Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Empirical Finance 11 (2004) 163–184
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/econbase

An empirical analysis of the role of the trading intensity in information dissemination on the NYSE

Laura Spierdijk

Department of Econometrics and CentER, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

Accepted 20 December 2002

Abstract

In this paper, we use high-frequency data on five frequently traded stocks listed on the New York Stock Exchange (NYSE) in the year 1999 to examine the price impact of trades and its relation to the trading intensity. We show that the distribution of the absolute price change with fast trading first-order stochastically dominates the distribution of the absolute price change with slow trading. Moreover, we find significant causality from the trade characteristics to the trading intensity. Large trades significantly increase the speed of trading, while large returns tend to decrease the trading intensity. We show that this feedback has little impact on the distribution of the price impact of trades.

JEL classification: C41; G14
Keywords: Price impact; Trading intensity; VAR-model; ACD-model

1. Introduction

An important component of market microstructure theories is the concept of asymmetric information. This phenomenon arises when both uninformed and informed traders are present at the market. Uninformed traders trade for liquidity reasons. Informed traders, however, have private information on the fundamental value of the security to be traded. They trade to take advantage of their superior knowledge. Due to the presence of informed traders, the transaction process itself potentially reveals information on the value of the security.

E-mail address: L.Spierdijk@TilburgUniversity.nl (L. Spierdijk).

0927-5398/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2002.12.004

Information dissemination through trading has been the subject of both theoretical and empirical research. Hasbrouck (1991a,b) uses a VAR-model to jointly model returns and trading volume. He shows that trades contain information, since they have persistent impact on prices. Recently, the information content of the trading intensity has been investigated. The trading intensity refers to the process of durations, where a duration is defined as the time that elapses between two consecutive transactions. The main question is whether the trading intensity conveys any information in addition to trading volume. According to Diamond and Verrecchia (1987), slow trading indicates bad news. In the model of Admati and Pfleiderer (1988), fast trading refers to an increased risk of informed trading. In the model of Easley and O'Hara (1992), slow trading is associated to the lack of news. In an empirical setting, Dufour and Engle (2000) model the trading intensity using the ACD-model proposed by Engle and Russell (1998). Dufour and Engle (2000) use a bivariate VAR-model for returns and trade sign to assess the effect of the trading intensity on the price adjustment process in both transaction and calendar time. The authors show that the price impact of a trade is larger the higher the trading intensity, implying that trades are more informative in periods of frequent trading.

This paper extends Hasbrouck (1991a,b) and Dufour and Engle (2000). Using a joint model for returns on the midprice, trade size, trading intensity, and volatility we investigate the price impact of large trades and its relation to the trading intensity for a sample of frequently traded stocks listed on the New York Stock Exchange (NYSE). We show the distribution of the price change with fast trading first-order stochastically dominates the distribution of the price change with slow trading. As in Engle and Lunde (1998), we establish significant causality from trade characteristics to the trading intensity. Large returns slow down trading, while large trades increase the speed of trading. However, we show that this feedback has little impact on the distribution of the price impact of trades, both in transaction and in calendar time.

The organization of this paper is as follows. In Section 2, we review some market microstructure underpinnings with the focus on the role of the trading intensity in information dissemination. Section 3 provides a description of the data and their sample properties. Section 4 is devoted to a multivariate model for returns and trading volume that ignores the possible role for the trading intensity. Section 5 discusses the modeling of the trading intensity, while Section 6 examines the impact of trades on prices in a VAR-model that takes the role of the trading intensity into account. In Section 7, we allow for feedback from the trade characteristics such as returns and trading volume to the trading intensity and we investigate the effects of taking into account this feedback on the price impact of trades. Finally, Section 8 summarizes the main results of this paper.

2. Trading intensity and information

In this section, we briefly review some market microstructure studies that establish a relation between the trading intensity and the underlying value of the asset.

In the model of Easley and O'Hara (1992), an information signal is released at the beginning of the day with a certain probability. The market maker is uncertain about the existence of an information signal. He does not know whether or not an information event has taken place and he does not know the direction of the possible news event (good or bad news). The market maker acts as a Bayesian and adjusts his prices by watching the order flow. Informed traders, who have knowledge on the signal that is possibly released at the beginning of the day, buy or sell their stock in case of good and bad news, respectively. Uninformed traders are allowed to refrain from trading.

When a news event has been released, a trade is more likely than a no trade outcome due to the presence of informed traders who want to trade to benefit from the private information they possess. Therefore, Easley and O'Hara (1992) associate fast trading to the existence of news and slow trading to the absence of news. Empirically, the model predicts that lagged durations are negatively correlated to the bid-ask spread. Since the market maker associates fast trading to an increased risk of informed trading, lagged durations will also be negatively correlated to price volatility.

Easley and O'Hara (1992) also conjecture a role for aggregated volume. This follows directly from the fact that each trade has unit size in the model. Therefore, aggregated volume equals the number of trades up to that moment. As a consequence, lagged aggregated volume is also positively related to the bid-ask spread and the volatility of prices. However, the assumption of a market with only unit size trades is unrealistic. It is therefore useful to consider the Easley and O'Hara (1987) model. The setting of the latter model is basically the same as in Easley and O'Hara (1992). However, although there is event uncertainty, uninformed traders are not allowed to refrain from trading. Therefore, durations do not play a role in this model. However, traders are allowed to trade either a small or a large quantity. When news has been released at the beginning of a trading day, it is more likely that a large quantity will be traded. Therefore, the bid-ask spread and price volatility are positively related to trading volume. It is straightforward to combine the Easley and O'Hara (1987, 1992) models, which yields a model in which durations and different trade sizes play a role. In the combined model, the absence of a trade is more likely when no news has been released and a large trade is more likely in case of a high information signal.

In a different framework, Diamond and Verrecchia (1987) also relate the trading intensity to the presence of news. Traders are either informed or uninformed and, moreover, own or do not own the stock. If they do not own the stock, they might wish to short-sell when there is an opportunity to trade. All trades fall into three groups: those who face no costs in short selling, those who are prohibited from short selling and finally, those who are restricted in short selling. In the latter case the proceeds from short-selling are delayed until the price of the asset falls. Neither the market maker, nor the traders can observe why there has been no trade and whether a sell is a short sell or not but every agent knows all relevant probabilities. When they observe a no trade outcome, they know that there are several possibilities. Either a trader did not want to trade, or he could not trade due to short-sell restrictions or prohibitions. The probability of a no trade outcome is higher in case of bad news, because of the informed traders who are constrained from selling short. Therefore, Diamond and Verrecchia (1987) associate slow trading to bad news. The empirical implications of this model are as follows. Lagged durations are positively correlated to the bid-ask spread. Moreover, lagged durations and price volatility are also positively correlated. Finally, lagged durations and (mid) prices as well as bid/ask quotes are negatively correlated. Admati and Pfleiderer (1988) distinguish informed and liquidity traders. Liquidity traders are either nondiscretionary traders who must trade a certain number of shares at a particular time or discretionary traders who time their trades such that the expected cost of their transactions are minimized. We consider the version of the model with endogenous information acquisition; i.e. private information is acquired at some cost and traders obtain this information if and only if their expected profit exceeds this cost. In this framework, the presence of informed traders lowers the cost of trading for liquidity traders. Moreover, informed traders prefer to trade when there are many liquidity traders at the market. Hence, both informed and uninformed traders want to trade when the market is 'thick'. This results in concentrated patterns of trading: informed traders and liquidity traders tend to clump together. Hence, according to Admati and Pfleiderer (1988) frequent trading is associated to news. This implies that prices are more informative in periods of frequent trading; i.e. the trading intensity positively affects volatility.

Table 1 summarizes the empirical implications of Easley and O'Hara (1987, 1992), Diamond and Verrecchia (1987), and Admati and Pfleiderer (1988). One of the crucial assumptions underlying the Easley and O'Hara (1987, 1992) model and Diamond and Verrecchia (1987) is the absence of feedback from the trade characteristics such as returns, bid-ask spread and trading volume to the trading intensity. Goodhart and O'Hara (1997) put forward that trade characteristics convey information on the value of the asset. Therefore, traders may learn from it and adjust their speed of trading in reaction to this. As indicated in Dufour and Engle (2000), for example, a large change in the market maker's midprice may be a signal to the informed traders that their information, initially unknown to other market participants, has been revealed to the market maker assuming that no new signal has been released thereafter. This means that their information is no longer superior. Therefore, the incentive to trade disappears, which decreases the trading intensity. However, from an inventory perspective, large quote changes would attract opposite-side traders, thus increasing the trading intensity. Similar effects occur when informed traders observed large trades. An additional complexity arises, however, when uninformed traders show strategic behavior as well, see O'Hara (1995). They will increase the probability they attach to the risk of informed trading when they notice large absolute returns or large trading volume. Consequently, they will down their trading intensity. The overall effect on the trading intensity is therefore unclear when both informed and uninformed traders show strategic behavior. This issue will be investigated empirically in the sequel.

Table 1
Implications for the correlation sign

Variables	EoH87	EoH92	AP88	DV87
Duration (y_t), spread (s_{t+1})	?	–	–	+
Duration (y_t), volatility (σ_{t+1})	?	–	–	+
Duration (y_t), midprice (m_{t+1})	?	?	?	–
Duration (y_t), bid/ask quote ($q_{t^{a,b}+1}$)	?	?	?	–
Volume ($\lvert x_t \rvert$), spread (s_t)	+	?	?	?

Summary of the implications of market microstructure models for the sign of the correlation between several trade-related variables. The studies are Easley and O'Hara (1987, 1992), Admati and Pfleiderer (1988), and Diamond and Verrecchia (1987), which are abbreviated by EoH87, EoH92, AP88, and DV87, respectively. A question mark indicates that the model does say anything on the sign of the correlation.

3. The data

We use high-frequency data on five of the most actively traded stocks listed on the NYSE, see Table 2. The data are taken from the *Trade and Quote* (TAQ) database. For each stock, the data consist of all transactions during the months August, September and October 1999, and covers 64 trading days.

We remove all trades that take place outside the opening hours, i.e. before 9:30 a.m. and after 4:00 p.m. Moreover, we also delete trades that take place before the first quotes are generated. For each stock, the associated characteristics of each trade are recorded: trade moment τ_t in seconds after midnight, unsigned log trade size x_t and transaction price p_t, where t indexes subsequent transactions (i.e. t indexes 'transaction time'). All data are measured in transaction time. The duration (in 'calendar time') between subsequent trades is defined as $y_t = \tau_t - \tau_{t-1}$. Overnight durations are removed from the data set.

To each trade, we also associate a prevailing bid and ask quote, denoted by q_t^b and q_t^a. To obtain the prevailing quotes, we use the '5-s rule' by Lee and Ready (1991) which associates each trade to the quote posted at least 5 s before the trade, since quotes can be posted more quickly than trades are recorded. The 5-s rule solves the problem of potential mismatching. The prevailing midprice m_t is the average of the prevailing bid and ask quotes: i.e. $m_t = (q_t^b + q_t^a)/2$. The log return over the prevailing and subsequent midprice is expressed in basis points (bp) and denoted by $r_t = \log(m_{t+1}/m_t)$. Overnight returns are excluded from the sample.

Since the transaction data provided by the NYSE are not classified according to the nature of a trade (buy or sell), we use the Lee and Ready (1991) 'midquote rule' to classify a trade. With this rule, the prevailing midprice corresponding to a trade is

Table 2
Ticker symbols, company names, and some sample statistics

Ticker symbol	IBM	MAT	MCD	SLB	WMT
Company name	Int. Business Machines	Mattel	McDonald's	Schlumberger	WalMart Stores
# transactions	140,013	41,822	57,860	63,905	90,042
Durations (s)					
Mean	11	36	26	23	16
Median	7	19	15	14	10
5% quantile	2	2	2	2	2
95% quantile	30	127	86	161	52
Volume (shares)					
Mean	3055	4305	2512	2187	2957
Median	1000	1000	900	1000	1000
5% quantile	100	100	100	100	100
95% quantile	11,000	20,000	10,000	9000	11,400
Returns (bp)					
Mean	− 0.0231	− 0.1729	0.0099	− 0.0238	0.0273
Median	0.0000	0.0000	0.0000	0.0000	0.0000
5% quantile	− 5.2562	− 22.8050	− 7.7131	− 9.4384	− 7.0646
95% quantile	5.2016	16.2101	7.7610	9.3721	7.0796

used to decide whether a trade is a buy, a sell, or undecided. If the transaction price is lower (higher) than the midprice, it is viewed as a sell (buy). If the price is exactly at the midprice, its nature (buy or sell) remains undecided. To each trade, we associate a trade indicator x_t^0, which indicates the nature of the trade: 1 (buy), -1 (sell), or 0 (undecided). From the trade size and the trade indicator, we can construct signed log trading volume x_t. If a trade is unclassified, signed trading volume will be zero.

It sometimes occurs that multiple trades take place at the same second. We follow Engle and Russell (1998) and treat multiple transactions at the same time as one single transaction and aggregate their trade volume and average prices.

As a first exploration of our data, we compute sample mean and median of several trade characteristics for each stock, see Table 2. This table shows that IBM is the most frequently traded stock in the sample, with the average duration equal to 11 s. Mattel is the least frequently traded stock of the sample with an average duration of 36 s. Average unsigned trading volume varies from 2187 shares (Schlumberger) to 4305 shares (Mattel) and average returns are close to zero.

For the McDonald's stock, we compute Spearman's rank correlations between the durations and several trade characteristics to get a notion of the possible dependence. We establish significantly negative correlation between lagged unsigned log trading volume $|x_{t-1}|$ and durations y_t, as well as between durations y_t and returns $|r_t|$. Moreover, we find significantly positive correlation between lagged returns $|r_{t-1}|$ and durations y_t and between durations y_t and unsigned trade size $|x_t|$. Table 3 reports the exact value of the sample correlations and provides standard errors corresponding to the correlations. This table also displays, for comparison, the autocorrelations in each variable. The correlations reported in Table 3 can be caused by asymmetric information or inventory effects as described in Section 2, but they can equally well be due to other factors such as time of the day periodicities. In order to separate these effects, we will explicitly model the relation between these variables in the next sections.

Table 3
Rank correlations

	Estimate	Std. error
Cross-correlations		
Return ($\|r_{t-1}\|$), duration (y_t)	0.2317	0.0043
Volume ($\|x_{t-1}\|$), duration (y_t)	-0.0942	0.0042
Duration (y_t), return ($\|r_t\|$)	-0.0861	0.0042
Duration (y_t), volume ($\|x_t\|$)	0.0112	0.0042
Autocorrelations		
Return ($\|r_t\|$), return ($\|r_{t-1}\|$)	0.0395	0.0042
Volume ($\|x_t\|$), volume ($\|x_{t-1}\|$)	0.1672	0.0042
Duration (y_t), duration (y_{t-1})	0.0578	0.0042

Spearman's rank correlation (with corresponding standard errors) between durations and several trade characteristics for the McDonald's stock.

4. The price impact of trades in transaction time

In this section, we discuss a two-dimensional VAR-model to capture the relation between returns and trade size in transaction time. This model does not take into account the possible role of the trading intensity. The approach is based upon on Hasbrouck (1991a,b). We specify the VAR-model for z_t=$(r_t,x_t)'$, expressed in transaction time, as

$$A(L)z_t = c + v_t, \tag{1}$$

where $A(L)$ is an m-th order (2×2) matrix polynomial in the lag operator L of the form $I - A_0 - A_1L - \ldots - A_mL^m$. The (k,l)-th element of the matrix A_j is denoted by $a_{j,(k,l)}$. The matrix A_0 can be normalized in various forms, which do not affect the properties of the model. We choose the formulation of Hasbrouck (1991a,b), such that trade size contemporaneously influences returns.[1] In expression (1), the variables v_t=$(v_{t,1},v_{t,2})'$ are (2×1) vectors of mean-zero disturbances that are jointly and serially uncorrelated; i.e.

$$\mathbb{E}v_{t,i} = \mathbb{E}v_{t,i}v_{s,i} = 0 \ \ (t{\neq}s;\ \ i = 1,2); \quad \mathbb{E}v_{t,1}v_{s,2} = 0.$$

We will measure the price impact of trades by means of the cumulative impulse response function. Given a certain history up to time τ_t, the cumulative impulse response function at time τ_{t+k} corresponding to an unexpected buy of M shares at time τ_t is defined as

$$\mathbb{E}_{t-1}(r_t + \ldots + r_{t+k} \mid v_{t,2} = \log(M)) - \mathbb{E}_{t-1}(r_t + \ldots + r_{t-k}). \tag{2}$$

Hence, the cumulative impulse response function represents the expected price impact of an unexpected trade, relative to the expected price impact conditional on the history only. See, for instance, Koop et al. (1996). Kraus and Stoll (1972) and Hasbrouck (1991a,b) point out that the persistent price impact of an unexpected trade is naturally interpreted as the information content of the trade. The persistent impact is obtained for $k \rightarrow \infty$ in expression (2).

4.1. Estimation results

In line with Hasbrouck (1991a,b) and Dufour and Engle (2000), we truncate the VAR-model at $m = 5$. We estimate the model by means of OLS. We use the method proposed by White (1980) to obtain heteroskedasticity-consistent standard errors. We verify the correctness of the truncation lag by testing for autocorrelation in the OLS-residuals using the Ljung–Box test. This test is asymptotically equivalent to the standard LM-test for serial correlation in the residuals of a regression, but computationally less demanding. The test does not lead to any evidence that more lags should be included in the VAR-model. The estimation results are given in Table 4. They show that for all stocks, trade size has a positive immediate impact on returns. This empirically confirms

[1] With this normalization, A_0 has one nonzero element, namely $a_{0,(1,2)}$.

Table 4
Estimation results for the return equation in the VAR-model

Coefficient	Lag	IBM		MAT		MCD		SLB		WMT	
		Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Const.	j	−0.2626	0.0092	−0.8063	0.0524	−0.1390	0.0203	−0.2736	0.0172	−0.3284	0.0165
$a_{j(1,1)}$	1	−0.0360	0.0036	−0.0817	0.0075	−0.0581	0.0046	−0.0333	0.0052	−0.0476	0.0040
	2	0.0355	0.0038	−0.0289	0.0064	−0.0140	0.0048	0.0230	0.0053	0.0097	0.0051
	3	0.0246	0.0037	−0.0203	0.0078	0.0083	0.0048	0.0228	0.0048	0.0205	0.0044
	4	0.0242	0.0038	−0.0023	0.0070	0.0108	0.0046	0.0138	0.0047	0.0135	0.0042
	5	0.0151	0.0038	0.0141	0.0061	0.0036	0.0046	0.0132	0.0056	0.0101	0.0045
$a_{j(1,2)}$	0	0.2279	0.0015	0.5127	0.0088	0.3733	0.0036	0.4030	0.0034	0.3067	0.0029
	1	0.0855	0.0017	0.1706	0.0091	0.0927	0.0038	0.0934	0.0040	0.1027	0.0030
	2	0.0003	0.0017	0.0623	0.0088	0.0111	0.0039	0.0059	0.0039	0.0193	0.0032
	3	−0.0086	0.0017	0.0303	0.0093	−0.0023	0.0039	−0.0057	0.0038	−0.0032	0.0031
	4	−0.0107	0.0017	0.0110	0.0091	−0.0133	0.0039	−0.0071	0.0039	−0.0039	0.0030
	5	−0.0082	0.0017	−0.0019	0.0089	−0.0049	0.0038	−0.0127	0.0039	−0.0072	0.0030
R^2		0.2215		0.1377		0.2190		0.2808		0.1798	

The return equation of the VAR-model defined in Eq. (1) is estimated using OLS. The standard errors in the columns on the right-hand side are computed using White's (1980) heteroskedasticity-consistent covariance matrix.

the results of Easley and O'Hara (1987) and Hasbrouck (1991a,b). We test for Granger-causality from returns to trade size and from trade size to returns. We do this by testing the null hypothesis that the corresponding coefficients in the VAR-model are jointly zero. For example, to test whether or not trade size Granger-causes returns, we use a Wald test and test the null hypothesis H_0: $a_{j,(1,2)}=0$ for $j=0,\ldots,5$. This null hypothesis is rejected at a 5% level for all stocks. Similarly, returns significantly Granger-cause trade size for all stocks in the sample. This emphasizes the importance of taking into account the feedback among the trade characteristics.

4.2. The price impact of trades

To investigate the short and long run price impact of a large trade on the McDonald's stock, we assume that the market is in a state of and 'equilibrium'. We define this as a situation in which past returns and trade sizes are equal to their sample average. We consider a buy consisting of 10,000 shares. This amount of shares corresponds to the 95% sample quantile of unsigned trading volume in our data. We compute the impulse response function for an unexpected trade of 10,000 shares. The two conditional expectations in Eq. (2) are obtained by iterating the VAR-model in expression (1) k periods ahead. A parametric bootstrap from the asymptotic distribution of the OLS-estimates can be used to obtain confidence intervals for the impulse response function. After 20 transactions, the expected price impact equals 6.7 bp. The corresponding 95% confidence interval equals (6.5, 6.9) bp and is based upon a bootstrap with $N=10{,}000$ draws. Note that the price impact is linear in log trading volume, so the impulse response functions for other trading volumes are easily derived from the price-impact function corresponding to a trade of 10,000 shares.

To estimate the model in Eq. (1) by means of OLS, we do not need the distribution of the disturbances $(v_t)_t$. Similarly, the distribution of $(v_t)_t$ is not needed for the estimation of the expected price change of a trade, since the disturbances have mean zero and thus cancel out in expression (2). However, not only the expected price change of a trade is of interest. To assess the entire distribution of the price impact caused by a large trade, we have to make explicit assumptions on the distribution of the VAR-disturbances. Since returns are likely to exhibit volatility clustering, we assume that the disturbances $(v_{t,1})_t$ follow a GARCH-process; see Engle (1982) and Bollerslev (1986). In Manganelli (2002), it is shown that time-varying conditional heteroskedasticity is not only present in returns, but also in trading volume. An ARCH LM-test (see Engle (1982)) provides significant evidence for autoregressive conditional heteroskedasticity in the VAR-disturbances $(v_{t,1}, v_{t,2})'$, since the null hypothesis of no ARCH-effects in $(v_{t,1})_t$ and $(v_{t,2})_t$ is rejected at each reasonable significance level. We therefore specify a bivariate GARCH-model in transaction time for $(v_t)_t$; i.e. $v_t = \Sigma_t \eta_t$. Here, Σ_t denotes a diagonal matrix with elements $\sigma_{t,1}$ and $\sigma_{t,2}$. Moreover, $(\eta_t)_t = (\eta_{t,1}, \eta_{t,2})_t$ is a bivariate sequence of mean zero, identically distributed random variables, with $\eta_{t,1}$ independent of the information known up to time τ_t and $\eta_{t,2}$ independent of the information set at time τ_{t-1} and $\eta_{t,1}$ and $\eta_{s,2}$ independent for all s, t. A specification search leads to a bivariate EGARCH(1,1) specification:

$$\log\sigma^2_{t,1} = \alpha_{r,1} + \alpha_{r,2}\,|\eta_{t-1,1}| + \alpha_{r,3}\eta_{t-1,1} + \alpha_{r,4}\log\sigma^2_{t-1,1} + \alpha_{r,5}\,|x_t| + \alpha_{r,6}\,|x_{t-1}|; \tag{3}$$

$$\log\sigma^2_{t,2} = \alpha_{x,1} + \alpha_{x,2}\,|\eta_{t-1,2}| + \alpha_{x,3}\eta_{t-1,2} + \alpha_{x,4}\log\sigma^2_{t-1,2}. \tag{4}$$

Hence, we allow for GARCH-effects in both returns and trading volume and, moreover, for feedback from trading volume to volatility. Since volatility is affected by the magnitude of the trade rather than its sign (buy or sell), we include unsigned trading volume in Eq. (3). Since we do not find significant evidence for feedback from the trade characteristics to the variance of trading volume, we not include any explanatory variables in Eq. (4). Using the OLS-residuals, we estimate the EGARCH-model given by Eqs. (3) and (4) by means of quasi-maximum likelihood (QML). The BHHH-algorithm of Berndt et al. (1974) is used for the numerical optimization. Furthermore, we estimate the robust standard errors of Bollerslev and Wooldridge (1992) to deal with any deviations from normality in $(\eta_t)_t$.

The estimation results are displayed in Table 5. The results show that returns are highly persistent, since the coefficients $\alpha_{r,4}$ are close to one. There is less persistence in trading volume, but the values of the coefficients $\alpha_{x,4}$ are still relatively high and thus indicate that large trades tend to clump together. For three out of five stocks (IBM, Mattel, and WalMart), the coefficients $\alpha_{r,3}$ are significantly different from zero, which means that positive and negative shocks have asymmetric impact on volatility. For all five stocks, the coefficients $\alpha_{x,3}$ are significantly different from zero, which indicates that there are asymmetric effects for trading volume as well. Note that $(\alpha_{r,5} + \alpha_{r,6})\,(1 - \alpha_{r,4})^{-1}$ represents the 'equilibrium multiplier' and thus reflects the change in the equilibrium value of

Table 5
Estimation results for the EGARCH-model for the VAR-residuals

	IBM		MAT		MCD		SLB		WMT	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Returns										
$\alpha_{r,1}$	−0.1089	0.0057	0.0657	0.0179	0.0668	0.0149	0.0291	0.0098	−0.0226	0.0050
$\alpha_{r,2}$	0.1287	0.0056	0.1542	0.0092	0.0718	0.0068	0.0744	0.0095	0.0670	0.0046
$\alpha_{r,3}$	0.0128	0.0034	0.0169	0.0065	−0.0010	0.0037	0.0008	0.0036	−0.0103	0.0025
$\alpha_{r,4}$	0.9702	0.0036	0.9694	0.0039	0.9620	0.0056	0.9530	0.0064	0.9912	0.0016
$\alpha_{r,5}$	0.1408	0.0032	0.0401	0.0051	0.0730	0.0037	0.1414	0.0047	0.0795	0.0037
$\alpha_{r,6}$	−0.1273	0.0031	−0.0455	0.0052	−0.0733	0.0036	−0.1324	0.0051	−0.0792	0.0038
Volume										
$\alpha_{x,1}$	0.7433	0.0306	0.7631	0.0459	1.2954	0.0857	1.4362	0.0838	1.4370	0.0617
$\alpha_{x,2}$	0.1747	0.0043	0.3095	0.0093	0.1690	0.0073	0.1893	0.0073	0.2173	0.0062
$\alpha_{x,3}$	−0.0143	0.0021	0.0192	0.0050	−0.0158	0.0035	−0.0185	0.0041	−0.0172	0.0031
$\alpha_{x,4}$	0.7502	0.0090	0.7060	0.0141	0.5922	0.0250	0.5287	0.0254	0.5390	0.0180

This table reports the results for the EGARCH-model given in Eqs. (3) and (4). The EGARCH-model is estimated using QML. The standard errors in the columns on the right-hand side are computed using the Bollerslev and Wooldrigde's (1992) robust covariance matrix.

volatility caused by a ceteris paribus change in trading volume. A Wald test shows that unsigned trading volume is significantly positively related to volatility for two out of five stocks (IBM and Schlumberger). The positive impact of trading volume on volatility is in line with the model of Easley and O'Hara (1987). It is also in line with the empirical conclusions of Lamoureux and Lastrapes (1990) and Manganelli (2002). Hence, following a large trade, the market maker updates his beliefs which leads to a persistent price change. In addition to this, large trades also have a positive impact on volatility. This can be explained by the fact that large trades are associated to an increased risk of informed trading, see Easley and O'Hara (1987). However, for the stocks McDonald's and WalMart, the volume multiplier does not significantly differ from zero and, for Mattel, the multiplier is significantly negative. Hence, the significantly positive effect of volume on volatility is restricted to two stocks only.

We estimate the distribution of the price impact using the bootstrap approach of Hasbrouck (1991b). This means that we consider an unexpected buy of M shares and simulate values of $(r_{t+k}, x_{t+k})'$ by drawing from the empirical distribution of the standardized VAR-disturbances $v_{t,i}/\sigma_{t,i}$ which are assumed iid. We do this $N=10{,}000$ times and, for each simulated sequence of $(r_{t+k}, x_{t+k})'$, we compute the corresponding price changes at time τ_{t+k}. We find that the 5% quantile of the price impact after 20 trades equals −0.6 bp. The 95% quantile of the price change after 20 trades is equal to 14.2 bp. Thus, with a probability of 90% the price change of a trade of 10,000 shares is in the interval (−0.6, 14.2) bp. Fig. 1 shows the expected price change corresponding to the unexpected trade of 10,000 shares, including the 5% and 95% quantiles of the distribution of the price change. The remaining quantiles of the distribution of the persistent price impact of a trade of 10,000 shares are reported in the column with the caption 'no durations' in Table 6. Up to now, we only considered impulse response

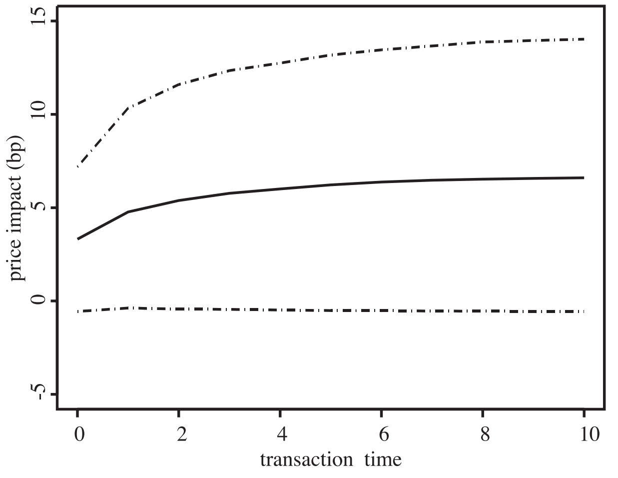

Fig. 1. Impulse response function and 90% prediction interval. This plot shows the expected price impact (solid line) and the 5% and 95% quantiles (dashed lines) of the distribution of the price impact corresponding to an unexpected trade of 10,000 shares of McDonald's stock, based on the VAR-model defined in Eq. (1).

functions in transaction time. From Fig. 1, we can see that it takes about 10 transactions before the new efficient price has been reached. Since the average duration for the McDonald's stock is 26 s, it takes slightly less than 4.5 min before the new efficient price has been attained.

Table 6
The distribution of the price impact of a trade of 10,000 shares of the McDonald's stock

Quantile (%)	No durations	No feedback	No feedback	Feedback	Feedback
		Fast trading	Slow trading	Fast trading	Slow trading
2.5	− 1.3	− 1.8	− 1.1	− 1.2	− 0.9
5	− 0.6	− 0.8	− 0.5	− 0.3	− 0.2
10	0.3	0.4	0.1	0.8	0.4
20	1.7	2.1	1.3	2.4	1.7
30	2.8	3.8	2.4	3.6	2.5
40	4.1	5.1	3.3	4.7	3.5
50	5.8	7.0	4.6	6.5	5.0
60	7.8	9.2	6.2	9.2	6.6
70	9.6	11.6	7.4	11.5	7.8
80	11.2	13.3	8.8	12.9	9.1
90	12.9	15.1	10.1	15.0	10.3
95	14.2	16.3	10.8	16.2	11.2
97.5	15.2	17.1	11.5	17.1	11.9

This table reports the estimated quantiles (in bp) of the distribution of the persistent price impact of trades implied by an unexpected trade of 10,000 shares of the McDonald's stock in several models: in the VAR-model defined in Eq. (1) without a role for the trading intensity, in the extended VAR-model with a role for the trading intensity, and in the extended VAR-model with feedback from the trade characteristics to the trading intensity.

5. A model for the trading intensity

In the VAR-model of Hasbrouck (1991a,b), the price impact of trades can only be measured in transaction time. It is often useful to have impulse responses in calendar time, since this allows, e.g., for the computation of the exact time it takes to reach a certain price level. In this section, we focus on the specification of the data generating process underlying the trading intensity. We use a version of Engle and Russell's (1998) ACD-model for this purpose, assuming that the duration process is strongly exogenous, cf. Engle et al. (1983), and that $(y_t)_t$ is generated by a log ACD (1,1)-model, cf. Bauwens and Giot (2000), i.e.

$$y_t = \psi_t \varepsilon_t, \quad \psi_t = \mathbb{E}_{t-1}(y_t), \tag{5}$$

with $(\varepsilon_t)_t$ a sequence of identically distributed variables with unit mean, independent of the information up to time τ_{t-1} and of $\upsilon_{i,s}$ for $i=1$, 2 and all s. The log conditional duration is specified recursively as

$$\log\psi_t = \beta_1 + \beta_2 \log\,\varepsilon_{t-1} + \beta_3 \log\psi_{t-1}. \tag{6}$$

The model is expressed in terms of diurnally corrected durations, which are also denoted by y_t as well, with some abuse of notation. The diurnally corrected durations are obtained as in Engle and Russell (1998). The expected duration given the time of the day is approximated by a piecewise linear and continuous spline with nodes set on 9:30–10:00, 10:00–11:00, ..., 14:00–15:00, and 15:30–16:00 h. We compute the diurnally corrected durations by dividing each duration by the diurnal correction factor.

5.1. Estimation results

We first estimate the diurnal component separately by means of a regression, cf. Engle and Russell (1998). Subsequently, we estimate the ACD(1,1)-model by means of QML, see Engle and Russell (1998) and Drost and Werker (2001). We use the BHHH-algorithm of Berndt et al. (1974) for the numerical optimization. We compute the Bollerslev and Wooldridge (1992) robust covariance matrix to obtain standard errors that are robust against deviations in exponentiality of ε_t. The row with the caption 'no feedback' in Table 7 shows the QML-estimation results of the ACD(1,1) model for each stock. As usual, the persistence parameter β_3 is close to one. It varies from 0.988 to 0.999. The estimation results for the diurnal correction factor are available upon request.

6. The price impact of trades and calendar-time effects

It is likely that the price impact of trades depends upon the trading intensity, cf. Diamond and Verrecchia (1987), Admati and Pfleiderer (1988), and Easley and O'Hara (1992). In this section, we proceed in the line of Dufour and Engle (2000).

Table 7
QML-estimation results for the ACD(1,1)-model

	IBM		MAT		MCD		SLD		WMT	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
No feedback										
β_1	– 0.0375	0.0014	– 0.0457	0.0015	– 0.0364	0.0016	– 0.0468	0.0016	– 0.0233	0.0010
β_2	0.0605	0.0022	0.0783	0.0026	0.0624	0.0027	0.0789	0.0027	0.0393	0.0017
β_3	0.9933	0.0006	0.9994	0.0003	0.9879	0.0011	0.9926	0.0007	0.9963	0.0004
With feedback										
β_1	– 0.0396	0.0017	– 0.0461	0.0021	– 0.0351	0.0028	– 0.0527	0.0023	– 0.0279	0.0016
β_2	0.0760	0.0026	0.0887	0.0030	0.0979	0.0036	0.1068	0.0033	0.0648	0.0024
β_3	0.9904	0.0008	0.9989	0.0004	0.9708	0.0018	0.9849	0.0010	0.9898	0.0008
$\xi_{\|r\|,1}$	0.1406	0.0021	0.0746	0.0024	0.1006	0.0026	0.1596	0.0028	0.1203	0.0020
$\xi_{\|r\|,2}$	– 0.1398	0.0021	– 0.0738	0.0024	– 0.0997	0.0025	– 0.1588	0.0028	– 0.1210	0.0020
$\xi_{\|x\|,1}$	– 0.1105	0.0058	– 0.1196	0.0084	– 0.1934	0.0082	– 0.2228	0.0062	– 0.1457	0.0062
$\xi_{\|x\|,2}$	0.1043	0.0059	0.1161	0.0084	0.1780	0.0081	0.2186	0.0063	0.1391	0.0062
ξ_Q	– 0.0002	0.0000	– 0.0002	0.0001	– 0.0005	0.0001	– 0.0005	0.0001	– 0.0003	0.0000

Estimation results for the ACD(1,1)-model with and without feedback as specified in Eqs. (12) and (13). The standard errors in this table are computed from the Bollerslev and Wooldridge (1992) robust covariance matrix.

As in Section 4, we specify a VAR-model in transaction time for the vector $z_t=(r_t, x_t)'$, but now $A(L)$ is allowed to depend upon the trading intensity; i.e.

$$A(L) = A(y_t)(L). \tag{7}$$

The impact of past trading volumes on returns and current trade size depends upon the trading intensity in the following way:

$$a_{j,(k,2)} = \gamma_{(j,k)} + \delta_{(j,k)} \cdot \log y_{t-j}, \tag{8}$$

similar to Dufour and Engle (2000). With this specification, the impact of a trade on returns depends upon the trading intensity. For example, when the coefficient $\delta_{(j,1)}$ is negative (positive), the impact of a trade on returns is lower (higher) when the corresponding duration is long (short). Moreover, with this specification, the correlation between consecutive trading volumes depends on the durations in a similar way.

6.1. The price impact of trades

As before, we estimate the VAR-model using OLS with truncation at $m=5$. The estimation results for the McDonald's stock are given in Table 8. Only the results for the return equation are displayed. Similar to the model of Hasbrouck (1991a,b) without durations, we test for Granger-causality. Again, we establish significant Granger-causality from returns to trade size and vice versa. Moreover, the null hypothesis that the impact of trades does not depend upon the trading intensity is rejected for all stocks.

As in the VAR-model without a role for the trading intensity, we estimate the impulse response functions to measure the price impact. Since the durations enter the model in a nonlinear fashion, we have to average out the durations. Therefore, we estimate the impulse

Table 8
Estimation results for the return equation in the VAR-model

		IBM		MAT		MCD		SLB		WMT	
Coefficient	Lag	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Const.	j	− 0.2617	0.0092	− 0.8232	0.0525	− 0.1428	0.0203	− 0.2742	0.0172	− 0.3305	0.0164
$a_{j,(1,1)}$	1	− 0.0308	0.0035	− 0.0821	0.0076	− 0.0545	0.0046	− 0.0257	0.0051	− 0.0453	0.0040
	2	0.0350	0.0038	− 0.0300	0.0064	− 0.0132	0.0048	0.0234	0.0053	0.0100	0.0050
	3	0.0238	0.0037	− 0.0213	0.0078	0.0083	0.0048	0.0216	0.0048	0.0203	0.0044
	4	0.0244	0.0038	− 0.0029	0.0070	0.0101	0.0046	0.0119	0.0047	0.0133	0.0042
	5	0.0154	0.0038	0.0128	0.0061	0.0032	0.0046	0.0113	0.0056	0.0103	0.0045
$\gamma_{(j,1)}$	0	0.3082	0.0051	0.7409	0.0237	0.5048	0.0102	0.5339	0.0093	0.3860	0.0080
	1	0.0681	0.0048	0.2244	0.0234	0.1022	0.0097	0.1161	0.0092	0.0958	0.0079
	2	− 0.0065	0.0045	0.1119	0.0224	0.0169	0.0097	0.0133	0.0092	0.0186	0.0078
	3	− 0.0059	0.0047	0.0546	0.0228	− 0.0087	0.0097	− 0.0093	0.0091	0.0036	0.0077
	4	− 0.0122	0.0046	− 0.0135	0.0222	− 0.0224	0.0097	− 0.0189	0.0089	0.0032	0.0078
	5	− 0.0128	0.0047	0.0161	0.0221	− 0.0081	0.0094	− 0.0081	0.0089	− 0.0053	0.0076
$\delta_{(j,1)}$	0	− 0.0355	0.0021	− 0.0757	0.0070	− 0.0463	0.0033	− 0.0468	0.0031	− 0.0312	0.0030
	1	0.0056	0.0020	− 0.0227	0.0068	− 0.0068	0.0031	− 0.0131	0.0029	0.0007	0.0029
	2	0.0029	0.0019	− 0.0185	0.0065	− 0.0031	0.0031	− 0.0043	0.0029	0.0000	0.0029
	3	− 0.0010	0.0019	− 0.0098	0.0065	0.0017	0.0030	0.0005	0.0029	− 0.0028	0.0028
	4	0.0006	0.0019	0.0066	0.0064	0.0032	0.0030	0.0042	0.0028	− 0.0031	0.0028
	5	0.0020	0.0020	− 0.0061	0.0064	0.0012	0.0030	− 0.0013	0.0028	− 0.0011	0.0028
R^2		0.2237		0.1411		0.2212		0.2842		0.1810	

This table reports the results for the VAR-model defined in Eq. (1) with duration dependence. The return equation is estimated using OLS. The standard errors in the columns on the right-hand-side are computed using White's (1980) heteroskedasticity-consistent covariance matrix.

response function by simulating $N = 10{,}000$ future paths of durations. For each path of durations, we compute price-impact functions as before and, finally, we average the impulse responses over the $N = 10{,}000$ simulations to obtain the final impulse response function.

To simulate future paths of durations, we need random values of the ACD-disturbances $(\varepsilon_t)_t$. Since we used QML to estimate the coefficients of the ACD-model, we did not make any additional distributional assumptions apart from some regularity conditions. We therefore assume that the ACD-disturbances are independent and identically distributed according to the corresponding empirical law. Hence, to obtain random values of the ACD-disturbances, we randomly draw from the empirical distribution of the ACD-residuals.

Again, we focus on the McDonald's stock. We compute impulse response functions for the model of Section 6 in two different situations: in a situation of 'low' and 'high' trading intensity. We compute the 99.5% and the 0.5% quantiles of the durations in our data. Subsequently, we initialize the ACD-model with these durations. As we compute the impulse response functions by simulating future paths of durations, we also need to compute the diurnal correction factor. Therefore, it is necessary that we specify explicitly the time at which the large trade takes place. Consistent with the daily periodicities observed in the trading intensity, we assume that the period of slow trading takes place at 12:30 p.m. and the fast trading at 10:00 a.m. By doing so, we capture the effect of different trading intensities on the impulse response functions. As in Section 4, we assume that the trade characteristics are in a state of equilibrium at the time of the unexpected trade.

Fig. 2 shows the impulse response functions for a trade of size 10,000 with 'slow' and 'fast' trading, as well as the impulse response function in the VAR-model in which the trading intensity does not play a role. We see that, 20 transactions after the trade of 10,000 shares, the impulse response equals 5.1 bp with slow trading and 7.8 bp with fast trading.

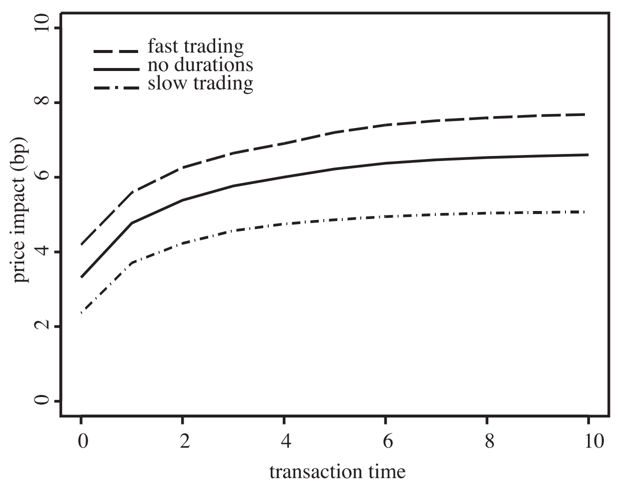

Fig. 2. Impulse response function: slow versus fast trading. This plot shows the price-impact functions corresponding to an unexpected trade of 10,000 shares of the McDonald's stock, based on the VAR-model defined in Eq. (1) without duration dependence and with fast and slow trading.

Note that the expected price impact of 6.7 bp as computed by the model of Hasbrouck (1991a,b) lies between these two values. The corresponding 95% confidence intervals equal (4.5, 5.6) and (7.4, 8.1) bp. A 95% upper one-sided confidence interval for the difference between the price impact with slow and fast trading is $(-\infty, -3.4)$ bp, so the price impact with slow trading is significantly lower than with fast trading.

To derive the entire distribution of the price impact, we proceed as in Section 4 and specify an EGARCH(1,1)-model for the VAR-disturbances $(v_{t,1}, v_{t,2})'$. Taking into account the durations in our specification search, we arrive at the specification

$$\log\sigma_{t,1}^2 = \alpha_{r,1} + \alpha_{r,2}|\eta_{t-1,1}| + \alpha_{r,3}\eta_{t-1,1} + \alpha_{r,4}\log\sigma_{t-1,1}^2 + \alpha_{r,5}|x_t| + \alpha_{r,6}|x_{t-1}| + \alpha_{r,7}\log y_t + \alpha_{r,8}\log y_{t-1}; \tag{9}$$

$$\log\sigma_{t,2}^2 = \alpha_{x,1} + \alpha_{x,2}|\eta_{t-1,2}| + \alpha_{x,3}\eta_{t-1,2} + \alpha_{x,4}\log\sigma_{t-1,2}^2 \tag{10}$$

We now also include feedback from the trading intensity to volatility, which is motivated by Easley and O'Hara (1992).

As before, we estimate the EGARCH-model by means QML, applying the Bollerslev and Wooldridge (1992) robust standard errors. The estimation results are displayed in Table 9. The equilibrium multiplier for trading volume indicates that unsigned trading volume has a positive impact on volatility for two out of five stocks (IBM and Schlumberger). Furthermore, the trading intensity is significantly positively related to volatility for all five stocks, since the equilibrium multiplier corresponding to the durations is negative in all five cases. The positive impact of the trading intensity on volatility was

Table 9
Estimation results for the EGARCH-model for the VAR-residuals

	IBM		MAT		MCD		SLB		WMT	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Returns										
$\alpha_{r,1}$	−0.0670	0.0078	0.3294	0.0348	0.2239	0.0308	0.0918	0.0169	0.0255	0.0089
$\alpha_{r,2}$	0.1352	0.0059	0.1918	0.0105	0.0937	0.0081	0.0799	0.0099	0.0806	0.0050
$\alpha_{r,3}$	0.0139	0.0034	0.0354	0.0072	0.0009	0.0044	−0.0015	0.0038	−0.0107	0.0029
$\alpha_{r,4}$	0.9670	0.0039	0.9385	0.0058	0.9341	0.0081	0.9471	0.0068	0.9843	0.0021
$\alpha_{r,5}$	0.1430	0.0033	0.0432	0.0047	0.0741	0.0038	0.1401	0.0047	0.0793	0.0038
$\alpha_{r,6}$	−0.1291	0.0033	−0.0505	0.0047	−0.0750	0.0037	−0.1320	0.0050	−0.0791	0.0039
$\alpha_{r,7}$	−0.1695	0.0113	−0.1446	0.0133	−0.1376	0.0095	−0.1110	0.0096	−0.0921	0.0110
$\alpha_{r,8}$	0.1498	0.0110	0.0962	0.0129	0.1065	0.0102	0.0935	0.0096	0.0771	0.0111
Volume										
$\alpha_{x,1}$	0.7169	0.0303	0.7531	0.0455	1.2954	0.0857	1.4206	0.0835	1.4116	0.0618
$\alpha_{x,2}$	0.1683	0.0043	0.3115	0.0094	0.1690	0.0073	0.1880	0.0073	0.2125	0.0061
$\alpha_{x,3}$	−0.0138	0.0021	0.0184	0.0050	−0.0158	0.0035	−0.0184	0.0041	−0.0170	0.0031
$\alpha_{x,4}$	0.7591	0.0089	0.7082	0.0140	0.5922	0.0250	0.5336	0.0253	0.5473	0.0180

This table reports the results for the EGARCH-model given in Eqs. (3) and (4). The EGARCH-model is estimated using QML. The standard errors in the columns on the right-hand side are computed using the Bollerslev and Wooldrigde (1992) robust covariance matrix.

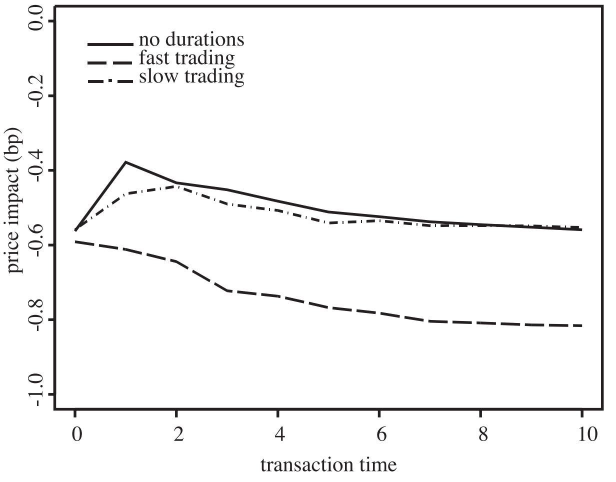

Fig. 3. Impulse response function: slow versus fast trading. This plot shows the 5% quantile of the distribution of the price change caused by a trade of 10,000 shares of the McDonald's stock, based on the VAR-model in Eq. (1) with and without a role for the trading intensity.

also found by Manganelli (2002) and implies that, in periods of frequent trading, volatility is higher. This can be explained within the model Easley and O'Hara (1992), in which fast trading is associated to an increased risk of informed trading.

Using Hasbrouck's (1991b) bootstrap approach as in Section 4, we find that the 5% quantiles of the expected price change after 20 transactions with slow and fast trading are -0.5 and -0.8 bp, respectively. The 95% quantiles of the price change are 10.8 and 16.3 bp. Thus, with 90% probability, the price change with slow trading is in the interval $(-0.5, 10.8)$ bp. With fast trading, the price change is with 90% probability in the interval $(-0.8, 16.3)$ bp. Hence, the entire distribution of the price change is different in periods of fast and slow trading and thus depends upon the trading intensity. The distribution of the absolute price change with fast trading first-order stochastically dominates the distribution of the absolute price change with slow trading. This means that trades have more impact on prices in periods of frequent trading, hence trades convey more information when durations are short. Figs. 3 and 4 show the 5% and the 95% quantiles of the distribution of the price change without durations and in periods of fast and slow trading. The remaining quantiles of the distribution of the persistent price impact of a trade of 10,000 shares are reported in the columns with the caption 'no feedback' in Table 6.

Finally, to gain insight into the adjustment process of the price following a large trade, we now consider the expected price-impact function in calendar time. The impulse response functions in calendar time[2] show that it takes approximately 3.5 min to reach the new efficient price that follows the unexpected trade in case of frequent trading[3] while

[2] Since we simulate paths of durations for the computation of the impulse response function, we can sample over each 5 s. We then obtain the impulse response function in calendar time.

[3] We measure the time it takes to reach 99.5% of the long-run impulse response.

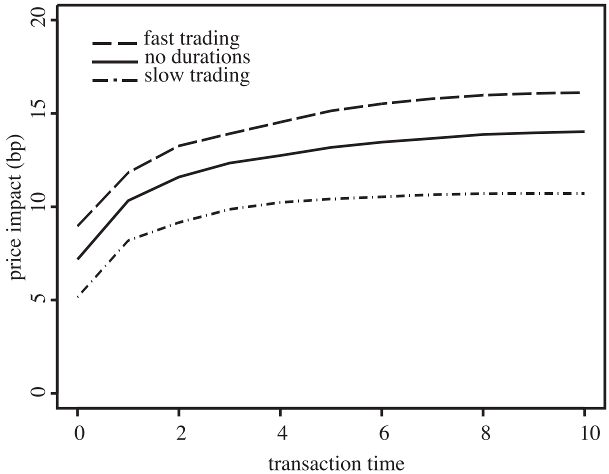

Fig. 4. Impulse response function: slow versus fast trading. This plot shows the 95% quantile of the distribution of the price change caused by a trade of 10,000 shares of the McDonald's stock, based on the VAR-model in Eq. (1) with and without a role for the trading intensity.

this takes about 10 min in case of slow trading. See Fig. 5. In the VAR-model without durations, we had estimated the time to reach the new efficient price to be approximately 4.5 min, which is in between the convergence time for fast and slow trading.

For the other stocks under consideration, we obtain similar results.

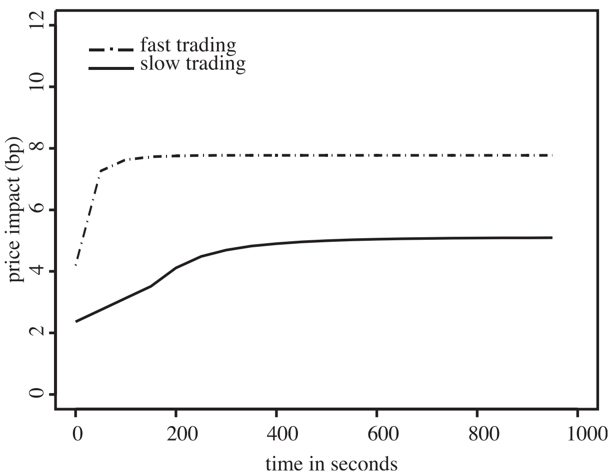

Fig. 5. Impulse response function: convergence time. This plot shows the price-impact function in calendar time following an unexpected trade of 10,000 shares of the McDonald's stock. The impulse response functions are based on the VAR-model in Eq. (1) with duration dependence, in periods of fast and slow trading. The horizontal axis displays the time in seconds starting at the time at which the trade has been initiated.

7. Feedback from trade characteristics to the trading intensity

In Section 6, we measured the impact of a transitory shock on prices, assuming that there is no feedback from the trade characteristics to the trading intensity. In Section 2, we made clear that trade characteristics are likely to have impact on the trading intensity. This additional feedback may affect the impulse response functions. In this section we investigate whether or not the trade characteristics affect the trading intensity and to what extent the impulse response functions are influenced by this feedback.

We specify the log ACD(1,1)-model with feedback as follows. Let again

$$y_t = \psi_t \varepsilon_t, \quad \psi_t = \mathbb{E}_{t-1}(y_t), \tag{11}$$

with ε_t iid with unit mean, independent of the information up to time τ_{t-1} and of $\upsilon_{i,s}$ for $i=1$, 2 and all s. The information known up to time τ_{t-1} now also includes the values of the trade characteristics up to that moment. The log conditional expectation is extended with a vector of trade characteristics:

$$\log\psi_t = \beta_1 + \beta_2 \log\varepsilon_{t-1} + \beta_3 \log\psi_{t-1} + \xi' v_{t-1}. \tag{12}$$

We include several variables in υ_t that may, according to Section 2, affect the trading intensity. We take

$$v_{t-1} = (|r_{t-1}|, |r_{t-2}|, |x_{t-1}|, |x_{t-2}|, Q_{t-1})', \quad Q_{t-1} = \left|\sum_{i=1}^{5} x_{t-i}\right|. \tag{13}$$

The variable Q_{t-1} represents the imbalance in signed volume over the five most recent transactions. This variable is included to investigate the effect of order imbalances on the trading intensity. The effect of trade size and absolute returns on durations is probably more related to the magnitude of these variables than their sign, so we take these variables unsigned. The reasons for including absolute returns have been pointed out by Dufour and Engle (2000). Large absolute returns may attract opposite side traders, which would increase the trading intensity. Alternatively, they may slow down trading since informed traders interpret large absolute returns as a signal that their private information has already been incorporated into prices. Unsigned trading volume is also included in the conditional expected duration given in expression (13). According to Easley and O'Hara (1987), large trading volume indicates an increased risk of informed trading. This may be a reason for uninformed traders to refrain from trading, which would slow down the trading intensity. However, it may also attract informed traders who want to benefit from their private information before others do so. This would lead to more trading activity.

Similar to durations, trade characteristics such as absolute returns and trading volume, also exhibit daily periodicities, cf. Engle and Lunde (1998). Therefore, they have to be diurnally corrected in the usual way to account this.

7.1. Estimation results

The estimation results for the diurnal components of the trade characteristics are available upon request. Again, we use QML to estimate the ACD-model. We use a Wald test to test for higher-order effects, for which there is no significant evidence. The row with the caption 'with feedback' in Table 7 displays the estimation results. The estimation results for the diurnal correction factor are available upon request. The null hypothesis of no Granger-causality from the trade characteristics to the trading intensity is rejected at each reasonable confidence level using a Wald test, making clear that there is significant feedback between the trading intensity and the various trade characteristics. To assess the effect of trade characteristics on the trading intensity, we investigate the sign and significance of the equilibrium multipliers $(\xi_{|r|,1}+\xi_{|r|,2})(1-\beta_3)^{-1}$ (absolute returns), $(\xi_{|x|,1}+\xi_{|x|,2})(1-\beta_3)^{-1}$(unsigned volume), and $\xi_Q(1-\beta_3)^{-1}$ (imbalance).

For three out of five stocks the long-term impact of absolute returns on durations is significantly positive. For McDonald's, the equilibrium multiplier is not significantly different from zero and for Schlumberger the effect is significantly negative. An explanation of the positive impact of absolute returns on durations is given in Dufour and Engle (2000), who note that a large change in the market maker's midprice may be a signal to the informed traders that their information has been revealed to the market maker. This means that their information is no longer superior. Therefore, their incentive to trade disappears, which decreases the trading intensity.

For all stocks, the long-term impact of trading volume on durations is significantly negative. The negative relation suggests that informed traders increase their trading intensity when they observe large trades. Since large trades are associated to an increased risk of informed trading, see e.g. Easley and O'Hara (1987), informed traders increase their speed of trading to quickly benefit from the private information they possess.

Finally, the coefficient of the imbalance in trading volume is significantly negative for all five stocks. The negative sign of the volume imbalance over the five most recent transactions suggests some effect of asymmetric information: when there is imbalance between the buy and the ask side of the market this may indicate the presence of news in one direction, either good or bad news.

This may force informed traders to increase their trading intensity to quickly benefit from the private information they possess.

7.2. The price impact of trades with feedback

As in the model without feedback, we focus on the price change of a large trade. To estimate the expected price impact of a large trade and the corresponding distribution of the price impact, we proceed as before and use the bootstrap approach of Hasbrouck (1991b).

We consider the McDonald's stock one more time. We estimate impulse response functions for a trade of 10,000 shares and compute the corresponding confidence and prediction intervals. With slow trading, the expected price change after 20 transactions equals 5.2 bp with the 5% and 95% quantiles equal to -0.2 and 11.2 bp, respectively. In case of fast trading, the expected price impact after 20 trades equals 7.5 bp with 5% and

95% quantiles equal to -0.3 and 16.2 bp. The estimates of the expected persistent price impact and of the quantiles of the persistent price impact are very close to the corresponding results in the model without feedback. See also Table 6. We obtain similar results for the price-impact function for other trading volumes, as well as for impulse response functions in calendar time. For the other stocks in our sample, we get comparable results.

Statistically speaking, the feedback from the trade characteristics is significant. Moreover, the effects are economically interpretable, using market microstructure theory. However, the impulse response functions show that economic importance of the feedback from the trade characteristics to the trading intensity is small, since it hardly affects the distribution of the price impact of a large trade.

8. Conclusions

In this paper, we investigated the price impact of trades and the relation to the trading intensity, using high-frequency data on five frequently traded stocks listed on the NYSE.

We showed that large trades lead to persistent price changes and for some stocks increase price volatility. Instead of focusing on the expected price change only, we modeled the entire distribution of the price change. In line with Dufour and Engle (2000) and Zebedee (2001), we showed that the price impact of an unexpected buy is larger in periods of frequent trading. The distribution of the absolute price change with fast trading first-order stochastically dominates the distribution of the absolute price change with slow trading. Furthermore, volatility is also higher when durations between trades are short. Hence, trades are more informative in periods of frequent trading.

We established significant causality from absolute returns, trade size, and trade imbalance to the trading intensity. Ceteris paribus, large trades, and large order imbalances increase the trading intensity, but large absolute returns slow down trading. We investigated the economic relevance of this feedback by comparing the distribution of the price change following a trade in the models with and without feedback. We show that there is hardly any difference, which suggests that the economic impact of the feedback from the trade characteristics to the trading intensity is small for the stocks considered in this paper.

Acknowledgements

The author would like to thank the participants of the CentER seminar at Tilburg University, the participants of the 2001-conference in Sandbjerg organized by the Center of Analytical Finance of the University of Arhus, and the participants of the 14th Australasian Finance and Banking Conference in Sydney, 2001. The author is grateful for financial support from the Netherlands Organization for Scientific Research.

References

Admati, A.R., Pfleiderer, P., 1988. A theory of Intraday patterns: volume and price variability. The Review of Financial Studies 1, 3–40.

Bauwens, L., Giot, P., 2000. The logarithmic ACD-model: an application to the bid-ask quote process of three NYSE-stocks. Annales d'Economie et de Statistique 60, 117–150.

Berndt, E.R., Hall, B.H., Hall, R.E., Hausman, J.A., 1974. Estimation and inference in nonlinear structural models. Annals of Economic and Social Measurement 3, 653–665.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric Reviews 11, 143–172.

Diamond, D.W., Verrecchia, R.E., 1987. Constraints on short-selling and asset price adjustments to private information. Journal of Financial Economics 18, 277–311.

Drost, F.C., Werker, B.J.M., 2001. Semiparametric duration models. Discussion paper, Tilburg University and CentER, nr. 2001-11. Journal of Business and Economic Statistics. In press.

Dufour, A., Engle, R.F., 2000. Time and the impact of a trade. Journal of Finance 55, 2467–2498.

Easley, D., O'Hara, M., 1987. Price, size and the information in security markets. Journal of Financial Economics 16, 69–90.

Easley, D., O'Hara, M., 1992. Time and the process of security price adjustment. Journal of Finance 47, 577–605.

Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. Econometrica 50, 987–1008.

Engle, R.F., Lunde, 1998. Trades and quotes: a bivariate point process, Unpublished working paper, University of California San Diego

Engle, R.F., Russell, J., 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. Econometrica 66, 1127–1162.

Engle, R.F., Hendry, D.F., Richard, J.F., 1983. Exogeneity. Econometrica 51, 277–304.

Goodhart, C.A.E., O'Hara, M., 1997. High-frequency data in financial markets: issues and applications. Journal of Empirical Finance 4, 73–114.

Hasbrouck, J., 1991a. Measuring the information content of stock trades. Journal of Finance 66, 179–207.

Hasbrouck, J., 1991b. The summary informativeness of stock trades: an econometric analysis. Review of Financial Studies 4, 571–595.

Koop, G., Pesaran, M.H., Potter, S.M., 1996. Impulse response functions in nonlinear multivariate models. Journal of Econometrics 74, 119–147.

Kraus, A., Stoll, H., 1972. Price impacts of block trading on the New York Stock Exchange. Journal of Finance 27, 569–588.

Lamoureux, C., Lastrapes, W., 1990. Heteroskedasticity in stock return data: volume versus GARCH-effects. Journal of Finance 45, 221–229.

Lee, C.M.C., Ready, M.J., 1991. Inferring trade direction from Intraday data. Journal of Finance 46, 733–746.

Manganelli, S., 2002. Time, Volume and the Price Impact of Trades, Working Paper European Central Bank nr. 125.

O'Hara, M., 1995. Market Microstructure Theory, Basil Blackwell. Blackwell Publishers Inc., Cambridge.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.

Zebedee, A.A., 2001. The impact of a trade on national best bid and offer quotes: a new approach to modeling irregularly spaced data. Journal of Multinational Management 11, 363–383.