Econometric analysis of financial trade processes by discrete mixture duration models

Reinhard Hujer [a,⋆,*], Sandra Vuletić [b]

[a] *University of Frankfurt/M., IZA Bonn, ZEW Mannheim*
[b] *University of Frankfurt/M.*

Revised version: 4 July 2005

Abstract

We propose a new framework for modelling the time dependence in duration processes being in force on financial markets. The pioneering *ACD* model introduced by Engle and Russell (1998) will be extended in a manner that the duration process will be accompanied by an unobservable stochastic process. The *Discrete Mixture ACD framework* provides us with a general methodology which puts the idea into practice. It is established by introducing a discrete-valued latent regime variable which can be justified in the light of recent market microstructure theories. The empirical application demonstrates its ability to capture specific characteristics of intraday transaction durations while alternative approaches fail.

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

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

⋆ The authors thank the anonymous referees who provided us with helpful comments and suggestions for improvement. Moreover we express our gratitude to Joachim Grammig who agreed to place the transaction data set at the disposal.
* Corresponding author. Johann Wolfgang Goethe-University, Department of Economics and Business Administration, Institute of Statistics and Econometrics, Mertonstrasse 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).

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 in many contributions. In this respect Easley et al. (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 study of Kokot (2004).

Innovative econometric methods appear rapidly and they experience an extensive application in the research field. 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 especially appropriate 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 et al. (2000) the periods of time elapsing between successive trades exhibit an idiosyncrasy which could not be captured by the original model. Therefore, recent endeavors give rise to a statistical superordinate framework which provides the researcher with extraordinary flexibility. There are two competing methods which bear resemblance to the general switching autoregression model introduced by Hamilton (1989). Integral part of both approaches is a latent discrete valued regime variable whose involvement can be justified by recent market microstructure theories. 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. They have the discrete mixture of distributions in common with the *Threshold* ACD *model* introduced by Zhang et al. (2001), which allows switches between different regimes to be driven by past realizations of the dependent variable. Alternative modifying *ACD* models are based on latent factors giving rise to a continuous mixture distribution, Bauwens and Veredas (2004) and Ghysels

et al. (2004) draw this line by introducing the *Stochastic Conditional Duration model* and the *Stochastic Volatility Duration model* model respectively. As a generalizing concept the discrete mixture *ACD* framework nests many existing models developed in the course of propagation of the ordinary *ACD* approach.

The *Markov Switching ACD model* developed by Hujer et al. (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 some parsimonious parameterization. It is called *Static Mixture ACD model* which has affinity to the duration model of De Luca and Gallo (2004) and is conducive to better performance too.

This paper is structured as follows: A brief review of the idea of discrete mixture modelling inside of the *ACD* framework is given in section 2. Techniques for estimation will be discussed and specification tests applicable to discrete mixture *ACD* models will be presented. Moreover we establish a relationship to market microstructure theory. In an empirical application in section 3 we present estimation results employing a transaction data set for the common share of Eastman Kodak traded on the New York Stock Exchange. Finally, in section 4 we summarize our main results and give a perspective on possible issues for future research.

2 The discrete mixture *ACD* framework

2.1 The methodological approaches

Let $x_n = t_n - t_{n-1}$ be the duration between the $(n-1)$-th and the n-th market event with deterministic conditional mean $\psi_n \equiv E(x_n|\mathcal{F}_{n-1};\theta)$ where the information set $\mathcal{F}_{n-1}$ consists of all preceding durations up to time t_{n-1} and θ is the set of parameters. The Autoregressive Conditional Duration approach (ACD) is an innovative method, introduced by Engle and Russell (1998), which is based on the assumption that the innovation process $\varepsilon_n \equiv x_n \cdot \psi_n^{-1}$ is independent in time across the N observations considered in the sample and $E(\varepsilon_n) = 1$ is a requirement which is produced by identical distributions. The recent research focusses on weak form ACD models which also allow for higher order dependence in the series of innovations.

A new concept is introduced by the *Discrete Mixture ACD framework* which allows for multifarious applications. The essence is that the duration process x_n is accompanied by an unobservable stochastic process s_n which 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 since time t_{n-1}. 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. The way of fixing the total number of regimes, given by the integer number J, can take place in different ways. To begin with, Lindsay and Roeder (1992) propose the use of diagnostic plots in order to detect the presence of mixing, clarifying the elementary suspicion of $J > 1$. Moreover, a theoretical background and also different residual measures or information criteria can prove to be helpful for limiting the scope of J.

The basic assumption of the *Static Mixture ACD model*, also referred to as *SMACD*, is that the innovation process has a known discrete mixture distribution with mean equal to unity and serial independence. Thus, the density of each innovation has the following general form

$$g(\varepsilon_n;\theta) = \sum_{j=1}^{J} \pi^{(j)} \cdot g(\varepsilon_n \mid s_n = j;\theta), \tag{2.1}$$

where each weight $0 \leq \pi^{(j)} \leq 1$ represents the corresponding probability for prevailing in state j. Any life distribution may be used in order to specify the regime-specific density of the innovation process. De Luca and Zuccolotto (2003) discuss the issue of mixtures for financial duration distributions, and in this regard 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 expected value of each innovation is constrained to be equal to one, and at the same time this expected value turns out to be a discrete mixture of regime-specific expectations. This implies the maintenance of the equality

$$1 = \sum_{j=1}^{J} \pi^{(j)} \cdot E\left(\varepsilon_n | s_n = j;\theta\right) \tag{2.2}$$

which does not require that all the regime-specific expectations are equal to one. By the change of variable technique we obtain the corresponding conditional duration density

$$f(x_n \mid \mathcal{F}_{n-1};\theta) = \sum_{j=1}^{J} \pi^{(j)} \cdot f\left(x_n \mid s_n = j;\theta\right) \tag{2.3}$$

which represents the relevant distribution for statistical inference. Consequently, the conditional regime-specific expectancy of x_n, denoted by $\psi_n^{(j)}$, turns out to be

$$\psi_n^{(j)} = \psi_n \cdot E\left(\varepsilon_n | s_n = j;\theta\right) \tag{2.4}$$

saying that the conditional regime-specific duration expectation is a linear function of the contemporaneous conditional regime-unspecific duration expectation. Equation (2.2) allows for different specifications with specific restrictiveness. Regardless to this possibility, all the conditional regime-specific duration distributions are entirely different. Even the most constricted variant, implied by the assumption $E(\varepsilon_n|s_n = j, \mathcal{F}_{n-1}; \theta) = 1$ for each regime $j \in \mathcal{J}$, makes different distributional features possible, i. e. the first moment of x_n is fix across all regimes but all higher moments are regime-variant. The exhaustive freedom of (2.2) provides a cut above in the sense that all moments are allowed to be regime-specific.

The *Markov Switching ACD model*, already gained recognition by Liu et al. (2004) and abbreviated by *MSACD* in the following, assumes that the innovation process follows a discrete mixture of distributions with conditional mean equal to unity but its higher moments are allowed to be time-varying. A dynamic evolution of the regime variable will come up to this. The regime switch is governed by a Markov chain which is characterized by a $(J \times J)$-dimensional transition matrix with typical element p_{ji} equal to the transition probability $p_{ji} = p(s_n = j \mid s_{n-1} = i)$. Thus, the state of the latent process at recent point of time depends on the state of the previous. Recall the validity of equation (2.4), and per contra each function $\psi_n^{(j)} = E(x_n \mid s_n = j, \mathcal{F}_{n-1}; \theta)$ of regime-specific conditional duration means is allowed to have a life of its own in the *MSACD* model. As a result of this stipulation, any regime-specific mean function will have an autoregressive specification with self-contained parameters. The combination of them takes place by the regime-unspecific conditional mean

$$\psi_n = \sum_{j=1}^{J} \pi_{n|n-1}^{(j)} \cdot \psi_n^{(j)} \tag{2.5}$$

which corresponds to the following marginal duration density

$$f(x_n \mid \mathcal{F}_{n-1}; \theta) = \sum_{j=1}^{J} \pi_{n|n-1}^{(j)} \cdot f(x_n \mid s_n = j, \mathcal{F}_{n-1}; \theta) \tag{2.6}$$

with $\pi_{n|n-1}^{(j)} \equiv P(s_n = j \mid \mathcal{F}_{n-1}; \theta)$ representing the probability that s_n retrieves the j-th state given the filtration. The time-varying measure $\pi_{n+1|n}^{(j)}$ represents the ex-ante probability for being in regime j at time t_{n+1}, conditional on information available up to time t_n and can be evaluated using the two-step recursion

$$\pi_{n|n}^{(j)} = \frac{\pi_{n|n-1}^{(j)} \cdot f\left(x_n \mid s_n = j, \mathcal{F}_{n-1}; \theta\right)}{\sum_{k=1}^{J} \pi_{n|n-1}^{(k)} \cdot f\left(x_n \mid s_n = k, \mathcal{F}_{n-1}; \theta\right)} \tag{2.7}$$

$$\pi_{n+1|n}^{(j)} = \sum_{i=1}^{J} p_{ji} \cdot \pi_{n|n}^{(i)}. \tag{2.8}$$

according to Hamilton (1994). Even though the transition probabilities p_{ji} are constant, the regime probabilities $\pi_{n|n}^{(j)}$ and $\pi_{n+1|n}^{(j)}$ are time-varying. A static specification may be regarded as a special case of the *MSACD* model based on a restricted transition matrix with $p_{j1} = \ldots = p_{jJ}$, this implies time-invariant forecasts of $\pi_{n+1|n}^{(j)}$ but $\pi_{n|n}^{(j)}$ is still varying in time.

Under certain circumstances the *SMACD* model coincides with the *MSACD* model. There is a trivial concordance for $J = 1$ and the corresponding one-regime models are special cases of the ordinary *ACD* model. Moreover, the extremity of $E\left(\varepsilon_n | s_n = j; \theta\right) = 1$ for all $j \leq J$ in the *DMACD* model is equivalent to a static *MSACD* model with $\psi_n^{(1)} = \ldots = \psi_n^{(J)}$.

2.2 Estimation and statistical inference

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

$$\mathcal{L}_I(\theta) = \sum_{n=1}^{N} \ln\left[f(x_n \mid \mathcal{F}_{n-1}; \theta)\right] \tag{2.9}$$

under the linear constraint $\sum_{j=1}^{J} \pi^{(j)} = 1$ for any specification of the *SMACD* model and $\sum_{k=1}^{J} p_{kj} = 1$ for all $j \leq J$ in the case of a *MSACD* model and additional restrictions for nonnegativity, 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 et al. (1977).

Diebold et al. (1998) propose a method to test the forecast performance of general dynamic models. The idea behind this specification test has been extensively used by Bauwens et al. (2004) to compare different types of *ACD* models. Denote by $\{f(x_n \mid \mathcal{F}_{n-1}; \hat{\theta})\}_{n=1}^{N}$ the sequence of density forecasts

evaluated using the parameter vector estimate $\hat{\theta}$ from some parametric model and denote by $\{f(x_n \mid \mathcal{F}_{n-1}; \theta)\}_{n=1}^{N}$ 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 $H_0 : \{f(x_n \mid \mathcal{F}_{n-1}; \hat{\theta})\}_{n=1}^{N} = \{f(x_n \mid \mathcal{F}_{n-1}; \theta)\}_{n=1}^{N}$, the sequence of empirical integral transforms

$$\hat{\zeta}_n = \int_{-\infty}^{x_n} f(u \mid \mathcal{F}_{n-1}; \hat{\theta})\, du \tag{2.10}$$

will be uniform i.i.d. on the unit interval. Any statistical test for uniformity in the sequence of integral transforms can be used to assess the forecast performance of the model under consideration. A first indication for some misspecified model is provided by simple tests on the mean and variance. Checks for quantiles being equal to the population counterpart implied by the standard uniform distribution can be conducted additionally. Referring to this, let $\mathcal{N}_p$ be the number of empirical integral transforms being less or equal than p with $0 \leq p \leq 1$, then the statistic

$$Q_{\zeta_p} = \frac{\mathcal{N}_p - N \cdot p}{\sqrt{N \cdot p \cdot (1-p)}} \tag{2.11}$$

follows approximately the standard normal distribution under the null hypothesis $H_0 : \zeta_p = p$. The histogram test is well suited for testing the respectability of any model specification. Consider partitioning the support of ζ into K equally spaced bins and denote the number of observations falling into the k-th bin by $\mathcal{N}_k$. The confrontation of theoretical frequencies $\varsigma_k = \frac{1}{K}$ with observed relative frequencies $\hat{\varsigma}_k = \frac{\mathcal{N}_k}{N}$ constitutes the fundament of the statistic

$$RT_\zeta = -2 \cdot \sum_{k=1}^{K} \mathcal{N}_k \cdot \ln\left[\frac{\varsigma_k}{\hat{\varsigma}_k}\right] \tag{2.12}$$

which has a χ^2 distribution with $(K-1)$ degrees of freedom under the null hypothesis. The statistical tests for uniformity may be supplemented by graphical tools. Departures from uniformity can easily be detected using a quantile-quantile plot or histogram plot based on the sequence of $\hat{\zeta}_n$.

2.3 Link to the market microstructure theory

The modern literature on the microstructure of financial markets, gradually widening in the style of Easley et al. (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 or 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. Consequently, instances without news events are exclusively infested with uninformed traders while news regime drum both trader types.

The instantaneous transaction rates turn out to be different across the trader categories and that is the design which we want to mimic primarily. Informed traders make transactions as a result of hasty information based decisions and this behavior dispose for a transaction rate that breaks out in a rash. In contrast, the phlegm of uninformed traders swears to a transaction rate with moderate progression. And possibly, the respective transaction rate is even flat.

The collectivity of transactions, carried out either by the large attendance of uninformed traders or by sporadic emersions of informed traders 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 regime-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.

This methodological advancement, nailing down the presented theoretical background, is excellently reflected in the general coverage of discrete mixture ACD models. Thereby the mixing variable is responsible for culling the presence of unobservable information regimes and the mixing parameters pose as fractions of different information regimes by which the trading days are obsessed cyclically. They also allow to appreciate the proportion of different trader types acting on the financial market. The level of discrepancy between regime-specific peculiarities in trading behavior can be easily regulated by adapting the parameters inside of equation (2.2) for the $SMACD$ model. A

regime $j^\star$ that belongs to the subset $\mathcal{J}_1 \equiv \{j^\star | E(\varepsilon_n | s_n = j^\star; \theta) = 1\}$ deals with some "normal" manner of expectations while other regimes come along with changed expectations as soon as relevant information pervade the market. The *MSACD* model allows to enrich the vision of the trading mechanism by a dynamic design of the regime switch. An important regulator of both models is the distributional assumption. For mixture models it is a routine to assume that each regime-specific distribution belongs to the same family. Contrary to this practice a mix of different distribution types seems to be reasonable which can be argued from a theoretical point of view. So, the exponential density with its constant hazard will accommodate to the group of uninformed traders while hunchbacked hazard rates will be caused by the informed traders which spring into action in the event of any reception of news. So, a comparison of regime-specific hazard functions

$$h(x_n \mid \mathcal{F}_{n-1}; \theta) = \frac{f(x_n \mid \mathcal{F}_{n-1}; \theta)}{\int_0^{x_n} f(u \mid \mathcal{F}_{n-1}; \theta)\ du} \tag{2.13}$$

helps to identify the trader-specific consuetudes.

Bauwens et al. (2000) report on the deficiency of different *ACD* models which is well founded by the inability of modelling observations in the tails of the 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 discrete mixture *ACD* framework. By doing this, we hope to succeed in overcoming the lack of satisfactory forecast performance of recent *ACD* models and we expect a clear answer from the empirical application.

3 Empirical application

3.1 The data set

The data used for empirical application consists of transactions of the common stock of Eastman Kodak, recorded on the New York stock exchange from the trades and quotes database provided by the NYSE Inc. The sampling period spans twenty trading days, covering all working days from Monday to Friday, during the month of March in the year 2004. We focus on all "regular" trades recorded during the six and a half trading hours lasting from 9:30 to

16:00.[1] The trading times have been recorded with a precision measured in seconds. Durations with length of zero seconds deserve closer attention. Since the *ACD* framework does not permit the inclusion of null durations, a twofold treatment will be our answer to this problem. The entire elimination of null durations is a common treatment, which relies on the argument that trades executed within the same second are splits of a big order block initiated by the same trader. But we limit this procedure to durations that arise from successive trades without any price change. This logic can be called into question when we deal with successive trades having a price change, because it is plausible to think of null durations that evolve from actions of retail traders. So, transactions occurring within the same second with price change have been transformed according to Veredas et al. (2002), the concept behind this idea is the artificial enlargement of null durations while the next positive duration shrinks. 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 the multiplicative form $\tilde{x}_n = x_n \cdot \Phi(t_{n-1})$. 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.[2] 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 efforts to look into the long run duration dynamic.[3]

Descriptive information about sample moments of the raw and the seasonally adjusted duration data is reported in the left block of Table 1.

< insert Table 1 about here >

As expected, the series of adjusted durations has a mean of approximately

[1] A transaction is said to be regular if its execution has been realized without any condition or correction.

[2] 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.

[3] Estimates of the time-of-day function were obtained by conducting a polynomial-trigonometric regression of the durations on the time-of-day according to Gallant (1981) and Eubank and Speckman (1990), with separate application to each day of the week.

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. The goodness of fit, associated with each model specification, will by inquired by a twofold exercise. Each column entitled "In"-sample in the right block of Table 1 contains a statistical description of seasonally adjusted durations that cover the first ten trading days and they are employed to estimate the model parameters. The rest of the total data set is used to compute out-of-sample forecasts based on the estimated parameters. Descriptive statistics for the second subsample are contained in the column named "Out"-sample. The duration distributions in both subsamples appear to be qualitatively similar. They are branded by overdispersion and skewness to the right (because the mean exceeds the median), but the very large durations tend to appear more concentrated in the second subsample.

An obvious characteristic of the data is the presence of strong positive autocorrelation in the series of raw and seasonally adjusted intertrade durations. The pronounced autocorrelation can be directly recognized from the cutouts of the autocorrelation functions displayed in Figure 1.

< insert Figure 1 about here >

The series of raw durations have a recurrent dependence structure for each trading day environed by dotted vertical lines (dashed lines separate a join of trading days which constitutes a week). The bathtub-shaped evolution of the autocorrelation function recurs every day and of course, this plausible feature is due to the seasonality. The bathtub-shaped episode of the autocorrelation function for the adjusted durations disappears almost completely. Anyway, remaining dependence structure is present in the series of seasonally adjusted durations as evidenced by the plentitude of autocorrelations that reside outside the confidence band of 95 percent. Note that the autocorrelation function associated with the seasonally adjusted durations decreases dramatically for the beginning band of lag orders. This notice allows to make an educated guess that the separation of the seasonality part and the part of autocorrelated dynamic due to the gradual processing of information does not success completely. The proper community of "orthogonal" components of the process of raw durations seems to collapse in the incipient minutes of *NYSE* trading. The pivotal causation for this fact may come from the batch-auction which is an exceptional circumstance, characterized by a monopolistic status of the market maker (specialist). As a measure of precaution, we initiate a clearing up of the data set which corresponds to the suggested treatment of Engle and

Russell (1998): Observations recorded up to fifteen minutes after the opening of the *NYSE* will be excluded from our estimation task, they are suspected of being parts of the batch auction which might cause a contamination of the model that will be used for description of the trading velocity.

The simultaneous disappearance of autocorrelations will checked by the Ljung and Box (1978) test statistic. We fix the relevant lag order, required for this autocorrelation test, onto the number of diurnal observations which is given at the top of Table 2.[4] The entries at the bottom evidence the following fact: Even after seasonal adjustment the joint hypothesis that all the autocorrelation coefficients are simultaneously equal to zero can be rejected at conventional significance levels, (see the p-value of the LB-test p_{LB} which is permanently equal to zero), although the shape of the autocorrelation function changes dramatically.

< insert Table 2 about here >

Therefore, an autoregressive approach appears to be appropriate as a model for the transaction durations.

3.2 Specifications of discrete mixture ACD models

We enrich the ordinary ACD model by allowing for the possibility of interchange between a limited number of regimes. The consideration of two and three regimes will be effectual. Our fixing onto $J = 2$ is well founded by the theoretical vision of the trading mechanism which is outlined in paragraph 2.3. So we think of a *news* and a *no news* regime mastering the trading process interchangeably during the course of a trading day. The consideration of three regimes can be motivated from theoretical point of view as well: The distinction between favorable and unfavorable signals, catched by the informed market participants, might be a reasonable amelioration of the trading process under the news regime. This breeds the application of the general philosophy of discrete mixture ACD models, becoming manifest either in a static or dynamic manner. First of all, we intend to estimate the model of De Luca and Gallo (2004) which develops behind the scene of the *SMACD* approach

[4] Exemplary, we find that Friday, 26 March (Monday, 8 March) records 1222 (1936) transactions representing the datum that has the lowest (highest) number of diurnal observations and the whole-numbered average of daily durations is equal to 1590. Hence, the Ljung and Box (1978) statistic is computed for a couple of moderate lag orders: $\ell_L = 1222$ represents the minimum lag order, $\ell_H = 1936$ illustrates the maximum lag order and $\ell_M = 1590$ is a compromise between the two extremities, demonstrating our procedure when considering the whole data set.

and other contrastable specifications will be of interest too. In order to make them comparable to the dynamic world we impose the restriction of regime-independence of conditional durations means when considering specifications belonging to the *MSACD* model.

We distinguish between two extreme specifications of the *SMACD* model. The restrictive variant, denoted by the character $\underline{R}$ in the following, comprises the fact that all regime specific expectations of the innovation process are forced to be equal to one, so that absolutely no care for equation (2.2) is needed. This variant may be estimated by employing the *EM*-algorithm, while the nonrestrictive variant, characterized by the negation of the requirement that $E(\varepsilon_n|s_n = j;\theta) = 1$ holds true for each $j \leq J$ and denoted by the character $\overline{R}$ in the following, has to be estimated by maximizing the incomplete log-likelihood function directly. A semirestrictive representative of the *SMACD* model, denoted by the symbol R, arises from a compromise between the restrictive and nonrestrictive variant. It is based on the fact that one regime-specific mean of the innovation process is fixed to be equal to one. [5]

Before running any estimation we have to make an important decision which decides on the mean function. The observed sequence of durations on a trading day will be treated independently of durations recorded on other trading days. This means that on every trading day a recursion determining the mean function starts anew. [6] This design circumvents any transmission of the trading dynamic levelled off at the end of a trading day on the subsequent trading day. The mean function ψ_n is assumed to be linear and the two relevant lag orders in the recursion are equal to one, i. e.

$$\psi_n = \omega + \alpha \cdot x_{n-1} + \beta \cdot \psi_{n-1} \tag{3.1}$$

where the index n includes a trade day counter $d \leq 20$ and also a counter for daily transactions $k \leq N_d$. The function of conditional duration means obeys an autoregressive recursion, demanding for an appropriate initializing value for each trading day. An assortment of durations, belonging to the data subset reserved for the empirical analysis, will render an adjuvant service. Observations, recorded on each trading day between the time of day 9:45 and 10:00, are predestinated candidates for this purpose and we assume that the corresponding arithmetic mean gives a reasonable description of the genesis of the mean function.

[5] This idea makes sense for models that incorporate at least three regimes, i. e. $J \geq 3$. Otherwise, a two-regime model will be swamped with restrictions and there is nothing else for it but to reproduce the restrictive variant.

[6] Consequently, the log-likelihood function considering all available durations can be expressed as the sum of twenty daily log-likelihoods.

As argued previously, the exponential distribution is a candidate which comes into question for our purposes. Consulting the distributional proposal of Grammig et al. (1998), some regime-specific distribution can be taken from the Burr (1942) family of distributions. We introduce an identifying notation in order to distinguish between the different specifications appearing as appropriate for framing a multi-regime model. The realization of the variable $D^{(j)}$ denotes the distribution assumed for the j-th regime. $D^{(j)} = \mathcal{E}$ indicates the application of the exponential distribution for the j-th regime, while the character $\mathcal{B}$ stands for the Burr (1942) distribution with regular time-invariant distribution parameters $\kappa^{(j)}$ and $\sigma^{(j)}$. To some extent additional time-invariant distribution parameters $\lambda^{(j)}$ and $\mu^{(j)}$ will come into existence when considering the nonrestrictive and semirestrictive specifications of the *SMACD* model. In return, the equality $\sum_{j=1}^{J} \pi^{(j)} \cdot m^{(j)} = 1$ with distributional case differentiation in the j-th regime

$$m^{(j)} = \begin{cases} \lambda^{(j)} & \text{if} \quad D^{(j)} = \mathcal{E} \\ \left[\mu^{(j)}\right]^{-\frac{1}{\kappa^{(j)}}} \cdot \frac{\Gamma\left(1+\frac{1}{\kappa^{(j)}}\right) \cdot \Gamma\left(\frac{1}{\sigma^{(j)}}-\frac{1}{\kappa^{(j)}}\right)}{\sigma^{(j)\left(1+\frac{1}{\kappa^{(j)}}\right)} \cdot \Gamma\left(\frac{1}{\sigma^{(j)}}+1\right)} & \text{if} \quad D^{(j)} = \mathcal{B} \end{cases} \tag{3.2}$$

has to be ensured in the course of estimation. [7] In contrast, the restrictive specification variant of the *SMACD* model and also the *MSACD* model incorporate corresponding distributional parameters which obey a parameterization according to

$$\lambda^{(j)} = 1 \tag{3.3}$$

and

$$\mu^{(j)} = \left[\frac{\sigma^{(j)\left(1+\frac{1}{\kappa^{(j)}}\right)} \cdot \Gamma\left(\frac{1}{\sigma^{(j)}}+1\right)}{\Gamma\left(1+\frac{1}{\kappa^{(j)}}\right) \cdot \Gamma\left(\frac{1}{\sigma^{(j)}}-\frac{1}{\kappa^{(j)}}\right)} \right]^{-\kappa^{(j)}} \tag{3.4}$$

respectively. [8] Bringing together the different variants, the regime-specific conditional density of the duration x_n turns out to be

[7] 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 (2.2), we abandon the estimation of $\pi^{(J)}$ and $\lambda^{(J)}$ or $\mu^{(J)}$ respectively.

[8] These parameter determinations imply that each regime-specific expectation of the innovation process is equal to one.

$$f_n\left(x_n \mid s_n = j, \mathcal{F}_{n-1}; \theta\right) = \begin{cases} \lambda_n^{(j)} \cdot \exp\left(\lambda_n^{(j)} \cdot x_n\right) & \text{if} \quad D^{(j)} = \mathcal{E} \\ \frac{\mu_n^{(j)} \cdot \kappa^{(j)} \cdot x_n^{\kappa^{(j)}-1}}{\left(1+\sigma^{(j)} \cdot \mu_n^{(j)} \cdot x_n^{\kappa^{(j)}}\right)^{\frac{1}{\sigma^{(j)}}+1}} & \text{if} \quad D^{(j)} = \mathcal{B} \end{cases} \tag{3.5}$$

with the time-variant parameter $\lambda_n^{(j)} = \psi_n^{-1}$ in the exponential case while $\mu_n^{(j)} = \psi_n^{-\kappa^{(j)}} \cdot \mu^{(j)}$ is the corresponding alternative.

We attend to the topic concerning the outclassing performance of our specifications compared to the reference model of De Luca and Gallo (2004), which conforms to the nonrestrictive instruction associated with the *SMACD* model with exponential distribution in each existent regime. Coming from this model there is no exorbitant increase in the number of parameters composing our preferred specifications of the *SMACD* and *MSACD* model. Another interesting issue becoming apparent is whether the predetermined arrangement of $\lambda^{(j)}$ and $\mu^{(j)}$ allows for flexibility which is sufficient to catch regime-specific characteristics hidden in the duration process. And the question for dynamics in the regime switch is also relevant. We expect answers from our estimation results.

3.3 Estimation results

For each two-regime specification, that arises from the general *SMACD* model, its parameter estimates and corresponding standard errors are given in the upper block of Table 3.[9] The subsequent entries comprehend the value of the incomplete log-likelihood function $\mathcal{L}_I$ and also the Bayesian information criterion *BIC* of Schwarz (1978), which is computed as $-2 \cdot \mathcal{L}_I + \ln(N) \cdot \sharp\theta$ with N denoting the number of durations exclusively used for the estimation purpose and $\sharp\theta$ representing the number of estimated parameters. By means of estimation results we carry out a couple of specification tests. 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 standard uniform distribution are given in the end.

< insert Table 3 about here >

At first, the *BIC* does not support the introducing specification $\{\mathcal{E}, \mathcal{E}\}$ which reflects the De Luca and Gallo (2004) model. Some specification tests produce results which are in favor of the uniformity hypothesis. In this concern, the test

[9] 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).

on the mean argues for the null hypothesis $H_0 : E(\zeta) = 0.5$, and the variance test does the same. In addition, the test on the median (third quartile) accepts the null hypothesis H_0: $\zeta_{0.5} = 0.5$ ($\zeta_{0.75} = 0.75$) from statistical point of view. But all these intercessional results do not have power of persuasion. The disappearing p-value obtained from the test associated with the first quartile is a sign of bad adaption in the tail of the distribution. Moreover, the alternative histogram specification test does not support this model specification, this conclusion can be recognized from the p-value of the ratio test which is equal to zero. Hence, the apparent defect of this reference specification stems from the improper choice of distribution.

Does any distributional modification, taking place either in one or both regimes, overcomes these serious problems ? In order to answer this question we consult the estimation results of the two supplementary specifications $\{\mathcal{E}, \mathcal{B}\}$ and $\{\mathcal{B}, \mathcal{B}\}$ which voice realizations of the replacement characters $\{D^{(1)}, D^{(2)}\}$. Both routes will be concretized by the restrictive and also by the nonrestrictive instruction. All the specifications, that evolve from the combination of the distributional aspect and the room for restrictiveness, call for more parameters than the inaugurating model of De Luca and Gallo (2004), but the fact of reduced BIC values assigns them as reasonable alternatives. And by assuming the comprehending Burr (1942) distribution for each existent regime best trade-offs between parsimony and likelihood benefit will be achieved. The commonness of the four widening specifications is that they adapt to the real data in some better manner, but this improved performance does not suffice to authenticate the superiority of them. In fact, they effectuate a dramatic reduction of the value associated with the RT_ζ-test, but the corresponding p-values are still equal to zero. So the restrictive variant $\underline{R}$ of the most extensive distribution composition generates a 85 percent retrenchment of the RT_ζ-value, compared to the reference specification $\{\mathcal{E}, \mathcal{E}\}$. The replacement of the exponential distribution by the Burr (1942) distribution in both regimes seems to get a grip on the distributional rigidity for very small durations, from which the reference model suffers. Anyhow, a distributional shortcoming relocates to the medial durations as the negligible p-value of the median test demonstrates.

This general result prompts us to combine the amenities of the exponential and Burr (1942) distribution, showing to advantage for our data situation, in a three-regime fashion of the $SMACD$ model. The twofold use of the exponential (Burr (1942)) distribution and the unique appearance of the Burr (1942) (exponential) distribution provides us with reasonable options. The corresponding specifications are transliterated by $\{\mathcal{E}, \mathcal{E}, \mathcal{B}\}$ and $\{\mathcal{E}, \mathcal{B}, \mathcal{B}\}$ respectively. And since the exponential distribution turns out to be a limiting case of the Burr

(1942) as in the j-th regime the parameter $\sigma^{(j)}$ tends to zero and $\kappa^{(j)} = 1$ holds true, we also consider the vision of the $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ specification. A preliminary investigation shows that restrictive variants of these specifications are not able to identify three diversified regimes, so that semirestrictive and nonrestrictive representatives quicken our interests. The estimation results are given in Table 4 which has the the same system as the last-mentioned and $\mathcal{J}_{=}$ gives the regime $j^{\star}$ that ensures the validity of $E(\varepsilon_n | s_n = j^{\star}; \theta) = 1$ for the semirestrictive variants.

< insert Table 4 about here >

Specifications, characterized by distributional heterogeneity, have problems to model edging durations appropriately. Thus, the dominance of the exponential (Burr (1942)) distribution comes along with some unsatisfactory modelling of short (large) durations, as the negatory results of the quartile tests give reason to believe. The specifications under $\{\mathcal{E}, \mathcal{E}, \mathcal{B}\}$ seem to be devoted to some "normal" durations ascribed to the uninformed traders, while the exceptional trade durations of the informed traders have profuse consideration in the $\{\mathcal{E}, \mathcal{B}, \mathcal{B}\}$ specifications. Loosely speaking, it is either the group of uninformed or informed traders which stands up to the other. The estimates for the regime probabilities support this impression: The former specifications seem to consider the uninformed traders twice as much as the informed traders, while the latter specifications do the converse. Therefore, we feel well advised to bring an equality of opportunity into being, attempted by the homogeneous specification $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$. Therein, the arithmetic mean of the empirical integral transforms draws near one half, the corresponding empirical variance 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 values of the RT_{ζ}-test register further increase and the BIC prefers these specification, even more than the corresponding two-regime specification.

Summarizing, the estimation results for *SMACD* models will induce us to cling to distributions that are more comprehensive than the exponential. The Burr (1942) distribution proves to be a good choice, placing at the disposal nonmonotonic shapes of the hazard rate. It is able to alleviate the distributional ailment from which thrifty models occasionally suffer. We polish the scaffolding of discrete mixture *ACD* models by fathering some closer connectivity between the existent regimes. The interdependent change of regimes, realized by the *MSACD* model, will make one's debut now. Table 5 renders an account of forecast performance of *MSACD* model specifications that have

been able to be carried forward from its static pendants.

< insert Table 5 about here >

The specification $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ is high in parameters but it brings out the best in terms of likelihood and data adaption. It will be favored by the BIC and even the RT_{ζ}-test convinces of its outstanding performance.

For purposes of comparison Figure 2 contains QQ-plots and histograms for the series of integral transforms implied by the promising three-regime homogeneity specification $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ of the general discrete mixture ACD framework.

< insert Figure 2 about here >

The charts clearly show that the different designs of the towering three-regime specification produce 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 of the corresponding two-regime model specifications, presented in Figure 3, show that the empirical integral transforms disagree sharply with the theoretical density, and that they tend to produce systematically biased forecasts of small and/or large x, as can be seen from some histogram bars that lie outside of the 95% confidence interval. Thereby, the exponential model $\{\mathcal{E}, \mathcal{E}\}$ of De Luca and Gallo (2004) comes off very badly.

< insert Figure 3 about here >

So, the distributional assumption is the most efficient tool by which the forecast performance of discrete mixture ACD models can be highly improved. The quantile points form increasingly a diagonal line and the steep incline associated with the histogram bars disappears. The number of regimes and the modelling of the regime variable via Markov chain are also important issues which account for some reasonable fine-tuning. These influencing factors have a share in a total harmonization of the histogram bars. Note that the subordinate specifications to the nonrestrictive variant of the $SMACD$ model do not sustain a serious loss in performance.

Summarizing, our assemblage of advanced regime switch model specifications displays a superior data fit relative to the simple De Luca and Gallo (2004) model. By our out-of-sample analysis we intend to form an opinion about their gain in prediction. Does a suasive forecast performance can be ev-

idenced as in the in-sample inquiry ? In order to answer this question we use a procedure which has been adventured by Bessec and Bouabdallah (2005). The forecast accuracy will be examined by rolling the origin of forecasting and a weekly horizon is an reasonable assumption for computing the forecasts. By using the parameter estimates, we calculate the arithmetic mean of the squared prediction errors $(x_{N+k} - \psi_{N+k})^2$ and absolute errors $|x_{N+k} - \psi_{N+k}|$ for $1 \leq k \leq 7100$, corresponding to the first quintet of out-sample trading days, and the relevant measures will be denoted by *MSE* and *MAE* respectively. The 6441 recordings associated with the next five out-sample trading days will experience the analogue treatment, thereby the estimated parameters will be recalibrated by considering the durations of the first forecast week. Additionally, we carry out the same tests as in the in-sample analysis. Its p-values and also the two conglomerates of prediction errors are given in Table 6 for specifications belonging to the group of *SMACD* models and Table 7 concerns the category of *MSACD* models.

< insert Table 6 and 7 about here >

In the category of two-regime models it is the reference specification $\{\mathcal{E}, \mathcal{E}\}$ of De Luca and Gallo (2004) which hands over the best predictions for the first future trading week, i. e. no other specification is able to underbid the value 1.441 (0.846) associated with the mean of squared (absolute) prediction errors. But on the other hand, the distributional problem is still present as the evanescent p-value of the histogram test indicates, and a humble improvement is provided by the homogeneous specification $\{\mathcal{B}, \mathcal{B}\}$ of the *MSACD* model. The view on the second future trading week gives rise to another conclusion: Indeed, the nonrestrictive competitors overtake the De Luca and Gallo (2004) model in terms of the *MSE*, but none of the two-regime specifications reduces the value 59.102 of the histogram test. Loosely speaking, for two-regime models we have to deal with a trade-off between distributional acclimatization and forecast accuracy. But the migration to a three-regime fashion can be convenient for both objectives. So, each nonrestrictive specification of the *SMACD* model with supernumerous consideration of the Burr (1942) distribution outperforms the addressed reference model for the first week of predictions: the distributional deficiency will be reduced and the accuracy of predictions will be upgraded. The ideal capacity of the $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ specification, affirmed by the in-sample analysis, does not correspond to some excellent forecast performance. This fact allows to make the following general conclusion: Although extensive discrete mixture *ACD* models display a superior in-sample fit, the gain in prediction is small, does not even emerge or deteriorates the predic-

tions. The same empirical finding has been noticed by Bessec and Bouabdallah (2005) and they argue that the main source of this flaw is due to the failure of forecasting the regime indicator.

3.4 Theoretical interpretation

We discuss the theory-related content of the $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ specification, evidenced as the most promising representative associated with the discrete mixture *ACD* framework. By applying both the *SMACD* model and the *MSACD* model the existence of three constitutively different streams, governing the process of intertrade durations, will be affirmed. They allow to visualize the different velocities from which trading evolves.

The estimation results show that the two regular distribution parameters $\kappa^{(j)}$ and $\sigma^{(j)}$ vary keenly across the regimes. Expectedly, several Wald-tests (see Buse (1982)) conform this disparity of regimes: The corresponding results are given in Table 8, rejecting any hypothesis of pairwise equality between regime-specific distribution parameters at the conventional significance level of five percent.

< insert Table 8 about here >

This stamping of regime-specificity has a strong impact on the shape of the hazard function considered for each regime separately. The distributional parameter $\kappa^{(j)}$ is the sole control instrument of the hazard shape in the j-th regime. For $\kappa^{(j)} \leq 1$ the Burr (1942) distribution implies a strong decreasing failure rate, while the case $\kappa^{(j)} > 1$ gives rise to a nonmontonic hazard shape. By conducting the t-test we can confirm the hypothesis of a hunchbacked hazard rate for each regime under consideration. Figure 4 displays the three regime-specific survivor functions $S(x|s_n = j, \mathcal{F}_{n-1}; \hat{\theta})$ and the corresponding hazard rates $h(x|s_n = j, \mathcal{F}_{n-1}; \hat{\theta})$ each evaluated for $\psi_n = 1$ and by taking the estimates of the distribution parameters into account.

< insert Figure 4 about here >

Note that the choice for one or the other mixture model specification does not change the qualitative nature of the functions which are relevant for survival analysis. The hazard rates for two regimes tend to rise rather quickly after a transaction has been observed, one is characterized by some moderate hazard amplitude while the other has some sharp peak. In contrast the hazard function under the remaining regime increases slightly and gives clearly

more weight to spells with a length of more than two units of time. This hazard tends to be constant for large durations, approximating the feature of memoryless implied by the exponential distribution. Thus, large durations are relative insensitive to an arbitrary change of the distributional assumption while contaminations will be predominantly caused by the small durations. The imagination of an overall constant hazard, implied by the exponential distribution, is far from reality and the model of De Luca and Gallo (2004) is not appropriate for the recent trading mechanism.

Let us look at the parameters estimates that determine the mixing variable. In order to compare the transition probabilities of the *MSACD* model with the regime probabilities of the *SMACD* model we throw a glance at the so called steady state distribution associated with the Markov process.[10] A solution providing us with the steady spate distribution $\{\pi_j | j \leq J\}$ is given in Hamilton (1994). Table 9 confronts the regime probabilities $\pi^{(j)}$ of the *SMACD* model with the ergodic probabilities π_j of the *MSACD* model.

< insert Table 9 about here >

It also contains the regime-specific innovation means $E\left(\varepsilon_n | s_n = j; \hat{\theta}\right)$ calculated by the parameter vector estimate $\hat{\theta}$. First of all, we have to mention the ordinal congruence between the regime probabilities and the regime-specific means, i. e. the relation $E\left(\varepsilon_n | s_n = i; \hat{\theta}\right) \geq E\left(\varepsilon_n | s_n = j; \hat{\theta}\right)$ comes along with the validity of $\pi^{(i)} > \pi^{(j)}$ and $\pi_i > \pi_j$ respectively. Obviously, any three-regime specification under consideration does the same job and the theoretical interpretation of this common feature is that small (medium and large) durations are present by a small (medium and large) fraction. There is always one regime which occupies round fifty percent of all regime emersions, and it comes along with some above-average duration expectancy if countenanced by the art of specification. An important note is that it corresponds to the regime that has a flat hazard rate. The other two regimes share the remaining fifty percent of regime fraction and in the case of $E\left(\varepsilon_n | s_n = j; \hat{\theta}\right) = 1$ there will be a relative equable proportioning.

All these stylized facts are in line with the theoretical background discussed in section 2.3. The inertial trading activity, adumbrated by some flatness of the hazard rate, predominates the whole trading process and can be associated with the vision of trading behavior ascribed to the group of uninformed

[10] An ergodic Markov chain implies that as k increases ad infinitum the k-step transition matrix P^k converges elementwise to a matrix in which each column is the unique steady state distribution, id est $\lim_{k\to\infty} p_{ji}^{(k)} = \pi_j$ for all $i \leq J$.

traders. Following the theory, the lack of news events gives rise to this constellation of characteristics and consequently, we have to think of a so called *no news regime* when dealing with an extremely large regime-specific innovation mean and regime probability. The other two regimes award the image of succinct trading which can be traced back to the informed traders. The hunchbacked hazard functions reflect their abrupt spooking on the financial market. So, a superior regime of news happenings is formed by these twin regimes. The considered two-regime specifications of the discrete mixture *ACD* framework identify the coexistence of the *no news regime* and *news regime* but the internal difference of the latter is preeminent, so that they experience a statistical refuse as discussed in section 3.3. The three-regime models allow to consider a case differentiation of the *news regime*, disclosing a *good news regime* for favorable signals and a *bad news regime* for unfavorable signals. But until now, it is impossible to identify them unambiguously. But we find a remedy by analyzing the trade direction, determined by the quote test of Lee and Ready (1991) and formalized by a indicator variable with valuation equal to one if the transaction was buyer initiated and minus one otherwise, and the log-ratios of smoothed regime inferences

$$r_n\left(\overline{k};\underline{k}\right) = \ln\left(\frac{\pi_{n|N}^{(\overline{k})}}{\pi_{n|N}^{(\underline{k})}}\right) \tag{3.6}$$

$$R_n\left(\overline{k},k;\underline{k}\right) = \ln\left(\frac{\pi_{n|N}^{(\overline{k})} + \pi_{n|N}^{(k)}}{\pi_{n|N}^{(\underline{k})}}\right) \tag{3.7}$$

considering available information up to the end of each trading day, which can be computed by using the algorithm of Kim (1994).[11] If the regime $\overline{k}$ has higher probability than the regime $\underline{k}$ then the corresponding log ratio $r_n(\overline{k};\underline{k})$ will be positive. Observing a sequence of buy orders allows to perceive a propitious situation for the informed traders, coming along with risen (fallen) probabilities for good (bad) news. For the first hundred transactions Figure 5 visualizes a comparison of the trade direction and the log-ratios, where k is chosen to be the no news regime characterized by the motionless hazard rate, and $\overline{k}$ denotes the regime that has the most sensitive hazard, so that $\underline{k}$ reflects the regime with the moderate hazard function.

< insert Figure 5 about here >

[11] The equality $\pi_{n|N}^{(j)} = \pi_{n|n}^{(j)}$ for $j \leq J$ holds true in the case of static *MSACD* models. The analoge measure for *SMACD* models can be obtained by stretching a matrix of probabilities with $p_{ji} = \pi^{(j)}$ for all $i, j \leq J$.

Obviously, a parallel progression of the trade direction and the corresponding log-ratios is certifiable. Consequently, it is the third (second) regime of the *SMACD* (*MSACD*) trinity model which may be presumed to mask the bad one. Note that the two competing *SMACD* model specifications imply a quasi proportional relationship between $r(1;3)$ and $R(1,2;3)$, while an oscillating proportion is due to the Markov property of the regime variable. It is an indication of varying presence of informed trades.

4 Conclusion

The *Discrete Mixture ACD framework* flows from the idea of an unobservable regime switch accompanying the process of durations. Encroaching upon its flexibility can be done in a couple of directions. Effective regulators are given by the number of regimes and the distributional assumption. The recursion of the mean function and the design of the residual expectancy, satisfying the obligatory demand for a unit mean, are also important starting points for altering the comprehension of the framework.

In this paper we present, challenge and put to the test two alternative approaches of this general framework. The empirical work allows to conclude that even the *SMACD* model, characterized by the constancy of the regime probabilities all along the trading time, is a promising new approach for modelling autocorrelated intraday durations obtained from high frequency data sets from stock and foreign exchange markets. As a static representative it is able to reduce the distributional problem from which the pioneering model of De Luca and Gallo (2004) racily suffers. The alternative *MSACD* model provides a better service by solving this problem, whereas strong restrictions on the function of conditional means do not get in the way of its performance. Beyond that, its allowance for time-varying regime probabilities makes advances to recent data reality which calls for extensive dynamic behavior of the regime variable. Although superior data adaption can be registered for copious mixture models the gain in forecast performance, if it exists, is poor. The misclassification of future regimes seems to be a crucial reason for eminent errors of prediction.

Under certain circumstances the *SMACD* model coincides with the *MSACD* model. There is a trivial concordance for one-regime specifications stooping to the ordinary *ACD* model introduced by Engle and Russell (1998). Moreover, the lack of regime-specificity associated with the conditional mean function makes the static *MSACD* model equivalent to the *SMACD* model.

An interesting asset of both is the interpretation in the context of recent

market microstructure models. The unobservable regime variable emulates the arrival or absence of private information harnessed by the informed traders acting on the financial market. The representatives of the *Discrete Mixture ACD framework* afford an excellent opportunity to look into the trade behavior of market participants influenced by the vicissitude of the information flow.

References

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.

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

Bauwens, L., Veredas, D., 2004. The stochastic conditional duration model: a latent variable model for the analysis of financial durations. Journal of Econometrics 119, 381–412.

Bessec, M., Bouabdallah, O., 2005. What causes the forecasting failure of Markov-switching models ? A Monte Carlo study. Tech. rep., Social Science Research Network.

Burr, I. W., 1942. Cumulative frequency functions. Annals of Mathematical Statistics 13, 215–232.

Buse, A., 1982. The likelihood ration, wald and lagrange multiplier tests: An expository note. American Statistician 36, 153–157.

De Luca, G., Gallo, Giampiero, M., 2004. Mixture processes for financial intradaily durations. Studies in Nonlinear Dynamics & Econometrics 8 (2), 1–18, article 8.

De Luca, G., Zuccolotto, P., 2003. Finite and infinite mixtures for financial durations. Metron - International Journal of Statistics LXI (3), 431–456.

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 irregulary spaced transaction data. Econometrica 66 (5), 1127–1162.

Eubank, R. L., Speckman, P., 1990. Curve fitting by polynomial-trigometric regression. Biometrica 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.

Ghysels, E., Gouriéroux, C., Jasiak, J., 2004. Stochastic volatility duration models. Journal of Econometrics 119, 413–433.

Grammig, J., Hujer, R., Kokot, S., Maurer, K. O., 1998. Modeling the Deutsche Telekom IPO using a new ACD specification - An application of the Burr-ACD model using high frequency IBIS data. Discussion Paper 55, Sonderforschungsbereich 373, Humboldt Universität zu Berlin.

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., 1994. Time Series Analysis. Princeton University Press, Princeton, New Jersey.

Hujer, R., Vuletić, S., Kokot, S., 2002. The markov switching ACD model. Working Paper Series: Finance and Accounting 90, University of Frankfurt/Main.

Kim, C.-J., 1994. Dynamic linear models with markov-switching. Journal of Econometrics 60 (1), 1–22.

Kokot, S., 2004. The Econometrics of Sequential Trade Models : Theory and Applications Using High Frequency Data. Springer Verlag.

Lee, C. M. C., Ready, M. J., 1991. Inferring trade direction from intraday data. Journal of Finance 46 (2), 733–746.

Lindsay, B. G., Roeder, K., 1992. Residual dignostics for mixture models. Journal of the American Statistical Association 87 (419), 785–794.

Liu, Y., Hong, Y., Wang, S., 2004. How well can autoregressive duration models capture the price durations dynamics of foreign exchanges. Tech. Rep. 2004007 EN, The China Center for Financial Research, Tsinhghua University.

Ljung, G. M., Box, G. E. P., 1978. On a measure of lack of fit in time series models. Biometrica 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

Statistic	Duration		Subsample	
	$\tilde{x}_n$	x_n	In	Out
Mean	14.630	1.000	0.930	1.082
Standard dev.	18.398	1.198	1.124	1.273
Overdispersion	23.137	1.434	1.359	1.498
Minimum	0.333	0.017	0.017	0.018
First quartile	3.000	0.194	0.175	0.212
Median	8.000	0.563	0.512	0.631
Third quartile	19.000	1.350	1.233	1.467
Maximum	231.000	13.806	11.594	13.806
Interquartile range	16.000	1.156	1.057	1.255
Number of obs.	31802	31802	17083	14719

Table 2
Number of observations and tests for autocorrelation

Day	In - sample		Out - sample	
	Nob.	Date	Nob.	Date
Mo	1605	Mar, 1	1365	Mar, 15
Tu	1902	Mar, 2	1804	Mar, 16
We	1543	Mar, 3	1428	Mar, 17
Th	1394	Mar, 4	1586	Mar, 18
Fr	1418	Mar, 5	1525	Mar, 19
Mo	1936	Mar, 8	1609	Mar, 22
Tu	1860	Mar, 9	1229	Mar, 23
We	1930	Mar, 10	1594	Mar, 24
Th	1873	Mar, 11	1357	Mar, 25
Fr	1622	Mar, 12	1222	Mar, 26

	Raw durations			Adj. durations		
	Lag	LB-test	p_{LB}	Lag	LB-test	p_{LB}
ℓ_L	1222	14553.15	0.00	1222	4672.23	0.00
ℓ_M	1590	18626.49	0.00	1590	5148.69	0.00
ℓ_H	1936	21001.65	0.00	1936	5570.52	0.00
	In - sample (adj.)			Out - sample (adj.)		
	Lag	LB-test	p_{LB}	Lag	LB-test	p_{LB}
ℓ_L	1394	2174.36	0.00	1222	1886.41	0.00
ℓ_M	1708	2526.00	0.00	1471	2151.69	0.00
ℓ_H	1936	2802.78	0.00	1804	2516.89	0.00

Table 3. Estimation results for two-regime specifications of the *SMACD* model

	$\{\mathcal{E},\mathcal{E}\}$		$\{\mathcal{E},\mathcal{B}\}$				$\{\mathcal{B},\mathcal{B}\}$			
	Variant $\overline{R}$		Variant $\underline{R}$		Variant $\overline{R}$		Variant $\underline{R}$		Variant $\overline{R}$	
Coeff.	Est	Stderr	Est	Stderr	Est	Stderr	Est	Stderr	Est	Stderr
ω	0.003	0.001	0.013	0.004	0.008	0.002	0.015	0.005	0.007	0.002
α	0.007	0.001	0.009	0.002	0.007	0.001	0.011	0.002	0.006	0.001
β	0.990	0.002	0.981	0.004	0.986	0.003	0.979	0.005	0.988	0.002
$\lambda^{(1)}$	0.754	0.010			0.912	0.019				
$\mu^{(1)}$									0.963	0.031
$\kappa^{(1)}$							1.294	0.015	1.303	0.015
$\kappa^{(2)}$			6.073	0.735	5.968	0.690	4.504	0.315	4.303	0.303
$\sigma^{(1)}$							0.284	0.021	0.294	0.021
$\sigma^{(2)}$			5.747	0.701	5.446	0.617	4.269	0.303	3.745	0.292
$\pi^{(1)}$	0.671	0.016	0.769	0.009	0.771	0.009	0.718	0.010	0.723	0.010
$\sharp\theta$		5		6		7		8		9
N		15801		15801		15801		15801		15801
$\mathcal{L}_I$		-14101.700		-13559.100		-13557.000		-13455.000		-13448.400
BIC		28251.800		27176.360		27181.740		26987.530		26983.870
	Variant $\overline{R}$		Variant $\underline{R}$		Variant $\overline{R}$		Variant $\underline{R}$		Variant $\overline{R}$	
Integral Tr.	Test	p-value	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0.500	0.873	0.504	0.110	0.504	0.104	0.500	0.815	0.500	0.907
Stand. dev.	0.286	0.055	0.280	*0.000*	0.280	*0.000*	0.287	0.438	0.287	0.417
First quart.	4.680	*0.000*	-6.490	*0.000*	-6.619	*0.000*	-2.338	0.019	-2.742	*0.006*
Median	0.485	0.628	1.599	0.110	1.679	0.093	3.874	*0.000*	3.922	*0.000*
Third quart.	-0.253	0.801	2.632	*0.009*	2.577	*0.010*	0.188	0.851	0.097	0.923
RT_ζ, $K = 10$	352.486	*0.000*	154.241	*0.000*	160.569	*0.000*	141.699	*0.000*	161.808	*0.000*

Table 4. Estimation results for three-regime specifications of the *SMACD* model

	$\{\mathcal{E},\mathcal{E},\mathcal{B}\}$				$\{\mathcal{E},\mathcal{B},\mathcal{B}\}$				$\{\mathcal{B},\mathcal{B},\mathcal{B}\}$			
	Variant R		Variant $\overline{R}$		Variant R		Variant $\overline{R}$		Variant R		Variant $\overline{R}$	
Coeff.	Est	Stderr	Est	Stderr	Est	Stderr	Est	Stderr	Est	Stderr	Est	Stderr
ω	0.008	0.002	0.008	0.002	0.007	0.002	0.008	0.002	0.008	0.002	0.007	0.001
α_1	0.007	0.001	0.007	0.001	0.006	0.001	0.006	0.001	0.006	0.001	0.005	0.001
β_1	0.986	0.003	0.986	0.003	0.987	0.002	0.986	0.003	0.987	0.002	0.988	0.002
$\lambda^{(1)}$			1.119	0.080	0.650	0.011	0.680	0.027				
$\lambda^{(2)}$	0.702	0.056	0.772	0.046								
$\mu^{(1)}$											5.378	1.311
$\mu^{(2)}$							2.096	0.927	0.611	0.059	0.459	0.037
$\kappa^{(1)}$									2.659	0.247	2.040	0.086
$\kappa^{(2)}$					1.494	0.036	1.670	0.181	1.417	0.020	1.540	0.026
$\kappa^{(3)}$	6.394	0.809	6.567	0.901	4.819	0.318	5.427	0.676	4.249	0.227	4.444	0.204
$\sigma^{(1)}$									1.914	0.295	0.814	0.082
$\sigma^{(2)}$					0.456	0.056	0.592	0.168	0.166	0.045	0.232	0.032
$\sigma^{(3)}$	5.757	0.694	5.943	0.779	3.947	0.270	4.316	0.419	3.039	0.160	3.137	0.157
$\pi^{(1)}$	0.574	0.057	0.373	0.106	0.297	0.019	0.449	0.103	0.280	0.037	0.320	0.026
$\pi^{(2)}$	0.207	0.053	0.411	0.105	0.457	0.018	0.331	0.082	0.465	0.048	0.446	0.031
$\mathcal{J}_=$		1				2				1		
$\sharp\theta$		8		9		10		11		12		13
N		15801		15801		15801		15801		15801		15801
$\mathcal{L}_I$		-13554.100		-13553.600		-13436.400		-13433.400		-13358.400		-13324.300
BIC		27185.550		27194.380		26969.480		26973.250		26832.900		26774.400
	Variant R		Variant $\overline{R}$		Variant R		Variant $\overline{R}$		Variant R		Variant $\overline{R}$	
Integral Tr.	Test	p-value	Test	p-value	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0.505	0.043	0.505	0.032	0.503	0.178	0.505	0.033	0.500	0.898	0.500	0.995
Stand. dev.	0.280	*0.000*	0.280	*0.000*	0.288	0.620	0.287	0.432	0.287	0.447	0.288	0.775
First quart.	-6.343	*0.000*	-6.380	*0.000*	-2.375	0.018	-1.750	0.080	-0.152	0.880	0.381	0.703
Median	1.376	0.169	1.058	0.290	1.710	0.087	-0.199	0.842	-0.581	0.561	-0.676	0.499
Third quart.	2.026	*0.043*	1.750	0.080	-2.457	0.014	-2.751	*0.006*	1.364	0.173	1.309	0.191
RT_ζ, $K=10$	148.823	*0.000*	142.695	*0.000*	107.953	*0.000*	77.424	*0.000*	55.411	*0.000*	52.463	*0.000*

Table 5. Estimation results for static specifications of the *MSACD* model

	2-regime				3-regime			
	$\{\mathcal{E}.\mathcal{B}\}$		$\{\mathcal{B},\mathcal{B}\}$		$\{\mathcal{E},\mathcal{B},\mathcal{B}\}$		$\{\mathcal{B},\mathcal{B},\mathcal{B}\}$	
Coeff.	Est	Stderr	Est	Stderr	Est	Stderr	Est	Stderr
ω	0.014	0.004	0.019	0.006	0.011	0.003	0.033	0.010
α	0.011	0.002	0.013	0.003	0.008	0.002	0.018	0.003
β	0.979	0.005	0.974	0.006	0.984	0.003	0.962	0.009
$\kappa^{(1)}$			1.298	0.027			1.323	0.038
$\kappa^{(2)}$	5.634	0.461	4.449	0.259	5.133	0.333	4.793	0.252
$\kappa^{(3)}$					1.366	0.046	2.658	0.235
$\sigma^{(1)}$			0.290	0.028			0.176	0.034
$\sigma^{(2)}$	5.331	0.440	4.217	0.249	4.870	0.319	4.589	0.243
$\sigma^{(3)}$					0.414	0.055	2.175	0.213
p_{11}	0.724	0.011	0.673	0.013	0.483	0.032	0.569	0.025
p_{12}	0.904	0.019	0.832	0.019	0.679	0.044	0.706	0.038
p_{13}					0.000	-	0.263	0.056
p_{21}					0.517	0.032	0.358	0.018
p_{22}					0.137	0.017	0.194	0.021
p_{23}					0.107	0.018	0.072	0.050
$\sharp\theta$		7		9		13		15
N		15801		15801		15801		15801
$\mathcal{L}_I$		-13537.500		-13428.100		-13372.600		-13346.600
BIC		27142.870		26943.340		26870.980		26838.400
	$\{\mathcal{E},\mathcal{B}\}$		$\{\mathcal{B},\mathcal{B}\}$		$\{\mathcal{E},\mathcal{B},\mathcal{B}\}$		$\{\mathcal{B},\mathcal{B},\mathcal{B}\}$	
Integral Tr.	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0.504	0.103	0.499	0.733	0.500	0.880	0.498	0.425
Stand. dev.	0.280	*0.000*	0.288	0.511	0.286	0.154	0.287	0.453
First quart.	-6.931	*0.000*	-2.650	0.008	-2.411	0.016	0.345	0.731
Median	1.440	0.150	3.667	*0.000*	2.872	*0.004*	0.979	0.328
Third quart.	2.981	*0.003*	0.409	0.683	0.097	0.923	1.603	0.109
RT_ζ. $K = 10$	124.906	*0.000*	100.951	*0.000*	64.357	*0.000*	16.422	0.059

Table 6
Out-sample performance of the *SMACD* model

	2-regime				3-regime			
					Semirestrictive $\{\mathcal{E},\mathcal{E},\mathcal{B}\}$			
					First week		Second week	
					Test	p-value	Test	p-value
Mean					0,522	0,000	0,520	0,000
Stand. Dev.					0,281	0,001	0,289	0,848
1-st quart.					-8,277	0,000	-5,734	0,000
Median					-4,439	0,000	-3,377	0,001
3-rd quart.					-2,494	0,013	-4,539	0,000
$RT,\ K=10$					102,273	0,000	89,793	0,000
N						7100		6441
MSE						1,441		1,812
MAE						0,858		0,953
	Nonrestrictive $\{\mathcal{E},\mathcal{E}\}$				Nonrestrictive $\{\mathcal{E},\mathcal{E},\mathcal{B}\}$			
	First week		Second week		First week		Second week	
	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0,512	0,000	0,509	0,016	0,522	0,000	0,521	0,000
Stand. Dev.	0,285	0,130	0,293	0,124	0,281	0,001	0,289	0,845
1-st quart.	-0,713	0,476	2,842	0,005	-8,277	0,000	-5,705	0,000
Median	-3,632	0,000	-1,782	0,075	-4,676	0,000	-3,427	0,001
3-rd quart.	-2,412	0,016	-3,619	0,000	-2,741	0,006	-4,597	0,000
$RT,\ K=10$	87,045	0,000	59,102	0,000	101,012	0,000	89,830	0,000
N		7100		6441		7100		6441
MSE		1,441		1,813		1,441		1,812
MAE		0,846		0,946		0,860		0,954
	Restrictive $\{\mathcal{E},\mathcal{B}\}$				Semirestrictive $\{\mathcal{E},\mathcal{B},\mathcal{B}\}$			
	First week		Second week		First week		Second week	
	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0,522	0,000	0,520	0,000	0,521	0,000	0,519	0,000
Stand. Dev.	0,281	0,001	0,289	0,828	0,288	0,879	0,296	0,003
1-st quart.	-8,387	0,000	-5,849	0,000	-6,222	0,000	-3,288	0,001
Median	-4,510	0,000	-3,302	0,001	-3,893	0,000	-2,878	0,004
3-rd quart.	-2,412	0,016	-4,453	0,000	-5,180	0,000	-6,899	0,000
$RT,\ K=10$	105,181	0,000	91,499	0,000	94,768	0,000	98,075	0,000
N		7100		6441		7100		6441
MSE		1,459		1,820		1,442		1,815
MAE		0,900		0,990		0,846		0,944
	Nonrestrictive $\{\mathcal{E},\mathcal{B}\}$				Nonrestrictive $\{\mathcal{E},\mathcal{B},\mathcal{B}\}$			
	First week		Second week		First week		Second week	
	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0,521	0,000	0,519	0,000	0,523	0,000	0,519	0,000
Stand. Dev.	0,280	0,001	0,289	0,925	0,288	0,713	0,296	0,003
1-st quart.	-8,332	0,000	-5,877	0,000	-6,057	0,000	-3,288	0,001
Median	-4,154	0,000	-3,028	0,003	-5,934	0,000	-3,277	0,001
3-rd quart.	-1,891	0,059	-3,964	0,000	-5,783	0,000	-7,187	0,000
$RT,\ K=10$	109,798	0,000	90,759	0,000	98,763	0,000	95,661	0,000
N		7100		6441		7100		6441
MSE		1,443		1,811		1,443		1,815
MAE		0,866		0,959		0,842		0,943
	Restrictive $\{\mathcal{B},\mathcal{B}\}$				Semirestrictive $\{\mathcal{B},\mathcal{B},\mathcal{B}\}$			
	First week		Second week		First week		Second week	
	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0,518	0,000	0,517	0,000	0,518	0,000	0,516	0,000
Stand. Dev.	0,288	0,813	0,296	0,003	0,288	0,702	0,296	0,005
1-st quart.	-6,359	0,000	-4,122	0,000	-4,879	0,000	-1,734	0,083
Median	-3,086	0,002	-2,181	0,029	-6,029	0,000	-4,349	0,000
3-rd quart.	-4,111	0,000	-5,892	0,000	-2,905	0,004	-4,885	0,000
$RT,\ K=10$	90,922	0,000	103,042	0,000	78,013	0,000	58,434	0,000
N		7100		6441		7100		6441
MSE		1,483		1,839		1,442		1,811
MAE		0,928		1,018		0,861		0,958
	Nonrestrictive $\{\mathcal{B},\mathcal{B}\}$				Nonrestrictive $\{\mathcal{B},\mathcal{B},\mathcal{B}\}$			
	First week		Second week		First week		Second week	
	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0,516	0,000	0,515	0,000	0,517	0,000	0,516	0,000
Stand. Dev.	0,288	0,713	0,296	0,004	0,288	0,926	0,296	0,002
1-st quart.	-6,276	0,000	-3,489	0,001	-4,550	0,000	-1,360	0,174
Median	-2,231	0,026	-1,782	0,075	-6,053	0,000	-4,847	0,000
3-rd quart.	-3,700	0,000	-5,345	0,000	-2,933	0,003	-4,741	0,000
$RT,\ K=10$	93,302	0,000	98,715	0,000	76,339	0,000	61,348	0,000
N		7100		6441		7100		6441
MSE		1,441		1,811		1,444		1,820
MAE		0,860		0,954		0,839		0,935

Table 7
Out-sample performance of the *MSACD* model

	2-regime				3-regime			
	$\{\mathcal{E}, \mathcal{B}\}$				$\{\mathcal{E}, \mathcal{B}, \mathcal{B}\}$			
	First week		Second week		First week		Second week	
	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0,523	0,000	0,521	0,000	0,517	0,000	0,517	0,000
Stand. Dev.	0,280	0,001	0,289	0,912	0,288	0,865	0,296	0,004
1-st quart.	-8,798	0,000	-4,957	0,000	-4,851	0,000	-2,511	0,012
Median	-4,771	0,000	-3,427	0,001	-3,299	0,001	-2,629	0,009
3-rd quart.	-2,330	0,020	-4,309	0,000	-4,358	0,000	-5,806	0,000
RT, $K = 10$	99,440	0,000	76,975	0,000	40,327	0,000	75,998	0,000
N		7100		6441		7100		6441
MSE		1,460		1,820		1,472		1,830
MAE		0,902		0,992		0,916		1,007
	$\{\mathcal{B}, \mathcal{B}\}$				$\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$			
	First week		Second week		First week		Second week	
	Test	p-value	Test	p-value	Test	p-value	Test	p-value
Mean	0,519	0,000	0,518	0,000	0,520	0,000	0,520	0,000
Stand. Dev.	0,288	0,842	0,296	0,003	0,288	0,653	0,296	0,007
1-st quart.	-6,139	0,000	-3,604	0,000	-5,290	0,000	-2,885	0,004
Median	-3,442	0,001	-2,505	0,012	-5,744	0,000	-4,249	0,000
3-rd quart.	-4,467	0,000	-5,978	0,000	-3,590	0,000	-5,662	0,000
RT, $K = 10$	83,373	0,000	104,791	0,000	45,031	0,000	62,571	0,000
N		7100		6441		7100		6441
MSE		1,484		1,838		1,554		1,889
MAE		0,929		1,018		0,987		1,065

Table 8
Testing parameter equalities for the $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ specification

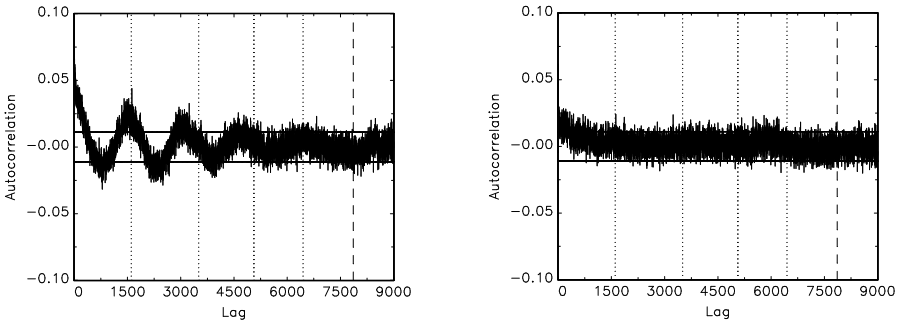
Null hypothesis	Regime (combinations)					
	$(i,j) = (1,2)$		$(i,j) = (1,3)$		$(i,j) = (2,3)$	
Linear equality	W-test	p-value	W-test	p-value	W-test	p-value
	Semirestrictive *SMACD* model					
$H_0 : \kappa^{(i)} = \kappa^{(j)}$	25.922	0.000	50.650	0.000	151.904	0.000
$H_0 : \sigma^{(i)} = \sigma^{(j)}$	45.367	0.000	18.544	0.000	388.828	0.000
$H_0 : \pi_i = \pi_j$	4.794	0.029				
	Nonrestrictive *SMACD* model					
$H_0 : \kappa^{(i)} = \kappa^{(j)}$	32.140	0.000	142.654	0.000	186.834	0.000
$H_0 : \sigma^{(i)} = \sigma^{(j)}$	83.133	0.000	188.367	0.000	334.562	0.000
$H_0 : \pi_i = \pi_j$	5.033	0.025				
	MSACD model					
$H_0 : \kappa^{(i)} = \kappa^{(j)}$	171.393	0.000	33.712	0.000	33.354	0.000
$H_0 : \sigma^{(i)} = \sigma^{(j)}$	314.271	0.000	92.911	0.000	50.759	0.000
$H_0 : p_{2i} = p_{2j}$	11.940	0.001	18.365	0.000	41.164	0.000
$H_0 : p_{1i} = p_{1j}$	60.464	0.000	26.686	0.000	4.221	0.040
	$j = 1$		$j = 2$		$j = 3$	
Linear inequality	t-test	p-value	t-test	p-value	t-test	p-value
	Semirestrictive *SMACD* model					
$H_0 : \kappa^{(j)} < 1$	6.717	0.000	20.850	0.000	14.313	0.000
	Nonrestrictive *SMACD* model					
$H_0 : \kappa^{(j)} < 1$	12.093	0.000	20.769	0.000	16.882	0.000
	MSACD model					
$H_0 : \kappa^{(j)} < 1$	8.500	0.000	15.052	0.000	7.055	0.000

Table 9
Regime characteristics for the $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ specification

	Regime		
	$j = 1$	$j = 2$	$j = 3$
	Semirestrictive *SMACD* model		
$\pi^{(j)}$	0.280	0.465	0.255
$E\left(\varepsilon_n \vert s_n = j; \hat{\theta}\right)$	1.000	1.434	0.211
	Nonrestrictive *SMACD* model		
$\pi^{(j)}$	0.320	0.446	0.234
$E\left(\varepsilon_n \vert s_n = j; \hat{\theta}\right)$	0.585	1.711	0.215
	MSACD model		
π_j	0.545	0.259	0.196
$E\left(\varepsilon_n \vert s_n = j; \hat{\theta}\right)$	1.000	1.000	1.000

Figures

Fig. 1. Autocorrelation function for durations



Raw durations $\tilde{x}_n$ Adjusted durations x_n

Fig. 2. QQ-plots and histograms for integral transforms implied by the $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ specification

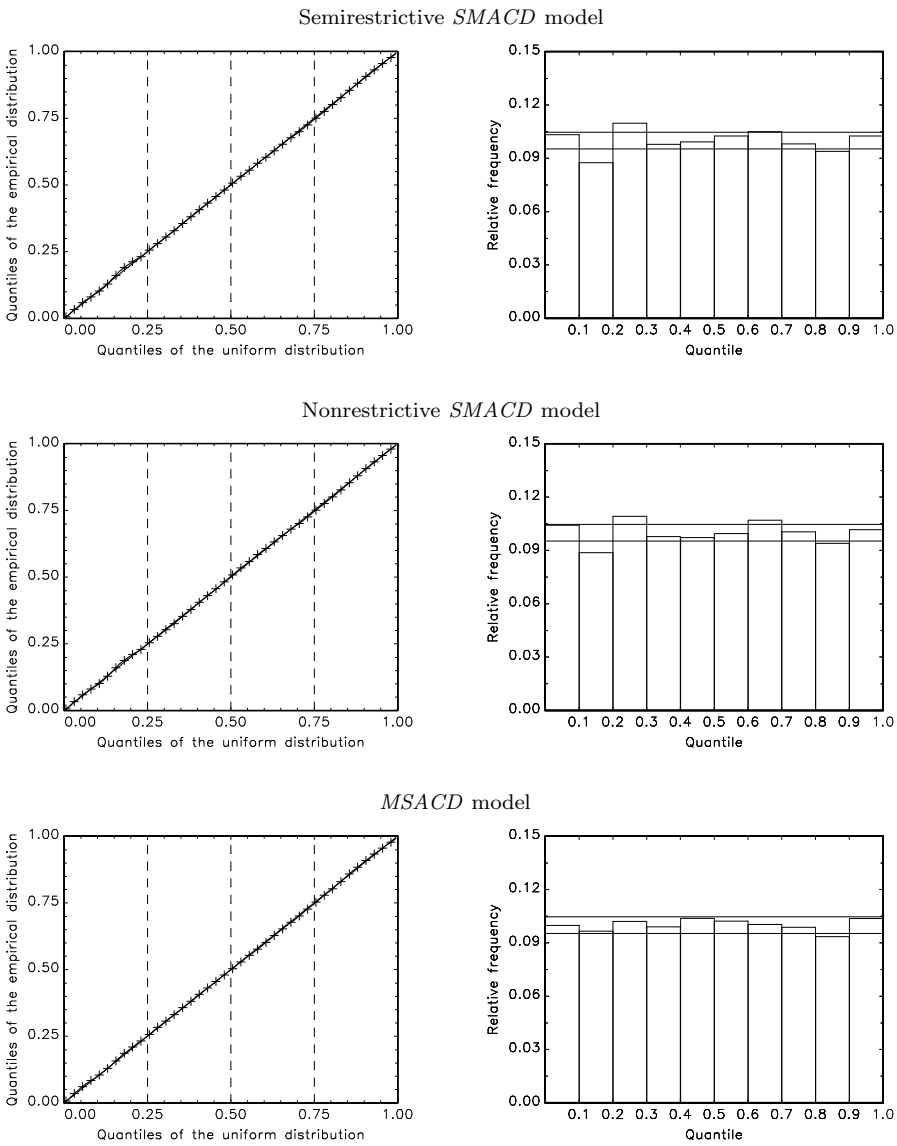
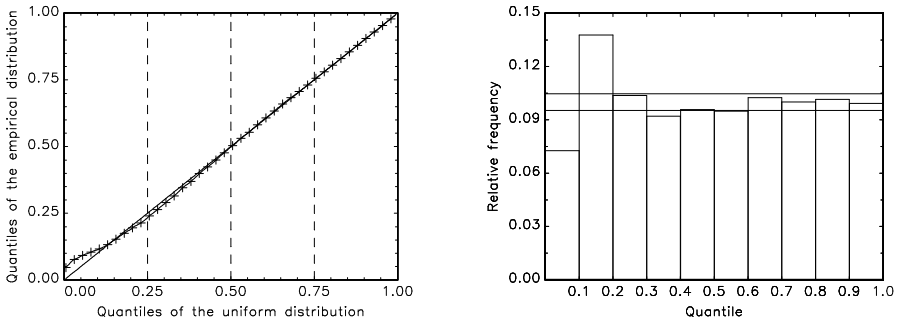
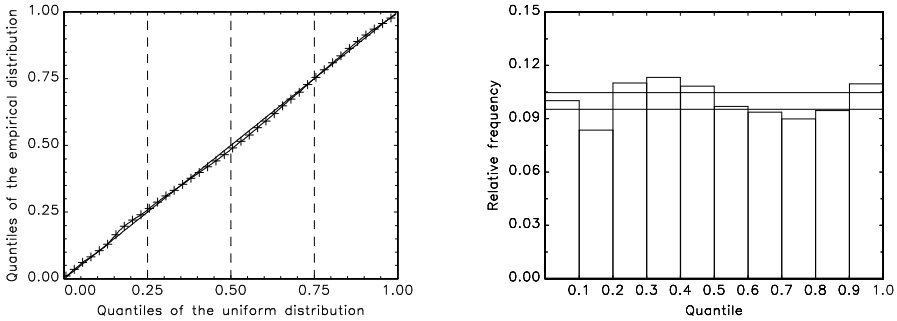


Fig. 3. QQ-plots and histograms for integral transforms implied by an assortment of two-regime specifications

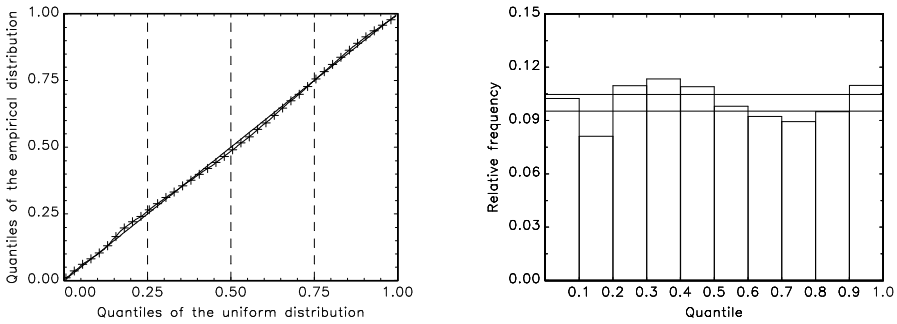
$\{\mathcal{E},\mathcal{E}\}$ specification of the nonrestrictive *SMACD* model



$\{\mathcal{B},\mathcal{B}\}$ specification of the restrictive *SMACD* model



$\{\mathcal{B},\mathcal{B}\}$ specification of the nonrestrictive *SMACD* model



$\{\mathcal{B},\mathcal{B}\}$ specification of the *MSACD* model

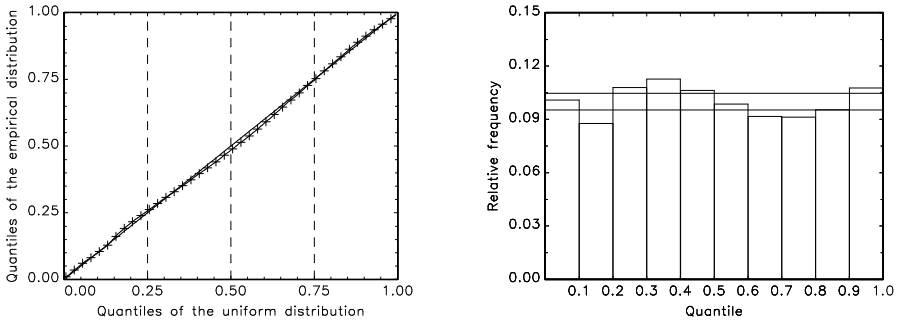


Fig. 4. Analysis of the $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ specification

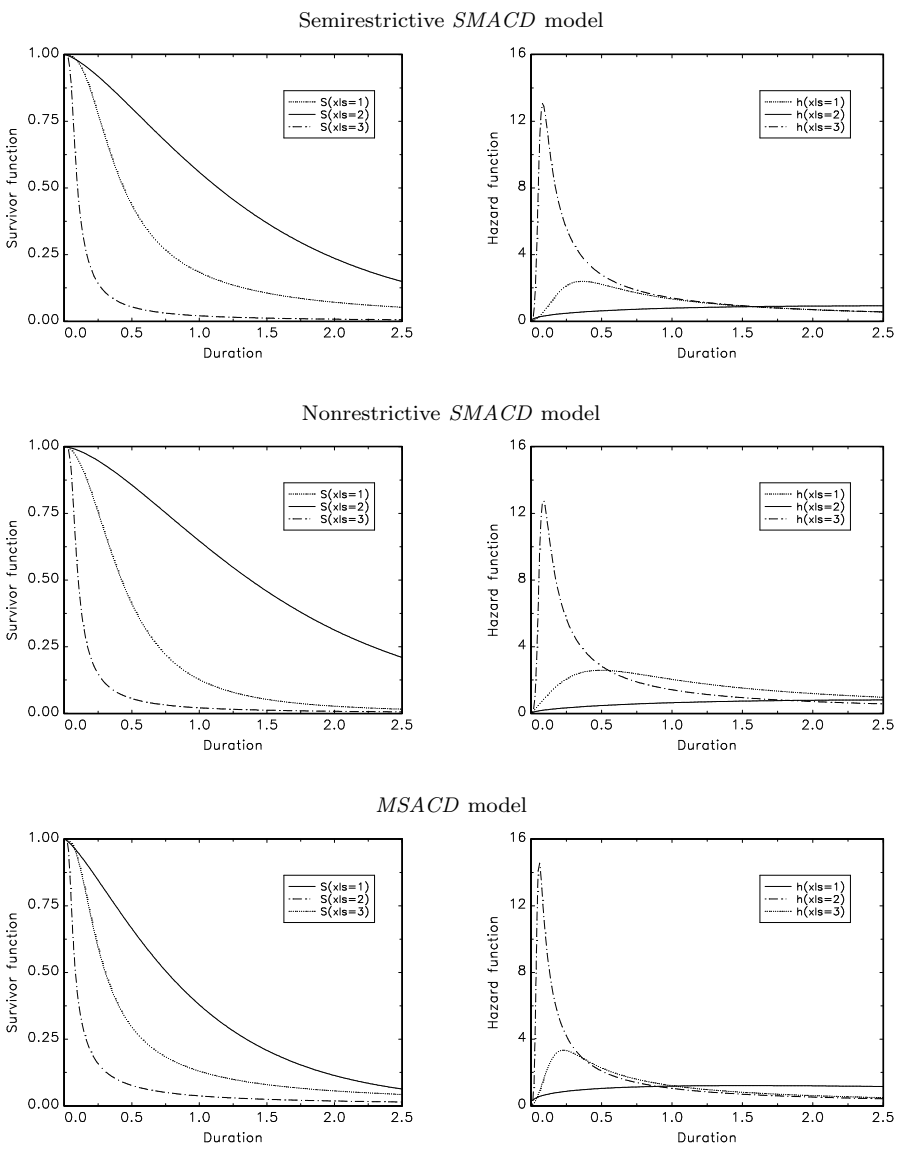
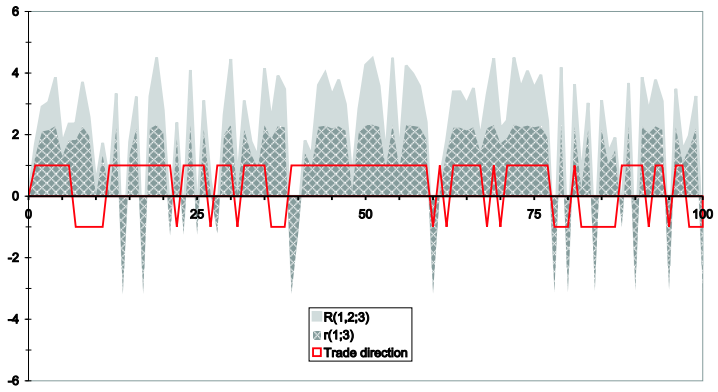
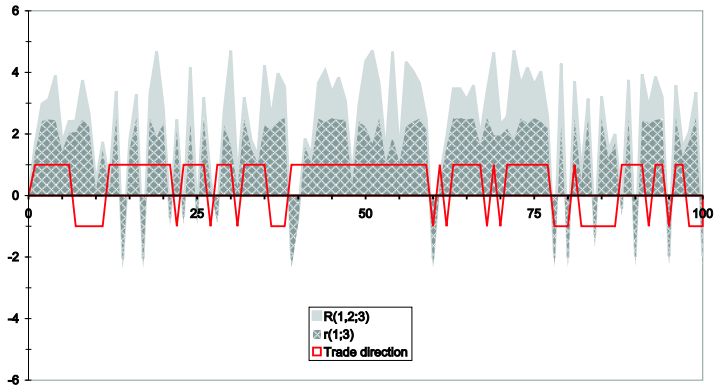


Fig. 5. Regime identification of the $\{\mathcal{B}, \mathcal{B}, \mathcal{B}\}$ specification

Semirestrictive *SMACD* model



Nonrestrictive *SMACD* model



MSACD model