

Durations, Volume and the Prediction of Financial Returns in Transaction Time

Christian M. Hafner*

September 1999

Abstract

Traditional microstructural theories of asset pricing emphasize the role of volume as a trend indicator. With the availability of large transaction data sets, one has started recently to incorporate more information of the trades, such as the time between trades, to describe the multivariate dynamics of transactions. Without knowing *a priori* the relation between the observed components of a trade – price, duration between trades, and volume – one may follow the principle of ‘letting the data speak for themselves’. The goal of this paper is to evaluate the informational content of both volume and durations to predict transaction returns using explorative nonparametric methods. The empirical results for transaction data of IBM stock prices confirm the role of volume as a trend indicator and suggest that the bid–ask bounce is smaller in highly active than in less active trading periods. That is, after a sell (buy) expected returns are decreasing (increasing) with volume and increasing (decreasing) with durations.

*Sonderforschungsbereich 373, Humboldt–Universität zu Berlin, Germany. Mailing address: Institut für Statistik und Ökonometrie, Wirtschaftswissenschaftliche Fakultät, Humboldt–Universität zu Berlin, Spandauer Str. 1, D-10178 Berlin, Germany; e-mail hafner@wiwi.hu-berlin.de; Financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged. Helpful discussion was provided by the participants at the ‘Symposium on Microstructure and High Frequency Data’, december 1998 in Paris. The author wants to thank Luc Bauwens and Pierre Giot for helpful discussions and for generously providing the IBM stock data.

1 Introduction

A standard branch of microstructure theory has been concerned with the informational content of volume to identify informed traders. Also empirical evidence suggests that the trading volume contains information that can be exploited for profitable trading strategies, see e.g. Acar and Lequeux (1996) and Ghysels, Gouriéroux and Jasiak (1998). These approaches are very close to technical trading rules used by market participants, one of which can be summarized as ‘volume goes with the trend’.

For transaction data, however, the role of the timing of trades has so far attracted little attention. Therefore, we would like to have a model framework that simultaneously deals with volume and durations to predict returns. First approaches in this direction have been provided (1) by Russell and Engle (1998) in their autoregressive conditional multinomial (ACM) model, where the parameters of a multinomial logit model follow a vector ARMA process, and (2) by Rydberg and Shephard (1998), who decompose the discrete price series into activity, direction and size, predicting each component by autologistic models. These approaches crucially rely on restrictive assumptions about the dynamic structure of trade variables. However, it is known that high frequency financial data often exhibit complex nonlinear patterns that render the task of model identification difficult. Without imposing a potentially restrictive model structure, one may thus want to follow the principle of ‘letting the data speak’ and use explorative non-parametric models. The visualization of function estimates then helps to construct parametric models that can be used, e.g., for prediction purposes.

In Hafner (1998), a related paper dealing with option prices, we found that the information contained in volume dissipates for longer durations between trades. For option markets, this is plausible because other sources of information such as the evolution of the underlying become more important than the marks of the last trade in the option as time elapses. If, on the other hand, durations are short so that no exogenous new information such as new transactions in the underlying can be expected for the new trade, volume retains its traditional role of a trend indicator. That is, when the trade is buy-initiated one expects a positive trend, when it is sell-initiated one expects a negative trend, conditional on high volume and short durations.

In this paper, we analyze transaction data for the IBM stock, as this is a data set frequently used in recent papers. In Section 2, we make use of the autoregressive conditional duration (ACD) model of Engle and Russell (1998) to predict durations. As Dufour and Engle (1999) emphasize, there is yet no satisfactory economically motivated model for durations, so one usually assumes that durations are exogenous, using standard time series models to predict durations. One could extend the approach followed here to approaches considering durations as endogenous. Indeed, many papers consider durations or the trading intensity, which is closely related to volatility, as the dependent variable for which microstructural theories are formulated and tested. Examples are Bauwens and Giot (1999a, 1999b), Engle (1996) and sections 7 and 8 of Engle and Russell (1998). As emphasized above, we take another perspective and consider the direction of price changes as the variable to be explained. This may have the advantage that not only the existence of news matters for the price updating behavior of agents, but also its quality: When traders learn about the existence of good (bad) news from observing the trade

marks, they will immediately update their beliefs about short-term positive (negative) trends. On the other hand, whether news is good or bad does not affect (or, at least, much less) the trading intensity or duration process. Thus, observing past price changes as well as trade marks may give a better understanding of how agents update their beliefs than considering the marks alone.

In Section 3, we estimate the expectation of returns conditional on current durations and lagged volume. To estimate the conditional expectation, we use nonparametric kernel methods and, based on this analysis, parametric threshold models. Our results indicate that unlike for option transactions volume keeps its informational content for longer durations. This reveals a substantial difference of the dynamics of transaction marks in option markets and in stock markets.

2 Durations and the ACD model

A well known property of transaction durations is the clustering of high respectively low transaction rates and thus positive autocorrelation of durations. See Figure 4 for the empirical autocorrelation function of IBM stock transaction durations, adjusted for deterministic seasonality effects. To take this duration clustering into account, a popular modelling approach is based on the autoregressive conditional duration (ACD) model of Engle and Russell (1998), which we also employ in this paper.

We analyze transaction data for the IBM stock traded at the New York Stock Exchange during september, october and november 1996, a total of $n = 61,063$ observations. The same data is used and described in more detail by Bauwens and Giot (1999a, 1999b). The data set consists of the trading time (t_i), the price (S_i) and the trading size (q_i). We define (log) volume $v_i = \log(S_i q_i)$, returns in transaction time, $r_i = \log S_i / S_{i-1}$ and durations $D_i = t_i - t_{i-1}$. To find a suitable model for the durations D_i , we first of all adjust for deterministic time-of-day effects, as it is well known that activity is larger in the morning and afternoon than it is over lunch. Thus, we estimate the conditional expectation $E[D_i | t_{i-1} = t] = \tau(t)$ via nonparametric regression methods and calculate the seasonally adjusted durations $d_i = D_i / \tau(t_{i-1})$. The estimated function $\tau(t)$ is displayed in Figure 1. Clearly visible are the larger durations over lunch-time. Summary statistics for seasonally adjusted durations d_i , returns r_i and log volume v_i are provided in Table 1. The transaction returns have a strong negative first order autocorrelation (-0.164), which may be explained by the usually observed bid-ask bounce for high frequency financial time series.

As is obvious from the ACF in Figure 4, the temporal dependence of durations remains strong after seasonal adjustments. Therefore, we fit a model for the expectation ψ_i of duration d_i conditional on the filtration \mathcal{F}_{i-1} that generates the processes q_i , S_i and d_i . We assume that ψ_i only depends on past realizations of the durations such that

$$\psi_i = \omega + \alpha d_{i-1} + \beta \psi_{i-1} \tag{1}$$

with constant parameters $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$. Including lagged volume as explanatory variable in (1) yielded an insignificant coefficient which justifies our assumption concerning non-causality

	mean	variance	skewness	kurtosis	ρ_1
durations	1.37	2.28	2.90	26.22	0.230
returns	5.6E-06	2.8E-07	-0.19	18.07	-0.164
log volume	11.80	1.85	0.06	2.55	0.171

Table 1: *Summary statistics for seasonally adjusted durations, returns, and log volume. ρ_1 is the first order autocorrelation.*

	Distribution	$\gamma < 1$	$\gamma = 1$	$\gamma > 1$
$\alpha = 0$		e. d.	none	l. d.
$\alpha > 0$	conditional	e. d.	none	l. d.
	unconditional	e. d.	e.d.	?

Table 2: *Excess dispersion (e.d.), less dispersion (l.e.) or none for alternative parameters of the Weibull ACD(1,1) model.*

of volume for durations. Denoting $\phi_i = \psi_i/\Gamma(1+1/\gamma)$ for a constant $\gamma > 0$ and Gamma function Γ , we can write the model for the durations as

$$d_i = \phi_i \varepsilon_i \quad (2)$$

with stochastic *i.i.d.* errors ε_i that are Weibull(1, γ) distributed. The case $\gamma = 1$ corresponds to the exponential distribution, for which we have $\phi_i = \psi_i$. Models of higher order than the ACD(1,1) model in (1) may also be considered.

One feature of financial duration data is the excess dispersion, that is, a standard deviation which is larger than the mean. This can be the case for the unconditional and the conditional distribution. For the Weibull ACD(1,1) model, six cases can be distinguished, which are summarized in Table 2. Note that for the case $\alpha > 0$ and $\gamma > 1$, it depends on the actual numerical value of the parameters (also of β) whether or not the unconditional distribution displays excess dispersion. Recall from Table 1 that the IBM stock durations exhibit some excess dispersion with a mean duration of 1.37 and a standard deviation of 1.51.

The parameters of the ACD model are estimated by maximum likelihood estimation. The log likelihood function is

$$\log L(\theta \mid d_1, \dots, d_n) = \sum_{i=1}^n \log \frac{\gamma}{d_i} + \gamma \sum_{i=1}^n \log \frac{d_i}{\phi_i} - \sum_{i=1}^n \left(\frac{d_i}{\phi_i} \right)^\gamma, \quad (3)$$

which can be maximized numerically with respect to the parameter vector $\theta = (\omega, \alpha, \beta, \gamma)'$. Estimation results for the ACD(1,1) model applied to the IBM stock durations are presented in Table 3. We note the usual result of high persistence, i.e., $\alpha + \beta$ is very close to one for both Exponential and Weibull ACD. All parameter estimates are highly significant. In particular, the parameter $\gamma = 1.14$ of the Weibull ACD is significantly larger than one. The exponential ACD

	Exponential	Weibull
ω	.0106 (.0014)	.0099 (.0013)
α	.0644 (.0037)	.0640 (.0037)
β	.9282 (.0045)	.9293 (.0045)
γ		1.1397 (.0031)
$\log L$	-74,987.2	-74,138.0

Table 3: *QMLE Estimation results for the ACD(1,1) model with (1) exponential, (2) Weibull conditional density. Heteroskedasticity-consistent standard errors are given in parentheses. $\log L$ denotes the log likelihood value.*

is strongly rejected by a likelihood ratio test in favor of the Weibull ACD. Engle and Russell (1998) find $\gamma < 1$ for IBM stock transaction data in november to january 1990/91, which may indicate a change of duration dynamics over recent years in this particular stock.

A way to characterize the instantaneous probability of an event at a given time t is the hazard function. Denote by $f(t)$ the (unconditional or conditional) density of durations, and by $F(t)$ the cumulative distribution function. Then the hazard is defined as

$$\lambda(t) = \frac{f(t)}{1 - F(t)}, \quad (4)$$

which is the probability of an event in the next small interval $[t, t + \Delta t]$ conditional on no event up to time t . For the exponential distribution, the hazard is a constant. For the Weibull distribution it decreases for $\gamma < 1$ whereas it increases for $\gamma > 1$. Estimates of the unconditional duration density and hazard are shown in Figure 2 and 3, respectively. A decreasing shape of the hazard can be caused by duration clustering ('ACD effect'), by a Weibull innovation with $\gamma < 1$, or both. An increasing shape of the hazard can be caused by a Weibull innovation with $\gamma > 1$ that weighs stronger than the duration clustering effect. Interestingly, the hazard estimate shows a non-monotone shape with three modes while the overall shape is increasing.

Extensions of the standard ACD model are possible. For example, Bauwens and Giot (1999a) model logarithmic expected durations with an ACD-type model. This does not constrain the parameters to be positive, which allows them to include additional explanatory variables. Gouriéroux, Jasiak and LeFol (1997) investigate the time varying shapes of duration densities and hazard functions over calendar time. Furthermore, in the light of the slow decay of the ACF in Figure 4 one may consider long memory ACD models.

3 Model identification for returns

In the previous section we have analyzed the conditional distribution of transaction durations based on observations of the past. In particular, we have used the ACD model to obtain a simple but powerful way to estimate the conditional expectation, based only on past durations.

Now, we introduce volume v_i and return r_i of transaction i to be dealt with. The object is to describe the distribution of returns conditional on the current durations (which can be predicted using the results of the previous section) and lagged volume. The joint likelihood of returns and durations can be written as

$$f(r_i, d_i | \mathcal{F}_{i-1}) = g(r_i | d_i, \mathcal{F}_{i-1})h(d_i | \mathcal{F}_{i-1}) \quad (5)$$

where $h(d_i | \mathcal{F}_{i-1})$ is derived from the ACD model. Now, we are concerned with finding a model for $g(r_i | d_i, \mathcal{F}_{i-1})$, i.e., the distribution of returns conditional on current durations and past returns and volumes. By assuming weak exogeneity of durations, one can model and estimate $g(r_i | d_i, \mathcal{F}_{i-1})$ without efficiency loss.

3.1 Nonparametric threshold models

Because we have no *a priori* conjecture of what the model for $g(r_i | d_i, \mathcal{F}_{i-1})$ in (5) might be, we may use an explorative nonparametric approach. To this end, we choose one lag of both volume and returns as explanatory variables. Since now we have three variables on which the distribution of r_i depends, i.e., d_i , v_{i-1} and r_{i-1} , a further restriction is necessary to visualize the results and to facilitate the analysis. Here we can exploit the discrete nature of price changes, being mostly plus or minus one or two tick sizes (1/8 of a dollar) or zero. As outlined in the introduction, market participants often consider volume related to some ‘trend’ measure. The principle trading rule is that volume reinforces the trend. This trend may be positive or negative, as primarily indicated by the sign of returns. Thus, we consider the restriction to the case of lagged returns being positive and negative, or more generally being discretized according to a partition with finitely many intervals. Moreover, in this framework we are not interested in the entire distribution of r_i but only in its expectation. To summarize, we will try to find a model for $m_j(d_i, v_{i-1}) = E[r_i | v_{i-1}, d_i]$ conditional on $r_{i-1} \in A_j$ where A_1, \dots, A_J is a partition of the real line. Such a threshold model may be written as

$$r_i = \sum_{j=1}^J m_j(d_i, v_{i-1})I(r_{i-1} \in A_j) + \varepsilon_i \quad (6)$$

where $I()$ denotes the indicator function. In (6), m_1, \dots, m_J are unknown smooth functions that can be estimated using standard nonparametric methods.

Consider the more general case of p regressors, $\mathbf{X} = (X_1, \dots, X_p)'$. Then, a simple Nadaraya-Watson estimate of the function m_j is

$$\hat{m}_j(x) = \sum_{i=2}^n \frac{\mathcal{K}_H(\mathbf{X}_i - x)I(r_{i-1} \in A_j)r_i}{\sum_{i=2}^n \mathcal{K}_H(\mathbf{X}_i - x)I(r_{i-1} \in A_j)} \quad (7)$$

with a $\mathbb{R}^p \rightarrow \mathbb{R}$ kernel function $\mathcal{K}(u)$, non-singular $(p \times p)$ bandwidth matrix H , and $\mathcal{K}_H(u) = \frac{1}{|H|}\mathcal{K}(H^{-1}u)$. As is well known, the Nadaraya-Watson estimator can be considered as a local constant estimator, i.e., it solves the weighted least squares regression problem

$$\min_{\beta_0} \sum_{i=2}^n (r_i - \beta_0)^2 I(r_{i-1} \in A_j) \mathcal{K}_H(\mathbf{X}_i - x).$$

More generally, one can fit local polynomials. For example, local linear estimates solve

$$\min_{\beta_0, \beta_1} \sum_{i=2}^n (r_i - \beta_0 - (\mathbf{X}_i - x)' \beta_1)^2 I(r_{i-1} \in A_j) \mathcal{K}_H(\mathbf{X}_i - x),$$

where β_0 is scalar and β_1 a p -dimensional vector. The weighted least squares solution is given by

$$\hat{\beta} = (\beta_0, \beta_1')' = (Z'WZ)^{-1}Z'Wr \quad (8)$$

with the regressor matrix

$$Z = \begin{pmatrix} 1 & X_{12} - x_1 & \cdots & X_{p2} - x_p \\ 1 & X_{13} - x_1 & \cdots & X_{p3} - x_p \\ \dots & \dots & \dots & \dots \\ 1 & X_{1n} - x_1 & \cdots & X_{pn} - x_p \end{pmatrix},$$

the weight matrix $W = \text{diag}(\mathcal{K}_H(\mathbf{X}_i - x)I(r_{i-1} \in A_j))$, and $r = (r_2, \dots, r_n)'$. For a survey on local polynomial estimators, see Fan and Gijbels (1996).

For the partition, we start with three intervals, $A_1 = (0, \infty)$, $A_2 = (-\infty, 0)$ and $A_3 = [0]$ that will give us some indication of short term price movements. The choice of this partition naturally corresponds to the partition concerning the initiation of a trade: If $r_i \in A_1$, the trade very likely was buy-initiated, if $r_i \in A_2$, it was very likely sell-initiated, and if $r_i \in A_3$, the trade took place at the last transaction price. One has to ensure that for each interval there are enough observations such that the principle of local averaging does not collapse. For our three cases, there are 8215, 7866 and 44923 observations, respectively, for positive, negative and zero returns, which is sufficient for the bivariate estimate. Note that due to the discrete character of returns the probability of a return falling in A_3 is not zero. In fact, by far the most returns, about 73.6%, are zero. Since there are few observations with very large durations, the function estimates might be deteriorated in those regions because of data sparseness. Therefore, for the estimation we have eliminated those cases of durations that are more than two standard deviations larger than the mean.

We use the simplifying assumption of a diagonal bandwidth matrix, $H = \text{diag}(h_1, h_2)$. Also, we employ the product kernel $\mathcal{K}(u) = \prod_{i=1}^p K(u_i)$ with univariate kernels K . Then the Nadaraya-Watson estimates are

$$\hat{m}_1(d, v) = \sum_{i=2}^n \frac{K_{h_1}(d_i - d)K_{h_2}(v_{i-1} - v)I(r_{i-1} > 0)r_i}{\sum_{i=2}^n K_{h_1}(d_i - d)K_{h_2}(v_{i-1} - v)I(r_{i-1} > 0)}, \quad (9)$$

$$\hat{m}_2(d, v) = \sum_{i=2}^n \frac{K_{h_1}(d_i - d)K_{h_2}(v_{i-1} - v)I(r_{i-1} < 0)r_i}{\sum_{i=2}^n K_{h_1}(d_i - d)K_{h_2}(v_{i-1} - v)I(r_{i-1} < 0)}, \quad (10)$$

and

$$\hat{m}_3(d, v) = \sum_{i=2}^n \frac{K_{h_1}(d_i - d)K_{h_2}(v_{i-1} - v)I(r_{i-1} = 0)r_i}{\sum_{i=2}^n K_{h_1}(d_i - d)K_{h_2}(v_{i-1} - v)I(r_{i-1} = 0)}, \quad (11)$$

with $K_h(u) = (1/h)K(u/h)$. The bandwidths h_1 and h_2 determine the degree of smoothness of the function estimate. We used standard rule of thumbs to determine the bandwidths to balance the bias problem of oversmoothing and the variance problem of undersmoothing.

The estimated functions \hat{m}_1 , \hat{m}_2 and \hat{m}_3 are shown in Figure 5, 7 and 9, respectively. To give an idea of the corresponding densities of the explanatory variables, we also plot the density estimates in Figure 6, 8 and 10.

First of all, recall from Table 1 that the returns have negative first order autocorrelation (-0.164) such that m_1 will tend to be negative, whereas m_2 will tend to be positive. Consider first the estimate of m_2 in Figure 7. The function is increasing in the direction of the durations. This could be interpreted as a reduced bid–ask bounce effect for small durations. In other words, the bid–ask bounce effect appears to be smaller in highly active trading periods than in less active periods. Moreover, we see a decrease of the function in the direction of volume. This may indicate that the belief in a trend signal coming from an informed trader who has sold the stock is higher for large volumes. Together, we have two effects: the bid–ask bounce and the trend signal effects. The latter is found to be more important than the former in case of high trading activity and large volumes, whereas the bid–ask bounce dominates for small volumes and low transaction rates.

The shape of the function in the volume direction appears to be somewhat nonlinear: for small volumes it remains flat or even increases, then declines strongly. This may be interpreted as volume being a trend indicator only for sufficiently large volumes.

This picture is mirrored when we turn to the case of a positive lagged return in Figure 5. Note that the shape of this function is quite linear. As the negative first order autocorrelation implies, the function should be negative on average, which it is. However, we see that there is an increase in the direction of the volume, which indicates that for large volumes one may have a strong belief in a signal by an informed trader who has bought the stock. In the direction of durations, the function declines due to the increasing bid–ask bounce. Recall that we have already taken care of outliers such that the function estimates are not shown in the figures for extremely large durations.

Finally, for zero returns (Figure 9) the function appears to be flat without any important structure. Hence, if the last trade occurred without changing the prevalent price it does not seem to be possible to exploit the information of volume and duration to predict the direction of prices. This is because in this case one cannot infer from the trade whether it was buy– or sell–initiated.

To summarize the results of this explorative study, we can state that the sign of lagged returns may indeed indicate a certain signal coming from an informed trader. This signal may be considered to be stronger if the trade occurred with a large volume. This is consistent with traditional microstructural theories, see e.g. Glosten and Milgrom (1985) and Easley and O’Hara (1987). The new aspect of our analysis is the additional information about the timing of the trade.

3.2 Linear threshold models

At a next step, we may now formulate a parametric model based on the nonparametric estimates. We have noticed a distinctly different shape of the functions depending on whether lagged returns are positive, negative or zero. Thus, one may introduce a linear threshold model where the

model	α_1	α_2	α_3	α_4	α_5	α_6
M1	0.033 (2.42)	-1.875 (-18.82)	1.977 (21.08)	0.185 (4.52)	0.122 (15.31)	-0.134 (-17.05)
M2	0.059 (4.27)	-0.331 (-11.36)	0.412 (13.92)	0.026 (4.09)		
M3	0.034 (2.47)	-1.659 (-16.33)	1.899 (19.85)	0.218 (5.23)	0.112 (13.94)	-0.130 (-16.53)
	α_7	α_8	α_9	α_{10}	R^2	$Q(10)$
M1	-0.015 (-4.26)				3.99%	130.42
M2		-0.099 (-12.33)	0.046 (5.82)	-0.010 (-3.27)	3.45%	111.92
M3	-0.016 (-4.66)	-0.085 (-10.62)	0.032 (4.01)	-0.011 (-3.78)	4.22%	132.02

Table 4: *Estimation results of the threshold model given in (12). t statistics are given in parentheses. The coefficient of determination for the regression is denoted by R^2 and the Box Ljung statistic for ten lags by $Q(10)$ with 1% critical value 18.47.*

threshold variable is the lagged return r_{i-1} . As in the previous section, define the partition $A_1 = (0, \infty)$, $A_2 = (-\infty, 0)$, and $A_3 = [0]$. Then, we introduce the threshold model

$$r_i = \alpha_1 r_{i-1} + \sum_{j=1}^3 (\alpha_{j+1} + \alpha_{j+4} v_{i-1} + \alpha_{j+7} d_i) I(r_{i-1} \in A_j) + \varepsilon_i. \quad (12)$$

Of particular interest are restrictions of the threshold model in (12) that we call ‘volume model’ or M1, where α_8 , α_9 and α_{10} are restricted to zero, and ‘duration model’ or M2, where α_5 , α_6 and α_7 are restricted to zero. The unrestricted model (12) is termed M3. The parameter estimates of models M1, M2 and M3 are presented in Table 4.

Note first that all parameter estimates are significant, which is not surprising for more than 60,000 observations. The negative sign of the parameter α_8 in M2/M3 corresponds to the decreasing shape of \hat{m}_2 in Figure 5 in the duration direction, indicating a loss of signalling effects due to longer durations. The positive sign of α_9 in M2/M3 indicates the reverse situation, that is, after a negative return subsequent trades have on average lower returns for short durations than for long durations.

Concerning the volume parameters in M1/M3, the parameters again take the expected signs: negative (α_6) after a sell and positive (α_5) after a buy. Both are even stronger significant than the duration parameters, and together with the slightly higher R^2 of the regression for M1 than for M2 one may conclude that volume has more explanatory power for IBM returns than durations. This is different from the results for the DAX call option by Hafner (1998) where durations were found to be more important than volume. Furthermore, recall from Figure 7 that

the function for a negative lagged return was quite nonlinear in the volume direction. Including a threshold at medium volume levels, we were able to still increase the R^2 slightly.

Finally, note that all parameters related to zero returns (α_4 , α_7 and α_{10}) are much less significant, which corresponds to the flat function shown in Figure 9.

The significance of the Box Ljung statistic for the residuals of all models is also due to the very large number of observations, where typically standard specification tests reject any parsimonious model.

4 Conclusions and outlook

Using explorative nonparametric methods, we have shown two basic effects of IBM stock transaction data: the first is the apparent belief of market participants in the importance of volume for revealing informational asymmetry, as standard microstructure theory suggests. Secondly, we find that this information revelation is moderated by a second variable, i.e., the time between trades (durations). If there is a buy (sell), expected returns tend to decrease (increase) with durations. The bid-ask bounce is thus found to be smaller for high transaction rates.

Our nonparametric approach has the potential to discover nonlinear structures of the data and is thus particularly fruitful in the application to financial transaction data, where the functional relationships between the involved variables are far from clear. To some extent (i.e., for the case of a sell) we found evidence for nonlinearities in the data.

Further research may include the information of the spread between bid and ask quotes to find out whether volume is informational only in periods of wide spreads, as the model of Admati and Pfleiderer (1988) suggests. If spreads are small, their model predicts that liquidity traders rather than informed traders provide large volumes. Our finding of the strong informational content of volume might thus be due only to the cases where spreads are wide. Again, this analysis has to be performed using durations, volume, and spreads simultaneously.

Appendix: Some remarks on asymptotic theory

Although this paper only deals with bivariate regression problems, consider the general case of p regressors, X_1, \dots, X_p . For the asymptotic analysis of estimates of model (6), one needs that the number n_j of observations falling into the interval A_j converges to infinity. This is established by the stationarity of returns and the fixed probability p_j of returns falling into interval A_j . Thus, we have $n_j = np_j(1 + o_p(1))$. For notational simplicity we use a scalar bandwidth h . Assume that $h \rightarrow 0$ and $h^p n \rightarrow \infty$. Define the kernel constants $\|\mathcal{K}\|_2^2 = \int \mathcal{K}^2(u) du$ and $\mu_2(\mathcal{K}) I_p = \int uu' \mathcal{K}(u) du$. One can prove along the lines of Ruppert and Wand (1994) for the regression case and Härdle, Tsybakov and Yang (1998) for the time series case that the bias of $\hat{m}_j(x)$ for the Nadaraya-Watson estimator (7) is given by

$$E[\hat{m}_j(x)] - m_j(x) = h^2 \mu_2(\mathcal{K}) \left(\frac{1}{2} \text{Tr}(\nabla^2 m_j(x)) + \frac{\nabla' m_j(x) \nabla f_j(x)}{f_j(x)} \right) + o_p(h^2),$$

and for the local linear estimator (8) by

$$E[\hat{m}_j(x)] - m_j(x) = h^2 \mu_2(\mathcal{K}) \left(\frac{1}{2} \text{Tr}(\nabla^2 m_j(x)) \right) + o_p(h^2),$$

where f_j denotes the marginal density of X conditional on $r_{i-1} \in A_j$, ∇ the gradient and ∇^2 the Hessian operators. Furthermore, the variance of both estimators is given by

$$\text{Var}(\hat{m}_j(x)) = \frac{1}{n_j h^p} \|\mathcal{K}\|_2^2 \frac{\sigma_j^2(x)}{f_j(x)} (1 + o_p(1)),$$

where $\sigma_j^2(x) = \text{Var}(r_i | X_i = x, r_{i-1} \in A_j)$. Using the results of Härdle, Tsybakov and Yang (1998), one can also show asymptotic normality of the estimates.

References

- Acar, E. and P. Lequeux (1996), Intra-day patterns and the role of volume: An application to the DAX Futures contract, in: Proceedings of the Third International Conference on “Forecasting Financial Markets”, London.
- Admati, A.R. and P. Pfleiderer (1988), A theory of intraday patterns: volume and price variability, *The Review of Financial Studies* 1, 3–40.
- Bauwens, L. and P. Giot (1999a), The logarithmic ACD model: An application to the bid-ask quote process of two NYSE stocks, CORE, Université catholique de Louvain, Louvain-la-Neuve.
- Bauwens, L. and P. Giot (1999b), Asymmetric ACD models: Introducing price information in ACD models with a two state transition model, CORE, Université catholique de Louvain, Louvain-la-Neuve.
- Dufour, A. and R.F. Engle (1999), Time and the price impact of a trade, Department of Economics, University of California at San Diego.
- Easley, D. and M. O’Hara (1987), Price, trade size and information in securities markets, *Journal of Financial Economics* 19: 69–90.
- Engle, R.F. (1996), The econometrics of ultra-high frequency data, UCSD DP 9615, San Diego.
- Engle, R.F. and J. Russell (1998), Autoregressive conditional duration: A new model for irregularly spaced transaction data, *Econometrica* 66: 1127–1162.
- Fan, J. and I. Gijbels (1996), *Local Polynomial Modelling and its Applications*, Chapman & Hall.
- Ghysels, E., C. Gouriéroux and J. Jasiak (1998), Causality in return and volume state transitions, working paper.
- Glosten, L. and P. Milgrom (1985), Bid, ask, and the transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 13: 71–100.

- Gouriéroux, C., J. Jasiak and G. LeFol (1997), Intraday market activity, CREST DP, Paris.
- Härdle, W., A. Tsybakov and L. Yang (1998), Nonparametric vector autoregression, *Journal of Statistical Planning and Inference* 68: 221–245.
- Hafner, C. M. (1998), Durations, volume and the prediction of financial returns in transaction time: Evidence from the German futures exchange, presented at the Symposium on ‘Microstructure and High Frequency Data’, december 1998, Paris.
- Ruppert, D. and M.P. Wand (1994), Multivariate locally weighted least squares regression, *Annals of Statistics* 22: 1346–1370.
- Russell, J. and R.F. Engle (1998), Econometric analysis of discrete-valued, irregularly spaced financial transactions data using a new autoregressive conditional multinomial model, Graduate School of Business, University of Chicago.
- Rydberg, T.H. and N. Shephard (1998), Dynamics and trade-by-trade price movements: decomposition and models, Nuffield College, Oxford.

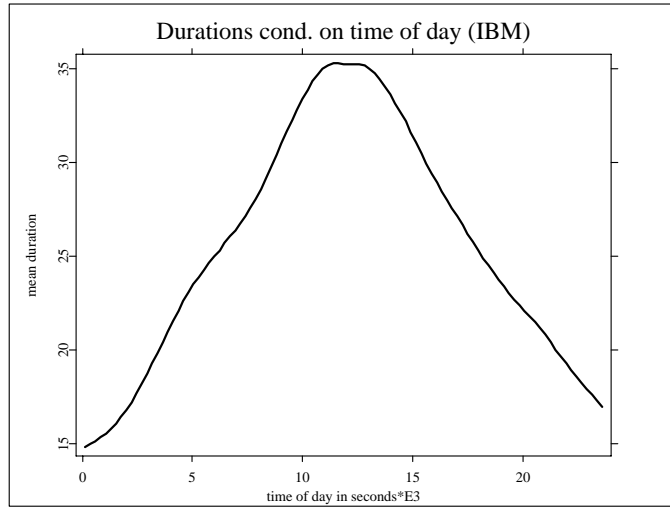


Figure 1: *Expected durations as a function of the time of day*

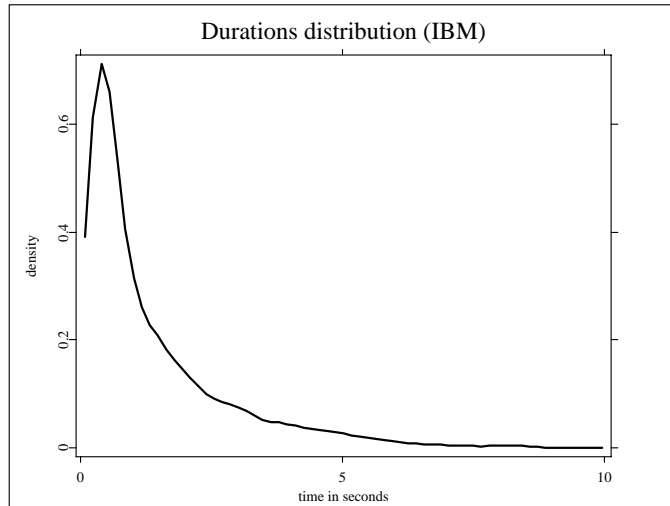


Figure 2: *Density estimate for the seasonally adjusted durations*

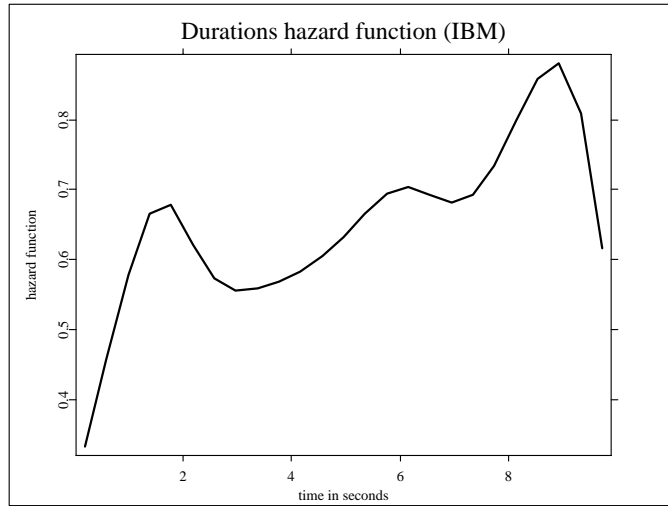


Figure 3: *Hazard function estimate for the seasonally adjusted durations*

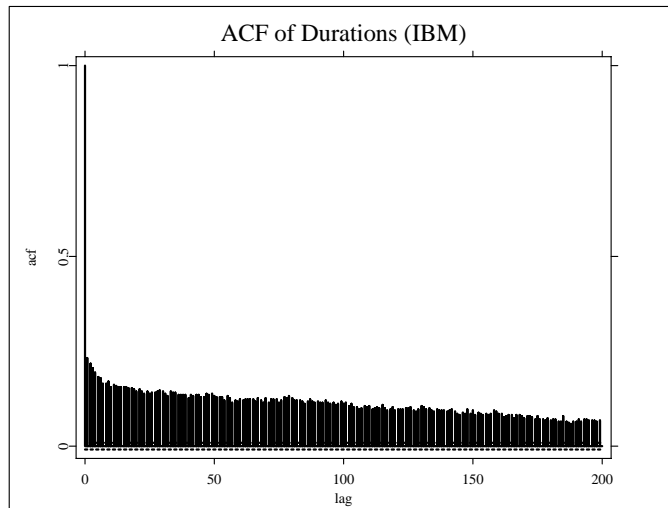


Figure 4: *The ACF of durations*

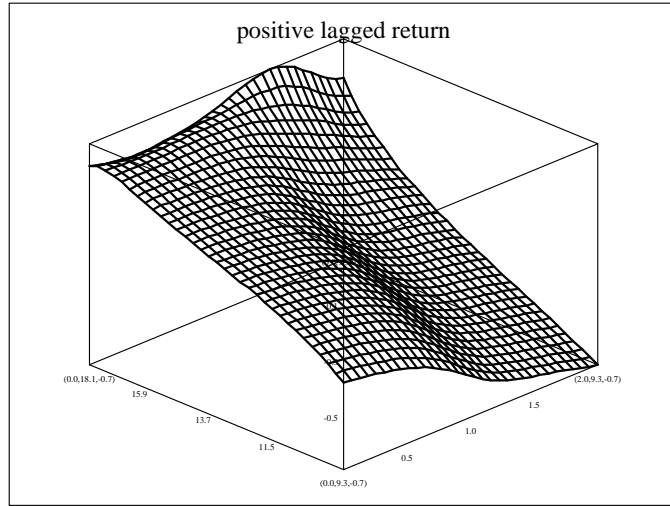


Figure 5: *Estimated return expectation conditional on current duration (right axis) and lagged volume (left axis) after a buy of the stock*

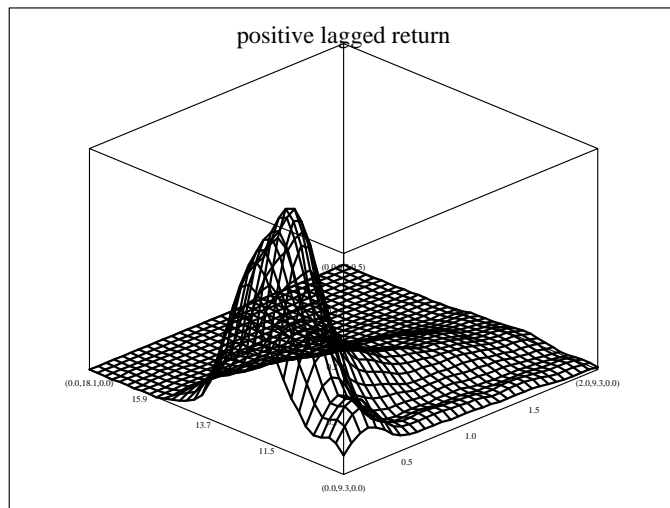


Figure 6: *Estimated density of current duration (right axis) and lagged volume (left axis) after a buy of the stock*

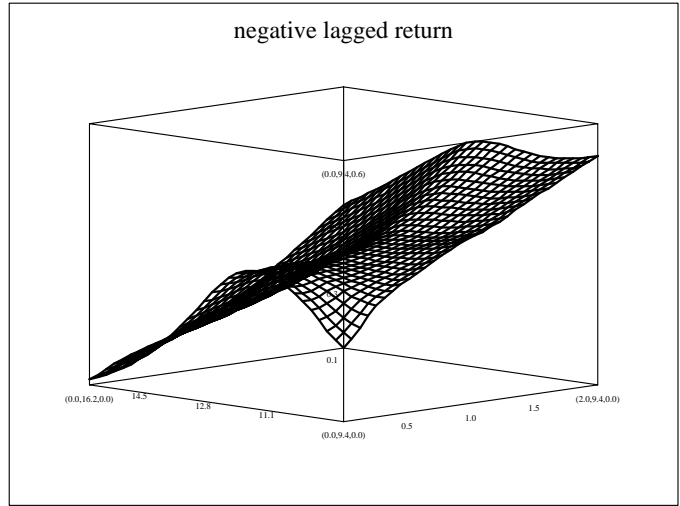


Figure 7: *Estimated return expectation conditional on current duration (right axis) and lagged volume (left axis) after a sell of the stock*

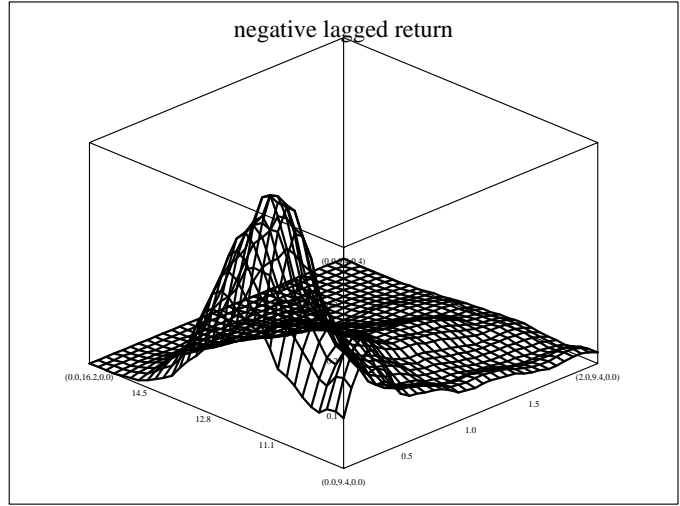


Figure 8: *Estimated density of current duration (right axis) and lagged volume (left axis) after a sell of the stock*

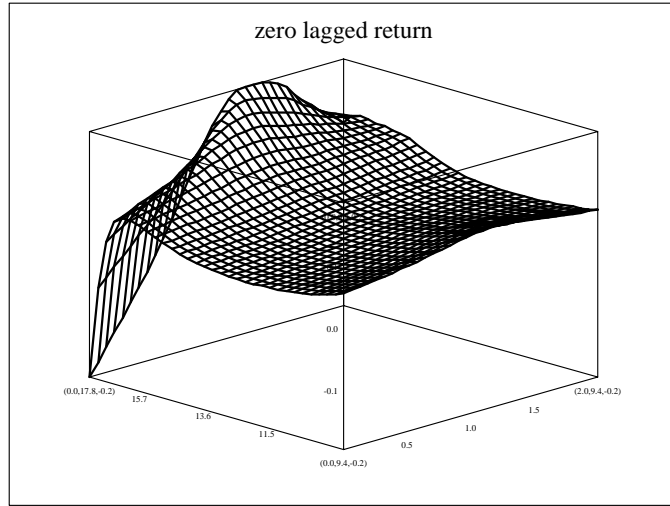


Figure 9: *Estimated return expectation conditional on current duration (right axis) and lagged volume (left axis) after a transaction without price change*

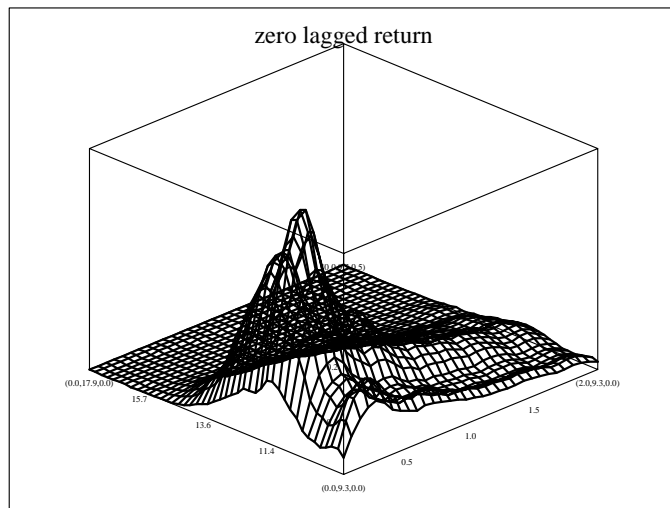


Figure 10: *Estimated density of current duration (right axis) and lagged volume (left axis) after a transaction without price change*