Managerial Economics
2016, vol. 17, no. 2, pp. 241–260
http://dx.doi.org/10.7494/manage.2016.17.2.241

Henryk Gurgul*, Robert Syrek**, Christoph Mitterer***

Price duration versus trading volume in high-frequency data for selected DAX companies

1. Introduction

The properties of the time series of durations between consecutive trades of a particular stock have been studied by many contributors in the literature of financial econometrics. Among them are highly prominent scientists like Engle (2000) and Gourieroux and Jasiak (2001). The importance of this topic, accompanied by the growing availability of (ultra-)high-frequency data, has prompted an increase of contributions in recent years. Intensive research based on high-frequency data has several financial motivations. First of all, it is linked with microstructure theory. Secondly, it contributes to the literature on stochastic time deformation. But the most important need for research on the dynamics of trade durations is the necessity to manage liquidity risk. The reason is that durations between the following trades are a widely accepted measures of market liquidity. In addition, their volatility reflects the liquidity risk.

The results of empirical investigations of trade durations suggest several stylized facts typical of high-frequency data. The knowledge of empirical facts is a precondition of the proper specification of econometric models. The most-important stylized facts include positive serial autocorrelations and clustering effects; i.e., the propensity of extremely long durations and extremely short-to-build clusters; the persistence of dependence in time; i.e., autocorrelations tend

* AGH University of Science and Technology in Krakow, Department of Applications of Mathematics in Economics, e-mail: henryk.gurgul@gmail.com
** Jagiellonian University in Krakow, Institute of Economics, Finance and Management, e-mail: robert.syrek@uj.edu.pl
*** Bruell Kallmus Bank AG, Institutional Banking, Graz, Austria, e-mail: mitterer@bkbank.at
Financial support for this paper from the National Science Center of Poland (Research Grant DEC-2012/05/B/HS4/00810) is gratefully acknowledged by Henryk Gurgul.

to decrease slowly, which indicates the possible existence of long memory. Further features are significant nonlinearities in the dynamics, reflected in nonlinear autocorrelograms. In addition, in high-frequency data, there is path-dependent (under-)overdispersion in the conditional distribution. Moreover, one can detect significant departures from an unconditional exponential distribution, negative duration dependence, and fat tails. In order to take into account these empirical facts, the researcher should assume flexible specifications for conditional mean and conditional variance. This is necessary for the proper management of liquidity risk. In many situations, extreme liquidity risks must be calculated. In this case, the first conditional moments may not be enough. Therefore, for some research questions, measures that reflect the entire conditional distribution are advisable. According to Ghysels et al. (2004), this situation may occur in the case of Time-at-Risk (TaR(t)). Time-at-Risk denotes the minimal time without a trade that can take place with a given probability. The mentioned measures need the most-flexible specifications possible for the entire conditional distribution of the duration process.

One the most-frequently used dynamic models for intertrade durations is the famous Autoregressive Conditional Duration (ACD) model formulated by Engle and Russell (1998). This model involves an accelerated hazard specification with conditional mean that underlines a deterministic autoregression. As demonstrated by Ghysels et al. (2004), the different stylized effects observed in the data can be replicated in the framework of the ACD. The drawback of this specification is the number of restrictive assumptions on the conditional distribution of the duration process. In this model, the dynamics of the conditional mean determines the dynamics of conditional moments of any order and of liquidity risk measures (e.g., TaR(t)). However, most of these restrictions are not reflected in empirical facts. The reason is that they imply a path-independent conditional dispersion. Moreover, they are not necessary for the management of liquidity risk. In order to avoid these problems, Ghysels et al. (2004) suggested a new specification of accelerated hazard. They derived the Stochastic Volatility Duration (SVD) model. In this model, the authors included two underlying factors; the conditional mean and conditional variance follow two independent dynamics.

The main goal of our paper is to analyze the dependence structure between duration and trading volume visible in high-frequency data.

The remaining part of the paper is scheduled as follows. In the following section, we give a brief literature overview that focuses on known empirical results concerning duration in the framework of the microstructure dynamics of tick-by-tick stock data. Section 3 outlines the basics of models of duration and dependence measures based on copulas. In the fourth section, we provide descriptive statistics of the intraday dataset and then present empirical results (especially on the dependence between duration and trading volume reflected in intraday data). Finally, we draw conclusions.

2. Literature overview

Over the last two decades, an essential part of the literature devoted to market microstructure has analyzed intraday prices and the process of their formation. De Jong and Rindi (2009), like many other authors, focused on theoretical deliberations especially concerned with market structure and market designs. The most important question was the impact of these factors on intraday price formation. In recent years, intraday high-frequency data has become increasingly available. Therefore, contributors started to empirically test some of the known theories of market microstructure. It was also possible to model the observed facts within the intraday price dynamics. Empirical studies on trading volume in the US equity markets (based on tick-by-tick data) showed the intraday behavior of stock prices. Engle (2000) found that the biggest increase in the volume of transactions takes place at the opening and closing of the market, so there is a U-shaped pattern of volatility over the day. The financial literature provides evidence that, for traditional stock price models, the size of time intervals is usually not important on long-time scales. However, for HF data modeling, this observation is not true. Diamond and Verrechia (1987), Easley and O'Hara (1992), Engle and Russell (1998), Engle (2000), Dufour and Engle (2000), Manganelli (2005), and Cartea and Meyer-Brandis (2010) show that, at high frequencies, the duration between trades supplies relevant information about the dynamics of tick-by-tick trades, including the behavior of the market, activity of uninformed or informed traders, volatility of price changes, and implied volatility from the option markets.

Therefore, duration (being a random variable) is one of the most important factors in stock-price behavior. It is extremely important over short periods of time. This random variable was frequently neglected in the past in most asset-pricing models with horizons of a few days or more. The reason for this was the widespread conviction that any effect of durations is dissipated very quickly. However, at present, the majority of trades are conducted by algorithmic trading processing information on a tick-by-tick level. Nowadays, duration is widely accepted as an essential random variable supplying important information about the behavior of the stock market over short-time intervals.

From a statistical point of view, the calendar-time distribution of stock price dynamics on small scales of time depends on both the distribution of price changes and the distribution of duration. The aim of trading strategies is to profit from recognized price patterns and behavior over ever-shrinking scales of time. Empirical observations show that the speed of trading has shortened by a factor of 10 in the last five years. Trading very quickly over short periods of time has become the main kind of trading (including algorithmic trading). There are many factors that support the expansion of algorithmic trading. One of them is the introduction of limit order markets. The second factor arises from changes in market structure. Both factors have lowered the entry barriers to new participants. In recent years, the capacity of computers has significantly increased. At the same time, its cost has significantly decreased. This has resulted in a rise in the number of market participants as well as a significant increase of the speed at which trading takes place.

The econometric literature on duration starts with the paper of Engle and Russell (1998), who derive the autoregressive conditional duration (ACD) model to capture the time of the arrival of financial data. Based on this seminal work, most contributors tried to generalize the ACD framework in different directions. The best-known of these are the logarithmic model of Bauwens and Giot (2000) and the extended class of models by Fernandes and Grammig (2005). Other extensions are based on regime-shifting and mixture ACD models, presented in Maheu and McCurdy (2000), Zhang et al. (2001), Meitz and Terasvirta (2006), and Hujer et al. (2002). A more-recent paper by Renault et al. (2012) suggests a structural model for the durations between events and associated marks. A detailed review of different ACD models is given in Bauwens and Hautsch (2009).

Cartea and Jaimungal (2013) stress the role of algorithmic trading (AT) and high-frequency (HF) trading (which is responsible for over 70% of the US stock trading volume). In the opinion of the contributors, both kinds of trading have greatly changed the microstructure dynamics of tick-by-tick stock data. The authors employ a hidden Markov model to examine changes in the intraday dynamics of the stock market. They try to find out how to exploit this information to develop the best trading strategies at high frequencies. The contributors demonstrate how to employ their model to submit limit orders and to profit from the bid–ask spread. They also provide evidence on how HF traders may profit from liquidity incentives (liquidity rebates). Based on data from between February 2001 and February 2008, they demonstrate that, while in 2001, the intraday states with the shortest average waiting times between trades (durations) were also the ones with very few trades; in 2008, the vast majority of trades took place in the states with the shortest average durations. In addition, the authors claim that, in 2008, the states with the shortest durations had the smallest price impact as measured by the volatility of price innovations.

3. Methodology

In our paper, we use the dynamic parametrization of the conditional mean function (Engle and Russell, 1998)

$$\psi_i := \psi_i(\theta) = E[x_i \mid \mathcal{F}_i; \theta]$$

where $\mathcal{F}_i$ stands for the information set including the observation from t_{i-1} (duration x_i between two events noticed at times t_{i-1} and t_i), and θ stands for vector of parameters.

The standardized durations

$$\varepsilon_i = \frac{x_i}{\psi_i}$$

are the sequence of independent and identically distributed random variables with $E[\varepsilon_i] = 1$. The reasons for the variation in autoregressive conditional duration models are different choices of functional form for the conditional mean function and the selection of the distribution of standardized durations.

The most-common specification suggested by Engle and Russell (1998) is linear parametrization.

Bauwens and Giot (2000) suggest two extensions of the linear ACD model. These models (known as logarithmic ACD) are of the forms

$$\ln \psi_i = \omega + \sum_{j=1}^{P} \alpha_j \ln \varepsilon_{i-j} + \sum_{j=1}^{Q} \beta_j \psi_{i-j}$$

and

$$\ln \psi_i = \omega + \sum_{j=1}^{P} \alpha_j \varepsilon_{i-j} + \sum_{j=1}^{Q} \beta_j \psi_{i-j}$$

In our contribution, we call these specifications $LACD_1$ and $LACD_2$, respectively. In these models, there are no sign restrictions on parameters to ensure the positivity of conditional duration.

We restrict our attention to cases where $P = Q = 1$, which is sufficient in our analysis. In this case, inequality $\alpha + \beta < 1$ ensures the existence of an unconditional mean of duration of the ACD model. The covariance-stationarity of $LACD_1$ is ensured by $|\beta| < 1$, whereas for $LACD_2$, we have $|\alpha + \beta| < 1$.

Another specification that researchers have to choose is the distribution for standardized durations. In their seminal paper, Engle and Russel (1998) study exponential and Weibull distributions (the exponential distribution is used in a quasi-maximum likelihood estimation).

In our contribution, we are going to fit generalized gamma, Burr distributions, and q-Weibull to the tick-by-tick data for selected German companies. The formula for the density of the generalized gamma distribution is given in Lunde (2000), whereas Gramming and Maurer (2000) consider properties of the Burr distribution. The exponential and Weibull distributions are special and limiting cases.

The q-Weibull distribution is considered by Vuorenmaa (2009) with density

$$f(\varepsilon)=(2-q)\frac{\alpha}{\beta^{\alpha}}\varepsilon^{\alpha-1}\left[1-(1-q)\left(\frac{\varepsilon}{\beta}\right)^{\alpha}\right]^{\frac{1}{1-q}}$$

with $\beta=\frac{(q-1)^{\frac{1+\alpha}{\alpha}}}{2-q}\frac{\alpha\Gamma\left(\frac{1}{q-1}\right)}{\Gamma\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{1}{q-1}-\frac{1}{\alpha}-1\right)}$ and for our purposes $1<q<2$ and $\alpha>0$.

When $q = 1$, the q – Weibull distribution includes the standard Weibull distribution, and for $\alpha = 1$, it is equivalent to an exponential distribution. Similar specifications of models and distributions are used by Gurgul and Syrek (2016).

The series of trading volumes have similar characteristics to the duration series. For this reason, trading volume series can be modeled with ACD-type models. Following Manganelli (2005), we call these models ACV- autoregressive conditional volume models.

We now turn our attention to the contemporaneous dependence between modeled variables.

The analysis of dependence can be performed with different tools. In our research, we use quantile dependence and copulas. The strength of dependence measured by "quantile dependence" (in the joint lower or upper tails) is defined as:

$$\lambda^{q}=\begin{cases}P(U_{1t}\leq q|U_{2t}\leq q), & 0<q\leq 0.5\\ P(U_{1t}>q|U_{2t}>q), & 0.5<q\leq 1\end{cases}$$

where U_{1t} and U_{2t} are probability integral transforms. The estimators of quantile dependence are as follows

$$\hat{\lambda}^{q}=\begin{cases}\frac{1}{Nq}\sum_{t=1}^{N}1\{U_{1t}\leq q\,|\,U_{2t}\leq q\}, & 0<q\leq 0.5\\ \frac{1}{N(1-q)}\sum_{t=1}^{N}1\{U_{1t}>q|U_{2t}>q\}, & 0.5<q\leq 1\end{cases}$$

Using a quantiles dependence function, it is possible to test for (under the null) symmetric dependence (Patton, 2012) $\lambda^q = \lambda^{1-q}$ for every $q \in [0,1]$. To perform the test, the estimated quantile dependence measures are stacked in vector $\hat{\lambda}$ with $q_{k=j} = 1 - q_j$, for $j = 1, 2, ..., k$. The test is

$$H_0 : R\lambda = 0$$

against

$$H_1 : R\lambda \neq 0$$

with $R \equiv [k \vdots - I_k]$.

The test statistics proposed by Rémillard (2010) is based on bootstrap. Under the null, we have

$$N(\hat{\lambda} - \lambda)' R' (RV_{\lambda,S} R')^{-1} R(\hat{\lambda} - \lambda) \xrightarrow{d} \chi^2_k$$

where $V_{\lambda,S}$ denotes the bootstrap estimate of V_λ (for more details, see Patton (2013)).

Copulas are multivariate distributions with uniform margins. Sklar's theorem states that every multivariate distribution can be decomposed into two parts: marginal distributions and copulas that describe the dependence structure. There are many functional forms of copulas (see Nelsen, 1999). The usfulness of copulas comes from the disadvanages of the classic measure of dependence; i.e., Pearson's correlation coefficient. This measure is appropriate only in the case of elliptical distributions and measures only linear dependence (which is rather rare in real-world applications). One of the alternativies is to use concordance measures; for example, Kendall's coefficient, which is the probabability of concordance minus the probability of disconcordance, and can be expressed as

$$\tau = 4\int_0^1\int_0^1 C(u_1,u_2)\, dC(u_1,u_2) - 1$$

Kendall's τ coefficient is invariant under strictly increasing transformations, and this is not true in general for a linear correlation coefficient.

Obtaining the limit of (population) quantile dependence, we get measures of the dependence between extreme events; that is, tail-dependence coefficients. Formally for any copula C, we have

$$\tau^L = \lim_{q \to 0^+} \frac{C(q,q)}{q}$$

$$\tau^U = \lim_{q \to 1^-} \frac{1 - 2q + C(q,q)}{1-q}$$

The specific copulas exhibit the degree of tail dependency. For example, a normal copula and Frank copula exhibit tail independence, whereas a t copula exhibits symmetric dependence. A Gumbel copula describes upper tail dependence and lower tail independence, whereas a Clayton copula describes the opposite pattern of dependence.

4. Empirical results

Our contribution is based on the tick-by-tick transactions of some DAX30 companies. The dataset includes the prices of companies from 2013-08-08 to 2013-09-24 (33 trading days). First, we calculate price durations whose threshold equals a 10-tick size. The overnight durations and durations corresponding to events recorded outside regular opening hours (9:00 to 17:30) are removed. We sum the number of shares traded within each price duration (trading volume, hereafter).

Table 1 shows the order statistics of the main descriptive statistics of plain-price durations and trading volumes. In addition, we include the values of Ljung-Box test statistics with 15 lags.

Table 1

The order statistics of plain-price durations and trading volumes (number of observations [*N*], mean, standard deviation, minimum, quantiles, maximum, and Ljung-Box test statistic)

Price durations					
statistics	**min**	**0.25q**	**median**	**0.75q**	**max**
N	1077	2509	3294	6361	10095
Mean	102.76	162.46	312.65	413.02	888.49
S.D.	140.43	241.79	434.57	524.06	1374.20
min	1	1	1	1	1
0.25q	21.0	27.0	52.0	80.8	133.0
Median	56.0	77.3	164.5	243.8	414.0
0.75q	129.0	198.5	392.0	539.5	1134.0
Max	1858.0	3548.8	5788.5	7763.5	18606.0
LB(15)	250.255	627.168	1312.157	2930.819	5816.247

Table 1 cont.

Trading volume					
statistics	**min**	**0.25q**	**median**	**0.75q**	**max**
N	1077	2509	3294	6361	10095
Mean	2351.86	8143.40	17903.64	56331.64	104757.02
S.D.	5689.24	14761.75	39404.04	82783.48	212405.76
min	1	3	21.5	42	880
0.25q	659	2301.625	4977	14682	36362
Median	1377	4949.75	10710.75	33119	70740
0.75q	2795.5	9782.375	21270	69070.06	131853
Max	273766	673582	1723156	2406367	8396198
LB(15)	12.84	32.64	65.00	119.86	1226.44

In both types of series, the results are in line with stylized facts about duration data. In the time series under study, both overdispersion and autocorrelation are shown. The series of price durations show a diurnal pattern. Many authors have noticed the intraday seasonality in a duration series called diurnality. Similar to Bauwens and Giot (2000) and Vuorenmaa (2009), we apply cubic splines to discover diurnal patterns. We set the nodes every 60 minutes. Two additional nodes are set ten minutes after the opening and ten minutes before close (in the case of non-positive adjusted durations, we introduced some modifications of node positions). In Table 2, we present the descriptive statistics of diurnally adjusted price durations and trading volumes (plain durations series divided by seasonal component).

Table 2

The order statistics of adjusted price durations and trading volumes (number of observations [*N*], mean, standard deviation, minimum, quantiles, maximum, and Ljung-Box test statistic)

Price durations					
Statistics	**Min**	**0.25q**	**Median**	**0.75q**	**Max**
N	1077	2509	3294	6361	10095
Mean	1.020	1.033	1.039	1.050	1.103

Table 2 cont.

Price durations					
statistics	**min**	**0.25q**	**median**	**0.75q**	**max**
S.D.	1.094	1.206	1.245	1.323	1.476
min	0.001	0.002	0.002	0.003	0.005
0.25q	0.182	0.241	0.255	0.268	0.303
Median	0.521	0.619	0.630	0.675	0.704
0.75q	1.283	1.338	1.358	1.385	1.426
Max	9.813	12.087	17.104	19.429	26.497
LB(15)	170.032	349.404	600.670	1346.109	3056.831
Trading volume					
statistics	**min**	**0.25q**	**median**	**0.75q**	**max**
N	1077	2509	3294	6361	10095
Mean	1.012	1.030	1.040	1.046	1.119
S.D.	1.043	1.145	1.241	1.419	1.714
min	0.001	0.001	0.001	0.004	0.017
0.25q	0.273	0.315	0.331	0.361	0.397
Median	0.591	0.651	0.674	0.716	0.759
0.75q	1.216	1.293	1.320	1.349	1.386
Max	11.827	17.548	23.852	39.406	76.195
LB(15)	36.980	78.002	268.750	378.102	1016.003

By construction, the mean of the adjusted series should be close to 1. The price durations have the well-known inverted U-shape type pattern for all days of the week, so we observe increasing activity at the beginning and end of the session. With the trading volumes series, we found instead an inverted V-shape pattern, but only on Friday (the peak is between 13:00 and 15:00). In Figures 1 and 2, we present typical shapes of diurnal patterns.

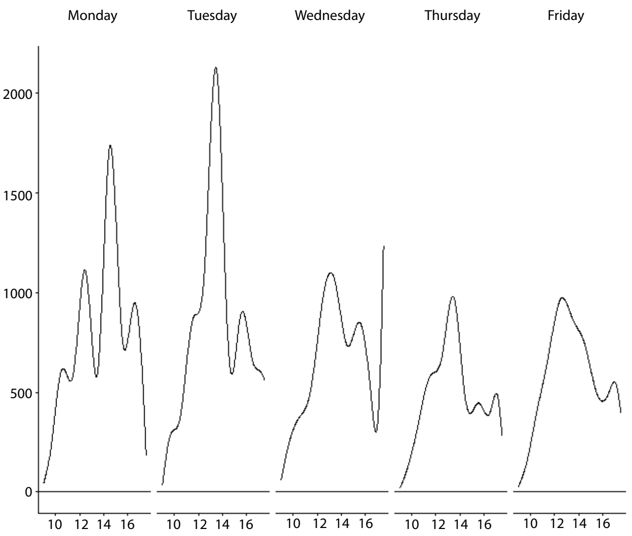

Figure 1. Diurnal pattern of price durations (Adidas)

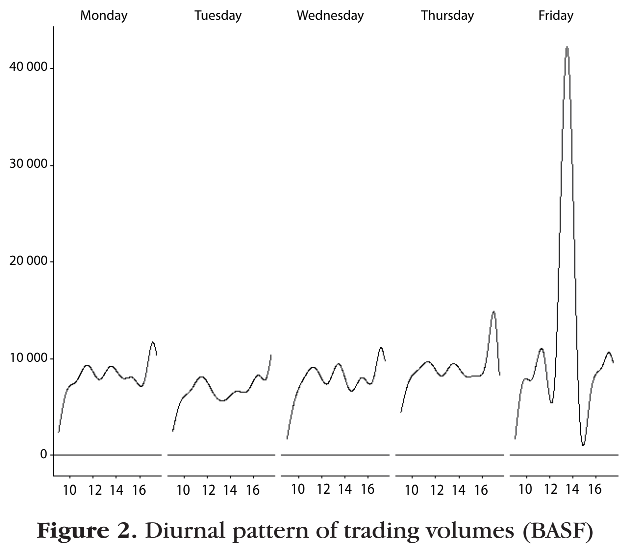

Figure 2. Diurnal pattern of trading volumes (BASF)

The autocorrelation of price duration series is now reduced (but not eliminated). In the case of trading volume series, there is no such reduction (surprisingly, we noticed a rise in some autocorrelation coefficients for some lags).

The estimation of the parameters of ACD and ADV models is carried out by the maximum likelihood method. We fitted different models combining conditional mean and distribution function. It turned out that the restriction $P = Q = 1$ is sufficient to describe the series characteristics mentioned above. The selection of models that best fit a given company is done by the Bayesian Information Criterion. We restricted our attention to the models that describe autocorrelation properly, and have uniformly distributed probability integral transforms (Diebold et. al. 1998). We checked this with the Ljung-Box test (applied to residuals and their squares) and the Anderson-Darling test (in testing for the uniformity of probability integral transform). In Table 3, we present the results of the estimation of conditional duration models.

For all of the series under study, the mean and standard deviation of residuals properly reflects characteristics from the descriptive statistics of adjusted durations. In most cases, the model that fits best is ; only in three cases is linear parametrization better. The sum of parameters α and β of the ACD model reflects the stationarity of the duration process, but the large value (at least 0.95) of this sum confirms the stylized fact of clustering of the durations. The same conclusions apply to the logarithmic model. Regarding the conditional distribution of residuals, there is no outstanding distribution. The test of parameter significance indicates a strong rejection of exponential and Weibull distributions.

Similar conclusions are drawn from the estimation results for trading volume series (Tab. 4). In four cases, linear parametrization fits better than logarithmic, and the small *p*-values in significance parameter testing reject exponential and Weibull distributions.

To obtain information about dependence structure st, we simply apply the sample Kendall's correlation coefficient for standardized residuals of price durations (p_t) and volumes (v_t). In Table 5, we also present the computational results for lagged variables. Numbers in bold indicate significance at a 5% level.

The dependence measured by the correlation coefficient is strong and significant only for contemporaneous variables. For the pair price duration – lagged trading volume, the dependence is significant but very weak in most cases. The results for the third pair indicate that the variables are uncorrelated. Next, we estimate quantile dependence and perform the test of symmetry (having in mind that this is only a test of the necessary condition for equality). Only in the case of companies Allianz, EON, Muenchner Rueck, and Thyssen do we fail to reject the null of symmetric dependence. The *p*-values based on 500 bootstrap samples equals at least 0.10. This concerns only contemporaneous dependence. If at least one variable is lagged, we fail the null for all companies. Figure 3 presents typical plot of price durations and associated trading volumes (transformed with estimated conditional distributions).

Table 3

Models for price durations

Company	Model	Dist.	LB(15)	A-D	ε	S_ε	ω	α	β	μ_1	μ_2
Adidas	ACD	G-G	0.27	0.48	1.00	1.15	0.04	0.13	0.84	1.73	0.69
Allianz	LACD1	Burr	0.81	0.63	1.02	1.11	0.08	0.17	0.82	1.22	0.23
BASF	LACD1	G-G	0.07	0.55	1.00	1.15	0.07	0.14	0.93	2.74	0.57
Beiersdorf	LACD1	q-Weibull	0.08	0.59	1.01	1.13	0.07	0.11	0.85	1.03	1.11
BMW	ACD	G-G	0.05	0.48	1.00	1.14	0.04	0.14	0.83	1.88	0.67
Commerzbank	LACD1	G-G	0.19	0.51	1.01	1.45	0.13	0.20	0.89	3.52	0.43
Deutsche_Lufthansa	ACD	G-G	0.08	0.80	1.02	1.10	0.06	0.17	0.78	2.05	0.66
EON	LACD1	Burr	0.78	0.51	1.02	1.28	0.07	0.13	0.90	1.04	0.17
Muenchner Rueck	LACD1	G-G	0.48	0.70	1.02	1.04	0.06	0.13	0.95	1.93	0.72
RWE	LACD1	Burr	0.10	0.77	1.01	1.17	0.11	0.18	0.72	1.16	0.25
SAP	LACD1	q-Weibull	0.27	0.54	1.01	1.25	0.08	0.15	0.92	1.20	1.25
Thyssen	LACD1	q-Weibull	0.10	0.33	1.08	1.66	0.20	0.34	0.84	1.07	1.23
VW	LACD1	q-Weibull	0.07	0.88	1.01	1.27	0.10	0.17	0.96	1.16	1.26

LB(15) denotes the value of the Ljung-Box test statistics applied to residuals. In A-D column is p-value in GOF testing, parameters μ_1 and μ_2 refer to κ and γ for generalized gamma distribution, κ and σ^2 for Burr distribution, and a and q for q-Weibull distribution, respectively; ε denotes mean of residuals, whereas S_ε is standard deviation of residuals

Table 4

Models of trading volumes

Company	Model	Dist.	LB(15)	A-D	ε	S_ε	ω	α	β	μ_1	μ_2
Adidas	LACD1	q-Weibull	0.90	0.31	1.00	1.85	0.07	0.11	0.89	1.34	1.35
Allianz	LACD1	G-G	0.84	0.95	1.01	1.02	0.05	0.12	0.68	8.75	0.38
BASF	LACD1	G-G	0.98	0.37	1.01	2.02	0.05	0.11	0.92	7.67	0.36
Beiersdorf	LACD1	Burr	0.75	0.28	1.00	1.30	0.05	0.09	0.88	1.42	0.50
BMW	LACD1	Burr	0.95	0.07	0.99	1.31	0.06	0.11	0.91	1.35	0.48
Commerzbank	LACD1	G-G	0.10	0.76	1.00	1.18	0.05	0.10	0.96	6.83	0.36
Deutsche_Lufthansa	ACD	q-Weibull	0.41	0.56	0.99	1.00	0.04	0.08	0.89	1.53	1.32
EON	LACD1	Burr	0.07	0.20	1.00	1.19	0.06	0.13	0.86	1.42	0.50
Muenchner Rueck	LACD1	q-Weibull	0.70	0.85	1.00	1.01	0.06	0.10	0.60	1.54	1.30
RWE	ACD	q-Weibull	0.80	0.96	1.00	1.22	0.33	0.16	0.54	1.44	1.33
SAP	LACD1	Burr	0.66	0.07	1.01	1.64	0.06	0.12	0.87	1.40	0.48
Thyssen	ACD	G-G	0.20	0.55	1.03	1.00	0.20	0.13	0.68	7.24	0.42
VW	ACD	q-Weibull	0.41	0.56	0.99	1.00	0.04	0.08	0.89	1.53	1.32

LB(15) denotes the value of the Ljung-Box test statistics applied to residuals. A-D shows p-values in GOF testing, parameters μ_1 and μ_2 refer to κ and γ for generalized gamma distribution, κ and σ^2 for Burr distribution, and a and q for q-Weibull distribution, respectively; ε denotes mean of residuals, whereas S_ε is standard deviation of residuals

Table 5
Sample Kendall's correlation coefficients

Company	$p_t - v_t$	$p_t - v_{t-1}$	$p_{t-1} - v_t$
Adidas	0.40	–0.03	0.02
Allianz	**0.57**	**–0.04**	0.02
BASF	**0.44**	**–0.04**	–0.01
Beiersdorf	**0.37**	**–0.04**	–0.01
BMW	0.41	–0.03	–0,02
Commerzbank	0.48	–0.05	0,05
Deutsche_Lufthansa	**0.41**	**–0.03**	0.00
EON	0.52	–0.05	–0.03
Muenchner Rueck	**0.49**	**–0.04**	–0.03
RWE	**0.41**	**–0.04**	–0.01
SAP	**0.47**	–0.01	0.00
Thyssen	0.48	–0.05	0.07
VW	**0.46**	**–0.07**	0.01

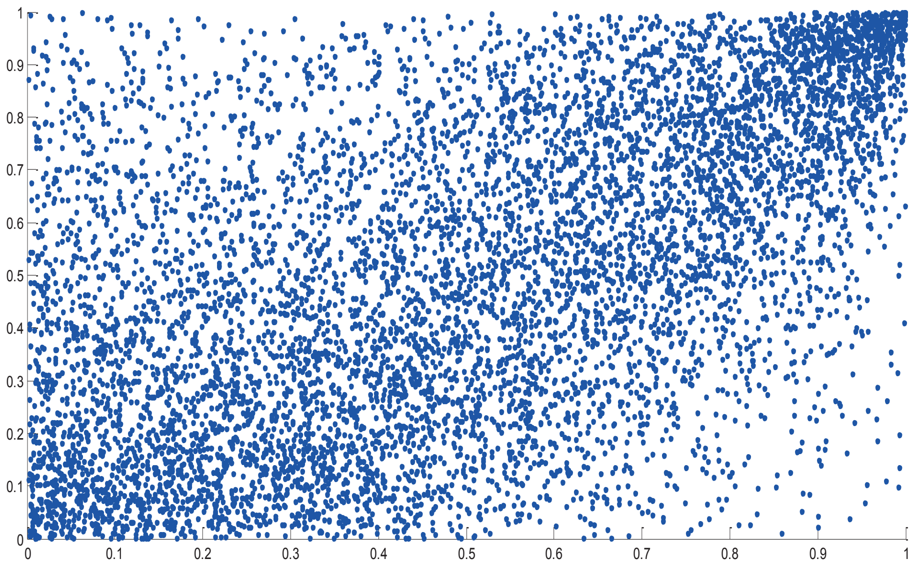

Figure 3. Price durations and trading volumes

We observe the concentration of points is in the upper-right corner, (that is, for simultaneously large values of both series). To gain more insight into the dependence structure, we use copula functions. Given the PIT series obtained from the ACD and ACV models, we use a maximum-likelihood method to estimate the parameters of copula functions. This one is the IFM method of Joe and Xu (1996). To select the copulas that fit best, we use the BIC criterion. The results of parameter estimation are in Table 6. In addition, we present dependence measures based upon the copula selected.

Table 6

Copula estimation results and dependence measures

Company	**Copula**	τ	τ^L	τ^U
Adidas	Gumbel	0.37	0.00	0.46
Allianz	*t*	0.55	0.31	0.31
BASF	Gumbel	0.41	0.00	0.50
Beiersdorf	Gumbel	0.34	0.00	0.42
BMW	Gumbel	0.38	0.00	0.47
Commerzbank	Gumbel	0.46	0.00	0.54
Deutsche_Lufthansa	Gumbel	0.39	0.00	0.47
EON	*t*	0.50	0.25	0.25
Muenchner Rueck	*t*	0.47	0.25	0.25
RWE	Gumbel	0.38	0.00	0.46
SAP	Gumbel	0.45	0.00	0.53
Thyssen	*t*	0.46	0.11	0.11
VW	Gumbel	0.44	0.00	0.53

Contemporaneous price durations and associated trading volumes are dependent (as can be seen from the values of the Kendall correlation coefficient). In addition, the Gumbel copula fits best for most cases that exhibit positive upper tail dependence and lower tail independence. The coefficient of upper tail dependence equals at least 0.46 (for Adidas and RWE). In the remaining cases, elliptical copula *t* fits the data best. These results are in line with results from testing for symmetry using quantile-dependence measures. In these cases, the lower and upper tail dependence coefficients are equal and relatively low as compared to a nonsymmetrical copula. For the case of the lagged variable, the independence copula fits the best.

5. Conclusions

In this paper, we show the usefulness of the copula function in the description of the dependence structure of specific unevenly spaced time series. The behavior of the time series of price durations and trading volumes under study are in line with common observation from other empirical findings. We observe clustering, overdispersion, and diurnality. In most cases, the seminal model (linear parametrization with exponential or Weibull distribution) is displaced by a logarithmic specification with more-flexible conditional distributions. The price duration and trading volume associated with this duration are dependent in the tails of distribution. We may conclude that high cumulative volumes are associated with long durations, but also that dependence between short durations and low cumulative volumes can be observed. This is concerned with contemporaneous variables. For the case where one of the variables is lagged, we conclude that the dependence (if any) is very weak.

References

[1] Admati, A.R. and Peiderer, P. (1988) 'A theory of intraday patterns: Volume and price variability', *The Review of Financial Studies*, vol. 1(1), pp. 3–40.

[2] Alfonsi, A. and Schied, A. (2010) 'Optimal trade execution and absence of price manipulations in limit order book models', *SIAM Journal on Financial Mathematics*, vol. 1, pp. 490–522.

[3] Almgren, R. (2003) 'Optimal execution with nonlinear impact functions and trading-enhanced risk', *Applied Mathematical Finance*, vol. 10(1), pp. 1–18.

[4] Almgren, R. (2009) 'Optimal trading in a dynamic market', Working Paper, New York University.

[5] Asmussen, S. (2003) 'Applied probability and queues', 2nd ed., Berlin: Springer.

[6] Avellaneda, M. and Stoikov, S. (2008) 'High-frequency trading in a limit order book', *Quantitative Finance*, vol. 8, pp. 217–224.

[7] Baum, L., Petrie, T., Soules, G. and Weiss, N. (1970) 'A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains', *The Annals of Mathematical Statistics*, vol. 41(1), pp. 164–171.

[8] Bauwens, L. and 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*, vol. 60, pp. 117–149.

[9] Bauwens, L. and Hautsch, N. (2009) 'Modelling financial high frequency data using point processes', in Mikosch, T., Kreiß, J.P., Davis R.A., and Andersen, T.G. (eds) Handbook of Financial Time Series, Berlin: Springer, pp. 953–979.

[10] Bayraktar, E. and Ludkovski, M. (2011) 'Liquidation in limit order books with controlled intensity', Working Paper, University of Michigan and UCSB.
[11] Biernacki, C., Celeux, G. and Govaert, G. (2001) 'Assessing a mixture model for clustering with the integrated completed likelihood', IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22(7), pp. 719–725.
[12] Bouchard, B., Dang, N.M. and Lehalle, C.A. (2011) 'Optimal control of trading algorithms: A general impulse control approach', *SIAM Journal on Financial Mathematics*, vol. 2, pp. 404–438.
[13] Cappé, O., Moulines, E. and Rydén, T. (2005) Inference in hidden Markov models, Berlin: Springer.
[14] Cartea, Á., Jaimungal. S. and Ricci, J. (2011) 'Buy low sell high: a high frequency trading perspective', SSRN eLibrary, [Online], Available: http://ssrn.com/abstract=1964781.
[15] Cartea, Á. and Jaimungal, S. (2012) 'Risk metrics and fine tuning of high frequency trading strategies', *Mathematical Finance*, [Online], Available: http://dx.doi.org/10.1111/mafi.12023.
[16] Cartea, Á. and Jaimungal, S. (2013) 'Modelling Asset Prices for Algorithmic and High-Frequency Trading', *Applied Mathematical Finance*, vol. 20 (6), pp. 512–547.
[17] Cartea, Á. and Meyer-Brandis, T. (2010) 'How duration between trades of underlying securities affects option prices', *Review of Finance*, vol. 14(4), pp. 749–785.
[18] Cartea, Á. and Penalva, J. (2012) 'Where is the value in high frequency trading?', *Quarterly Journal of Finance*, vol. 2(3), pp. 1–46.
[19] Celeux, G. and Durand, J.B. (2008) 'Selecting hidden Markov model state number with cross-validated likelihood', *Computational Statistics*, vol. 23(4), pp. 541–564.
[20] Cvitanic, J. and Kirilenko, A.A. (2010) 'High frequency traders and asset prices', SSRN eLibrary, [Online], Available: http://ssrn.com/abstract=1569067.
[21] Diamond, D.W. and Verrechia, R.E. (1987) 'Constraints on short-selling and asset price adjustment to private information', *Journal of Financial Economics*, vol. 18, pp. 277–311.
[22] Diebold, F.X., Gunther, T.A. and Tay, A.S. (1998) 'Evaluating density forecasts with applications to financial risk management', *International Economic Review*, vol. 39, pp. 863–883.
[23] Dufour, A. and Engle, R.F. (2000) 'Time and the price impact of a trade', *The Journal of Finance*, vol. LV(6), pp. 2467–2498.
[24] Easley, D. and O'Hara, M. (1992) 'Time and the process of security price adjustment', *The Journal of Finance*, vol. XLVII(2), pp. 577–605.
[25] Engle, R.F. (2000) 'The econometrics of ultra-high-frequency data', *Econometrica*, vol. 68(1), pp. 1–22.

[26] Engle, R.F. and Russell, J.R. (1998) 'Autoregressive conditional duration: A new model for irregularly spaced transaction data', *Econometrica*, vol. 66(5), pp. 1127–1162.

[27] Fernandes, M. and Grammig, J. (2005) 'Nonparametric specification tests for conditional duration models', *Journal of Econometrics*, vol. 127(1), pp. 35–68.

[28] Ghysels, E., Gourieroux C. and Jasiak, J. (2004), 'Stochastic Volatility Duration Models', *Journal of Econometrics*, vol. 119(2), pp. 413–433.

[29] Gourieroux, C. and Jasiak, J., 2001, Financial Econometrics: Problems, Models and Methods, New Jersey: Princeton University Press.

[30] Gramming, J. and Maurer, K.O. (2000) 'Non-Monotonic Hazard Functions and the Autoregressive Conditional Duration Model', *The Econometrics Journal*, vol. 3, pp. 16–38.

[31] Gurgul, H. and Syrek, R. (2016) 'The logarithmic ACD model: The microstructure of the German and Polish stock markets', *Managerial Economics*, vol. 17(1), pp. 77–92.

[32] Hujer, R., Vuletic, S. and Kokot, S. (2002) 'The Markov switching ACD model', SSRN eLibrary, [Online], Available: http://ssrn.com/abstract=332381.

[33] Jaimungal, S. and Kinzebulatov, D. (2012) 'Optimal execution with a price limiter', SSRN eLibrary, [Online], Available: http://ssrn.com/abstract=2199889.

[34] de Jong, F. and Rindi, B. (2009) 'The microstructure of financial markets', 1st ed., Cambridge: Cambridge University Press.

[35] Kharroubi, I. and Pham, H. (2010) 'Optimal portfolio liquidation with execution cost and risk', *SIAM Journal on Financial Mathematics*, vol. 1, pp. 897–931.

[36] Latza, T., Marsh, I. and Payne, R. (2012) 'Computer-based trading in the cross-section', Working Paper, Cass Business School.

[37] Lorenz, J. and Almgren, R. (2011) 'Meanvariance optimal adaptive execution', *Applied Mathematical Finance*, vol. 18, pp. 395–422.

[38] Lunde, A. (2000) 'A Generalized Gamma Autoregressive Conditional Duration Model', Discussion paper, Aarlborg University.

[39] Maheu, J.M. and McCurdy, T.H. (2000) 'Volatility dynamics under duration-dependent mixing', *Journal of Empirical Finance*, vol. 7(3–4), pp. 345–372.

[40] Manganelli, S. (2005) 'Duration, volume and volatility impact of trades', *Journal of Financial Markets*, vol. 8(4), pp. 377–399.

[41] Meitz, M. and Terasvirta, T. (2006) 'Evaluating models of autoregressive conditional duration', *Journal of Business & Economic Statistics*, vol. 24, pp. 104–124.

[42] Mongillo, G. and Deneve, S. (2008) 'Online learning with hidden Markov models', *Neural Computation*, vol. 20(7), pp. 1706–1716.

[43] Rémillard, B. (2010) 'Goodness-of-fit tests for copulas of multivariate time series', Working paper, HEC Montreal.

[44] Renault, E., van der Heijden, T. and Werker, B.J.M. (2012) 'The dynamic mixed hitting-time model for multiple transaction prices and times', Working Paper, [Online] Available: http://dx.doi.org/10.2139/ssrn.2146220.
[45] Patton, A. (2013) 'Copula Methods for Forecasting Multivariate Time Series, in Elliott, G. and Timmermann, A. (eds) Handbook of Economic Forecasting, vol. 2, London: Springer Verlag.
[46] SEC (2010) 'Concept release on equity market structure', Concept Release No. 34–61358, File No. S7-02-10, SEC. 17 CFR PART 242.
[47] Viterbi, A. (1967) 'Error bounds for convolutional codes and an asymptotically optimum decoding algorithm', IEEE Transactions on Information Theory, vol. 13(2), pp. 260–269.
[48] Vuroenmaa, T.A. (2009) 'A q-Weibull Autoregressive Conditional Duration Model with an application to NYSE and HSE Data', [Online], Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1952550.
[49] Zhang, M.Y., Russell, J.R. and Tsay, R.S. (2001) 'A nonlinear autoregressive conditional duration model with applications to financial transaction data', *Journal of Econometrics*, vol. 104(1), pp. 179–207.