

# Econometric analysis of financial trade processes by mixture duration models

Reinhard Hujer<sup>a,\*</sup>, Sandra Vuletić<sup>b</sup>

<sup>a</sup>*University of Frankfurt/M., IZA Bonn, ZEW Mannheim*

<sup>b</sup>*University of Frankfurt/M.*

Version: 18 October 2004

---

## Abstract

We propose a new framework for modelling the time dependence in duration processes. The well known ACD approach introduced by Engle and Russell (1998) will be extended so that an unobservable stochastic process accompanies the duration process. Our creation is called *Mixture ACD model* (MACD) which puts the conjunction into practice. It is a moderate tool for description of financial duration processes. The introduction of a latent regime variable can be justified in the light of recent market microstructure theories. In an empirical application we show that the MACD approach is able to capture specific characteristics of intraday transaction durations while alternative ACD models fail.

*Key words:* Duration models, time series models, mixture models, financial transaction data, market microstructure.

*JEL classification:* C41, C22, C25, C51, G14.

---

## 1 Introduction

Investigating the microstructure of financial markets has become very popular over the last twenty years. Theoretical assertions concerning the behavior of market participants in the presence of asymmetric information are discussed

---

\* Corresponding author. Johann Wolfgang Goethe-University, Department of Economics and Business Administration, Institute of Statistics and Econometrics, Meritonstrasse 17, 60054 Frankfurt on the Main, Germany. Tel.: +49 69 798 28115; Fax: +49 69 798 23673. E-mail: hujer@wiwi.uni-frankfurt.de (R. Hujer).

<sup>1</sup> The authors thank Dr. Stefan Kokot for valuable preparatory work.

in many contributions. In this respect Easley, Kiefer, O'Hara, and Paperman (1996) deliver a prominent approach. Statistical methodology will be employed in order to check empirically the validity of the implications of market microstructure models. Since rich transaction data sets are available containing detailed information about the timing of trades, prices, volume and other relevant characteristics for a wide range of financial securities, it is possible to get to the bottom of financial markets. Theory and the application of a tailor-made statistical instrument are combined in the elaboration of Kokot (2004).

New econometric methods appear rapidly and they experience an extensive application in the branch of finance. The autoregressive conditional duration model (ACD) introduced by Engle and Russell (1998) is an auspicious approach which couples the spirit of time series models with econometric tools for the analysis of transition data. Ultra high frequency data, stemming from transaction data sets and having the characteristic of irregular spacing in time, are ideal actuality for the use of the innovative framework. The ACD model is perfectly suitable for the analysis of dynamics of arbitrary events associated with the trading process along time, and the durations between successive occurrences of interesting market events are object of investigation.

As demonstrated by Bauwens, Giot, and Grammig (2000) the periods of time elapsing between successive trades exhibit an idiosyncrasy which could not even be captured by extensions of the original model. For the first time the flexible Markov switching ACD model developed by Hujer, Vuletić, and Kokot (2002) is capable of higher forecast accuracy of the trading process itself, but it requires much effort and computing power in estimation. We intend to introduce an alternative model with a parsimonious parameterization, called the *Mixture* ACD model (MACD), which also attains to good performance. Integral part of the MACD model is a latent discrete valued regime variable whose involvement can be justified in the light of recent market microstructure models. The unobservable regime can be associated with the presence (or absence) of private information about an asset's value that is initially available exclusively to a subset of informed traders and only eventually disseminates through the mere process of trading to the broader public of all market participants.

The manageable MACD model bears a resemblance to the general switching autoregression model introduced by Hamilton (1989) and nests many of the existing autoregression duration models as special cases. There are several

models that are closely related to our approach as well. Despite the affinity to the duration model given by De Luca and Gallo (2004), the MACD model differs substantially in the distributional assumption. It has the discrete mixture in common with the threshold ACD model introduced by Zhang, Russell, and Tsay (2001).

This paper is structured as follows: A brief review of the idea of ACD modeling is given in Section 2. In Section 3 the MACD model will be introduced and compared to related work on duration models. Moreover we discuss estimation procedures and specification tests for MACD models. In an empirical application in Section 4 we present estimation results employing a transaction data set for the common share of Boeing traded on the New York Stock Exchange. Finally, in Section 5 we summarize our main results and give a perspective on possible issues for future research.

## 2 The ACD model

Autoregressive conditional duration (ACD) models, introduced by Engle and Russell (1998), are designed to account for patterns of autocorrelation typically observed in time series of intervals between successive occurrences of market events associated with the trading process. The definition of the market event depends on the specific aim of the study.

Let  $x_n = t_n - t_{n-1}$  be the duration between the recordings of the  $(n-1)$ -th and the  $n$ -th market event with the deterministic conditional mean function

$$\psi_n = E(x_n | \mathcal{F}_{n-1}; \theta_\psi), \quad (2.1)$$

where the information set  $\mathcal{F}_{n-1}$  consists of all preceding durations up to time  $t_{n-1}$  and  $\theta_\psi$  is the corresponding set of parameters. The ACD model is defined by some parameterization of this conditional mean and by the decomposition

$$\varepsilon_n = \frac{x_n}{\psi_n}, \quad (2.2)$$

where the residual process  $\varepsilon_n$  is assumed to be i. i. d. with density  $g(\varepsilon_n; \theta_\varepsilon)$  depending on a set of distributional parameters  $\theta_\varepsilon$ , support on the positive real line and an unconditional expectation equal to one. The flexibility of the ACD model can be altered by modifying the distributional assumption of the residuals and/or by changing the specification of the conditional mean function. The distributional assumption of the residuals determines the density

of the durations  $f_n(x_n | \mathcal{F}_{n-1}; \theta)$ , where  $\theta = (\theta_\psi, \theta_\varepsilon)$  represents the whole parameter set. A list of common choices for  $g(\varepsilon_n; \theta_\varepsilon)$  includes the exponential, the Weibull, the Burr (1942), and the generalized gamma distribution, all of them nested in the comprehensive family of distributions.

In a standard ACD( $p, q$ ) model the parameterization of the conditional mean is linear according to

$$\psi_n = \omega + \sum_{k=1}^p \beta_k \cdot \psi_{n-k} + \sum_{k=1}^q \alpha_k \cdot x_{n-k}, \quad (2.3)$$

and it can be transformed into an ARMA( $\max(p, q), p$ ) representation from which expressions for the unconditional moments of  $x_n$  may be derived easily. In order to ensure non-negativity for the conditional mean the parameters  $\omega$ ,  $\alpha_k$ , and  $\beta_k$  are forced to be non-negative. Computational problems due to this strong restriction may be circumvented by using logarithmic versions. Bauwens and Giot (2002) propose the following LACD( $p, q$ ) specification

$$\ln(\psi_n) = \omega + \sum_{k=1}^p \beta_k \cdot \ln(\psi_{n-k}) + \sum_{k=1}^q \alpha_k \cdot \ln(x_{n-k}) \quad (2.4)$$

and the corresponding analytical expressions for the unconditional moments are given by Bauwens, Galli, and Giot (2003). In both specifications stationarity depends on the magnitudes of the parameters  $\alpha_k$ , and  $\beta_k$ .

### 3 The Mixture ACD model

#### 3.1 The basic framework

The basic assumption of the *Mixture ACD model*, also referred to as MACD, is that the duration process  $x_n$  is accompanied by an unobservable stochastic process  $s_n$ . The stochastic process  $s_n$  is characterized by a discrete valued random variable with countable support  $\mathfrak{J} = \{j \mid 1 \leq j \leq J, J \in \mathbb{N}\}$  and has the task to represent the regime in which the duration process  $x_n$  prevails at time  $t_n$ . In financial applications the existence of different trading regimes may provide evidence on the presence of agents with private information about an asset's value.

Decomposition (2.2) holds in the sense that the innovation process  $\varepsilon_n$  has a known discrete mixture distribution with  $E(\varepsilon_n) = 1$  and invariant higher moments across the  $N$  observations considered in the sample. The density of

each innovation  $\varepsilon_n$  has the following formal appearance

$$g(\varepsilon_n; \theta_\varepsilon, \theta_\pi) = \sum_{j=1}^J \pi^{(j)} \cdot g(\varepsilon_n \mid s_n = j; \theta_\varepsilon^{(j)}), \quad (3.1)$$

where each weight  $0 \leq \pi^{(j)} \leq 1$  represents the corresponding long run probability for prevailing in state  $j$  and  $\theta_\varepsilon^{(j)}$  is the corresponding parameter vector characterizing the conditional density of the innovation process driven in the  $j$ -th regime. Consequently, the unconditional density of the innovation process as given in equation (3.1) depends on all regime specific distributional parameters gathered into the vector  $\theta_\varepsilon = (\theta_\varepsilon^{(1)}, \dots, \theta_\varepsilon^{(J)})'$  and on  $\theta_\pi = (\pi^{(1)}, \dots, \pi^{(J)})'$ . Any of the densities mentioned in Section 2 may be used in order to specify the regime specific distributions of the innovation process. De Luca and Gallo (2004) build up a duration model where the innovation process follows the Schuhl distribution, being simply a discrete mixture of exponential distributions. The MACD model can be recognized as a generalization which allows for more flexibility.

On the one hand the expected value of each innovation  $E(\varepsilon_n)$  is constrained to be equal to one and on the other hand this expected value turns out to be a discrete mixture of regime specific expectations  $E(\varepsilon_n \mid s_n = j; \theta_\varepsilon^{(j)})$ . This implies the maintenance of the equality

$$1 = \sum_{j=1}^J \pi^{(j)} \cdot E(\varepsilon_n \mid s_n = j; \theta_\varepsilon^{(j)}) \quad (3.2)$$

which does not require that all the regime specific expectations are equal to one. In the case of  $E(\varepsilon_n \mid s_n = j; \theta_\varepsilon^{(j)}) = 1$  for all  $j \in \mathcal{J}$ , the MACD model coincides with a special case of the static variant of the Markov switching ACD model developed by Hujer, Vuletić, and Kokot (2002).

By the change of variable technique with  $x_n = \varepsilon_n \cdot \psi_n$ , the relevant density for statistical inference is the duration's marginal density

$$f_n(x_n \mid \mathcal{F}_{n-1}; \theta) = \sum_{j=1}^J \pi^{(j)} \cdot f_n(x_n \mid s_n = j; \theta_\varepsilon^{(j)}, \theta_\psi) \quad (3.3)$$

which depends on the parameter vector  $\theta = (\theta_\varepsilon, \theta_\psi, \theta_\pi)'$ . The mean function  $\psi_n = E(x_n \mid \mathcal{F}_{n-1}; \theta_\psi)$  is assumed to capture the whole persistence of the duration process by an appropriate recursion.

Note, that the MACD model does not allow for different regime specific

mean functions. This feature may induce scathing criticism, especially from a theoretical point of view. But the empirical experience with strongly restricted Markov switching ACD models can be used as a vindicative argument. Hujer, Kokot, and Vuletić (2003) conclude that even the static variant of the Markov switching ACD model with regime independent dynamics in the mean function and regime specific distributional parameters performs reasonably well in terms of forecast accuracy.

### 3.2 Estimation of the Mixture ACD model

For discrete mixture models there are two ways by which maximum likelihood estimates of the parameter vector  $\theta$  may be obtained. The direct numerical maximization of the incomplete log-likelihood function

$$\mathcal{L}_I(\theta) = \sum_{n=1}^N \ln [f_n(x_n \mid \mathcal{F}_{n-1}; \theta)] \quad (3.4)$$

under the linear constraint  $\sum_{j=1}^J \pi^{(j)} = 1$  and additional restrictions for non-negativity, stationarity and eventually for distributional parameters is the standard approach. Log-likelihood functions of mixture models are characterized by the existence of multiple local maxima. In order to catch the global maximum, the repetition of the parameter estimation with different start values is strongly recommended. Since standard maximization algorithms often fail or produce nonsensical results, maximum likelihood estimates for discrete mixture models are often obtained by the use of the robust Expectation-Maximization (EM) algorithm introduced by Dempster, Laird, and Rubin (1977).

In the hypothetical situation where we can observe the realizations of the regime variable the complete log-likelihood function is given by

$$\mathcal{L}_C(\theta) = \sum_{n=1}^N \sum_{j=1}^J z_n^{(j)} \left( \ln[f_n(x_n \mid s_n = j, \mathcal{F}_{n-1}; \theta_\varepsilon^{(j)}, \theta_\psi)] + \ln[\pi^{(j)}] \right), \quad (3.5)$$

where  $z_n^{(j)} = 1$  if  $s_n = j$  and zero otherwise. The expectation of (3.5) conditional on all observed data  $\mathcal{X}_N = (x_1, \dots, x_N)$  leads to the expected complete log-likelihood function  $\mathcal{L}_{EC}(\theta, \theta_0) = E(\mathcal{L}_C(\theta) \mid \mathcal{X}_N; \theta_0)$  which is simply obtained by replacing  $z_n^{(j)}$  by the probabilistic inference

$$\xi_{n|n}^{(j)} = \frac{\pi_0^{(j)} f_n(x_n | s_n = j, \mathcal{F}_{n-1}; \theta_{0\varepsilon}^{(j)}, \theta_{0\psi})}{\sum_{k=1}^J \pi_0^{(k)} f_n(x_n | s_n = k, \mathcal{F}_{n-1}; \theta_{0\varepsilon}^{(k)}, \theta_{0\psi})} \quad (3.6)$$

evaluated for some parameter vector guess  $\theta_0$ . Evaluation of  $\mathcal{L}_{EC}(\theta, \theta_0)$  constitutes the first part of the EM-algorithm and is commonly referred to as the E-step. The associated M-step consists of maximizing  $\mathcal{L}_{EC}(\theta, \theta_0)$  in respect of the parameter vector  $\theta$  and can be conducted separately with respect to the regression parameters and the regime probabilities if  $\frac{\partial f_n(x_n | s_n = j, \mathcal{F}_{n-1}; \theta_{\varepsilon}^{(j)}, \theta_{\psi})}{\partial \pi^{(k)}} = 0$  for all  $j, k \in (1, \dots, J)$ . The estimates for the regime probabilities are given by

$$\hat{\pi}^{(j)} = \frac{1}{N} \sum_{n=1}^N \xi_{n|n}^{(j)} \quad (3.7)$$

and the remaining parameters may be obtained from the solution to

$$\sum_{n=1}^N \sum_{j=1}^J \xi_{n|n}^{(j)} \cdot \left( \frac{\partial \ln f_n(x_n | s_n = j, \mathcal{F}_{n-1}; \theta_{\varepsilon}^{(j)}, \theta_{\psi})}{\partial \theta} \right) = 0. \quad (3.8)$$

By repeating the two steps of the EM-algorithm until the absolute change of the parameter vector is smaller than some prespecified convergence criterion, estimates of the parameter vector are obtained. Hamilton (1990) shows that the final estimates  $\hat{\theta}$  maximize the incomplete log-likelihood function.

### 3.3 Statistical inference

Diebold, Gunther, and Tay (1998) propose a method which can be applied to test the forecast performance of general dynamic models. The idea behind this specification test has been extensively used by Bauwens, Giot, and Grammig (2000) to compare different types of ACD models. Denote by  $\{f_n(x_n | \mathcal{F}_{n-1}; \hat{\theta})\}_{n=1}^N$  the sequence of density forecasts evaluated using the parameter vector estimate  $\hat{\theta}$  from some parametric model and denote by  $\{f_n(x_n | \mathcal{F}_{n-1}; \theta)\}_{n=1}^N$  the sequence of densities corresponding to the true but unobservable data generating process of  $x_n$ . As shown by Rosenblatt (1952), under the null hypothesis

$$H_0 : \{f_n(x_n | \mathcal{F}_{n-1}; \hat{\theta})\}_{n=1}^N = \{f_n(x_n | \mathcal{F}_{n-1}; \theta)\}_{n=1}^N, \quad (3.9)$$

the sequence of empirical integral transforms defined by

$$\hat{\zeta}_n = \int_{-\infty}^{x_n} f_n(u \mid \mathcal{F}_{n-1}; \hat{\theta}) du \quad (3.10)$$

will be uniform i.i.d. on the unit interval. Any test for uniformity of the sequence of integral transforms can be used to assess the forecast performance of the model under consideration. Consider partitioning the support of  $\zeta_n$  into  $K$  equally spaced bins and denote the number of observations falling into the  $k$ -th bin by  $N_k$ . The test statistic  $RT_\zeta$

$$RT_\zeta = -2 \cdot \sum_{k=1}^K N_k \cdot \ln \left[ \frac{\varsigma_k}{\hat{\varsigma}_k} \right] \quad (3.11)$$

compares the theoretical frequency  $\varsigma_k = \frac{1}{K}$  to the observed relative frequency  $\hat{\varsigma}_k = \frac{N_k}{N}$  and has a  $\chi^2$  distribution with  $(K - 1)$  degrees of freedom under the null hypothesis. The independence feature may be checked by computing the Ljung and Box (1978) test for the sequence of empirical integral transforms. The statistical tests for i. i. d. uniformity may be supplemented by graphical tools. Departures from uniformity can easily be detected using a histogram plot or quantile-quantile plot based on the sequence of  $\hat{\zeta}_n$ , while the autocorrelogram for  $\hat{\zeta}_n$  can be used in order to assess the independence property.

### 3.4 *Link to microstructure models*

The modern literature on the microstructure of financial markets, gradually widening in the style of Easley, Kiefer, O'Hara, and Paperman (1996), picks out the presence of diverse types of market participants (traders) as a central theme. The intercommunity of the broad literature is the initial position that the market participants are differentiated by the level of information which they harness privately and consequently the trading mechanism will be discussed under the aspect of asymmetric information. Concerning this matter it is easy to imagine that some traders exist who catch a signal indicating that an asset is either overpriced or underpriced while other traders do not notice anything. So, the market development can be easily characterized by the coexistence and interaction of just two categories of traders: informed traders and uninformed traders, also called liquidity traders or followers. The informed trader's strategy consists of making purchases and sales of assets in the immediate aftermath of the recognition of favorable and unfavorable signals. The informed traders encroach upon the market development conjunctly



and trigger heaped transactions as soon as they bushwhack relevant news. Uninformed traders are insensible in regard to the information processing and retain the habitual trading activity.

The collectivity of transactions, carried out either by the large attendance of uninformed traders or by sporadic emersions of informed traders as a result of information based decisions, can be seen as a realization of a point process and the corresponding probability law that governs the occurrence of trades can be specified by a duration statistic. The presence of different traders acting on the financial market makes the embedding of a conglomerate of trader specific characteristics into the ordinary ACD framework adjacent. Because a specific transaction does not reveal by which type of trader it has been induced, the introduction of an underlying unobservable mixing variable with discrete distribution is reasonable. The mixing parameters represent the corresponding probabilities that a transaction arises from a specific type of trader.

This simple theoretical background is excellently reflected in the MACD framework which bases upon an arbitrary mixture distribution for the stochastic process of innovations. Thereby the regime variable is in the capacity of the mixing variable and the mixing parameters can be interpreted as fractions of the different trader types acting on the market. The level of discrepancy between trader specific peculiarities in trading behavior can be easily regulated by adapting the parameters inside of equation (3.2). The instantaneous transaction rates turn out to be different across the trader categories and this is what we want to achieve primarily.

Bauwens, Giot, and Grammig (2000) report on the deficiency of ordinary ACD models which is well founded by the inability of modelling observations in the tails of their distributions appropriately. This arouses the suspicion that the duration process is mulcted of some facts with fundamental importance. The thoughts stimulated by the market microstructure theory justify an advanced approach for duration data which is materialized in the concise MACD framework. By doing this, we hope to succeed in overcoming the lack of satisfactory forecast performance of ordinary ACD models and we expect a clear answer from the empirical application given in the following section.

## 4 Empirical application

### 4.1 The data set

The data used in our empirical application consists of transactions of the common stock of Boeing, recorded on the New York stock exchange from the trades and quotes database provided by the NYSE Inc. The sampling period spans 19 trading days from November 1 to November 27, 1996. We used all trades observed during the regular trading day (9:30 - 16:00). The trading times have been recorded with a precision measured in seconds. Observations occurring within the same second have been aggregated to one trade. In the final data set we removed censored observations: durations from the last trade of the day until the close and durations from the open until the first trade of the day.

It is well known that the length of the durations varies in a deterministic manner during the trading day that resembles an inverted U-shaped pattern. Engle and Russell (1997) propose to decompose the duration series into a deterministic time of day function  $\Phi(t_{n-1})$  and a stochastic component  $x_n$ , so that the raw durations are generated from  $\tilde{x}_n = x_n \cdot \Phi(t_{n-1})$ . In order to remove the deterministic component we apply the two step method proposed by Engle and Russell (1997) in which the time of day function is estimated separately from other model parameters.<sup>2</sup> Dividing each raw duration  $\tilde{x}_n$  in the sample by an estimate of the time of day function  $\Phi(t_{n-1})$ , a sequence of deseasonalized durations  $x_n$  is obtained which is used in all subsequent analyses.<sup>3</sup>

Descriptive information about sample moments and Ljung Box statistics of the raw and the seasonally adjusted duration data is reported in Table 1.

< insert Table 1 about here >

---

<sup>2</sup> Simultaneous ML-estimation as in Engle and Russell (1998) and Veredas et al. (2002) is also feasible. Engle and Russell (1998) report that both procedures give similar results if sufficient data is available.

<sup>3</sup> Estimates of the time of day function were obtained by conducting a semi-nonparametric regression of the durations on the time of day according to Gallant (1981) and Eubank and Speckman (1990). Details on the seasonality adjustment step are available from the authors upon request.

As expected, the series of adjusted durations has a mean of approximately one. Both time series exhibit overdispersion relative to the exponential distribution which has standard error equal to mean. A mixture of distributions will accommodate well to the stylized fact of overdispersion. Another eye-catching characteristic of the data is the presence of strong positive autocorrelation in the trade durations as can be seen in Figure 1.

< insert Figure 1 about here >

Even after seasonal adjustment, the Ljung-Box tests reject the hypothesis of no autocorrelation up to 50 lags at the 5% significance level, although the shape of the autocorrelation function changes slightly. Therefore, an autoregressive approach appears to be appropriate as a model for the transaction durations.

#### 4.2 Specification of the Mixture ACD Model

We estimate an ordinary ACD model and also two contrastable specifications of the MACD model with consideration of two regimes, i. e.  $J = 2$ . The mean function  $\psi_n$  is logarithmic and both lag orders  $p$  and  $q$  in the recursion are equal to one, i. e.

$$\psi_n = \exp(\omega) \cdot \psi_{n-1}^{\beta_1} \cdot x_{n-1}^{\alpha_1}. \quad (4.1)$$

Concerning the demand for a unit mean of the innovation process  $\varepsilon_n$  we distinguish between two different cases. The restrictive variant, denoted by the character  $R$  in the following, comprises the fact that all regime specific expectations of the innovation process  $E(\varepsilon_n | s_n = j; \theta_\varepsilon^{(j)})$  are forced to be equal to one, so that absolutely no care for equation (3.2) is needed. This variant may be estimated by employing the EM-algorithm, while the nonrestrictive variant, denoted by the character  $\bar{R}$  in the following, has to be estimated by maximizing the incomplete log-likelihood function directly.

Each regime specific distribution of the innovation process  $\varepsilon_n | s_n = j$  is taken from the Burr (1942) family of distributions with regular time-invariant distributional parameters  $\kappa^{(j)}$  and  $\sigma^{(j)}$  which are associated with each of the two regimes of interest. The introduction of additional time-invariant distributional parameters  $\mu^{(j)}$  is considered in the nonrestrictive case where the equality

$$\sum_{j=1}^J \pi^{(j)} \cdot [\mu^{(j)}]^{-\frac{1}{\kappa^{(j)}}} \cdot \frac{\Gamma\left(1 + \frac{1}{\kappa^{(j)}}\right) \cdot \Gamma\left(\frac{1}{\sigma^{(j)}} - \frac{1}{\kappa^{(j)}}\right)}{\sigma^{(j)\left(1 + \frac{1}{\kappa^{(j)}}\right)} \cdot \Gamma\left(\frac{1}{\sigma^{(j)}} + 1\right)} = 1 \quad (4.2)$$

has to be ensured in the course of estimation. Because of the need to consider two constrictive facts in estimation, i. e. the sum of all regime probabilities is equal to one and the requirement given in (4.2), one has to estimate  $\mu^{(1)}$  and  $\pi^{(1)}$  beside the regular distributional parameters and the parameters of the mean function only. In contrast, the restrictive case incorporates corresponding distributional parameters which obey a parameterization according to

$$\mu^{(j)} = \left[ \frac{\sigma^{(j)\left(1 + \frac{1}{\kappa^{(j)}}\right)} \cdot \Gamma\left(\frac{1}{\sigma^{(j)}} + 1\right)}{\Gamma\left(1 + \frac{1}{\kappa^{(j)}}\right) \cdot \Gamma\left(\frac{1}{\sigma^{(j)}} - \frac{1}{\kappa^{(j)}}\right)} \right]^{-\kappa^{(j)}} \quad (4.3)$$

so that they are exempted from estimation. This parameter determination implies that each regime specific expectation of the innovation process is equal to one. Bringing together the restrictive and the unrestrictive variant, each  $\varepsilon_n | s_n = j$  follows the Burr distribution with the three distributional parameters  $\mu^{(j)}$ ,  $\kappa^{(j)}$ ,  $\sigma^{(j)}$  and the regime specific density of the duration  $x_n$  turns out to be

$$f_n(x_n | s_n = j, \mathcal{F}_{n-1}; \theta_\varepsilon^{(j)}, \theta_\psi) = \frac{\mu_n^{(j)} \cdot \kappa^{(j)} \cdot x_n^{\kappa^{(j)}-1}}{\left(1 + \sigma^{(j)} \cdot \mu_n^{(j)} \cdot x_n^{\kappa^{(j)}}\right)^{\frac{1}{\sigma^{(j)}}+1}} \quad (4.4)$$

with the time variant parameter  $\mu_n^{(j)} = \psi_n^{-\kappa^{(j)}} \cdot \mu^{(j)}$ . Regardless to the inner constitution of equation (4.2) which makes room for the restrictive and unrestrictive variant of the MACD model, the regime specific distributions of a selective duration  $x_n$  turn out to be entirely different. Even the restrictive variant which implies  $E(x_n | s_n = j, \mathcal{F}_{n-1}; \theta_\varepsilon^{(j)}, \theta_\psi) = \psi_n$  for every regime  $j \in \mathcal{J}$ , gives leeway to different regime specific distributional features, i. e. the first moment of  $x_n$  is fix across all regimes but all higher moments are regime variant. The unrestrictive variant provides a cut above in the sense that all moments are allowed to be regime specific. But for all that, both specification variants definitively imply the fact  $E(x_n | \mathcal{F}_{n-1}; \theta) = \psi_n$ . An interesting issue becoming apparent is whether the restrictive variant is sufficiently flexible to catch regime specific characteristics hidden in the duration process.

But first of all, we attend to the topic concerning the outclassing perfor-

mance of MACD models. Coming from the standard ACD approach there is no exorbitant increase in the number of parameters composing a MACD model. In comparison to the ordinary Burr ACD model, the corresponding two regime MACD model, which conforms to the instruction of the  $R$  variant ( $\bar{R}$  variant), requires the estimation of three (four) additional parameters only.

Parameter estimates, standard errors<sup>4</sup>, values of the log-likelihood function and information criterion; descriptive statistics for the series of empirical integral transforms and p-values of statistical tests for the corresponding parameters being equal to the population counterpart implied by the uniform distribution on the unit interval; and also results of the specification tests for all of the model specifications we estimated are presented in Table 2.

< insert Table 2 about here >

At first, the Bayesian information criterion  $BIC$  proposed by Schwarz (1978) does not support the ordinary logarithmic ACD model which is nested as a special case in the MACD framework with logarithmic mean specification and  $J = 1$ . The test on the mean argues for the null hypothesis  $\bar{\zeta} \equiv E(\zeta_n) = 0.5$ , but the result of the variance test is not in favor of the null hypothesis  $\sigma_\zeta^2 \equiv Var(\zeta_n) = \frac{1}{12}$ . The low p-values obtained from the quantile tests are a sign of bad adaption in the tail of the distribution. Moreover, the specification test that we performed does not support the one regime model. This can be seen from the low p-value of the ratio test which is equal to zero. Hence, the apparent defect of the ordinary logarithmic ACD model stems from the improper choice of distribution for the innovation process. However, the ordinary logarithmic ACD model is able to capture the autocorrelation pattern of the intertrade durations adequately as indicated by the high p-value of the Ljung Box statistic for the series of empirical integral transforms.

The present results for proper mixture models indicate a significant improvement on the performance of the ordinary logarithmic ACD model. They admit for the general conclusion: for  $J$  greater than one, first order MACD models are able to eliminate the distributional problem of ordinary ACD models and the autocorrelation pattern in the duration data will be considered

---

<sup>4</sup> Standard errors have been computed based on numerical derivatives of the incomplete log likelihood function using the quasi - maximum likelihood estimates of the information matrix as suggested by White (1982).

adequately. Even the two regime case breeds best results, as can be seen from the last four columns of Table 2. For each of the two variants we estimated, the arithmetic mean of the empirical integral transforms, denoted by  $\bar{\zeta}$ , draws near one half, the corresponding empirical variance  $s_{\bar{\zeta}}^2$  becomes significantly one twelfth and the first, second and third quartile does not differ significantly from 0.25, 0.50 and 0.75 - these facts express the extraordinary conformance to the uniform distribution on the unit interval. The p-values of the  $RT_{\zeta}$  test increase by leaps and bounds, they rise to over 10%. The hypothesis of no autocorrelation in the integral transforms will be statistical significant at conventional significance levels.

For purposes of comparison Figure 2 contains histogram plots and QQ - plots for the series of integral transforms for the 1-regime and the nonrestrictive 2-regime model specification.

< insert Figure 2 about here >

The plots clearly show that the estimated MACD model produces empirical integral transforms that match the implied theoretical density very well and tends to give accurate forecasts over the whole range of observed values of  $x$ . In contrast, the plots for the one regime model show that the empirical integral transforms disagree sharply with the theoretical density, and that it tends to produce systematically biased forecasts of small  $x$ , the histogram for the first four quantiles is outside of the 95% confidence interval.

The parameter estimates for  $\omega$ ,  $\alpha_1$  and  $\beta_1$ , which determine the evolution of the duration's conditional mean in time, differ only marginally across the three models we estimated totally. The same may be noticed for the distributional parameters. The estimation results obtained from the multiple regime models show that the two regular distributional parameters  $\kappa^{(j)}$  and  $\sigma^{(j)}$  vary keenly across the regimes, each with larger value in the second regime than in the first. This has a strong impact on the shape of the hazard function considered for each regime separately. For both variants of the MACD model Figure 3 displays the two regime specific hazard functions  $\lambda_n(x_n|s_n = j, \mathcal{F}_{n-1}; \hat{\theta})$  and also the regime unspecific hazard rate  $\lambda_n(x_n|\mathcal{F}_{n-1}; \hat{\theta})$ , each evaluated for  $\psi_n = 1$  and by taking the parameter vector estimate into account.

< insert Figure 3 about here >

Note in the first instance, that the choice for the one or the other variant does not change the qualitative nature of the hazard rates. The hazard rate assigned to the second regime tends to rise rather quickly after a transaction has been observed. In contrast the hazard function under the first regime increases moderately and gives clearly more weight to spells with a length of more than two units of time. This corresponds nicely to the fact that the first regime has higher probability  $\pi^{(1)}$  than the second regime. Roughly three fourths of all transactions were generated in the first regime.

So, the application of the MACD model affirms the existence of two constitutively different streams governing the process of intertrade durations and visualizes the different velocities from which trading evolves. The inertial trading activity, adumbrated by the hazard rate of the first regime, predominates the whole trading process and can be associated with the theoretical vision of trading behavior ascribed to the uninformed traders. The second regime awards the image of succinct trading which can be traced back to informed traders participating on the financial market.

## 5 Conclusions

Mixture models are frequently used in econometrics. This motivates us to combine the basic idea of mixture models with the art of ACD modelling originally introduced by Engle and Russell (1998). The fusion is realized by the Mixture ACD model (MACD) which we present, challenge and put to the test in this paper. We can conclude from our research work that the MACD model turns out to be a promising new framework for modelling autocorrelated durations obtained from high frequency data sets from stock and foreign exchange markets.

In the first instance the MACD model emerges as a successful tool for forecasting time series of intraday transaction durations, as such it is able to remove the distributional problem from which ordinary ACD models occasionally suffer. Since the pompous Markov switching ACD model of Hujer, Vuletić, and Kokot (2002) and the slender discrete mixture exponential ACD model of De Luca and Gallo (2004) are seen as rivals, the creation of the MACD model can be recognized as a compromise solution between the two extremes. As a smart generalization the MACD model enhances the prestige of the discrete

mixture exponential ACD model and as a manageable special case it belittles the pride of the Markov switching ACD model. The amount of flexibility of the MACD model can be regulated in four directions: the number of regimes, the regime specific distributional assumptions, the mean function and finally the condition for unit mean in the residual process are starting points for altering the comprehension.

A further asset of the MACD model is its interpretation in the context of recent market microstructure models. The weights  $\pi^{(j)}$  can be perspicuously regarded as fractions of informed and uninformed traders acting on the financial market, but the imagination of constant proportions all along the time may be questionable. Therefore, an interesting extension of the MACD model would be to make it fit for time varying regime probabilities.



## References

- Bauwens, L., Galli, F., Giot, P., 2003. The moments of log-ACD models. Discussion Paper 11, CORE, Université Catholique de Louvain.
- Bauwens, L., Giot, P., 2002. The logarithmic ACD model: An application to the bid-ask quote process of three NYSE stocks. *Annales d'Economie et Statistique* 60, 117–149.
- Bauwens, L., Giot, P., Grammig, J. Veredas, D., 2000. A comparison of financial duration models via density forecasts. Discussion Paper 60, CORE, Université Catholique de Louvain and University of Frankfurt, forthcoming in: *International Journal of Forecasting*.
- Burr, I. W., 1942. Cumulative frequency functions. *Annals of Mathematical Statistics* 13, 215–232.
- De Luca, G., Gallo, Giampiero, M., 2004. Mixture processes for financial intradaily durations. *Studies in Nonlinear Dynamics & Econometrics* 8 (2), 1–18, article 8.
- Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* 39, 1–38, series B.
- Diebold, F. X., Gunther, T. A., Tay, A. S., 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review* 39 (4), 863–883.
- Easley, D., Kiefer, N., O'Hara, M., Paperman, J. P., 1996. Liquidity, information and infrequently traded stocks. *Journal of Finance* 51 (4), 1405–1436.
- Engle, R. F., Russell, J. R., 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.
- Engle, R. F., Russell, J. R., 1998. Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* 66 (5), 1127–1162.
- Eubank, R. L., Speckman, P., 1990. Curve fitting by polynomial-trigonometric regression. *Biometrika* 77 (1), 1–9.
- Gallant, A. R., 1981. On the bias in flexible functional forms and an essentially unbiased form. *Journal of Econometrics* 20 (2), 285–323.
- Hamilton, J. D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57 (2), 357–384.
- Hamilton, J. D., 1990. Analysis of time series subject to changes in regime. *Journal*

- of Econometrics 45, 39–70.
- Hujer, R., Kokot, S., Vuletić, S., 2003. Comparison of MSACD models, Johann Wolfgang Goethe University, Frankfurt am Main.
- Hujer, R., Vuletić, S., Kokot, S., 2002. The markov switching ACD model. Working Paper Series: Finance and Accounting 90, University of Frankfurt.
- Kokot, S., 2004. The Econometrics of Sequential Trade Models : Theory and Applications Using High Frequency Data. Springer Verlag.
- Ljung, G. M., Box, G. E. P., 1978. On a measure of lack of fit in time series models. *Biometrika* 65 (2), 297–303.
- Rosenblatt, M., 1952. Remarks on a multivariate transformation. *Annals of Mathematical Statistics* 23 (3), 470–472.
- Schwarz, G., 1978. Estimating the dimension of a model. *Annals of Statistics* 6 (2), 461–464.
- Veredas, D., Rodriguez-Poo, J., Espasa, A., 2002. On the intradaily seasonality and dynamics of a financial point process: A semiparametric approach. Discussion Paper 23, CORE, Université Catholique de Louvain.
- White, H., 1982. Maximum likelihood estimation of misspecified models. *Econometrica* 50 (1), 1–25.
- Zhang, M. Y., Russell, J. R., Tsay, R. S., 2001. A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics* 104 (1), 179–207.

## Tables

Table 1  
Descriptive Statistics for trade durations

Statistic	Raw durations $\tilde{x}_n$	Adj. durations $x_n$
Arithmetic mean	48.3248	1.0007
Standard deviation	61.8416	1.1933
Minimum	1.0000	0.0141
First Quartile	10.0000	0.2323
Median	27.0000	0.5875
Third Quartile	61.0000	1.2980
Maximum	894.0000	16.1672
Sample size	9092	9092
Ljung Box statistic <sup>a</sup>	3815.6633	1362.7593

<sup>a</sup> The Ljung Box statistic is based on 50 lags. For a significance level of 5% the tabulated critical value is 67.1671.

Table 2  
Estimation results and specification tests

Parameter	Ordinary		$R$ variant		$\bar{R}$ variant	
	Estimate	Stderr	Estimate	Stderr	Estimate	Stderr
$\omega$	0.0147	0.0021	0.0184	0.0028	0.0139	0.0020
$\alpha_1$	0.0248	0.0035	0.0224	0.0032	0.0233	0.0033
$\beta_1$	0.9715	0.0047	0.9725	0.0045	0.9713	0.0047
$\mu^{(1)}$	-	-	-	-	0.9241	0.0438
$\kappa^{(1)}$	1.1699	0.0182	1.4220	0.1801	1.4652	0.0513
$\kappa^{(2)}$	-	-	2.7822	0.0448	2.4410	0.1546
$\sigma^{(1)}$	0.3333	0.0284	0.3542	0.1750	0.3887	0.0425
$\sigma^{(2)}$	-	-	2.5387	0.0395	1.5921	0.2256
$\pi^{(1)}$	-	-	0.7252	0.0228	0.7238	0.0249
$N$	9092.00		9092.00		9092.00	
$\mathcal{L}_I$	-8691.82		-8518.40		-8510.35	
$BIC$	17429.21		17109.73		17102.75	
$\hat{\zeta}, p(T_{E(\zeta)})$	0.4963	0.2217	0.4984	0.5972	0.4994	0.8429
$s_{\hat{\zeta}}, p(T_{Var(\zeta)})$	0.2960	0.0006	0.2880	0.7618	0.2883	0.8706
$\hat{\zeta}_{0.25}, p(T_{Q25(\zeta)})$	0.2219	0.0000	0.2543	0.3433	0.2564	0.1577
$\hat{\zeta}_{0.5}, p(T_{Q50(\zeta)})$	0.4903	0.0631	0.4887	0.0310	0.4915	0.1050
$\hat{\zeta}_{0.75}, p(T_{Q75(\zeta)})$	0.7654	0.0007	0.7500	0.9915	0.7507	0.8846
$RT_{\zeta}, p(RT_{\zeta})$	248.4424	0.0000	25.5817	0.1423	30.7584	0.0429
$LB_{\zeta}, p(LB_{\zeta})$	54.2217	0.3166	53.4368	0.3437	52.4001	0.3810

$\mathcal{L}_I$  is the value of the incomplete log-likelihood function.  $BIC$  is the Bayesian information criterion computed as  $-2 \cdot \mathcal{L}_I + \ln(N) \cdot k$  where  $k$  denotes the number of estimated parameters. A couple of descriptive statistics is given for the series of empirical integral transforms:  $\hat{\zeta}$  is the arithmetic mean and  $p(T_{E(\zeta)})$  is the p-value of a test for  $E(\zeta) = 0.5$ .  $s_{\hat{\zeta}}$  is the standard deviation and  $p(T_{Var(\zeta)})$  is the p-value of a test for  $Var(\zeta) = 12^{-1}$ .  $\hat{\zeta}_{0.25}$  is the 25 percent quantile and  $p(T_{Q25(\zeta)})$  is the p-value of a test for  $\zeta_{0.25} = 0.25$ ; the analogous computations are done for the 50 and 75 percent quantile.  $RT_{\zeta}$  is the value of the ratio test for i. i. d. uniformity of  $\zeta$  using 20 equal bins and  $p(RT_{\zeta})$  is the corresponding p-value.  $LB_{\zeta}$  is the value of the Ljung-Box statistic for 50 lags and  $p(LB_{\zeta})$  is the corresponding p-value.

## Figures

Fig. 1. Autocorrelation function for durations

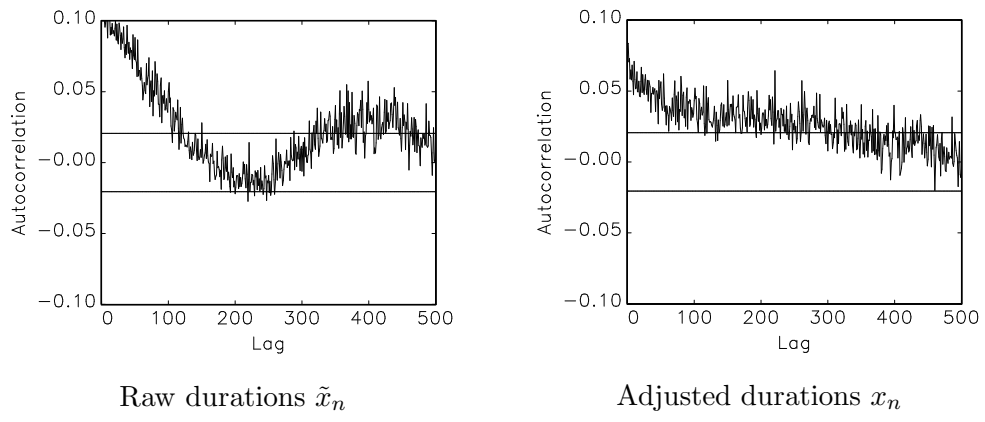


Fig. 2. Histograms and QQ-plots for integral transforms

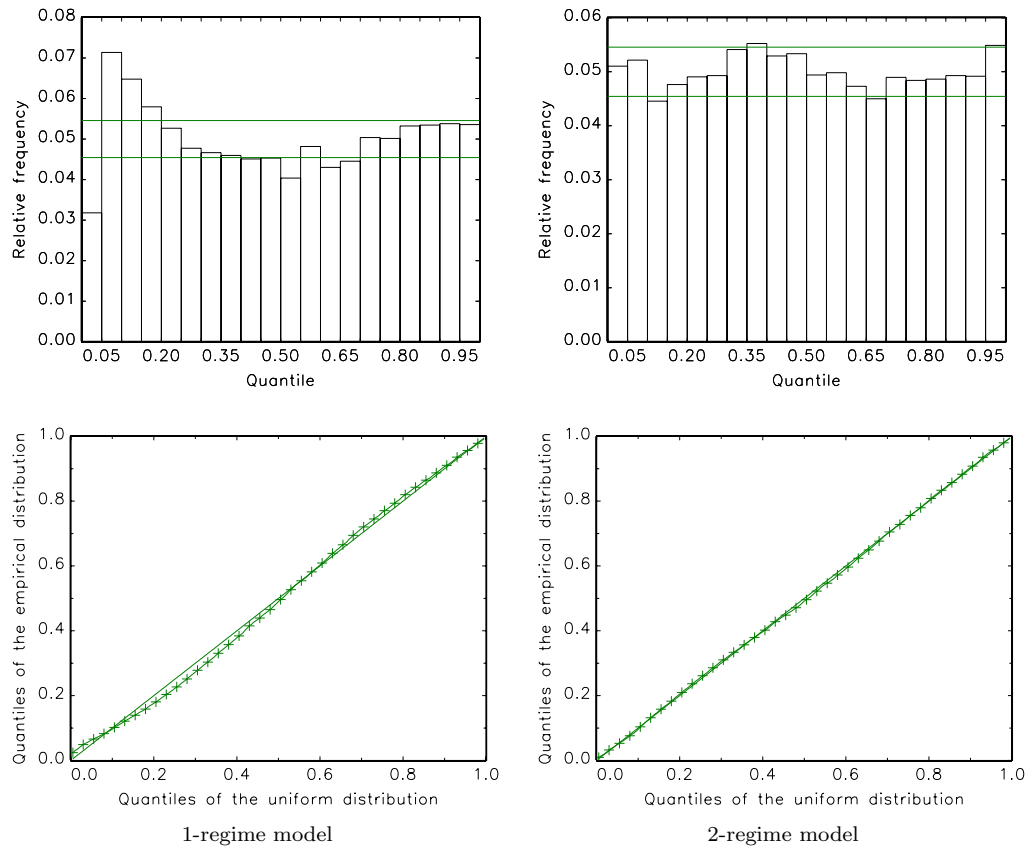


Fig. 3. Hazard function

