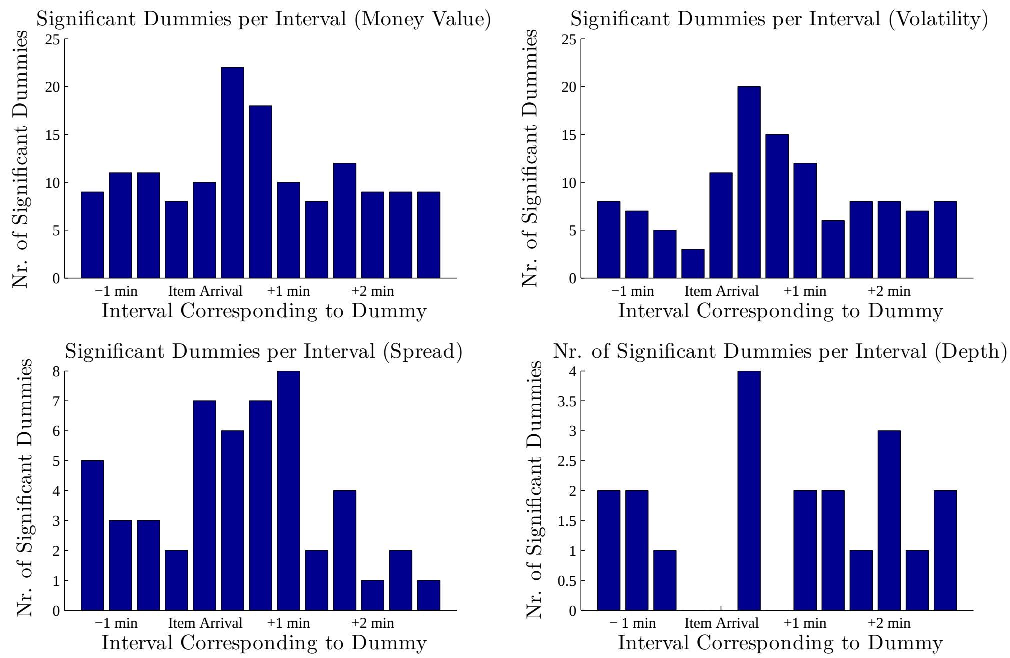

Figure 20: Numbers of significant dummy variables in the intervals around the news disclosure (relevant noEA partition).

trading activity and volatility have been high. Fourth, virtually no causalities from spreads and market depth on volatilities and volumes are observed. While liquidity demand and volatility stimulate liquidity supply, the converse relationship is thus not necessarily true.

Quantifying the average impact of news, we observe that the market reaction starts immediately after news disclosures. Due to the persistence in market dynamics information effects are carried over to subsequent periods. It is therefore not surprising that the direct impact of news as captured by the dummy variables dies out relatively quickly. It turns out that only the volatility and the trading volume are significantly (directly) affected by news. Conversely, we do not find corresponding effects for spreads and depths. These results are different to the unconditional estimates obtained in Section 3 and indicate that reactions of these variables during announcement periods are strongly induced by spill-overs from volatility and volume but do not necessarily arise from news in sentiments solely. Moreover, as shown by Table 3 depicting the corresponding results for EA news and noEA news with low relevance (indicator ≤ 0.5), we conclude that significant responses are generally only observable after the occurrence of relevant news.

Estimation results for the probit model widely confirm those for the VAR model. However, the fact that all news dummy variables are insignificant indicates that the probability for the occurrence of a trade in a 20-sec interval is not driven by news arrivals.

Though the averaged estimates capture the major features common to all assets, most stocks still reveal idiosyncratic responses to news. Figure 19 depicts the proportions of (5%) significant spread reactions to relevant noEA news for each stock in the sample. Though the *average* spread reaction is insignificant, we still observe significant individual spread responses for 27 out of 35 stocks in the sample. Similar results are shown (not depicted here) for market depth, whereas stock-specific effects for volatility and money value traded are more stable and in line with the average results shown above. Figure 20 reflects that the significant (positive) dummies for most stocks center around the item arrival interval. Accordingly, we can conclude that there is evidence for news-implied reactions in spreads and depth, which are, however, diffuse across the stock universe.

4.3 Impulse Response Analysis

To quantify the long-run market response to the arrival of a news item we perform an impulse response analysis. A 'news shock' is defined by a change in the news dummies.

As the arrival of news generally stimulates trading activity, it is sufficient to conduct the analysis given there is trading activity throughout the post-announcement periods, i.e. $\mathbf{y}_j | y_{1j} > 0$ for all $j = t, \dots, t+s$.

Then, the response after s periods to a news arrival in t is computed as

$$\Delta_s(\boldsymbol{\theta_1}) := E[\mathbf{y}_{t+s}|\Omega_{t-1}, d_t^x = 1; \boldsymbol{\theta}_1] - \underbrace{E[\mathbf{y}_{t+s}|\Omega_{t-1}, d_t^x = 0; \boldsymbol{\theta}_1]}_{(*)}, \quad x \in \{r, l, ea\}, \qquad (8)$$

where Ω_{t-1} represents the history of the multivariate process at t and the second term $(*)$ removes the effect of constants and initial values on the response function. Let $p_1 = 0$, $p_2 > 0$ and $\widehat{\boldsymbol{\Xi}}_{\cdot i}$ denote the i-th column of $\widehat{\boldsymbol{\Xi}}$. Coefficients in the second to p_2-th columns of $\widehat{\boldsymbol{\Xi}}$ that are not significantly different from zero at the 5% level are assumed to be zero throughout. Initially we have

$$\begin{aligned}\Delta_0 =& E[\mathbf{y}_t|\Omega_{t-1}, d_t^x = 1; \boldsymbol{\theta}_1] - E[\mathbf{y}_t|\Omega_{t-1}, d_t^x = 0; \boldsymbol{\theta}_1] \\ =& \widehat{\mathbf{c}} + \sum_{i=1}^{p}(\widehat{\boldsymbol{\Gamma}}_i \mathbf{y}_{t-i} + \widehat{\boldsymbol{\Psi}}_i Z_{t-i}) + \widehat{\boldsymbol{\Xi}}_{\cdot 1} - \left(\widehat{\mathbf{c}} + \sum_{i=1}^{p}(\widehat{\boldsymbol{\Gamma}}_i \mathbf{y}_{t-i} + \widehat{\boldsymbol{\Psi}}_i Z_{t-i})\right) = \widehat{\boldsymbol{\Xi}}_{\cdot 1}.\end{aligned}$$

Since the initial conditions, constants and Z_t cancel out, the responses in $t+s$, $s = 1, 2, ...$, to the dummy impulse in t are given as

$$\Delta_1 = \widehat{\boldsymbol{\Gamma}}_1 \Delta_0 + \widehat{\boldsymbol{\Xi}}_{\cdot 2}, \quad \Delta_2 = \widehat{\boldsymbol{\Gamma}}_1 \Delta_1 + \widehat{\boldsymbol{\Gamma}}_2 \Delta_0 + \widehat{\boldsymbol{\Xi}}_{\cdot 3}, \ ...$$

Standard errors of the response function are derived using the delta method. Accordingly, Δ_s is asymptotically distributed as

$$\Delta_s(\widehat{\boldsymbol{\theta_1}}) \xrightarrow{d} N(\Delta_s(\boldsymbol{\theta}_1), (1/T)\mathbf{G}_s(\boldsymbol{\Omega} \otimes \mathbf{Q}^{-1})\mathbf{G}_s'),$$

where $\mathbf{Q} = E[\mathbf{x}_t \mathbf{x}_t']$ and $\mathbf{G}_s = \frac{\partial \Delta_s(\boldsymbol{\theta_1})}{\partial \boldsymbol{\theta_1}'}$. Estimates for $\boldsymbol{\Omega}$ and $\mathbf{Q}$ are readily available from the VAR estimates. Following Hamilton (1994), we construct the columns of $\mathbf{G}_s$ based on finite differences according to

$$\frac{\partial \Delta_s(\widehat{\boldsymbol{\theta_1}})}{\partial \theta_{1i}} \approx \frac{\Delta_s(\widehat{\boldsymbol{\theta_1}} + \mathbf{e}_i h) - \Delta_s(\widehat{\boldsymbol{\theta_1}})}{h},$$

where h is some small number, θ_{1i} denotes the i-th element of $\boldsymbol{\theta_1}$ and e_i is the i-th unity vector.

Figures 21 to 23 show the impulse response to news-induced dummy variable changes based on the averaged VAR estimates. The depicted reaction to relevant

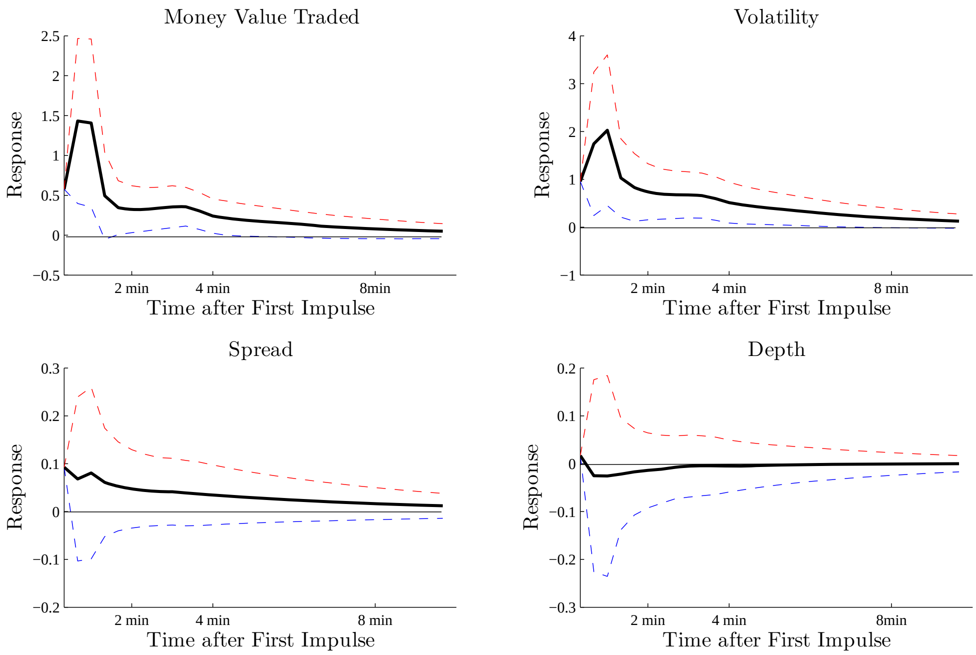

Figure 21: Response Analysis of a Change in the highly relevant noEA News Dummies (95% confidence intervals as dotted lines)

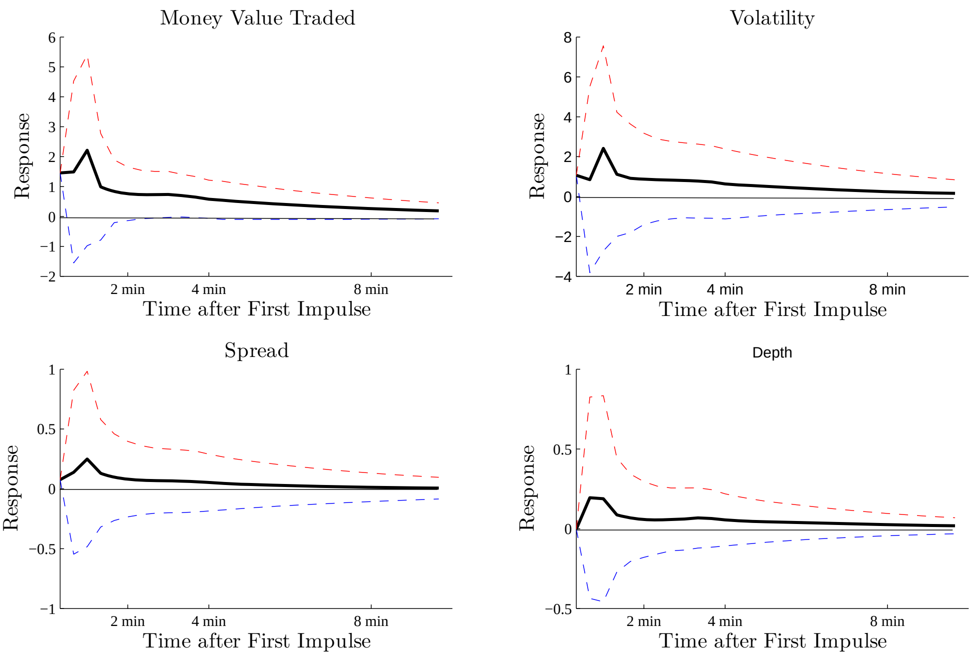

Figure 22: Response Analysis of a Change in the EA News Dummies (95% confidence intervals as dotted lines)

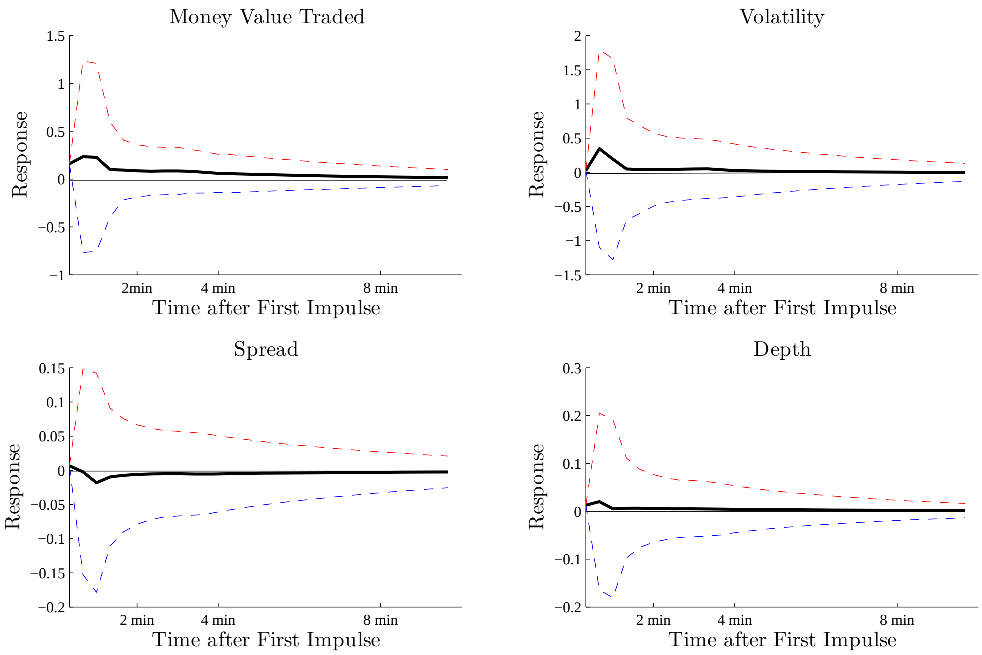

Figure 23: Response Analysis of a Change in the less relevant noEA News Dummies (95% confidence intervals as dotted lines)

noEA news mimics the unconditional market responses of volatility and money value traded quite well (cf. Figures 5 and 6). Nevertheless, while the reaction of money value traded is barely significant after the first minute after news arrival, the volatility response is more persistent and lasts until the fifth minute after the event. Moreover, as shown in Figures 22 and 23, market reactions to less relevant news and EA news are not statistically different from zero.

Overall, we can conclude that the dynamic analysis widely confirms the unconditional effects shown above. Obviously, volatility and trading volume are most sensitive to news arrival. Weaker reactions and a stronger impact of idiosyncratic effects are observed in spreads and depth. In order to check the robustness of our results, we have estimated several alternative specifications, in particular (i) a simple VAR model based on 20 second aggregates (without explicitly accounting for zero observations), (ii) the corner-solution model by Cragg (1971) for the conditional density based on 20 second aggregates, and (iii) simple VAR specifications based on 5 minute aggregates. For sake of brevity we refrain from reporting the corresponding estimates in the paper. It turns out that our findings are qualitatively quite stable across the individual specifications.

5 Conclusions

Motivated by the ongoing surge in the amount of electronic news, this study analyzes the impact of firm-specific news flow on the trading activity at the London Stock Exchange (LSE). The arrival of stock-specific news items is linked to liquidity, volatility and returns for a representative sample of stocks. While previous studies dominantly focus only on a part of published firm-specific news (typically earnings announcements), this study attempts covering the complete information flow provided by a news vendor.

Recording and analyzing the overall news flow for a specific asset is challenging since the amount of news, the number of news sources and the speed of information dissemination is rapidly increasing over time. Induced by the huge amount of information permanently published in all modern media, news are overlaid by substantial noise caused by irrelevant information. To reduce the impact of noise, we make use of data provided by an automated news analytics tool of the Reuters company which allows us to disentangle relevant news from irrelevant ones and to identify the sign of news. These identifications are based on indicators from linguistic pattern recognition algorithms. Until now this kind of news data has never been systematically studied in the literature. Consequently, the induced effects on intraday trading activity, volatility and liquidity are widely unknown. This paper addresses this question and explores the impact of news on high-frequency returns, trading volume, volatility, depth and spreads by means of a high-frequency VAR model.

Based on our empirical results we can summarize the following results. First, we find significant unconditional reactions in returns, volatility and liquidity. For trading volumes and volatilities these effects remain stable even if dynamics and cross-dependencies between the variables are taken into account. For market depth and spreads, news implied effects deteriorate and are less distinct in a multivariate framework. Second, market responses to information can only be identified for relevant news items. Conversely, less relevant news seems to be overlaid by noise. In this sense, our analysis confirms the usefulness of an automated linguistic pattern analysis. Third, it turns out that news impacts for individual stocks are influenced by considerable stock-specific noise. This is particularly true for the response of spreads and depth for which we find varying effects across the market. Fourth, the news impact on days of earnings announcements is different from the impact on other trading days. On these days, headlines on quarterly company earnings seem to be the dominating news reducing the importance of other information. Finally, we find evidence for market participants employing also other (sometimes more timely) news sources and for a general clustering of information. This is reflected by market activity being already significantly above average *before* the arrival of news on sentiments.

References

Andersen, T. G., T. Bollerslev, C. Diebold, and C. Vega (2003): "Micro effects of macro announcements: Real-time price discovery in foreign exchange?" *American Economic Review*, 93, 38–62.

Beaver, W. H. (1968): "The information content of annual earnings announcements," *Journal of Accounting Research*, 6, 67–92.

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

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997): *The econometrics of financial markets*, Princeton, New Jersey: Princeton University Press.

Clark, P. K. (1973): "A subordinate stochastic process model with finite variance for speculative prices," *Econometrica*, 41, 135–155.

Cragg, J. (1971): "Some statistical models for limited dependent variables with application to the demand for durable goods," *Econometrica*, 39, 829–844.

DeGennaro, R. P. and R. E. Shrieves (1997): "Public information releases, private information arrival and volatility in the foreign exchange market," *Journal of Empirical Finance*, 4, 295–315.

Ederington, L. H. and J. H. Lee (1993): "How markets process information: News releases and volatility," *The Journal of Finance*, 48, 1161–1191.

Fleming, M. J. and E. M. Remolona (1999): "Price formation and liquidity in the U.S. Treasury Market: The response to public information," *The Journal of Banking and Finance*, 28, 1441–1467.

Graham, J. R., J. Koski, and U. Loewenstein (2006): "Information flow and liquidity around anticipated and unanticipated dividend announcements," *The Journal of Business*, 79, 2301–2335.

Hamilton, J. D. (1994): *Time Series Analysis*, Princeton, New Jersey: Princeton University Press.

Hautsch, N. (2008): "Capturing common components in high frequency time series: a multivariate stochastic multiplicative error model," *Journal of Economic Dynamics and Control*, 32, 3978–4009.

Hautsch, N. and D. Hess (2002): "The processing of non-anticipated information in financial markets: Analyzing the impact of surprises in the employment report," *European Finance Review*, 6, 133–161.

Kalev, P. S., W. S. Liu, P. K. Pham, and E. Jarnecic (2004): "Public information arrival and volatility of intraday stock returns," *The Journal of Banking and Finance*, 28, 1441–1467.

Landsman, W. R. and E. L. Maydew (2002): "Beaver (1968) revisited: Has the information content of quarterly earnings announcements declined in the past three decades?" *Journal of Accounting Research*, 40, 797–808.

Malatesta, P. H. and R. Thompson (1985): "Partially Anticipated Events: A Model of Stock Price Reactions with an Application to Corporate Acquisitions," *Journal of Financial Economics*, 14, 237–250.

Manganelli, S. (2005): "Duration, Volume and Volatility Impact of Trades," *Journal of Financial Markets*, 8, 377–399.

Mitchell, M. L. and J. H. Mulherin (1994): "The impact of public information on the stock market," *The Journal of Finance*, 49, 923–950.

Ranaldo, A. (2008): "Intraday market dynamics around public information disclosures," in *Stock Market Liquidity*, ed. by F.-S. Lhabitant and G. Gregoriou, New Jersey: John Wiley and Sons, chap. 11, 199–226.

Tauchen, G. E. and M. Pitts (1983): "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, 51, 485–505.

6 Appendix

A Note on the Computation of Standard Errors of Across-Market Averages

In the following we describe two ways of computing the mean reactions and their standard errors. The pooled average used in Section 3 is based on the model

$$X_i = \mu + \epsilon_i, \qquad \epsilon_i \sim i.i.d.\ N(0, \sigma^2), \quad i = 1, .., n, \tag{9}$$

where we have suppressed the I_j index for the respective interval around the news item. Inference is based on the pooled estimator for the mean, $\overline{X} = 1/n \sum_{i=1}^{n} X_i$, where 95% confidence intervals are given as $\overline{X} \pm 2 * \widehat{\sigma}/\sqrt{n}$ with $\widehat{\sigma}^2 = \mathbf{e}'\mathbf{e}/(n-1)$.

To account for the fact that the stocks have very different numbers of news items (see Table 1), we alternatively used group-specific means. Let n_s denote the number of news for stock s and let X_{sj} be the reaction of a certain (trading) variable of stock s to item j. For the average reaction of each of the n_n stocks (the group mean), $\overline{X}_s = 1/n_s \sum_{j=1}^{n_s} X_{sj}$, we assume

$$\overline{X}_s = \mu + \epsilon_s, \qquad \epsilon_s \sim i.i.d.\ N(0, \sigma^2), \quad s = 1, .., n_n. \tag{10}$$

Then, inference is based on the estimator for the mean, $\overline{\overline{X}} = 1/n_n \sum_{s=1}^{n_n} \overline{X}_s$, where 95% confidence intervals are given as $\overline{\overline{X}} \pm 2 * \widehat{\sigma}/\sqrt{n_n}$ with $\widehat{\sigma}^2 = \mathbf{e}'\mathbf{e}/(n_n - 1)$.

Both approaches have their advantages. While the latter smoothes out the effect of a large number of news, it does not account for the within-group variation, which is captured by (9). Hence, confidence intervals are slightly more conservative using (10). Nevertheless, all results of Section 3 hold using both procedures. Plots of the group-means are available upon request from the authors.

Figures

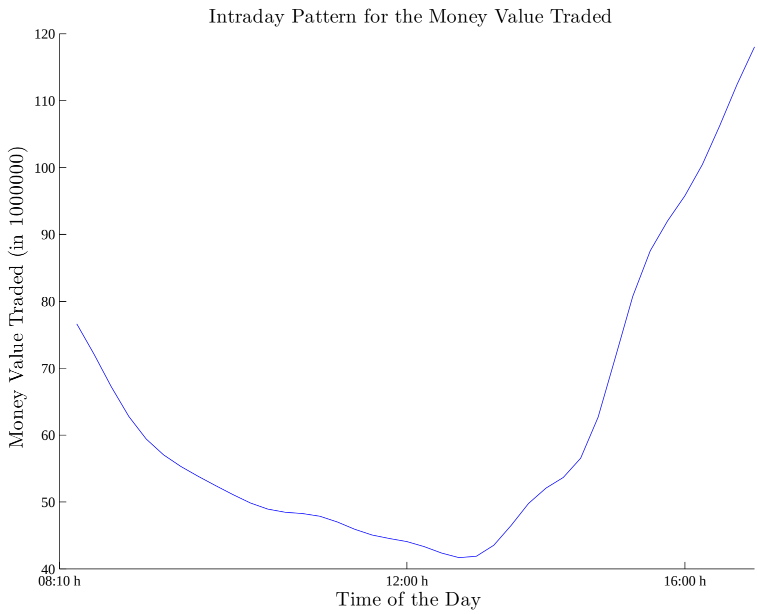

Figure 24: Intraday Pattern of Money Value Traded. Smoothed via kernel regression.

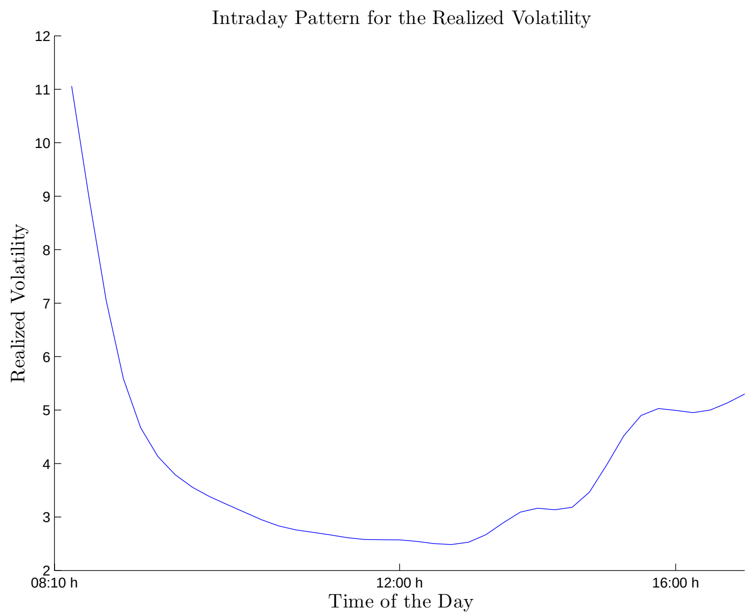

Figure 25: Intraday Pattern of Volatility. Smoothed via kernel regression.

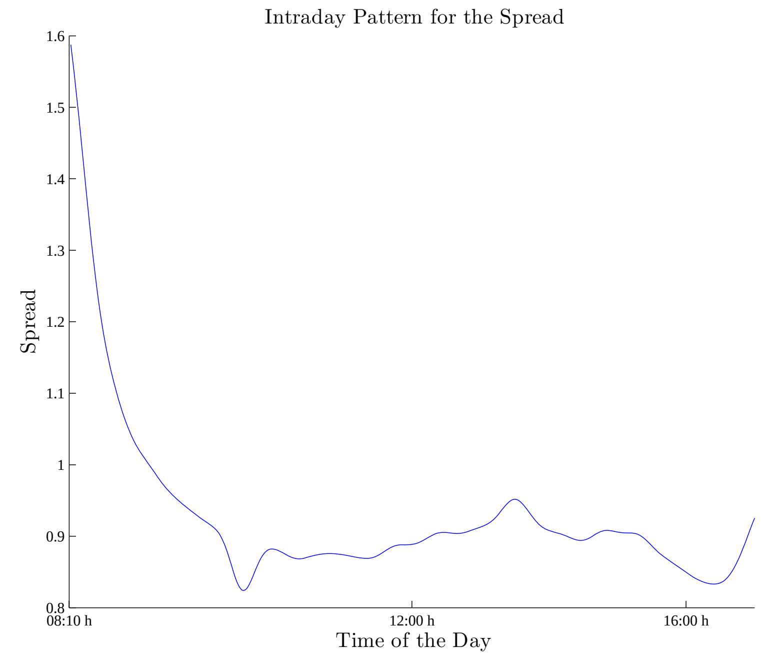

Figure 26: Intraday Pattern of Bid-Ask Spreads. Smoothed via kernel regression.

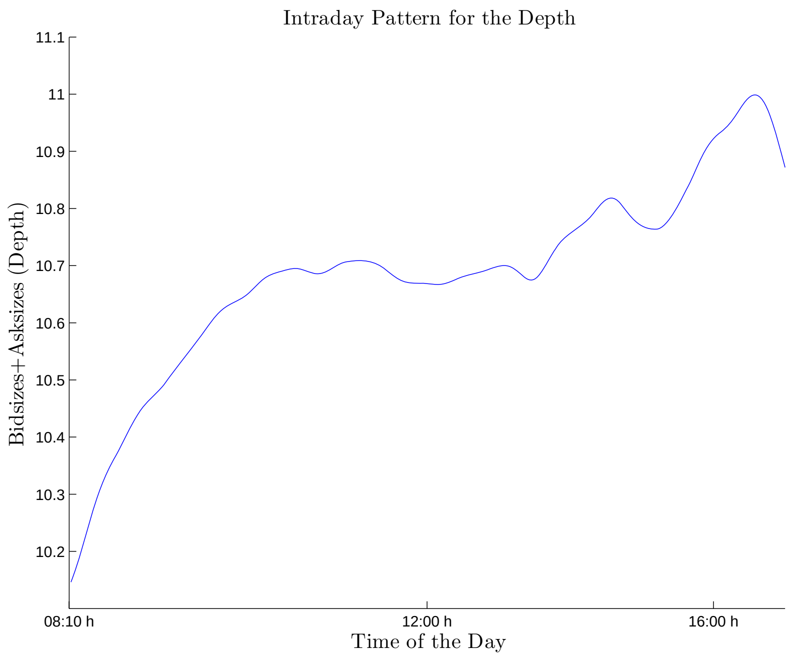

Figure 27: Intraday Pattern of Market Depth. Smoothed via kernel regression.

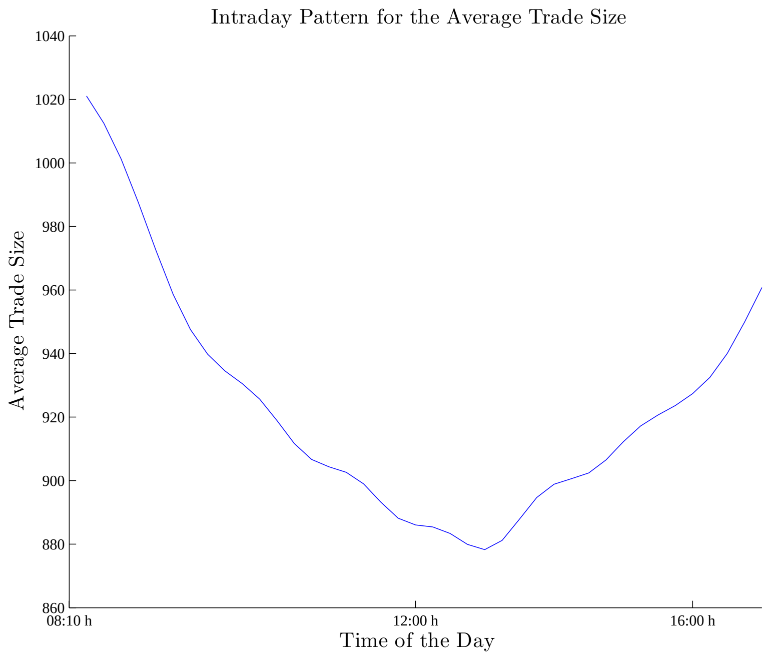

Figure 28: Intraday Pattern of Average Trade Sizes. Smoothed via kernel regression.

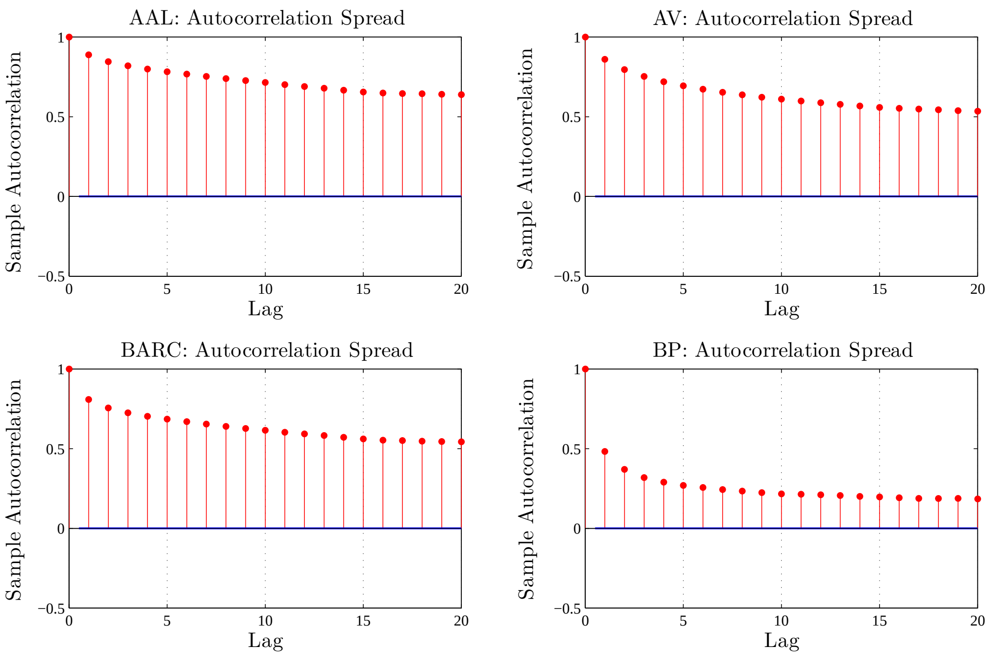

Figure 29: Autocorrelation Pattern of Spreads.

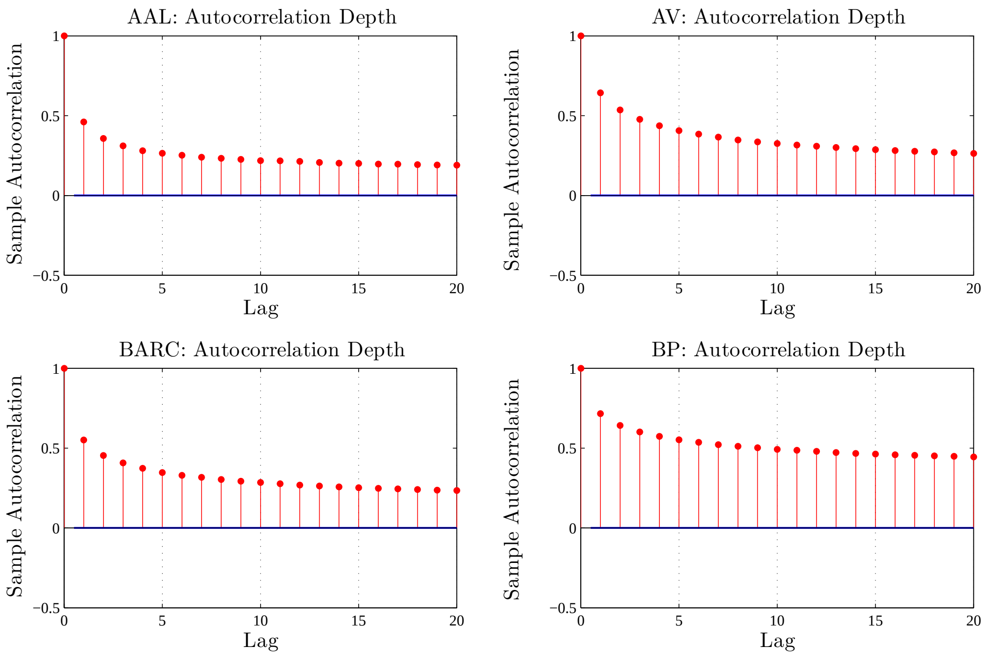

Figure 30: Autocorrelation Pattern of Depth.

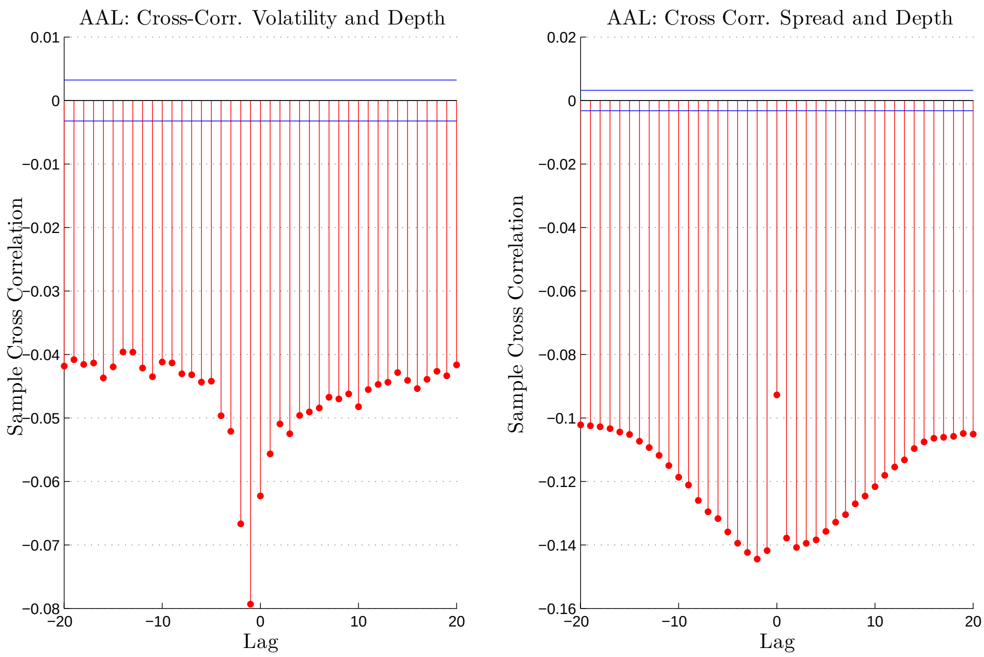

Figure 31: Cross-Correlations for the AAL stock

CFS Working Paper Series:

No.	Author(s)	Title
2009/30	**Volker Wieland**	Quantitative Easing: A Rationale and Some Evidence from Japan
2009/29	**Dimitris Georgarakos** **Giacomo Pasini**	Trust, Sociability and Stock Market Participation
2009/28	**Dimitris Christelis** **Dimitris Georgarakos**	Investing at Home and Abroad: Different Costs, Different People?
2009/27	**Erik Theissen**	Price Discovery in Spot and Futures Markets: A Reconsideration
2009/26	**Volker Wieland**	Fiscal Stimulus and the Promise of Future Spending Cuts: A Comment
2009/25	**Tobias Cwik** **Volker Wieland**	Keynesian government spending multipliers and spillovers in the euro area
2009/24	**Otmar Issing**	Politischer Wille oder ökonomisches Gesetz? - Einige Anmerkungen zu einem großen Thema -
2009/23	**Nikolaus Hautsch** **Ruihong Huang**	The Market Impact of a Limit Order
2009/22	**Christian Laux** **Christian Leuz**	Did Fair-Value Accounting Contribute to the Financial Crisis?
2009/21	**John B. Taylor** **Volker Wieland**	Surprising Comparative Properties of Monetary Models: Results from a New Data Base

Copies of working papers can be downloaded at http://www.ifk-cfs.de