View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

SFB 649 Discussion Paper 2008-047

Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models

Nikolaus Hautsch*
Vahidin Jeleskovic*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models*

Nikolaus Hautsch† Vahidin Jeleskovic‡

June 26, 2008

Abstract

In this paper, we study the dynamic interdependencies between high-frequency volatility, liquidity demand as well as trading costs in an electronic limit order book market. Using data from the Australian Stock Exchange we model 1-min squared mid-quote returns, average trade sizes, number of trades and average (excess) trading costs per time interval in terms of a four-dimensional multiplicative error model. The latter is augmented to account also for zero observations. We find evidence for significant contemporaneous relationships and dynamic interdependencies between the individual variables. Liquidity is causal for future volatility but not vice versa. Furthermore, trade sizes are negatively driven by past trading intensities and trading costs. Finally, excess trading costs mainly depend on their own history.

Keywords: Multiplicative error models, volatility, liquidity, high-frequency data.

JEL Classification: C13, C32, C52

1 Introduction

Due to the permanently increasing availability of high-frequency financial data, the empirical analysis of trading behavior and the modelling of trading processes has become a major theme in modern financial econometrics. Key variables in empirical studies of high-frequency data are price volatilities, trading volume, trading intensities, bid-ask spreads and market depth as displayed by an open limit order book. A common characteristic of these variables is that they are positive-valued and persistently clustered over time.

*This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

†Institute for Statistics and Econometrics and CASE - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Quantitative Products Laboratory (QPL), Berlin, and Center for Financial Study (CFS), Frankfurt. Address: Spandauer Str. 1, D-10178 Berlin, Germany. *Email:* nikolaus.hautsch@wiwi.hu-berlin.de

‡Institute for Statistics and Econometrics, Humboldt-Universität zu Berlin, and Quantitative Products Laboratory (QPL), Berlin. Address: Alexanderstrasse 5, 10178 Berlin, Germany. *Email:* vahidin.jeleskovic@db.com

To capture the stochastic properties of positive-valued autoregressive processes, so-called (MEMs) have become popular. The basic idea of modelling a positive-valued process in terms of the product of positive-valued (typically i.i.d.) innovation terms and an observation-driven and/or parameter driven dynamic function is well-known in financial econometrics and originates from the model structure of the *autoregressive conditional heteroscedasticity* (ARCH) model introduced by Engle (1982) or the *stochastic volatility* (SV) model proposed by Taylor (1982). Engle and Russell (1997, 1998) introduced the *autoregressive conditional duration* (ACD) model to model autoregressive (financial) duration processes in terms of a multiplicative error process and a GARCH-type parameterization of the conditional duration mean. The term 'MEM' is ultimately introduced by Engle (2002) who discusses this approach as a general framework to model any kind of positive-valued dynamic process. Manganelli (2005) proposes a multivariate MEM to jointly model high-frequency volatilities, trading volume and trading intensities. Hautsch (2008) generalizes this approach by introducing a common latent dynamic factor serving as a subordinated process driving the individual trading components. The resulting model combines features of a GARCH type model and an SV type model and is called *stochastic* MEM. Engle and Gallo (2006) apply MEM specifications to jointly model different volatility indicators including absolute returns, daily range, and realized volatility. Recently, Cipollini *et al.* (2006) extend the MEM by a copula specification in order to capture contemporaneous relationships between the variables.

Given the growing importance of MEMs for the modelling of high-frequency trading processes, liquidity dynamics and volatility processes, this paper gives an introduction to the topic and an overview of the current literature. Given that the ACD model is the most popular specification of a univariate MEM, we will strongly rely on this string of the literature. Finally, we will present an application of the MEM to jointly model the multivariate dynamics of volatilities, trade sizes, trading intensities, and trading costs based on limit order book data from the Australian Stock Exchange (ASX).

The paper is organized as follows: Section 2 presents the major principles and properties of univariate MEMS. In Section 3, we will introduce multivariate specifications of MEMs. Estimation and statistical inference is illustrated in Section 4. Finally, Section 5 gives an application of the MEM to model high-frequency trading processes using data from the ASX.

2 The Univariate MEM

Let $\{Y_t\}$, $t = 1,\dots,T$, denote a non-negative (scalar) random variable. Then, the univariate MEM for Y_t is given by

$$\begin{aligned} Y_t &= \mu_t \varepsilon_t, \\ \varepsilon_t | \mathscr{F}_{t-1} &\sim \text{i.i.d. } \mathbb{D}(1,\sigma^2), \end{aligned}$$

where $\mathscr{F}_t$ denotes the information set up to t, μ_t is a non-negative conditionally deterministic process given $\mathscr{F}_{t-1}$, and ε_t is a unit mean, i.i.d. variate process defined on

non-negative support with variance σ^2. Then, per construction we have

$$\mathsf{E}[Y_t|\mathscr{F}_{t-1}] \stackrel{\text{def}}{=} \mu_t, \tag{1}$$

$$\text{Var}[Y_t|\mathscr{F}_{t-1}] = \sigma^2\mu_t^2. \tag{2}$$

The major principle of the MEM is to parameterize the conditional mean μ_t in terms of a function of the information set $\mathscr{F}_{t-1}$ and parameters θ. Then, the basic linear MEM(p,q) specification is given by

$$\mu_t = \omega + \sum_{j=1}^{p} \alpha_j Y_{t-j} + \sum_{j=1}^{q} \beta_j \mu_{t-j}, \tag{3}$$

where $\omega > 0$, $\alpha_j \geq 0$, $\beta_j \geq 0$. This specification corresponds to a generalized ARCH model as proposed by Bollerslev (1986) as long as Y_t is the squared (de-meaned) log return between t and $t-1$ with μ_t corresponding to the conditional variance. Accordingly, the process (3) can be estimated by applying GARCH software based on $\sqrt{Y_t}$ (without specifying a conditional mean function). Alternatively, if Y_t corresponds to a (financial) duration, such as, e.g., the time between consecutive trades (so-called trade durations) or the time until a cumulative absolute price change is observed (so-called price durations), the model is referred to an ACD specification as introduced by Engle and Russell (1997, 1998).

The unconditional mean of Y_t is straightforwardly computed as

$$\mathsf{E}[Y_t] = \omega/(1 - \sum_{j=1}^{p} \alpha_j - \sum_{j=1}^{q} \beta_j). \tag{4}$$

The derivation of the unconditional variance is more cumbersome since it requires the computation of $\mathsf{E}[\mu_t^2]$. In the case of an MEM(1,1) process, the unconditional variance is given by (see, e.g., Hautsch (2004))

$$\text{Var}[Y_t] = \mathsf{E}[Y_t]^2\sigma^2(1-\beta^2-2\alpha\beta)/(1-(\alpha+\beta)^2-\alpha^2\sigma^2) \tag{5}$$

corresponding to

$$\text{Var}[Y_t] = \mathsf{E}[Y_t]^2(1-\beta^2-2\alpha\beta)/(1-\beta^2-2\alpha\beta-2\alpha^2) \tag{6}$$

in case of $\sigma^2 = 1$ which is, e.g., associated with a standard exponential distribution. Correspondingly, the model implied autocorrelation function is given by

$$\rho_1 \stackrel{\text{def}}{=} \text{Corr}[Y_t, Y_{t-1}] = \alpha(1-\beta^2-\alpha\beta)/(1-\beta^2-2\alpha\beta), \tag{7}$$

$$\rho_j \stackrel{\text{def}}{=} \text{Corr}[Y_t, Y_{t-j}] = (\alpha+\beta)\rho_{j-1} \quad \text{for } j \geq 2. \tag{8}$$

Similarly to the GARCH model, the MEM can be represented in terms of an ARMA model for Y_t. Let $\eta_t \stackrel{\text{def}}{=} Y_t - \mu_t$ denote a martingale difference, then the MEM(p,q) process can be written as

$$Y_t = \omega + \sum_{j=1}^{max(p,q)} (\alpha_j + \beta_j) Y_{t-j} - \sum_{j=1}^{q} \beta_j \eta_{t-j} + \eta_t. \tag{9}$$

The weak stationarity condition of a MEM(1,1) process is given by $(\alpha+\beta)^2-\alpha^2\sigma^2<1$ ensuring the existence of $\text{Var}[Y_t]$.

Relying on the GARCH literature, the linear MEM specification can be extended in various forms. A popular form is a logarithmic specification of a MEM ensuring the positivity of μ_t without imposing parameter constraints. This is particularly important whenever the model is augmented by explanatory variables or when the model has to accommodate negative (cross-) autocorrelations in a multivariate setting. Two versions of logarithmic MEM's have been introduced by Bauwens and Giot (2000) in the context of ACD models and are given (for simplicity for $p=q=1$) by

$$\log\mu_t = \omega + \alpha g(\varepsilon_{t-1}) + \beta\log\mu_{t-1}, \tag{10}$$

where $g(\cdot)$ is given either by $g(\varepsilon_{t-1})=\varepsilon_{t-1}$ (type I) or $g(\varepsilon_{t-1})=\log\varepsilon_{t-1}$ (type II). The process is covariance stationary if $\beta<1$, $\mathsf{E}[\varepsilon_t\exp\{\alpha g(\varepsilon_t)\}]<\infty$ and $\mathsf{E}[\exp\{2\alpha g(\varepsilon_t)\}]<\infty$. For more details, see Bauwens and Giot (2000). Notice that due the logarithmic transformation, the news impact function, i.e., the relation between Y_t and ε_{t-1} is not anymore linear but is convex in the type I case and is concave in the type II parameterization. I.e., in the latter case, the sensitivity of Y_t to shocks in ε_{t-1} is higher if ε_{t-1} is small than in the case where it is large.

A more flexible way to capture nonlinear news responses is to allow for a kinked news response function

$$\log\mu_t = \omega + \alpha\{|\varepsilon_{t-1}-b| + c(\varepsilon_{t-1}-b)\}^\delta + \beta\log\mu_{t-1}, \tag{11}$$

where b gives the position of the kink while δ determines the shape of the piecewise function around the kink. For $\delta=1$, the model implies a linear news response function which is kinked at b resembling the EGARCH model proposed by Nelson (1991). For $\delta>1$, the shape is convex while it is concave for $\delta<1$. Such a specification allows to flexibly capture asymmetries in responses of Y_t to small or large lagged innovation shocks, such as, e.g., shocks in liquidity demand, liquidity supply or volatility. A similar specification is considered by Cipollini *et al.* (2006) to capture leverage effects if Y_t corresponds to a volatility variable. For more details on extended MEM specifications in the context of ACD models, see Hautsch (2004) or Bauwens and Hautsch (2008).

The error term distribution of ε_t is chosen as a distribution defined on positive support and standardized by its mean. If Y_t is the squared (de-meaned) log return, then $\sqrt{\varepsilon_t}\sim\mathbb{N}(0,1)$ yields the Gaussian GARCH model. If Y_t denotes a liquidity variable (such as trade size, trading intensity, bid-ask spread or market depth), a natural choice is an exponential distribution. Though the exponential distribution is typically too restrictive to appropriately capture the distributional properties of trading variables, it allows for a quasi maximum likelihood (QML) estimation yielding consistent estimates irrespective of distributional misspecifications. For more details, see Section 4. More flexible distributions are, e.g., the Weibull distribution, the (generalized) gamma distribution, the Burr distribution or the generalized F distribution. The latter is proposed in an ACD context by Hautsch (2003) and is given in standardized form (i.e., with unit mean) by the p.d.f.

$$f_\varepsilon(x) = [a\{x/\zeta(a,m,\eta)\}^{am-1}[\eta+\{x/\zeta(a,m,\eta)\}]^{(-\eta-m)}\eta^\eta]/\mathscr{B}(m,\eta), \tag{12}$$

where a, m, and η are distribution parameters, $\mathscr{B}(m,\eta) = \Gamma(m)\Gamma(\eta)/\Gamma(m+\eta)$, and

$$\zeta(a,m,\eta) \stackrel{\text{def}}{=} \{\eta^{1/a}\Gamma(m+1/a)\Gamma(\eta-1/a)\}/\{\Gamma(m)\Gamma(\eta)\}. \tag{13}$$

The generalized F-distribution nests the generalized gamma distribution for $\eta \to \infty$, the Weibull distribution for $\eta \to \infty$, $m = 1$, the log-logistic distribution for $m = \eta = 1$, and the exponential distribution for $\eta \to \infty$, $m = a = 1$. For more details, see Hautsch (2004).

3 The Vector MEM

Consider in the following a k-dimensional positive-valued time series, denoted by $\{\mathbf{Y}_t\}$, $t = 1\ldots,T$, with $\mathbf{Y}_t \stackrel{\text{def}}{=} (Y_t^{(1)},\ldots,Y_t^{(k)})$. Then, the so-called vector MEM (VMEM) for $\mathbf{Y}_t$ is defined by

$$\begin{aligned}\mathbf{Y}_t &= \mu_t \odot \varepsilon_t \\ &= \text{diag}(\mu_t)\varepsilon_t,\end{aligned}$$

where $\odot$ denotes the Hadamard product (element-wise multiplication) and ε_t is a k-dimensional vector of mutually and serially i.i.d. innovation processes, where the j-th element is given by

$$\varepsilon_t^{(j)}|\mathscr{F}_{t-1} \sim \text{i.i.d. } \mathbb{D}(1,\sigma_j^2), \quad j = 1,\ldots,k.$$

A straightforward extension of the univariate linear MEM as proposed by Manganelli (2005) is given by

$$\mu_t = \omega + \mathscr{A}_0\mathbf{Y}_t + \sum_{j=1}^{p}\mathscr{A}_j\mathbf{Y}_{t-j} + \sum_{j=1}^{q}\mathscr{B}_j\mu_{t-j}, \tag{14}$$

where ω is a $(k\times 1)$ vector, and $\mathscr{A}_0$, $\mathscr{A}_j$, and $\mathscr{B}_j$ are $(k\times k)$ parameter matrices. The matrix $\mathscr{A}_0$ captures contemporaneous relationships between the elements of $\mathbf{Y_t}$ and is specified as a matrix where only the upper triangular elements are non-zero. This triangular structure implies that $Y_t^{(i)}$ is predetermined for all variables $Y_t^{(j)}$ with $j < i$. Consequently, $Y_t^{(i)}$ is conditionally i.i.d. given $\{Y_t^{(j)},\mathscr{F}_{t-1}\}$ for $j < i$.

The advantage of this specification is that contemporaneous relationships between the variables are taken into account without requiring multivariate distributions for ε_t. This eases the estimation of the model. Furthermore, the theoretical properties of univariate MEMs as discussed in the previous section can be straightforwardly extended to the multivariate case. However, an obvious drawback is the requirement to impose an explicit ordering of the variables in $\mathbf{Y}_t$ which is typically chosen in accordance with a specific research objective or following economic reasoning. An alternative way to capture contemporaneous relationships between the elements of $\mathbf{Y}_t$ is to allow for mutual correlations between the innovation terms $\varepsilon_t^{(j)}$. Then, the innovation term vector follows a density function which is defined over non-negative k-dimensional support $[0,+\infty)^k$ with unit mean ι and covariance matrix Σ, i.e.,

$$\varepsilon_t|\mathscr{F}_{t-1} \sim \text{i.i.d. } \mathscr{D}(\iota,\Sigma)$$

implying

$$\begin{aligned}\mathsf{E}[\mathbf{Y}_t|\mathscr{F}_{t-1}] &= \mu_t,\\ \mathrm{Var}[\mathbf{Y}_t|\mathscr{F}_{t-1}] &= \mu_t\mu_t^\top \odot \Sigma = \mathrm{diag}(\mu_t)\Sigma\ \mathrm{diag}(\mu_t).\end{aligned}$$

Finding an appropriate multivariate distribution defined on positive support is a difficult task. As discussed by Cipollini *et al.* (2006), a possible candidate is a multivariate gamma distribution which however imposes severe restrictions on the contemporaneous correlations between the errors $\varepsilon_t^{(i)}$. Alternatively, copula approaches can be used as, e.g., proposed by Heinen and Rengifo (2006) or Cipollini *et al.* (2006).

In correspondence to the univariate logarithmic MEM, we obtain a logarithmic VMEM specification by

$$\log\mu_t = \omega + \mathscr{A}_0 \log\mathbf{Y}_t + \sum_{j=1}^{p}\mathscr{A}_j g(\varepsilon_{t-j}) + \sum_{j=1}^{q}\mathscr{B}_j \log\mu_{t-j}, \tag{15}$$

where $g(\varepsilon_{t-j}) = \varepsilon_{t-j}$ or $g(\varepsilon_{t-j}) = \log\varepsilon_{t-j}$, respectively. Generalized VMEMs can be specified accordingly to Section 2.

A further generalization of VMEM processes has been introduced by Hautsch (2008) and captures mutual (time-varying) dependencies by a subordinated common (latent) factor jointly driving the individual processes. The so-called *stochastic* MEM can be compactly represented as

$$\mathbf{Y}_t = \mu_t \odot \lambda_t \odot \varepsilon_t, \tag{16}$$

where λ_t is a $(k\times 1)$ vector with elements $\{\lambda_t^{\delta_i}\}$, $i = 1,\ldots,k$,

$$\log\lambda_t = a\log\lambda_{t-1} + \nu_t,\ \ \nu_t \sim \text{i.i.d. } \mathbb{N}(0,1), \tag{17}$$

and ν_t is assumed to be independent of ε_t. Hence, λ_t serves as a common dynamic factor with process-specific impacts δ_i. Then, the elements of μ_t represent 'genuine' (trade-driven) effects given the latent factor. They are assumed to follow (15) with $g(\varepsilon_t) = \mathbf{Y}_t \odot \mu_t^{-1}$. The model corresponds to a mixture model and nests important special cases, such as the SV model by Taylor (1982) or the stochastic conditional duration model by Bauwens and Veredas (2004). Applying this approach to jointly model high-frequency volatilities, trade sizes and trading intensities, Hautsch (2008) shows that the latent component is a major driving force of cross-dependencies between the individual processes.

4 Statistical Inference

Define $f(Y_t^{(1)}, Y_t^{(2)}, \ldots, Y_t^{(k)}|\mathscr{F}_{t-1})$ as the joint conditional density given $\mathscr{F}_{t-1}$. Without loss of generality the joint density can be decomposed into

$$\begin{aligned}f(Y_t^{(1)}, Y_t^{(2)}, \ldots, Y_t^{(k)}|\mathscr{F}_{t-1}) =& f(Y_t^{(1)}|Y_t^{(2)},\ldots,Y_t^{(k)};\mathscr{F}_{t-1}) &(18)\\ &\times f(Y_t^{(2)}|Y_t^{(3)},\ldots,Y_t^{(k)};\mathscr{F}_{t-1}) &(19)\\ &\times f(Y_t^{k}|\mathscr{F}_{t-1}). &(20)\end{aligned}$$

Then, the log likelihood function is defined by

$$\mathscr{L}(\theta) \stackrel{\text{def}}{=} \sum_{t=1}^{T} \sum_{j=1}^{k} \log f(Y_t^{(j)} | Y_t^{(j+1)}, \ldots, Y_t^{(k)}; \mathscr{F}_{t-1}). \tag{21}$$

For instance, if $Y_t^{(j)} | Y_t^{(j+1)}, \ldots, Y_t^{(k)}; \mathscr{F}_{t-1}$ follows a generalized F distribution with parameters $a^{(j)}$, $m^{(j)}$ and $\eta^{(j)}$, the corresponding log likelihood contribution is given by

$$\log f(Y_t^{(j)} | Y_t^{(j+1)}, \ldots, Y_t^{(k)}; \mathscr{F}_{t-1}) \tag{22}$$

$$= \log[\Gamma(m^{(j)} + \eta^{(j)}) / \{\Gamma(m^{(j)})\Gamma(\eta^{(j)})\}] + \log a^{(j)} - a^{(j)} m^{(j)} \log \tilde{\mu}_t^{(j)} \tag{23}$$

$$+ (a^{(j)} m^{(j)} - 1) \log Y_t^{(j)} - (\eta^{(j)} + m^{(j)}) \log \left(\eta^{(j)} + Y_t^{(j)} / \tilde{\mu}_t^{(j)} \right) \tag{24}$$

$$+ \eta^{(j)} \log(\eta^{(j)}), \tag{25}$$

where $\tilde{\mu}_t^{(j)} = \mu_t^{(j)} / \zeta(a^{(j)}, m^{(j)}, \eta^{(j)})$ and $\zeta(\cdot)$ defined as above.

Constructing the likelihood based on an exponential distribution leads to the quasi likelihood function with components

$$\log f(Y_t^{(j)} | Y_t^{(j+1)}, \ldots, Y_t^{(k)}; \mathscr{F}_{t-1}) = -\sum_{t=1}^{T} \left(\log \mu_t^{(j)} + Y_t^{(j)} / \mu_t^{(j)} \right),$$

where the score and Hessian are given by

$$\frac{\partial \log f(Y_t^{(j)} | Y_t^{(j+1)}, \ldots, Y_t^{(k)}; \mathscr{F}_{t-1})}{\partial \theta^{(j)}} = -\sum_{t=1}^{T} \frac{\partial \mu_t^{(j)}}{\partial \theta^{(j)}} \frac{1}{\mu_t^{(j)}} \left(\frac{Y_t^{(j)}}{\mu_t^{(j)}} - 1 \right),$$

$$\frac{\partial^2 \log f(Y_t^{(j)} | Y_t^{(j+1)}, \ldots, Y_t^{(k)}; \mathscr{F}_{t-1})}{\partial \theta^{(j)} \partial \theta^{(j)^\top}} = \sum_{t=1}^{T} \left\{ \frac{\partial}{\partial \theta^{(j)^\top}} \left(\frac{1}{\mu_t^{(j)}} \frac{\partial \mu_t^{(j)}}{\partial \theta^{(j)}} \right) \left(\frac{Y_t^{(j)}}{\mu_t^{(j)}} - 1 \right) \right.$$

$$\left. - \frac{1}{\mu_t^{(j)}} \frac{\partial \mu_t^{(j)}}{\partial \theta^{(j)}} \frac{\partial \mu_t^{(j)}}{\partial \theta^{(j)^\top}} \frac{Y_t^{(j)}}{\mu_t^{(j)^2}} \right\}.$$

Building on the results by Bollerslev and Wooldridge (1992) and Lee and Hansen (1994), Engle (2000) shows the consistency and asymptotic normality of $\hat{\theta}$, where the asymptotic covariance corresponds to the Bollerslev and Wooldridge (1992) QML covariance matrix.

Model evaluation can be straightforwardly performed by testing the dynamic and distributional properties of the model residuals

$$e_t^{(j)} \stackrel{\text{def}}{=} \hat{\varepsilon}_t^{(j)} = Y_t^{(j)} / \hat{\mu}_t^{(j)}. \tag{26}$$

Under correct model specification, the series $e_t^{(j)}$ must be i.i.d. with distribution $\mathbb{D}(\cdot)$. Portmanteau statistics such as the Ljung-Box statistic (Ljung and Box (1978)) based on (de-meaned) MEM residuals can be used to analyze whether the specification is able to capture the dynamic properties of the process. The distributional properties

can be checked based on QQ-plots. Engle and Russell (1998) propose a simple test for no excess dispersion implied by an exponential distribution using the statistic

$$\sqrt{n}\left\{(\hat{\sigma}^2_{e^{(j)}}-1)/\tilde{\sigma}^{(j)}\right\},$$

where $\hat{\sigma}^2_{e^{(j)}}$ is the sample variance of $e_t^{(j)}$ and $\tilde{\sigma}^{(j)}$ is the standard deviation of $(\varepsilon_t^{(j)}-1)^2$. Under the null hypothesis of an exponential distribution, the test statistic is asymptotically normally distributed with $\widehat{\sigma}^2_{e^{(j)}}=1$ and $\left(\tilde{\sigma}^{(j)}\right)^2=\sqrt{8}$.

Alternatively, probability integral transforms can be used to evaluate the in-sample goodness-of-fit, see, e.g., Bauwens *et al.* (2004). Building on the work by Rosenblatt (1952), Diebold *et al.* (1998) show that

$$q_t^{(j)} \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f_{e^{(j)}}(s)ds$$

must be i.i.d. $\mathbb{U}[0,1]$. Alternative ways to evaluate MEMs are Lagrange Multiplier tests as proposed by Meitz and Teräsvirta (2006), (integrated) conditional moment tests as discussed by Hautsch (2006) or nonparametric tests as suggested by Fernandes and Grammig (2006).

5 High-Frequency Volatility and Liquidity Dynamics

In this section, we will illustrate an application of the VMEM to jointly model return volatilities, average trade sizes, the number of trades as well as average trading costs in intra-day trading. We use a data base extracted from the electronic trading of the Australian Stock Exchange (ASX) which is also used by Hall and Hautsch (2006, 2007). The ASX is a continuous double auction electronic market where the continuous trading period between 10:09 a.m. and 4:00 p.m. is preceded and followed by a call auction. During continuous trading, any buy (sell) order entered that has a price that is greater than (less than) or equal to existing queued buy (sell) orders, will execute immediately and will result in a transaction as long as there is no more buy (sell) order volume that has a price that is equal to or greater (less) than the entered buy (sell) order. In case of partial execution, the remaining volume enters the limit order queues. Limit orders are queued in the buy and sell queues according to a strict price-time priority order and may be entered, deleted and modified without restriction. For more details on ASX trading, see Hall and Hautsch (2007).

Here, we use data from completely reconstructed order books for the stocks of the National Australian Bank (NAB) and BHP Billiton Limited (BHP) during the trading period July and August 2002 covering 45 trading days. In order to reduce the impact of opening and closure effects, we delete all observations before 10:15 a.m. and after 3:45 p.m. To reduce the complexity of the model we restrict our analysis to equidistant observations based on one-minute aggregates. For applications of MEMs to irregularly spaced data, see Manganelli (2005) or Engle (2000).

Table 1 shows summary statistics for log returns, the average trade size, the number of trades, and the average (time-weighted) trading costs. The log returns correspond to the residuals of an MA(1) model for differences in log transaction prices. This preadjustment removes the effects of the well-known bid-ask bounce causing negative

first-order serial correlation, see Roll (1984). The trading costs are computed as the hypothetical trading costs of an order of the size of $10,000$ shares in excess to the trading costs which would prevail if investors could trade at the mid-quote. They are computed as a time-weighted average based on the average ask and bid volume pending in the queues and yield natural measures of transaction costs induced by a potentially lacking liquidity supply. Conversely, trade sizes and the number of trades per interval indicate the liquidity demand in the market.

We observe that high-frequency log returns reveal similar stochastic properties as daily log returns with significant overkurtosis and slight left-skewness. For the average trade size and the number of trades per interval we find a highly right-skewed distribution with a substantial proportion of observations being zero. These observations stem from tranquil trading periods, where market orders do not necessarily occur every minute. As illustrated below, these periods typically happen around noon causing the well-known 'lunch-time dip'. On the other hand, we also find evidence for very active trading periods resulting in a high speed of trading and large average trade sizes. On average, the number of trades per one-minute interval is around 2.5 and 3.5 for NAB and BHP, respectively, with average trade sizes of approximately $2,300$ and $5,800$ shares, respectively. The excess trading costs associated with the buy/sell transaction of $10,000$ shares are on average around 60 ASD for BHP and 188 ASD for NAB. Hence, on average, excess trading costs for NAB are significantly higher than for BHP which is caused by a higher average bid-ask spread and a lower liquidity supply in the book. The Ljung-Box statistics indicate the presence of a strong serial dependence in volatilities and all liquidity variables, and thus reveal the well-known clustering structures in trading processes. The significant Ljung-Box statistics for log returns are induced by the bid-ask bounce effect causing significantly negative first order autocorrelation. Obviously, the MA(1) filter does not work very well in the case of NAB data. Alternatively, one could use higher order MA-filter. The empirical autocorrelations (ACFs) shown in Figure 1 confirm a relatively high persistence in volatilities and liquidity variables indicated by the Ljung-Box statistics. A notable exception is the process of trade sizes for NAB revealing only weak serial dependencies. Figure 2 displays the cross-autocorrelation functions (CACFs) between the individual variables. It turns out that squared returns are significantly positively (cross-)autocorrelated with the number of trades and excess trading costs, and – to less extent – with the average trade size. This indicates strong dynamic interdependencies between volatility and liquidity demand as well as supply. Similarly, we also observe significantly positive CACFs between trade sizes and the speed of trading. Hence, periods of high liquidity demand are characterized by both high trade sizes and a high trading intensity. Conversely, the CACFS between trading costs and trade sizes as well as between trading costs and the trading intensity are significantly negative. Ceteris paribus this indicates that market participants tend to exploit periods of high liquidity supply, i.e. they trade fast and high volumes if the trading costs are low and thus liquidity supply is high.

Figure 1: Sample ACF of squared log returns (SR), trade size (TS), number of trades (NT), and trade costs (TC)(from top to bottom) for BHP (left) and NAB (right). The x-axis shows the lags. The broken line shows the asymptotic 95% confidence intervals.

A typical feature of high-frequency data is the strong influence of intra-day sea-

	BHP				NAB			
	LR	TS	NT	TC	LR	TS	NT	TC
Obs.	14520	14520	14520	14520	14503	14503	14503	14503
Mean	6.81E-7	5811.52	3.53	60.20	-3.19E-4	2295.24	2.69	188.85
S.D.	7.41E-2	8378.09	3.20	18.47	3.83E-2	7228.38	2.72	97.37
Min	-0.50	0	0	2.99	-0.31	0	0	16.52
Max	0.44	250000	24	231.38	0.38	757500.50	23	1043.35
q10	-0.10	0	0	5.00	-0.04	0	0	84.48
q90	0.10	13475	8	8.80	0.04	5150	6	317.48
Kurtosis	5.23	-	-	-	9.85	-	-	-
LB_{20}	29.61	1585.04	34907.77	22422.32	939.05	95.94	22825.72	23786.09
LB_{20}(SR)	2073.77	-	-	-	2808.75	-	-	-

Table 1: Descriptive statistics of log returns (LR), trade sizes (TS), number of trades (NT), and trade costs (TC) for BHP and NAB. Evaluated statistics: mean value, standard deviation (S.D.), minimum and maximum, 10%- and 90%-quantile (q10 and q90, respectively), kurtosis, and the Ljung-Box statistic (associated with 20 lags). LB_{20}(SR) represents the Ljung-Box statistic computed for the squared log returns (SR).

Figure 2: Sample CACF for BHP (top) and NAB (bottom). The solid, dash-dotted and dashed lines show the CACF between TC and SR, TC and TS, TC and NT, respectively, on the left side and between SR and TS, SR and NT, TS and NT, respectively, on the right side. The dotted line shows the asymptotic 95% confidence interval. The x-axis shows the lags.

sonalities which is well documented by a wide range of empirical studies. For detailed illustrations, see Bauwens and Giot (2001) or Hautsch (2004). One possibility to account for intra-day seasonalities is to augment the specification of μ_t by appropriate regressors. An alternative way is to adjust for seasonalities in a first step. Though the effect of a pre-adjustment on the final parameter estimates is controversially discussed in the literature (see e.g. Veredas *et al.* (2001)), most empirical studies prefer the two-stage method since it reduces model complexity and the number of parameters to be estimated in the final step. Here, we follow this proceeding and adjust the individual variables $Y_t^{(i)}$ for deterministic intraday-seasonalities based on cubic spline regressions with 30-minute nodes between 10:30 and 15:30. Figure 3 shows the resulting estimated seasonality components. Confirming empirical findings from other markets, we observe that the liquidity demand follows a distinct U-shape pattern with a clear dip around lunch time. However, a clearly different picture is revealed for the trading costs. Obviously, the liquidity supply is lowest during the morning and around noon inducing high trading costs. Then, (excess) trading costs decline during the afternoon and reach a minimum before market closure. This indicates that not only liquidity demand but also liquidity supply is highest before the end of the trading period. For volatility, we observe a rather typical picture with the highest volatility after the opening of the market and (but to less extent) before closure. The high volatility at morning is clearly associated with information processing during the first minutes of trading.

Conceptual difficulties are caused by the relatively high number of zeros in the liquidity demand variables which cannot be captured by a standard MEM requiring

Figure 3: Deterministic intra-day seasonality patterns for SR, TS, NT and TC (from top to bottom) for BHP (left) and NAB (right). The seasonality components are estimated using cubic spline functions based on 30-minute nodes. The x-axis gives the time of the day.

positive random variables. In order to account for zeros, we augment a Log-VMEM by corresponding dummy variables:

$$\log \mu_t = \omega + \mathscr{A}_0[(\log \mathbf{Y_t}) \odot 1_{\{Y_t>0\}}] + \mathscr{A}_0^0 \odot 1_{\{Y_t=0\}} \tag{27}$$

$$+ \sum_{j=1}^{p} \mathscr{A}_j[g(\varepsilon_{t-j}) \odot 1_{\{Y_{t-1}>0\}}] + \sum_{j=1}^{p} \mathscr{A}_j^0 \odot 1_{\{Y_{t-1}=0\}} \tag{28}$$

$$+ \sum_{j=1}^{q} \mathscr{B}_j \log \mu_{t-j}, \tag{29}$$

where $1_{\{Y_t>0\}}\}$ and $1_{Y_t=0}$ denote $k \times 1$ vectors of indicator variables indicating non-zeor and zero realizations, respectively, and $\mathscr{A}_j^0$, $j = 0, \ldots, p$, are corresponding $k \times k$ parameter matrices.

Then, the log likelihood function is split up into two parts yielding

$$\mathscr{L}(\theta) = \sum_{t=1}^{T} \sum_{j=1}^{k} \log f(Y_t^{(j)} | Y_t^{(j+1)}, \ldots, Y_t^{(k)}; Y_t^{(j)} > 0, \mathscr{F}_{t-1}) \tag{30}$$

$$\times \log \mathrm{P}[Y_t^{(j)} > 0 | Y_t^{(j+1)}, \ldots, Y_t^{(k)}; \mathscr{F}_{t-1}]. \tag{31}$$

If both likelihood components have no common parameters, the second part can be maximized separately based on a binary choice model including past (and contemporaneous) variables as regressors. Then, the first log likelihood component is associated only with positive values and corresponds to the log likelihood given by (27).

We estimate a four-dimensional Log-VMEM for squared log returns, trade sizes, the number of trades and transaction costs standardized by their corresponding seasonality components. For simplicity and to keep the model tractable, we restrict our analysis to a specification of the order $p = q = 1$. The innovation terms are chosen as $g(\varepsilon_t) = \varepsilon_t$. For the process of squared returns, $Y_t^{(1)} = r_t^2$, we assume $Y_t^{(1)} | Y_t^{(2)}, \ldots, Y_t^{(4)}, \mathscr{F}_{t-1} \sim \mathbb{N}(0, \mu_t^{(1)})$. Accordingly, for $Y_t^{(j)}$, $j \in \{2,3,4\}$, we assume $Y_t^{(j)} | Y_t^{(j+1)}, \ldots, Y_t^{(4)}, \mathscr{F}_{t-1} \sim Exp(\mu_t^{(j)})$. Though it is well-known that both the normal and the exponential distribution are not flexible enough to capture the distributional properties of high-frequency trading processes, they allow for a QML estimation of the model.

Hence, the adjustments for zero variables have to be done only in the liquidity components but not in the return component. Moreover, note that there are no zeros in the trading cost component. Furthermore, zero variables in the trade size and the number of trades per construction always occur simultaneously. Consequently, we can only identify the $(2,3)$-element in $\mathscr{A}_0^0$ and one of the two middle columns in $\mathscr{A}_1^0$, where all other parameters in $\mathscr{A}_0^0$ and $\mathscr{A}_1^0$ are set to zero.

For the sake of brevity we do not show the estimation results of the binary choice component but restrict our analysis to the estimation of the MEM. Figure 2 shows the estimation results for BHP and NAB based on a specification with fully parameterized matrix $\mathscr{A}_1$ and diagonal matrix $\mathscr{B}_1$.

	BHP		NAB	
Coeff.	Coeff.	Std. err.	Coeff.	Std. err.
ω_1	-0.0673	0.0663	0.0023	0.0302
ω_2	0.1921	0.0449	0.1371	0.0254
ω_3	-0.4722	0.1009	-0.1226	0.0432
ω_4	-0.4914	0.1066	-0.5773	0.0485
$A_{0,12}$	0.0549	0.0092	0.1249	0.0056
$A_{0,13}$	0.3142	0.0173	0.6070	0.0122
$A_{0,14}$	0.4685	0.0489	0.7876	0.0094
$A_{0,23}$	0.0673	0.0074	0.0531	0.0070
$A_{0,24}$	-0.1002	0.0289	0.0176	0.0093
$A_{0,34}$	-0.2181	0.0618	-0.0235	0.0123
$A^0_{0,12}$	-3.8196	0.0402	-1.5086	0.0176
$A_{1,11}$	0.1446	0.0080	0.0804	0.0038
$A_{1,12}$	0.0043	0.0090	0.0804	0.0041
$A_{1,13}$	-0.0939	0.0173	0.2036	0.0125
$A_{1,14}$	0.1487	0.0602	-0.0833	0.0214
$A_{1,21}$	0.0004	0.0034	-0.0002	0.0015
$A_{1,22}$	0.0488	0.0049	0.0259	0.0025
$A_{1,23}$	-0.0377	0.0115	-0.0116	0.0093
$A_{1,24}$	-0.1911	0.0398	-0.1329	0.0226
$A_{1,31}$	0.0100	0.0053	-0.0022	0.0020
$A_{1,32}$	0.0095	0.0071	0.0045	0.0031
$A_{1,33}$	0.1088	0.0152	0.0894	0.0109
$A_{1,34}$	0.3420	0.0932	0.0341	0.0377
$A_{1,41}$	0.0064	0.0113	0.0044	0.0067
$A_{1,42}$	0.0091	0.0163	0.0081	0.0081
$A_{1,43}$	0.0524	0.0321	0.0537	0.0249
$A_{1,44}$	0.4256	0.0898	0.5105	0.0431
$A^0_{1,21}$	1.1467	0.0911	-0.5181	0.0204
$A^0_{1,22}$	0.1497	0.0212	0.0341	0.0134
$A^0_{1,23}$	0.0946	0.0318	0.0985	0.0132
$A^0_{1,24}$	-0.0006	0.0755	0.0115	0.0579
$B_{1,11}$	0.4027	0.0252	0.2616	0.0078
$B_{1,22}$	0.7736	0.0179	0.9109	0.0081
$B_{1,33}$	0.9731	0.0074	0.9673	0.0070
$B_{1,44}$	0.5369	0.1024	0.7832	0.0374
Log Likelihood	-60211		-58622	
BIC	-60378		-58790	

Table 2: Quasi-maximum likelihood estimation results of a MEM for seasonally adjusted (i) squared (bid-ask bounce adjusted) log returns, (ii) average trade sizes, (iii) number of trades, and (iv) average trading costs per one-minute interval. Standard errors are computed based on the OPG covariance matrix.

Descriptive statistics of seasonally adjusted data								
	BHP				NAB			
	SR	TS	NT	TC	SR	TS	NT	TC
Mean	1.000	1.001	1.000	1.000	1.002	1.001	1.000	0.999
S.D.	1.963	1.528	0.834	0.300	3.152	2.644	0.991	0.507
LB_{20}	1159.456	202.001	8782.762	19210.412	800.808	124.806	3775.762	19707.831
Descriptive statistics of MEM-residuals								
	BHP				NAB			
	SR	TS	NT	TC	SR	TS	NT	TC
Mean	1.000	1.000	1.000	1.000	1.000	0.999	1.001	1.000
S.D.	1.568	1.348	0.629	0.228	2.595	2.280	0.675	0.386
LB_{20}	63.559	61.388	519.348	1568.428	63.455	14.201	751.317	163.426

Table 3: Summary statistics of the seasonality adjusted time series and the corresponding MEM residuals for BHP and NAB.

We can summarize the following major findings: First, we observe significant mutual correlations between nearly all variables. Confirming the descriptive statistics above, volatility is positively correlated with liquidity demand and liquidity supply. Hence, active trading as driven by high volumes and high trading intensities is accompanied by high volatility. Simultaneously, as indicated by significantly negative estimates of $\mathscr{A}^0_{24}$ and $\mathscr{A}^0_{34}$, these are trading periods which are characterized by low transaction costs.

Second, as indicated by the diagonal elements in $\mathscr{A}_1$ and the elements in $\mathscr{B}_1$, all trading components are strongly positively autocorrelated but are not very persistent. As also revealed by the descriptive statistics, the strongest first order serial dependence is observed for the process of trading costs. The persistence is highest for trade sizes and trading intensities.

Third, we find Granger causalities from liquidity variables to future volatility. High trade sizes predict high future return volatilities. However, the impact of trading intensities and trading costs on future volatility is less clear. Here, we find contradictive results for both stocks. Conversely, we do not observe any predictability of return volatility for future liquidity demand and supply. For both stocks all corresponding coefficients are insignificant.

Fourth, trade sizes are significantly negatively driven by past trading intensities and past trading costs. This finding indicates that a high speed of trading tends to reduce trade sizes over time. Similarly, increasing trading costs deplete the incentive for high order sizes but on the other hand increase the speed of trading. Hence, market participants observing a low liquidity supply reduce trade sizes but trade more often. A possible explanation for this finding is that investors tend to break up large orders into sequences of small orders.

Fifth, (excess) transaction costs depend only on their own history but not on the lagged volatility or liquidity demand. This indicates that liquidity supply is difficult to predict based on the history of the trading process.

Sixth, as shown by the summary statistics of the MEM residuals, the model captures a substantial part of the serial dependence in the data. This is indicated by a sig-

nificant reduction of the corresponding Ljung-Box statistics. Nevertheless, for some processes, there is still significant remaining serial dependence in the residuals. This is particularly true for the trading cost and trading intensity components for which obviously higher order dynamics have to be taken into account. For the sake of brevity we refrain from showing results of higher parameterized models. Allowing for more dynamic and distributional flexibility further improves the goodness-of-fit, however, makes the model less tractable and less stable for out-of-sample forecasts.

6 Conclusion

In summary, we find strong dynamic interdependencies and causalities between high-frequency volatility, liquidity demand, and liquidity supply. In particular, the high trade sizes are able to predict high future volatilities whereas the return volatility appears not to give rise to future liquidity demand and supply dynamic. The effects of the trading intensities and trading costs on future volatilities could not be uniformly concluded although these effects seem to be significant. An interesting finding is that the high trade costs, associated with low liquidity supply, lead to a decrease of the trade sizes and simultaneously to an increase of the trade intensities. However, the dynamic of the trade costs seems to be mostly driven by its own history. Last but not at least we find a higher persistence by liquidity variables than by return volatilities. Hence, these results might serve as valuable input for trading strategies and (automated) trading algorithms.

References

Bauwens, L. and N. Hautsch (2008). Modelling financial high frequency data with point processes. In: *Handbook of Financial Time Series* (T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch, Eds.). Springer Verlag.

Bauwens, L., P. Giot, J. Grammig and D. Veredas (2004). A comparison of financial duration models via density forecasts. *International Journal of Forecasting* **20**, 589–609.

Bauwens, Luc and David Veredas (2004). The stochastic conditional duration model: A latent factor model for the analysis of financial durations. *Journal of Econometrics* **119**, 381–412.

Bauwens, Luc and Pierre Giot (2000). The logarithmic ACD model: An application to the Bid/Ask quote process of two NYSE stocks. *Annales d'Economie et de Statistique* **60**, 117–149.

Bauwens, Luc and Pierre Giot (2001). *Econometric Modelling of Stock Market Intraday Activity*. Kluwer.

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

Bollerslev, Tim and Jeffrey Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews* **11**, 143–172.

Cipollini, F., R.F. Engle and G.M. Gallo (2006). Vector multiplicative error models: Representation and inference. Technical Report 2006/331. NBER.

Diebold, F. X., T. A. Gunther and A. S. Tay (1998). Evaluating density forecasts, with applications to financial risk management. *International Economic Review* **39**, 863–883.

Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica* **50**, 987–1008.

Engle, R.F. (2000). The econometrics of ultra-high frequency data. *Econometrica* **68**, 1–22.

Engle, R.F. (2002). New frontiers for arch models. *Journal of Applied Econometrics* **17**, 425–446.

Engle, R.F. and G.M. Gallo (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics* **131**, 3–27.

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

Engle, Robert F. and Jeffrey R. Russell (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* **66**, 1127–1162.

Fernandes, M. and J. Grammig (2006). A family of autoregressive conditional duration models. *Journal of Econometrics* **130**, 1–23.

Hall, A. D. and N. Hautsch (2006). Order aggressiveness and order book dynamics. *Empirical Economics* **30**, 973–1005.

Hall, A. D. and N. Hautsch (2007). Modelling the buy and sell intensity in a limit order book market. *Journal of Financial Markets* **10(3)**, 249–286.

Hautsch, N. (2003). Assessing the risk of liquidity suppliers on the basis of excess demand intensities. *Journal of Financial Econometrics* **1(2)**, 189–215.

Hautsch, N. (2004). *Modelling Irregularly Spaced Financial Data - Theory and Practice of Dynamic Duration Models*. Springer, Berlin.

Hautsch, N. (2006). Testing the conditional mean function of autoregressive conditional duration models. Technical Report 2006-06. Finance Research Unit, Department of Economics, University of Copenhagen.

Hautsch, N. (2008). Capturing common components in high-frequency financial time series: A multivariate stochastic multiplicative error model. *Journal of Economic Dynamics and Control, forthcoming*.

Heinen, Andréas and Erick Rengifo (2006). Multivariate autoregressive modelling of time series count data using copulas. *Journal of Empirical Finance* **14**, 564–583.

Lee, S. and B. Hansen (1994). Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory* **10**, 29–52.

Ljung, G. M. and G. E. P. Box (1978). On a measure of lack of fit in time series models. *Biometrika* **65**, 297–303.

Manganelli, Simone (2005). Duration, volume and volatility impact of trades. *Journal of Financial Markets* **8**, 377–399.

Meitz, M. and T. Teräsvirta (2006). Evaluating models of autoregressive conditional duration. *Journal of Business and Economic Statistics* **24**, 104–124.

Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Journal of Econometrics* **43**, 227–251.

Roll, Richard (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance* **39**, 1127–1139.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annals of Mathematical Statistics* **23**, 470–472.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes - a study of daily sugar prices. In: *Time Series Analysis: Theory and Practice* (O. D. Anderson, Ed.). North-Holland, Amsterdam.

Veredas, D., J. Rodriguez-Poo and A. Espasa (2001). On the (intradaily) seasonality, dynamics and durations zero of a financial point process. Technical Report WS013321. Universidad Carlos III, Departamento de Estad´istica y Econometr´ia.

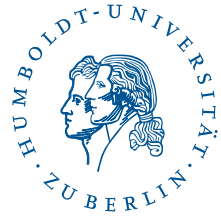
SFB 649 Discussion Paper Series 2008

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Testing Monotonicity of Pricing Kernels" by Yuri Golubev, Wolfgang Härdle and Roman Timonfeev, January 2008.
002 "Adaptive pointwise estimation in time-inhomogeneous time-series models" by Pavel Cizek, Wolfgang Härdle and Vladimir Spokoiny, January 2008.
003 "The Bayesian Additive Classification Tree Applied to Credit Risk Modelling" by Junni L. Zhang and Wolfgang Härdle, January 2008.
004 "Independent Component Analysis Via Copula Techniques" by Ray-Bing Chen, Meihui Guo, Wolfgang Härdle and Shih-Feng Huang, January 2008.
005 "The Default Risk of Firms Examined with Smooth Support Vector Machines" by Wolfgang Härdle, Yuh-Jye Lee, Dorothea Schäfer and Yi-Ren Yeh, January 2008.
006 "Value-at-Risk and Expected Shortfall when there is long range dependence" by Wolfgang Härdle and Julius Mungo, Januray 2008.
007 "A Consistent Nonparametric Test for Causality in Quantile" by Kiho Jeong and Wolfgang Härdle, January 2008.
008 "Do Legal Standards Affect Ethical Concerns of Consumers?" by Dirk Engelmann and Dorothea Kübler, January 2008.
009 "Recursive Portfolio Selection with Decision Trees" by Anton Andriyashin, Wolfgang Härdle and Roman Timofeev, January 2008.
010 "Do Public Banks have a Competitive Advantage?" by Astrid Matthey, January 2008.
011 "Don't aim too high: the potential costs of high aspirations" by Astrid Matthey and Nadja Dwenger, January 2008.
012 "Visualizing exploratory factor analysis models" by Sigbert Klinke and Cornelia Wagner, January 2008.
013 "House Prices and Replacement Cost: A Micro-Level Analysis" by Rainer Schulz and Axel Werwatz, January 2008.
014 "Support Vector Regression Based GARCH Model with Application to Forecasting Volatility of Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang Härdle, January 2008.
015 "Structural Constant Conditional Correlation" by Enzo Weber, January 2008.
016 "Estimating Investment Equations in Imperfect Capital Markets" by Silke Hüttel, Oliver Mußhoff, Martin Odening and Nataliya Zinych, January 2008.
017 "Adaptive Forecasting of the EURIBOR Swap Term Structure" by Oliver Blaskowitz and Helmut Herwatz, January 2008.
018 "Solving, Estimating and Selecting Nonlinear Dynamic Models without the Curse of Dimensionality" by Viktor Winschel and Markus Krätzig, February 2008.
019 "The Accuracy of Long-term Real Estate Valuations" by Rainer Schulz, Markus Staiber, Martin Wersing and Axel Werwatz, February 2008.
020 "The Impact of International Outsourcing on Labour Market Dynamics in Germany" by Ronald Bachmann and Sebastian Braun, February 2008.
021 "Preferences for Collective versus Individualised Wage Setting" by Tito Boeri and Michael C. Burda, February 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".



022 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycles" by Fang Yao, February 2008.
023 "Family Management, Family Ownership and Downsizing: Evidence from S&P 500 Firms" by Jörn Hendrich Block, February 2008.
024 "Skill Specific Unemployment with Imperfect Substitution of Skills" by Runli Xie, March 2008.
025 "Price Adjustment to News with Uncertain Precision" by Nikolaus Hautsch, Dieter Hess and Christoph Müller, March 2008.
026 "Information and Beliefs in a Repeated Normal-form Game" by Dietmar Fehr, Dorothea Kübler and David Danz, March 2008.
027 "The Stochastic Fluctuation of the Quantile Regression Curve" by Wolfgang Härdle and Song Song, March 2008.
028 "Are stewardship and valuation usefulness compatible or alternative objectives of financial accounting?" by Joachim Gassen, March 2008.
029 "Genetic Codes of Mergers, Post Merger Technology Evolution and Why Mergers Fail" by Alexander Cuntz, April 2008.
030 "Using R, LaTeX and Wiki for an Arabic e-learning platform" by Taleb Ahmad, Wolfgang Härdle, Sigbert Klinke and Shafeeqah Al Awadhi, April 2008.
031 "Beyond the business cycle – factors driving aggregate mortality rates" by Katja Hanewald, April 2008.
032 "Against All Odds? National Sentiment and Wagering on European Football" by Sebastian Braun and Michael Kvasnicka, April 2008.
033 "Are CEOs in Family Firms Paid Like Bureaucrats? Evidence from Bayesian and Frequentist Analyses" by Jörn Hendrich Block, April 2008.
034 "JBendge: An Object-Oriented System for Solving, Estimating and Selecting Nonlinear Dynamic Models" by Viktor Winschel and Markus Krätzig, April 2008.
035 "Stock Picking via Nonsymmetrically Pruned Binary Decision Trees" by Anton Andriyashin, May 2008.
036 "Expected Inflation, Expected Stock Returns, and Money Illusion: What can we learn from Survey Expectations?" by Maik Schmeling and Andreas Schrimpf, May 2008.
037 "The Impact of Individual Investment Behavior for Retirement Welfare: Evidence from the United States and Germany" by Thomas Post, Helmut Gründl, Joan T. Schmit and Anja Zimmer, May 2008.
038 "Dynamic Semiparametric Factor Models in Risk Neutral Density Estimation" by Enzo Giacomini, Wolfgang Härdle and Volker Krätschmer, May 2008.
039 "Can Education Save Europe From High Unemployment?" by Nicole Walter and Runli Xie, June 2008.
042 "Gruppenvergleiche bei hypothetischen Konstrukten – Die Prüfung der Übereinstimmung von Messmodellen mit der Strukturgleichungs-methodik" by Dirk Temme and Lutz Hildebrandt, June 2008.
043 "Modeling Dependencies in Finance using Copulae" by Wolfgang Härdle, Ostap Okhrin and Yarema Okhrin, June 2008.
044 "Numerics of Implied Binomial Trees" by Wolfgang Härdle and Alena Mysickova, June 2008.
045 "Measuring and Modeling Risk Using High-Frequency Data" by Wolfgang Härdle, Nikolaus Hautsch and Uta Pigorsch, June 2008.
046 "Links between sustainability-related innovation and sustainability management" by Marcus Wagner, June 2008.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

047 "Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models" by Nikolaus Hautsch and Vahidin Jeleskovic, July 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".