Centre for Computational Finance and Economic Agents

Working Paper Series

WP012-07

Wing Lon Ng

Analyzing Liquidity and Absorption Limits of Electronic Markets with Volume Durations

October 2007

University of Essex

CCFEA

www.ccfea.net

Analyzing Liquidity and Absorption Limits of Electronic Markets with Volume Durations

Wing Lon Ng
(wlng@essex.ac.uk) [1]
University of Essex - CCFEA

[1] The author is very grateful to Mark Trede for his valuable and insightful comments. He is also grateful to the Editor and two anonymous referees for their suggestions that led to the improvement of this paper. He is particularly indebted to conference participants at the Asian Financial Association (2005 Annual Meeting), the European Meeting of Statisticians (EMS 2005) and the International Workshop on Quantitative Finance (2006) for the extensive discussions. Financial support from the Institute of Econometrics and Economic Statistics (IÖW), the German Research Foundation (DFG) and the EU (Marie Curie) is also gratefully acknowledged.

Abstract

This paper focuses on the liquidity of electronic stock markets applying a sequential estimation approach of models for volume durations with increasing threshold values. A modified ACD model with a Box-Tukey transformation and a flexible generalized Beta distribution is proposed to capture the changing cluster structure of duration processes. The estimation results with the German XETRA data reveal the market's absorption limit for high volumes of shares, expanding the time costs of illiquidity when trading these quantities.

Key Words: Ultra high frequency transaction data, limit order book, volume duration, autoregressive conditional duration, Box-Tukey transformation, temporal aggregation.

JEL Classifications: $C22, C41, G12$.

1 Introduction

The literature on financial econometrics and quantitative finance has been used to focus on the stochastic process of daily prices or returns on assets and their volatility. Generally, most studies only pays little attention to other essential variables in financial markets and, thus, lose sight of the transaction volume – the second relevant factor when trading assets. It is often overlooked that the investigation of the trading quantity is also important, because it is one of the key factors responsible for market efficiency.

For decades, the volume was studied only in relationship with returns and volatility, serving as a explanatory variable for the analysis of the stochastic process. The literature goes back to Clark (1973), who introduced the use of subordinators in finance. In his model the daily price change represents a sum of within-day price changes. Since the trading volume is positively related to number of transactions, it helps explaining the variability of the price change. In another model proposed by Epps and Epps (1976) the change of the transaction price in the market is the mean of the changes of all trader's reservation price, and they found a positive relationship between the absolute of the change in the market price and the trading volume. A generalization of these models are derived in Tauchen and Pitts (1983). Likewise, Gallant, Rossi, and Tauchen (1992) also found a positive correlation between volatility and volume, applying a semi-nonparametric method to estimate the joint density of price change and volume. A survey can be found in Karpoff (1987) and Jones, Kaul, and L. (1994). With the results of Anè and Geman (2000), the volume variable has become less important again (as in Jones, Kaul, and L. (1994)), because in their study with the substitution of subordinators with stochastic time changes, the number of trades represents a better stochastic clock than the volume for inducing normal returns.

In contrast to this wide stream of literature, this paper analyzes the time-varying liquidity of electronic markets by focusing on the relationship between time and volume. Liquidity has been approved as an important factor characterizing the market efficiency. The concept of liquidity has various definitions. Usually, it has three dimensions and is understood as the capability to trade (a) a large amount of shares (b) in a short time (c) at low cost. For example, Engle and Lange (2001) introduced the VNET measure that captures the *net* directional one-sided volume over the *price duration*, that is, the amount of excess volume that can be sustained before prices are adjusted. However, only regarding the absolute one-sided excess quantity, this approach does not take the total traded volume into account, which is essential when investigating the "entire" absorptive ability of the market. Hence, in this paper, a market is considered as liquid if unlimited volumes can be traded immediately, to exploit the overall absorption limits independent of the price impact. An adequate measurement is based on volume durations that considers the time and volume dimension of the trading process. It is defined as the time elapsed until a certain quantity of shares is absorbed by the market (Bauwens and Giot (2001)) and, thus, provide a good indicator for the time costs of (il)liquidity (Hautsch (2004)).

The analysis of volume durations is important especially for liquidity traders. Short volume durations imply that threshold amount of stocks can be traded very quickly, whereas long ones signalize a (temporarily) illiquid market. Since the empirical cluster structures of volume durations indicate different degrees of market liquidity, they become important decision factors influencing traders' order-placement strategies. Thus, modeling volume durations can help finding the absorption limits of the market, saving the (time) costs of liquidity for traders.

Therefore, the main objective of this paper is to study the dynamics of volume durations by applying the ACD framework, an appropriate econometric tool for analyzing financial duration data (for a survey, see Bauwens, Giot, Grammig, and Veredas (2004) or Engle and Russell (2004)). With the increased availability of ultra-high-frequency transaction data in the last few years, many researchers have pushed the further development of the ACD model in order to describe limit order book activities more accurately. Recently, Fernandes and Grammig (2006) have introduced an augmented version that encompasses all specifications discussed in the literature by adopting the results of Hentschel (1995) into the ACD methodology. However, since the asymmetry effects of volume durations are assumed to be marginal, this paper concentrates on a Box-Tukey specification that is parsimonious and flexible as well (Box and Cox (1964)). Similar to Fernandes and Grammig (2006), the modified ACD model with this transformation is also able to induce various non-monotonic shapes of the news-impact curve to tailor the data more precisely. Interestingly, since the issue addressed here concerns the market's liquidity given an predetermined cumulated volume, the shifting parameters in the Box-Tukey specification can be used to identify the absorption limit.

Furthermore, a generalized Beta distribution is proposed that nests more than 30 distributions as limiting or special cases, including those already discussed in the ACD literature. It can be shown that the generalized Beta distribution allows for hazard functions with diverse slopes implying different duration dependence of the data. As extension to several studies in this research field (see, for example, Hautsch (2002) and Hautsch (2003)), investigates the entire evolution of volume duration processes by successively increasing the threshold value, resulting in a sequential estimation approach considering "temporal" aggregation effects.

The paper is structured as follows: In Section 2, the Box-Tukey-ACD model and concept of volume durations will be introduced. Section 3 describes the smoothing technique for deseasonalization and deals with the ML estimation. In Section 4, the data and results are presented. Section 5 concludes.

2 Methodology

Since ultra-high-frequency transaction data arrive at irregular time intervals, researchers must not only consider the key variables themselves (price, return, volume, etc.), but also their stochastic arrival times (Hafner (2005)). Hence, a duration analysis is the common approach to describe these time stamped data. Let $X_i = t_i - t_{i-1}$ (with $t_0 = 0$), where X_i is the i^{th} trade duration between the i^{th} and $(i-1)^{th}$ transaction. Further, define $\Psi_i \equiv E\left(X_i|\mathcal{F}_{i-1}\right)$, that is, the conditional expectation of the duration, given $\mathcal{F}_{i-1}$ i.e. the information available at t_{i-1}. The common ACD framework, a popular tool in recent financial econometrics, models a dynamic point process in which the conditional expectation is written as a linear function of past durations

$$\Psi_i = \omega + \sum_{j=1}^{p} \alpha_j X_{i-j} + \sum_{k=1}^{q} \beta_k \Psi_{i-k} \tag{1}$$

(see Engle and Russell (1998) and Engle (2000)). It is assumed that

$$X_i = \Psi_i \cdot \varepsilon_i \tag{2}$$

with (i.i.d.-)innovations ε_i. The $\varepsilon_i \equiv \frac{X_i}{\Psi_i}$ are also called standardized durations. Their density function $f\left(\varepsilon_i\right)$ is parametrized with (a) normalization $E\left(\varepsilon_i\right) = 1$ by construction, and (b) a non-negative support to avoid negative durations. Since the assumption of linearity is often too restrictive to capture the duration process, several models have been developed and modified with other dependence structures of the conditional mean in order to account for nonlinear impacts. Generally, new and different types of ACD models can be created by varying the functional form of Ψ_i in the model's mean equation.

Based on the *Box-Cox-ACD*, first introduced by Dufour and Engle (2000), modified and extended by Hautsch (2002), a Box-Tukey transformation of the durations is proposed in this study

$$\underbrace{\frac{\left(\Psi_i + \eta_1\right)^{\kappa_1} - 1}{\kappa_1}}_{\Psi_i^{BT}} = \sum_{j=1}^{p} \alpha_j \underbrace{\left(\frac{\left(\varepsilon_{i-j} + \eta_2\right)^{\kappa_2} - 1}{\kappa_2}\right)}_{\varepsilon_{i-j}^{BT}} + \sum_{k=1}^{q} \beta_k \underbrace{\left(\frac{\left(\Psi_{i-k} + \eta_1\right)^{\kappa_1} - 1}{\kappa_1}\right)}_{\Psi_{i-k}^{BT}}. \tag{3}$$

Here, Ψ_i^{BT} and ε_i^{BT} are the Box-Tukey-transformed durations. Instead of the common intercept ω, two new shifting parameters η_1 and η_2 are introduced to tailor the data and to allow a more flexible adjustment of the news-impact curve which measures the influence of innovations on the conditional

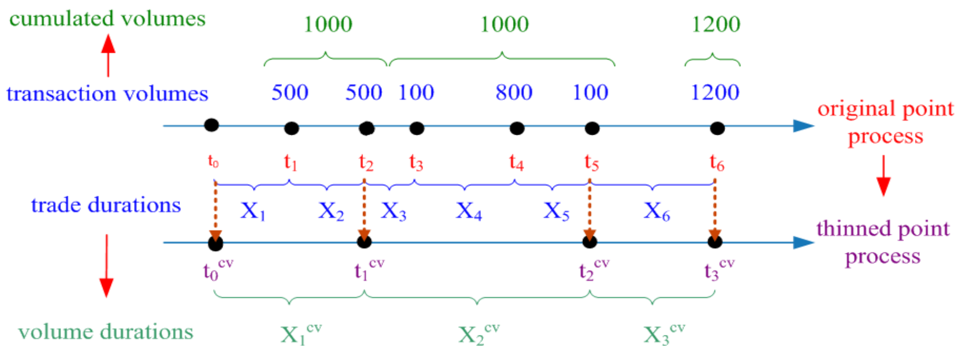

Figure 1: Generating volume durations with $cv = 1000$.

mean. This extended Box-Cox version also provides a accurate description of the conditional mean, although the parametrization is very parsimonious. In contrast to the basic ACD model, it allows additive innovation shocks with $\kappa_1 = \kappa_2 = 1$, as well as nonlinear ones for $\kappa_1, \kappa_2 \neq 1$. In the case of nonlinearity, i.e. $\kappa_1, \kappa_2 \neq 1$, stationarity in both cases is ensured for $\sum_{k=1}^{q} \beta_k < 1$ (see Dufour and Engle (2000) and Hautsch (2002)). Due to the double-shifting of durations in the conditional mean equation, it has similar statistical properties to more sophisticated ACD models focusing on the kink in the innovations' news-impact curves, such as the augmented ACD model of Fernandes and Grammig (2006).

Volume durations are defined as the period elapsed until a given amount cv of shares is traded on the market (Bauwens and Giot (2001)). In contrast to trade durations that stand for the time elapsed between two consecutive transactions, volume durations correspond to the time interval required to observe a certain cumulation of traded shares. In order to model the stochastic process of volume durations, the ACD framework can be simply adapted to the new thinned point process $(t_j^{cv})_{j\in\mathbb{N}}$, where t_j^{cv} now denotes the points each time the market has absorbed (at least) the pre-determined volume cv. As discernible in Figure 1, the resulting volume duration process $(X_j^{cv})_{j\in\mathbb{N}}$ with $X_j^{cv} = t_j^{cv} - t_{j-1}^{cv}$ is generated by cumulating single trade durations until the sum of single transaction volumes vol_i has been achieved or exceeded cv, i.e. $\sum vol_{t_i} \geq cv$, $\forall t_i \in (t_{j-1}^{cv}; t_j^{cv}]$. Subsequently, the model's new filtration is the modified σ-field $\mathcal{F}_j^{cv} = \sigma\left(t_j^{cv}, t_{j-1}^{cv}, ..., t_1^{cv}, t_0^{cv}\right)$.

Note that the single transaction volume vol_i does not represent the relevant random variable, but the duration required to absorb the predetermined cumulated volume cv. In contrast to the literature, where the duration process is analyzed for only one or a few "suitable" threshold values (see, for example, Bauwens and Giot (2001), Hautsch (2002) or Fernandes and Grammig (2006)), this paper focuses on the development of the models' parameters by successively increasing cv_k representing $k \cdot 500$ units. In the following, a sequence of 100 model estimations of successively aggregated sub-datasets is run with $k = 1, 2, ...100$. The first model concentrates on the points signaling that the market has absorbed cv_1, i.e. $1 \cdot 500$ shares, the second one analyzes the duration of trading $cv_2 \hat{=} 2 \cdot 500$ shares, etc. To keep the computational burdens of the 100 estimation procedures within manageable limits and in order to achieve a more reliable comparison of the different model's parameters, this paper concentrates only on a *Box-Tukey-ACD(1,1)* specification.

Taking temporal aggregation effects into account, the formulas of Drost and Nijman (1993) for computing low-frequency GARCH models were applied to get better start values for the estimation (see also Drost and Werker (1996) and Ghysels and Jasiak (1998)). Due to the irregular spacing in time and the existence of extreme long durations for high cv values, the "frequency parameter" m (originally for equidistant data) in their model was replaced by

$$m_k = \frac{\frac{1}{n_k}\sum_{i=1}^{n_k} X_i^{cv_k}}{\frac{1}{n_1}\sum_{i=1}^{n_1} X_i^{cv_1}}$$

to consider the non-constant temporal aggregation of the duration models. n_k stands for the number of observations in the k-th sub-dataset. Having the estimated parameters for the first aggregation level, all following start values for the α and β parameter of the remaining 99 models are obtained by

$$\alpha_{(m_k)} = (\alpha + \beta)^{m_k} - \beta_{(m_k)}$$

and

$$\beta_{(m_k)} = \frac{1}{2const} \pm \sqrt{\frac{1}{4const^2} - 1}$$

with

$$const = \frac{\beta\,(\alpha+\beta)^{m_k - 1}}{1 + \alpha^2 \frac{1-(\alpha+\beta)^{2m_k-2}}{1-(\alpha+\beta)^2} + \beta^2\,(\beta + a)^{2m_k - 2}}$$

and $\beta_{(m_k)} \in (0; 1)$. The start values for the Box-Cox and the Box-Tukey parameters are reset to $\kappa = 0.99999$ and $\eta = 0.00001$ for each estimation, beginning from a quasi-linear specification of the mean equation.

3 Estimation and Diagnostics

It is well-known that intraday data have a consistent diurnal pattern of trading activities over the course of a trading day, due to certain institutional characteristics of organized financial markets, such as opening and closing hours or intraday auctions. Since it is necessary to take the daily deterministic seasonality into account, smoothing techniques are required to get

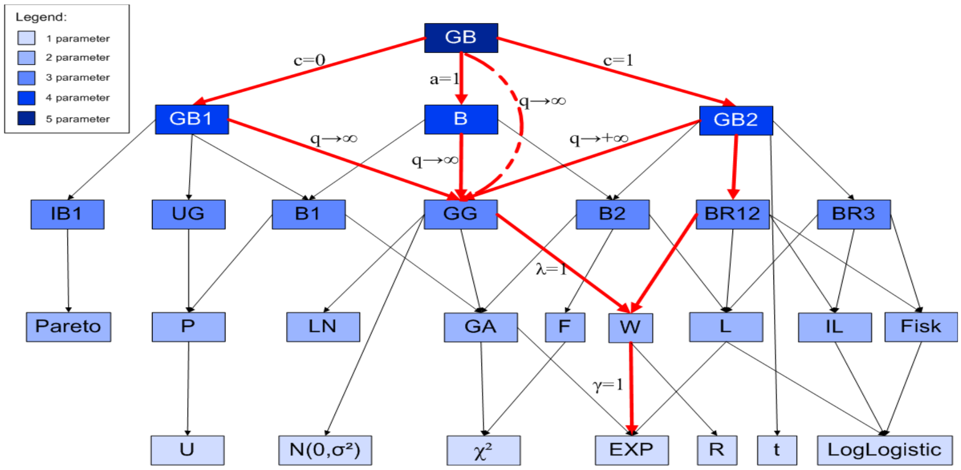

Figure 2: Distribution tree of McDonald and Xu (1995)

deseasonalized observations (Veredas, Rodriguez-Poo, and Espasa (2002)). Let $\tilde{X}_i$ denote the observed duration. Instead of applying cubic splines, where the positions of the nodes have to be carefully set, a kernel regression with an cross-validated bandwidth h_c was performed. Using the Gaussian density as kernel function $K(.)$, the diurnal periodic component (dependent on daytime d) can be computed by the Nadaraya-Watson estimator

$$nw(d_i) = \frac{\sum_{i=1}^{n} \tilde{X}_i \cdot K\left(\frac{d-d_i}{h_c}\right)}{\sum_{i=1}^{n} K\left(\frac{d-d_i}{h_c}\right)} \quad . \tag{4}$$

Thus, $X_i \equiv \frac{\tilde{X}_i}{nw(d_i)}$ is the deseasonalized duration and should have no diurnal pattern and a unit mean.

The ACD model is estimated by maximum-likelihood. Originally, Engle and Russell (1997) used an Exponential (EXP) and a Weibull (W) distribution for the residuals (see Figure 2), whereas other authors favoured more flexible alternatives (see, for example Bauwens, Giot, Grammig, and Veredas (2004)). In this study, the density of the innovations $f(.)$ is modelled by the generalized Beta (GB) distribution proposed by McDonald and Xu (1995)

$$f(\varepsilon_i; a, b, c, p, q) = \left(\frac{|a| \cdot \Gamma(p+q)}{\Gamma(p)\Gamma(q)}\right) \frac{\varepsilon_i^{ap-1}}{b^{ap}} \frac{\left(1-(1-c)\left(\frac{\varepsilon_i}{b}\right)^a\right)^{q-1}}{\left(1+c\left(\frac{\varepsilon_i}{b}\right)^a\right)^{p+q}}$$

with the parameter a controlling the peakedness of the density, the scale parameter b, the smoothing parameter c (with $0 \le c \le 1$) and the parameters p and q influencing the shape and the skewness (for statistical properties, estimation and the calculation of the moments of ε_i, see also Kleiber and

Kotz (2003)). Due to its flexibility the generalized Beta distribution nests all distributions already discussed in the ACD-literature as limiting or special cases, such as the Burr-distributions type III and type XII ($BR12$, see Fernandes and Grammig (2006)), the generalized Gamma (GG, see Lunde (2000)) or the generalized Beta of the second kind ($GB2$). A variation of this distribution is also called "generalized F-distribution" (see Hautsch (2004) and Kalbfleisch and Prentice (2002)). The moments of ε_i can be obtained by

$$E\left(\varepsilon_i^m\right) = b^m \frac{B\left(p + \frac{m}{a}, q\right)}{B\left(p, q\right)} \cdot \left({}_2F_1 \left[\begin{array}{c} p + \frac{m}{a}, \frac{m}{a} \\ p + \frac{m}{a} + q \end{array} ; c \right] \right)$$

with Beta function $B\left(.\right)$ and the Gaussian hypergeometric function

$${}_2F_1 \left[\begin{array}{cc} u_1 & u_2 \\ v_1 & \end{array} ; z \right] = \sum_{n=0}^{\infty} \frac{\left(u_1\right)_n \left(u_2\right)_n}{\left(v_1\right)_n} \frac{z^n}{n!} \quad .$$

Hence, the log-likelihood function of the model with a generalized Beta distribution for the innovations is

$$\begin{aligned} \mathcal{L}\left(x_1, ..., x_n\right) &= n \ln \left(\frac{|a| \cdot \Gamma\left(p + q\right)}{\Gamma\left(p\right) \Gamma\left(q\right)} \right) - apn \ln\left(b\right) \\ &\quad - \left(ap - 1\right) \sum_{i=1}^{n} \ln\left(x_i\right) - ap \sum_{i=1}^{n} \ln\left(\Psi_i\right) \\ &\quad + \left(q - 1\right) \sum_{i=1}^{n} \ln \left(1 - \left(1 - c\right) \left(\frac{x_i}{b\Psi_i} \right)^a \right) \\ &\quad - \left(p + q\right) \sum_{i=1}^{n} \ln \left(1 + c \left(\frac{x_i}{b\Psi_i} \right)^a \right) \end{aligned}$$

with

$$b = \frac{\Gamma\left(p + \frac{1}{a} + q\right) \Gamma\left(p\right)}{\Gamma\left(p + \frac{1}{a}\right) \Gamma\left(p + q\right)} \cdot \left({}_2F_1 \left[\begin{array}{c} p + \frac{1}{a}, \frac{1}{a} \\ p + \frac{1}{a} + q \end{array} ; c \right] \right)^{-1} \quad ,$$

yielding a unit expectation as required.

To check all 100 ACD model's diagnostics, three tests are applied: First, one can examine the properties of the residuals such as their autocorrelation structure which is i.i.d. under correct model specification. Hence, a Ljung-Box-test with modified Portmanteau statistic proposed by Li and Li (2005) is performed for the lag orders 5 and 10. Second, another general test is based on the integration of intensity $h\left(.\right)$ over duration $\left(t_i - t_{i-1}\right)$

$$\Lambda_i = \int_{s=t_{i-1}}^{t_i} h\left(s|\mathcal{F}_s\right) ds \tag{5}$$

that follows a standard exponential distribution under the correct model specification, as illustrated in Russell (1999). Thus, Engle and Russell (2004)

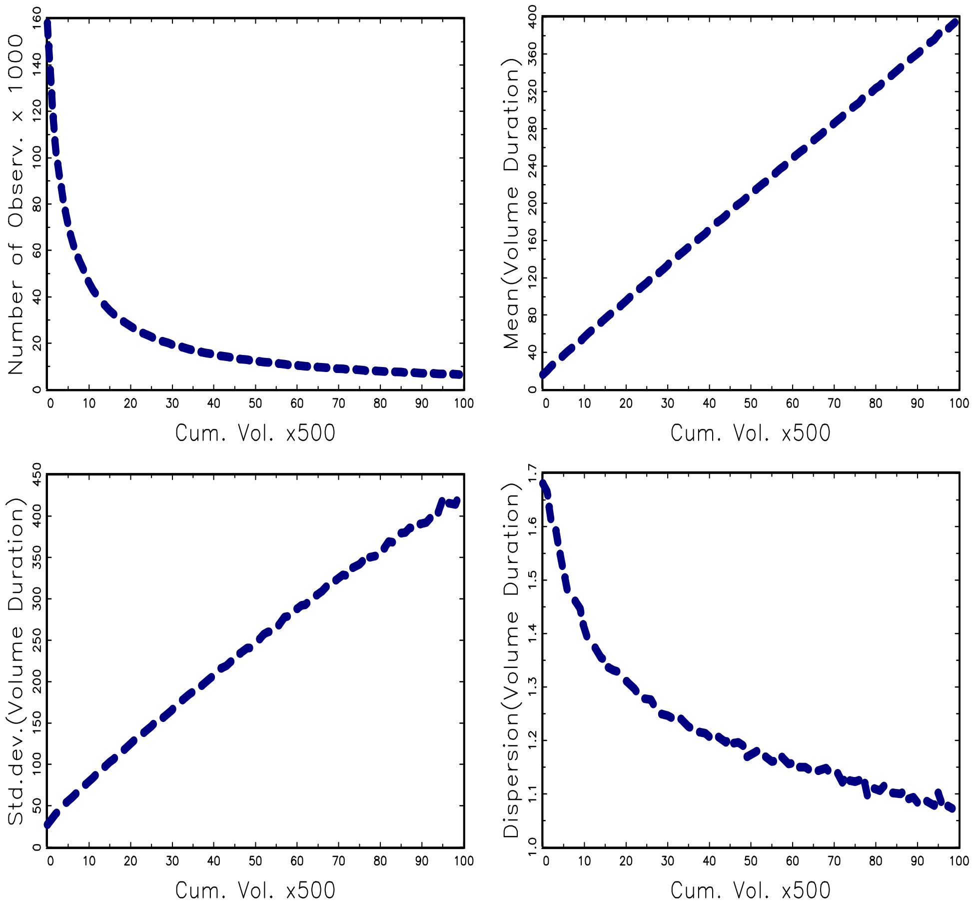

Figure 3: Description of volume durations (in sec.) of all 100 aggregated sub-datasets.

suggested testing the overdispersion of $\hat{\Lambda}_i$ (henceforth ER-Test). Finally, the theoretical distribution of the residuals implied by the estimated parameters $f(\varepsilon, \hat{\boldsymbol{\theta}})$ can be compared with the empirical one $\hat{f}(\hat{\varepsilon})$. For this purpose, Fernandes and Grammig (2006) developed the D-test to quantify the difference between them, which should be zero under correct model specification.

4 Dataset and Empirical Results

The transaction data of the Deutsche Telekom stock was extracted from the open order book of the German XETRA system, which is an order-driven market. The sample includes 225905 single transactions from 3^{rd} July until 6^{th} October 2000, observed for 69 trading days over 14 weeks. The continuous trading phase starts after the opening auction at 9 a.m.

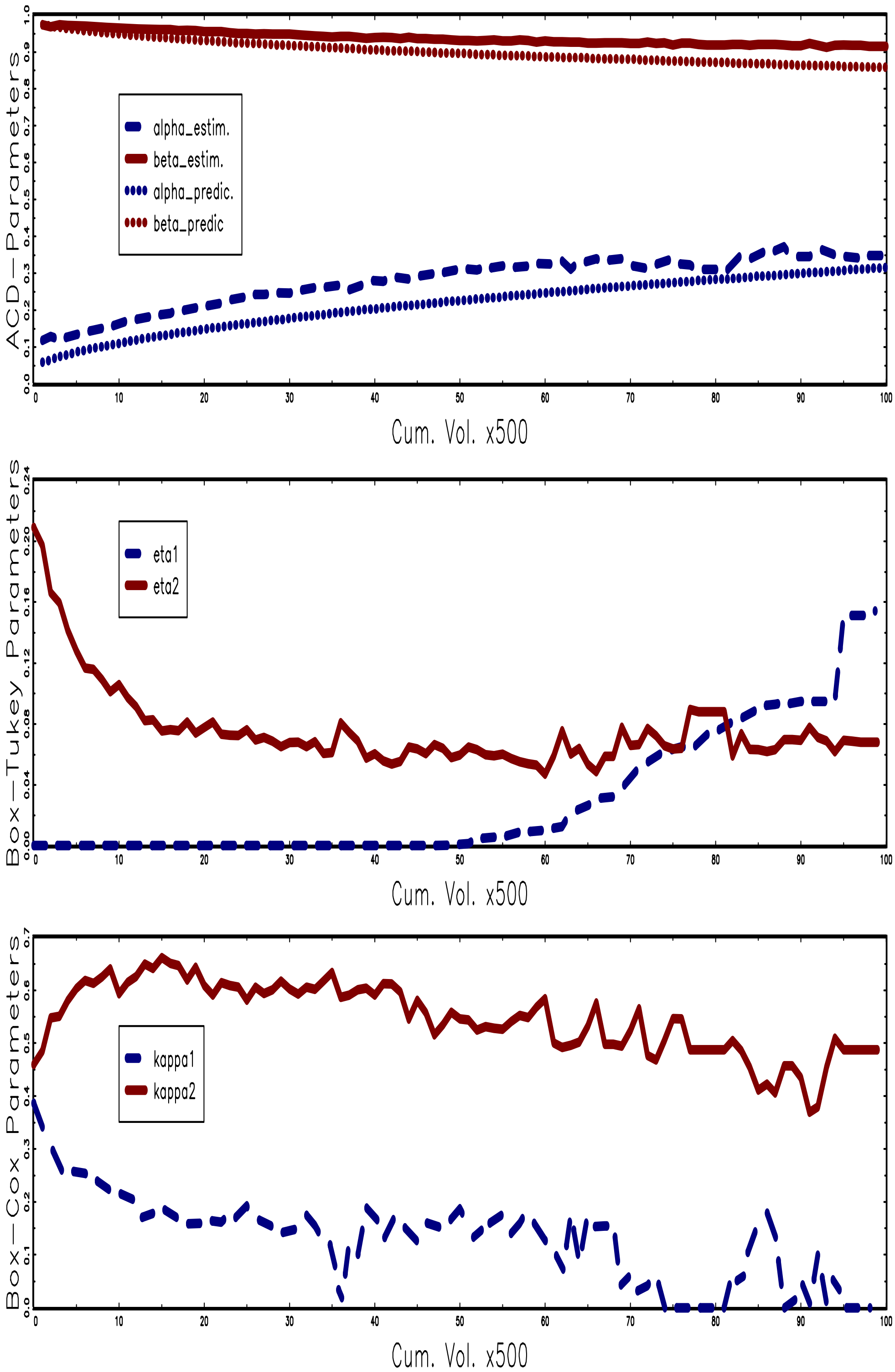

Figure 4: Estimated parameters of the ACD models for all 100 sub-datasets.

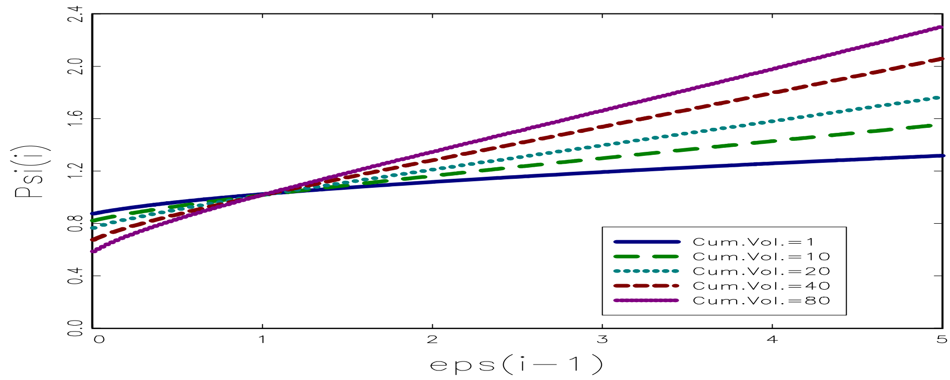

Figure 5: News-impact curve for different threshold values

and ends before the closing auction at 8 p.m. Further, it is interrupted by (at least) two intraday auctions at 1 p.m. and 5 p.m., each lasting at most 120 seconds. The electronic trading is based on an automatic matching algorithm, generally following a strict price-time priority of orders.

Figure 3 describes the volume durations data and shows how the mean and the standard deviation vary according to the threshold values. As expected, these measures increase, whereas the number of observations decreases in respect to a rising aggregation level. In contrast to the findings of Bauwens and Giot (2001) and Hautsch (2004), the dispersion declines as well, but still indicate overdispersion of the volume durations for high threshold values. Since the standard deviation exceeds the mean, it implies a negative duration dependence, that is, the waiting time to absorb cv is less likely to end the longer it lasts (Winkelmann (1998)).

In this study, the ACD models detect a strong cluster structure of volume durations, signaling a certain behavioral pattern of traders (see Figure 4). Generally, long volume durations tend to be followed by long ones and short durations by short ones, $\hat{\alpha}_1, \hat{\beta}_1 > 0$, revealing strong serial dependence. These findings are in line the common hypothesis of information-based market microstructure theory, where the trading and ordering process represents a source of information (for overview, see O'Hara (1997)). According to the theory, long durations suggests that uninformed traders (still) believe that the underlying value of the asset has not changed and only trade because of their own portfolio optimization. Contrary, short durations and, hence, intensive trading signalize the presence of informed traders who are assumed to make money by capitalizing on their informational advantage. The more information that is available in the market, the faster they have to react.

As evident in the upper panel of Figure 4, stationarity is always ensured ($\hat{\beta}_1 < 1$), but the persistence of the process diminishes slightly, whereas $\hat{\alpha}_1$ increases for a growing cv. Economically speaking, there is no (more)

	Parameters	$cv = 1$	$cv = 20$	$cv = 40$	$cv = 80$
	Box-Tukey-ACD model				
α_1	last standardized duration	0.1181 (0.0015)	0.2083 (0.0063)	0.2811 (0.0096)	0.3094 (0.0165)
β_1	last conditional duration	0.9740 (0.0009)	0.9557 (0.0028)	0.9382 (0.0045)	0.9196 (0.0058)
η_1	Box-Tukey$_1$	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0793 (0.0105)
η_2	Box-Tukey$_2$	0.2093 (0.0044)	0.0737 (0.0104)	0.0573 (0.0106)	0.0879 (0.0009)
κ_1	Box-Cox$_1$	0.3884 (0.0210)	0.1720 (0.0364)	0.1893 (0.0513)	0.0000 (.)
κ_2	Box-Cox$_2$	0.4588 (0.0099)	0.6436 (0.0355)	0.6041 (0.0476)	0.4871 (0.0200)
	Gen. Beta distribution				
a	peakedness of GBeta distr.	0.5972 (0.0055)	1.6981 (0.0912)	2.0549 (0.1280)	1.9258 (0.1365)
p	shape of GBeta distr.	2.0795 (0.0380)	0.8189 (0.0567)	0.8761 (0.0716)	1.2975 (0.1258)
q	shape of GBeta distr.	65.7257 (2.3774)	3.4026 (0.4695)	2.8113 (0.4024)	3.8918 (0.6904)
c	smooth of GBeta distr.	0.8537 (0.0282)	0.8651 (0.0419)	0.9562 (0.0825)	0.8602 (0.0913)

Table 1: The Box-Tukey-ACD Model

need for a rapid speed of updating the information in the market when regarding orders with high volumes, which necessarily take a long time to be absorbed by the market. According to the methodology of Easley and O'Hara (1992), the information efficiency of the market moves from a semi-strong to a strong level. Interestingly, the predicted α and β parameters by the Drost and Nijman formula that were used as start values for the estimation seem to serve as a lower bound.

Since both Box-Cox parameters $\hat{\kappa}_1, \hat{\kappa}_2 < 1$ (see lower panel of Figure 4), the dependence structure is nonlinear, implying concave news-impact curves, similar to the results of Fernandes and Grammig (2006). This means that the conditional duration must be adjusted more extensively during hectic periods (short durations) than in calm ones (long durations). The parameter $\hat{\kappa}_2$ is between 0.35 and 0.65, whereas $\hat{\kappa}_1$ is slightly decreasing, going to zero and, thus, converging to the so-called Log-ACD model for a few high cv values. Furthermore, both Box-Tukey parameters $\hat{\eta}_1, \hat{\eta}_2 \geq 0$ (see middle panel of Figure 4), which indicates that the durations must be shifted in

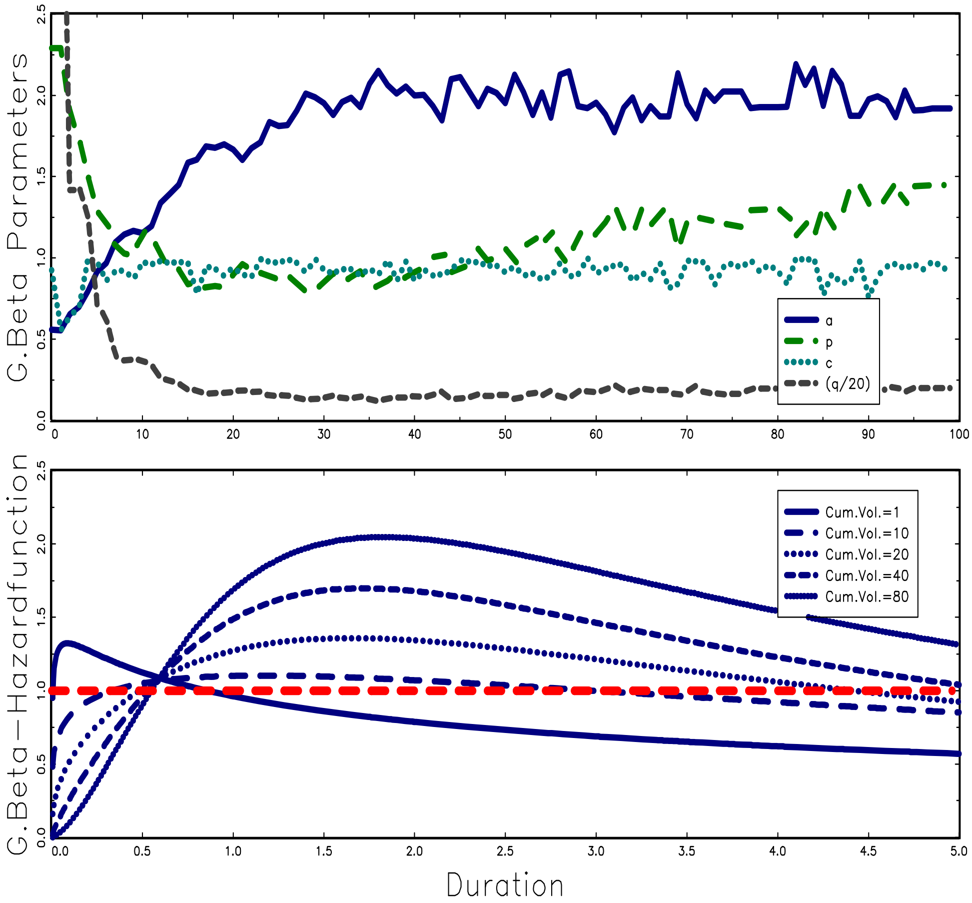

Figure 6: Estimated parameters of the generalized Beta distribution of all 100 ACD models and the implied hazard functions for different threshold values.

the overall adjustment process. As expected, the standardized duration ε always has to be enlarged with $\hat{\eta}_2$, whereas the conditional mean duration Ψ only needs to be shifted with $\hat{\eta}_1$ when regarding a high cv. In these cases, the model requires an additional intercept in order to capture the dynamics adequately. The Box-Tukey parameter $\hat{\eta}_1$ implies that the market is not liquid enough to absorb large amounts of shares ($cv \geq 50$) in a relatively short time, thus inducing additional time costs of liquidity.

In respect of reducing these costs, the results suggest that traders should either stop submitting new orders when observing these large amounts of traded shares or not exceed these thresholds values in case of fast (liquidity) block trading. A comparison of different news-impact curves for different threshold values is shown in Figure 5, displaying the market's reaction

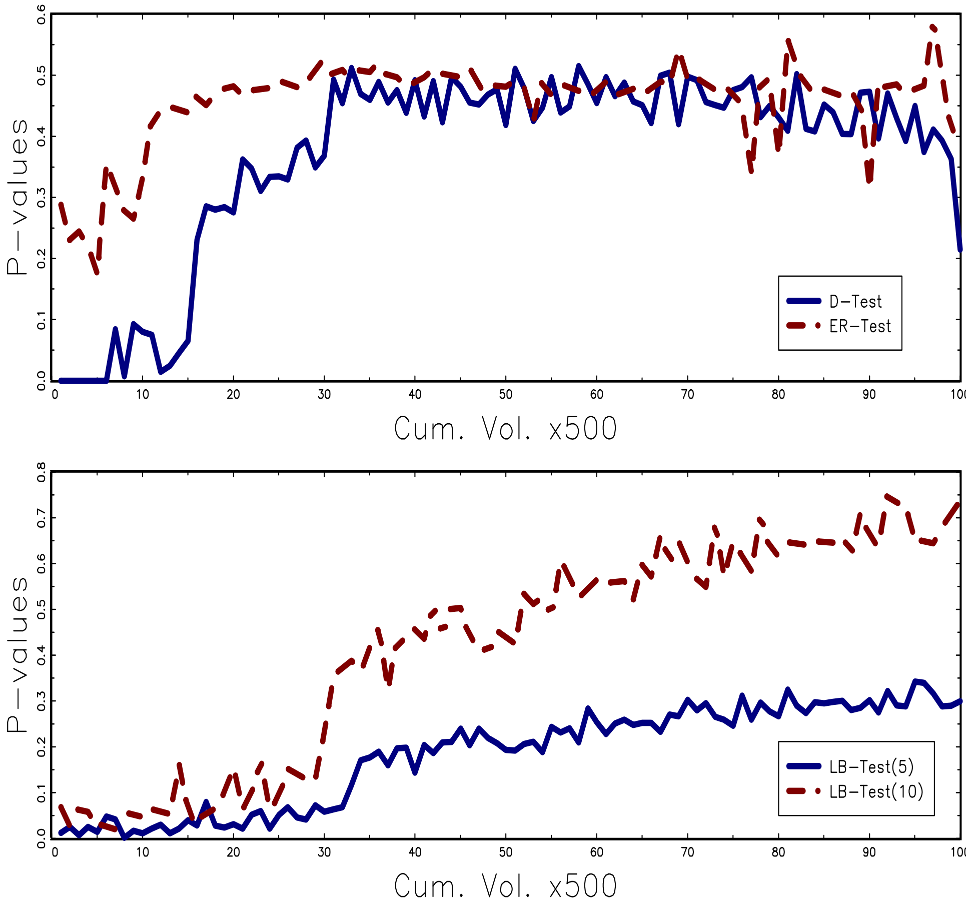

Figure 7: P-values of the Ljung-Box test, the D-test and the ER-test for all 100 ACD models

time. The influence of a high (low) duration shock ε_{i-1} will be the greater (smaller), the higher cv. The estimates and their standard errors (in parenthesis) of four duration models are reported in Table 1.

The errors $\hat{\varepsilon}_i = x_i/\hat{\Psi}_i$ are generalized Beta distributed. The development of the distribution parameters $\hat{a}, \hat{p}, \hat{q}$ and $\hat{c}$ are depicted in the upper panel of Figure 6, different implied hazard functions at the respective threshold levels are shown in the lower panel. As discernible, all hazard functions have a similar form and follow a non-monotonous concave slope, where the position of the maximum of the curve – determining the switch of the duration dependence – is closer to zero the smaller the threshold value. For example, the hazard function of the residuals for $cv = 1$ immediately decreases after a short jump, whereas other curves first show a moderate rise and then decline. This indicates that small quantities are most probably to be absorbed "immediately", while large amounts are more likely to be traded, the longer the time that has elapsed. In general, the hazard functions only differ in the time, where the positive duration ends and the negative begins, signalling that the spell to absorb cv is less likely to end the longer it lasts.

Figure 7 shows the p-values of the D-Test, the ER-Test and the Ljung-Box test (with the modified portmanteau statistic of Li and Li (2005)) of all estimated ACD models. Except the test of overdispersion by Engle and Russell (2004), all other tests reject a correct specification of the models for small aggregation levels. The D-test rejects the null for $cv \leq 15$ and the Ljung-Box test rejects for $cv \leq 33$, indicating that the model of lag orders $(1, 1)$ with the Box-Tukey transformation is not able to capture the dynamics of volume durations in the lower threshold levels. In contrast, the same model specification seem to show a good performance for all higher cv values, similar to the findings of Hautsch (2004).

5 Conclusion

This paper investigates the time varying liquidity of limit order books by analyzing the dynamics of volume durations with increasing threshold values. The main objective was to study the development of the models' parameters in order to reveal the market's absorption speed at different aggregation levels. Applying a flexible generalized Beta distribution to capture the extreme durations, and a Box-Tukey-transformation to account for a nonlinear news-impact curve, a sequence of ACD models was estimated to examine how the time series properties change as the threshold number of shares traded rises. Using detailed transaction data from the German XETRA system, the empirical results show that the cluster structure of the process and the market's absorption speed decreases continuously for highly cumulated volumes.

References

Anè, T., and H. Geman (2000): "Order Flow, Transaction Clock, and Normality of Asset Returns," *The Journal of Finance*, 5, 2259–2284.

Bauwens, L., and P. Giot (2001): *Econometric Modelling of Stock Market Intraday Activity.* Kluwer Academic, Boston et al.

Bauwens, L., P. Giot, J. Grammig, and D. Veredas (2004): "A Comparison of Financial Duration Models Via Density Forecasts," *International Journal of Forecasting*, 20(4), 589–609.

Box, G. E. P., and C. R. Cox (1964): "An Analysis of Transformations," *Journal of the Royal Statistical Society, Series B*, 26(2), 211–243.

Clark, P. K. (1973): "A subordinated stochastic process model with finite variance for speculative prices," *Econometrica*, 41(1), 135–155.

Drost, F. C., and T. E. Nijman (1993): "Temporal Aggregation of GARCH Processes," *Econometrica*, 61(1), 909–927.

Drost, F. C., and B. J. M. Werker (1996): "Closing the GARCH gap: Continous time GARCH modeling," *Journal of Econometrics*, 74, 31–57.

Dufour, A., and R. F. Engle (2000): "The ACD-Modell: Predictability of the Time Between Consecutive Trades," Discussion paper, ISMA Centre, University of Reading.

Easley, D., and M. O'Hara (1992): "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47(2), 577–605.

Engle, R. F. (2000): "The Econometrics of Ultra-High Frequency Data," *Econometrica*, 68(1), 1–22.

Engle, R. F., and J. Lange (2001): "Predicting VNET: A model of the dynamics of market depth," *Journal of Financial Markets*, 4(2), 113–142.

Engle, R. F., and J. R. Russell (1997): "Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model," *Journal of Empirical Finance*, 4(2-3), 187–212.

——— (1998): "Autoregressive Conditional Duration: A New Model for Irregularly-Spaced Financial Transactions Data," *Econometrica*, 66(5), 1127–1162.

——— (2004): "Analysis of High Frequency Financial Data," Forthcoming in Handbook of Financial Econometrics.

Epps, T. W., and M. L. Epps (1976): "The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis," *Econometrica*, 44(2), 305–321.

Fernandes, M., and J. Grammig (2006): "A Family of Autoregressive Conditional Durations Models," *Journal of Econometrics*, 130(1), 1–23.

Gallant, A. R., P. E. Rossi, and G. Tauchen (1992): "Stock Prices and Volume," *The Review of Financial Studies*, 2(2), 199–242.

Ghysels, E., and J. Jasiak (1998): "GARCH for Irregularly Spaced Financial Data: The ACD-GARCH Model," *Studies in Nonlinear Dynamics & Econometrics*, 2(4), 133–149.

Hafner, C. M. (2005): "Durations, volume and the prediction of financial returns in transaction time," *Quantitative Finance*, 5(2), 145–152.

Hautsch, N. (2002): "Modelling Intraday Trading Activity Using Box-Cox ACD Models," Working paper, CoFE.

——— (2003): "Assessing the Risk of Liquidity Suppliers on the Basis of Excess Demand Intensities," *Journal of Financial Econometrics*, 1(2), 189–215.

——— (2004): *Modelling Irregularly Spaced Financial Data*. Springer, Heidelberg.

Hentschel, L. (1995): "All in the family - Nesting symmetric and asymmetric GARCH models," *Journal of Financial Economics*, 39, 71–104.

Jones, C. M., G. Kaul, and L. M. L. (1994): "Transactions, volume and volatility," *Review of Financial Studies*, 7(4), 631–651.

Kalbfleisch, J. D., and R. L. Prentice (2002): *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, New York et al., 2nd edn.

Karpoff, J. M. (1987): "The Relation between Price Changes and Volume," *Journal of Financial and Quantitative Finance*, 22(1), 109–126.

Kleiber, C., and S. Kotz (2003): *Statistical Size Distributions in Economics and Actuarial Sciences*. John Wiley & Sons, New York et al.

Li, W. K., and G. Li (2005): "Diagnostic checking for time series models with conditional heteroscedasticity estimated by the least absolute deviation approach," *Biometrika*, 92(3), 691–701.

Lunde, A. (2000): "A Generalized Gamma Autoregressive Conditional Duration Model," Discussion paper, Aalborg University.

McDonald, J. B., and Y. J. Xu (1995): "A generalization of the beta distribution with applications," *Journal of Econometrics*, 69(2), 427–428.

O'Hara, M. (1997): *Market Microstructure Theory*. Blackwell, Oxford.

Russell, J. R. (1999): "Econometric Modelling of Multivariate Irregularly-Spaced High Frequency Data," Working paper, University of Chicago.

Tauchen, G. E., and M. Pitts (1983): "The Price Variability-Volume Relationsship on speculative Markets," *Econometrica*, 51(2), 485–502.

Veredas, D., J. Rodriguez-Poo, and A. Espasa (2002): "On the (intradaily) Seasonality and Dynamics of a Financial Point Process: A semiparametric Approach," Discussion paper 2002/23, CORE.

Winkelmann, R. (1998): *Econometric Analysis of Count Data*. Springer, Heidelberg, 2nd edn.