A COMPARISON OF FINANCIAL DURATION MODELS VIA DENSITY FORECASTS

Luc Bauwens

CORE and Department of Economics, Université catholique de Louvain

Voie du Roman Pays, 34, B-1348 Louvain-La-Neuve, Belgium

Ph: +32 10 474336. Fax: +32 10 474301

e-mail: bauwens@core.ucl.ac.be

Pierre Giot

CORE and University of Namur

Rempart de la Vierge, 8, B-5000 Namur, Belgium

Ph: +32 81 724887. Fax: +32 81 724840

email: pierre.giot@fundp.ac.be

Joachim Grammig[1]

University of Tübingen

Faculty of Economics

Mohlstr. 36, D-72074 Tübingen, Germany

Ph: +49 7071 29 78159. Fax: +49 7071 29 5546

e-mail: joachim.grammig@uni-tuebingen.de

David Veredas

CORE and CentER, Tilburg University

Department of Econometrics and OR

P.O. Box 90153, 5000 Tilburg, The Netherlands

Ph: +31 13466 3151. Fax: +31 13466 3280

e-mail: d.veredas@uvt.nl

[1]Corresponding author

A COMPARISON OF FINANCIAL DURATION MODELS VIA DENSITY FORECASTS

Abstract

Using density forecast evaluation techniques we compare the predictive performance of econometric specifications that have been developed for modeling duration processes in intra-day financial markets. The model portfolio encompasses various variants of the Autoregressive Conditional Duration (ACD) model and recently proposed dynamic factor models. The evaluation is conducted on time series of trade, price and volume durations computed from transaction data of NYSE listed stocks. The results show that simpler approaches perform at least as well as more complex methods. With respect to modeling trade duration processes, standard ACD models successfully account for duration dynamics whilst none of the models provides an acceptable specification for the conditional duration distribution. We find that the Logarithmic ACD, if based on a flexible innovation distribution, provides a quite robust and useful framework for the modeling of price and volume duration processes.

Keywords: Duration processes, transactions data, intra-day financial markets, density forecast evaluation.

JEL classification: C41, C52, C53, G14.

1 Introduction

Several contributions to the market microstructure literature emphasize that the waiting times between intra-day market events like trades, quote updates, price changes, and order arrivals play a key role for understanding the processing of private and public information in financial markets.[1] Hence, the accessibility of financial data at a micro level, which, in an ideal case, includes real time recordings of trades, order arrivals and quote updates, as well as the corresponding prices, volumes and time stamps, opened new perspectives for the empirical analysis of market microstructure. By appropriately 'thinning' the data, it is possible to define almost any event of interest, and compute the corresponding durations. In particular, Engle and Russell (1997) show that *price durations* are closely linked to the instantaneous volatility of the mid-quote price process and thus play an interesting role in option pricing (Pringent, Renault and Scaillet, 2001). (Inter-)*trade* and *volume durations* mirror features such as market liquidity and the information arrival rate (Gouriéroux, Jasiak and LeFol, 1999). The ability to predict those financial duration processes is therefore of crucial importance for empirical market microstructure analysis.

In this paper we assess the adequacy of recently proposed econometric approaches that we classify as financial duration models for the analysis of trade, volume and price duration processes. For this purpose we employ density forecasting techniques which enable the researcher to compare non-nested models and detect possible specification problems.

An econometric framework for the modeling of intertemporally correlated event arrival times was provided by Engle and Russell (1998), who introduced the Autoregressive Conditional Duration (ACD) model. The ACD approach combines elements from transition analysis (Lancaster, 1990) and Engle's (1982) Autoregressive Conditional Heteroskedasticity (ARCH) model. Indeed, upon closer inspection, the motivation behind the ACD and the ARCH model appears similar: financial market events, like trades and quote changes, occur in clusters. Following the contribution of Engle and Russell (1998), several modifications of the basic models have been put forward. Bauwens and Giot (2000) introduced a logarithmic version of the ACD model, which implies a nonlinear relation between the duration and its lags. The Log-ACD is especially useful when conditioning variables are included in the model in order to test implications of microstructure theories. As an alternative to the exponential and Weibull distributions used in Engle and Russell's (1998) seminal paper, Lunde (1999) and Grammig and Maurer (2000) considered ACD specifications based on the generalized gamma and the Burr distribution (both nest Weibull and exponential as special cases). Zhang, Russell and Tsay (2001) advocated the Threshold

[1] See O'Hara's book (1995) for a survey and references, and Goodhart and O'Hara (1997) and Madhavan (2000) for more recent surveys.

ACD model, formulated in the spirit of threshold autoregressive models, in order to capture a possible nonlinear relation between the duration and predetermined variables and to account for regime switches. Ghysels, Gouriéroux, and Jasiak (1998) developed the Stochastic Volatility Duration (SVD) model, a dynamic two factor model, which is designed to account both for mean and variance dynamics in financial duration processes. Another (single) factor model, the Stochastic Conditional Duration model (SCD), was put forth by Bauwens and Veredas (1999). Since the main motivation behind their development is the modeling of high frequency financial data and duration processes in financial markets we refer to those approaches as financial duration models.

Given the availability of a plethora of competing specifications, it becomes relevant to examine whether some perform better than others, and which model type is particularly suited, or inadequate, for those financial duration processes mentioned above. However, despite the recent boom in empirical applications, the literature has devoted so far little attention to testing the specification of financial duration models. The practice is to perform simple diagnostic tests to check whether the standardised residuals in models of ACD type are independent and identically distributed (iid). Most papers adopting Engle and Russell's ACD approach or one of its descendants use $Q-$statistic to test for serial dependence in the estimated residuals series or check the first and second moments of the duration innovations with particular attention to measuring excess dispersion, whilst others investigate QQ-plots (Bauwens and Veredas, 1999) and use Bartlett identity tests (Pringent, Renault and Scaillet, 1999) for specification testing. More recently, Fernandes and Grammig (2000) have proposed a nonparametric specification testing framework for ACD models.

Our aim is to contribute to this issue using the criterion of predictive ability. Diebold, Gunther and Tay (1998) have proposed tools to evaluate density forecasts produced by a given model, and we apply these tools to several financial duration models estimated on the same data. The procedure has the advantage of being easy to implement, of providing hints to sources of misspecification and of being suitable for comparing models which are not necessarily nested. This is the case for the financial duration models considered in this paper. Moreover, even if a model nests all the others, this model still needs to be evaluated and may be a bad performer in some respects. In our comparison we include both Engle and Russell's standard ACD models and its extensions, like the Burr-ACD and the Generalized Gamma ACD (both in linear and in logarithmic versions), as well as more complex approaches, like the Threshold ACD and the dynamic factor models, the SCD and the SVD.

Our results show that the more elaborated specifications (which also demand considerably more computer resources and more sophisticated estimation techniques) do not offer real improvements over more basic models in terms of predictive evaluation. On the contrary, the simpler approaches seem to do at least as well as more complex methods. With respect to modeling trade durations, standard ACD models do a good job of accounting for duration dynamics, yet none of the financial duration models considered in this paper is able provide an acceptable specification for the conditional duration distribution. Regarding the modeling of price and volume duration processes the situation is better: Our results indicate that a logarithmic ACD based on a flexible innovation distribution (generalized gamma or Burr) can be recommended as a useful modeling framework. We also conclude that the factor models, especially the SVD, should be endowed with less restrictive autoregressive dynamics and also more flexible marginal distributions.

The paper is organized as follows. In Section 2, we explain the use of density forecasts for model evaluation. In Section 3, we review specification and estimation of the financial duration models that are considered in this paper and we also explain the density forecast that is implied by each model. The results of the comparison using NYSE transactions data are discussed in Section 4. The last section contains our conclusions and an outlook for further research.

2 Density forecasts

In order to provide a comparative assessment of the financial duration models reviewed in the next section we employ the methods for evaluating density forecasts revived and formalized by Diebold, Gunther and Tay (1998), henceforth referred to as DGT. Their method relies on a result that dates back at least to Rosenblatt (1952). It allows comparing non-nested models by evaluating their forecasting performance regardless of the users' loss functions. Applications of the method can also be found in Shephard (1994) and Kim *et al.* (1998).

A density forecast is a density defined for the next observation of the variable of interest. In the context of this paper this is a price, trade or volume duration. The density forecast is usually implied by the model that is used for estimation, but it may also be an informal prediction issued by a forecaster. In a dynamic context, the forecast density is computed conditional on the information available at the most recent occurrence of the event (i.e. encompassing the sequence of past durations).

The basic idea behind the DGT approach is straightforward. In the context of the present paper we deal with an ordered sequence of durations x_i, $i = 1, \ldots, n$. Let us denote by $\{f_i(x_i \mid \mathcal{H}_i)\}_{i=1}^n$ a sequence of one-step-ahead density forecasts produced by some financial duration model and by $\{p_i(x_i \mid \mathcal{H}_i)\}_{i=1}^n$ the sequence of densities defining the data generating process governing the duration series x_i. $\mathcal{H}_i$ denotes the conditioning information set generated by the durations preceding x_i. DGT show that the correct density is weakly superior to all other forecasts, i.e. will be preferred, in terms of expected loss, by all forecast users regardless of their loss functions. This suggests that forecasts should be evaluated by assessing whether the forecasting densities are correct, i.e. whether

$$\{f_i(x_i \mid \mathcal{H}_i)\}_{i=1}^{n} = \{p_i(x_i \mid \mathcal{H}_i)\}_{i=1}^{n}. \tag{1}$$

At first sight, testing whether this is true appears difficult because $p_i(x_i \mid \mathcal{H}_i)$ is never observed. However, the properties of the probability integral transform

$$z_i = \int_{-\infty}^{x_i} f_i(u)du \tag{2}$$

provide the solution to this problem. Rosenblatt (1952) derived that the distribution of the probability integral transform under the null hypothesis (1) is uniform $U(0,1)$. DGT extended this result and showed that under the null hypothesis, the distribution of the sequence of probability transforms $\{z_i\}_{i=1}^{n}$ of $\{x_i\}_{i=1}^{n}$ with respect to $\{f_i(x_i \mid \mathcal{H}_i)\}_{i=1}^{n}$ is iid $U(0,1)$. This means that the empirical sequence of probabilty integral transforms implied by a model's or an individual's forecast can be used for testing.

For the purpose of this paper we follow DGT and emphasize visual inspection of autocorrelograms and the histogram implied by the z sequence and the informal nature of the approach.[2] Formal tests are also conducted, but, as DGT argue, they are not constructive: When rejection occurs the tests do not provide guidance to the origins of the problems. The autocorrelograms will reveal potential deficiencies of a model to account for the dynamics of the duration process. Serial correlation in the z series may indicate that the mean dynamics have been inadequately modeled by the forecaster. To detect nonlinear forms of dependence one may not only examine the z correlogram but also those of powers of z. The Ljung-Box Q-statistic computed on the z and z^2 series or the recently proposed test for independence by Brock, Dechert, Scheinkman and LeBaron (1996) can be straightforwardly applied to test departures from the iid assumption. By plotting a histogram based on an empirical z sequence, departures from iid $U(0,1)$ can easily be detected. Visual inspection will often provide obvious hints to the reasons for model failure. As an example, a humped shape of the z-histogram would indicate that the issued forecasts are too narrow and that the tails of the true density are not accounted for. This is the classic result obtained when predicting notoriously fat-tailed asset return distributions based on a model with Gaussian innovations. On the other hand, a U-shape of the histogram would suggest that the model issues forecasts that either under- or overestimate too frequently. Confidence bands for the $z-$histogram and a simple test for iid uniformity can be derived from the fact that under the null

[2]Berkowitz (2001) suggests an alternative, although rather similar, approach to the DGT method.

of iid $U(0,1)$ behavior of the $z-$sequence the joint distribution of the heights of the $z-$histogram is multinomial, i.e.

$$p(n_i) = \binom{n}{n_i} p^{n_i} (1-p)^{n-n_i} \tag{3}$$

where n is the sample size, n_i is the number of observations in the i^{th} bin, and $p = 1/m$ with m the number of histogram bins. One can therefore compute Pearson's goodness of fit statistic

$$\chi^2 = \sum_{j=1}^{m} \frac{(n_i - np)^2}{np} \tag{4}$$

which is under the null hypothesis asymptotically distributed chi-square with $m-1$ degrees of freedom.

The great advantage of the DGT method is its flexibility and visualization of possible specification problems. For the purpose of this paper it is important that non-nested models can be compared. Furthermore, the method does not require that the data generating process remains the same over time. Instead, the true data generating process $\{p_i(x_i \mid \mathcal{H}_i)\}_{i=1}^{n}$ may exhibit all kinds of structural change, as indicated by its time subscript. The important point is that the model's (or forecaster's) density forecasts $\{f_i(x_i \mid \mathcal{H}_i)\}_{i=1}^{m}$ are able to account for these features. Some drawbacks and limitations of the density evaluation method should be acknowledged, however. First, one may feel uncomfortable with its informal nature (which is also its strength, however). Second, the DGT method does not take into account the effect of parameter estimation uncertainty when evaluating density forecasts. Extending the method would be beyond the scope of the present paper, because the drawback is less severe in the context of our study: The high frequency data sets that are available to estimate financial duration models regale the econometrician with large sample sizes. Hence, the effect of parameter estimation uncertainty is alleviated. Third, as pointed out by a referee, one has to be aware that the idea of the DGT approach is to assess whether a conditional distribution is dynamically specified correctly. The true conditional distribution is the best predictive density, regardless of the chosen loss function. This implies, however, that the DGT method is not useful for a comparison of misspecified models: If one wants to rank misspecified densities, then the choice of the loss function becomes important.

3 Model Review

In this section, we review specification and estimation of the financial duration models that we intend to compare. Five classes of financial duration models are distinguishable: Autoregressive Conditional Duration, Logarithmic ACD, Threshold ACD, Stochastic Conditional Duration, and Stochastic Volatility Duration models. For each model class, we explain the specification, show the parameter estimation method and describe the implied conditional density forecast. Table 1 summarizes convenient model acronyms for ease of reference.

3.1 ACD Models

Model specification

The basic reference for ACD models is Engle and Russell (1998) (henceforth ER). ER specify the duration x_i as

$$x_i = \Psi_i \epsilon_i, \tag{5}$$

where ϵ_i (for $i = 1, \ldots, n$) are iid innovations with $\mathbb{E}(\epsilon_i) = \mu$, such that $\mathbb{E}(x_i|\mathcal{H}_i) = \Psi_i \mu$. Ψ_i is specified as a linear function of past durations and conditional durations and is called the conditional duration. For simplicity we restrict our attention to the ACD(1,1) case,

$$\Psi_i = \omega + \alpha x_{i-1} + \beta \Psi_{i-1}, \tag{6}$$

where $\omega > 0$, $\alpha \geq 0$, and $\beta \geq 0$ (but $\beta = 0$ if $\alpha = 0$). Although not necessary, these sign restrictions are convenient to ensure the positivity of Ψ_i in estimation.

A parametric model is obtained when the distribution of ϵ_i is specified up to a finite number of parameters. Engle and Russell (1998) proposed the standard exponential distribution, and as an extension the Weibull distribution with shape parameter equal to γ and scale parameter equal to 1. The exponential specification provides the quasi-likelihood function for parametric models, such that if $\mathbb{E}(x_i|\mathcal{H}_i)$ is correctly specified, the QML estimator of (ω, α, β) is consistent and asymptotically normal (under suitable regularity conditions). The Weibull distribution is more flexible in that it nests the exponential one (for $\gamma = 1$), and allows a non-flat hazard function. However, the Weibull hazard function is necessarily monotone: increasing if $\gamma > 1$, decreasing if $\gamma < 1$. As documented independently by Bauwens and Veredas (1999), and by Grammig and Maurer (2000), the hazard function of several types of financial durations may be increasing for small durations and decreasing for long durations (see Section 4). To account for this stylized fact, Grammig and Maurer (2000) proposed to use the Burr distribution which can have a hump-shaped hazard and nests the Weibull distribution as a particular case. The Burr distribution has two shape parameters, so that there is no one-to-one correspondence (contrary to the case of the Weibull distribution) between the properties of overdispersion (underdispersion) and of decreasing (respectively, increasing) hazard. Another distribution which has also two shape parameters and breaks this one-to-one correspondence is the generalized gamma.[3] It has been used by Lunde (1999) who has formulated a Generalized Gamma ACD model. Burr and Generalized Gamma nest ER's Exponential and Weibull ACD as special cases

Estimation and density forecast

Estimation of the parameters of ACD models can be done by maximizing the likelihood function, which is the product of the n densities $f(x_i\Psi_i^{-1})\Psi_i^{-1}$, where Ψ_i is defined by (6), and $f(\cdot)$ is the appropriate density, i.e. exponential, Weibull, Burr or generalized gamma, respectively, depending on the innovation distribution. The density forecast of ACD models is the conditional density of x_i given the past information, and it is easily computed as

$$f(x_i\Psi_i^{-1})\Psi_i^{-1}.$$

3.2 Log-ACD Models

Model specification

Bauwens and Giot (2000) argue that when additional explanatory variables implied by market microstructure are added linearly to the right-hand side of equation (6), and theory suggests negative slope coefficients of those variables, then the conditional duration Ψ_i may become negative which is not admissible as durations have to be non-negative. Imposing non-negativity restrictions on the slope coefficients, however, is tantamount to delete the corresponding variables, which is self-destructive. This led Bauwens and Giot (2000) to consider logarithmic versions of the ACD models. In the Log-ACD specification, equation (5) is therefore written as

$$x_i = \exp(\psi_i)\,\epsilon_i, \tag{7}$$

such that ψ_i is the logarithm of the conditional duration $\Psi_i = \exp(\psi_i)$. The difference compared to ACD models is that the autoregressive equation bears on the logarithm of the conditional duration rather than on the conditional duration itself. Two possible specifications of this equation are

$$\psi_i = \omega + \alpha \log x_{i-1} + \beta\psi_{i-1} \quad \text{(type 1)} \tag{8}$$

and

$$\psi_i = \omega + \alpha\epsilon_{i-1} + \beta\psi_{i-1} \qquad \text{(type 2)}. \tag{9}$$

[3]The generalized gamma distribution nests the Weibull and the gamma distribution as special cases.

Notice that no sign restrictions are needed on the parameters to ensure the positivity of the conditional duration. The range of possible distributional assumptions regarding the innovation ϵ_i is the same as for standard ACD models, i.e. the most general models in this class would be a Burr and a Generalized Gamma Log-ACD models.

Estimation and density forecast

Estimation and density forecasts are analogously computed as for the linear ACD case, taking into account that Ψ_i has been redefined, i.e. the forecast density is given by

$$f(x_i(\exp(\psi_i))^{-1})(\exp(\psi_i))^{-1}$$

where $f(\cdot)$ denotes, as above, the innovation density.

3.3 Threshold ACD Models

Model specification

In the standard ACD framework, the conditional mean dynamics are determined by the linear specifications in equations (6), (8) or (9). Zhang, Russell, and Tsay (2001), henceforth referred to as ZRT, argue that financial duration processes require a more flexible specification. ZRT's Threshold ACD model (TACD) is a three regime model where the regimes are allowed to have different duration persistence and innovation distributions. In a logarithmic version of the TACD, the logarithm of the conditional duration ψ_i evolves as

$$\psi_i = \begin{cases} \omega_1 + \alpha_1 \log x_{i-1} + \beta_1 \psi_{i-1} & \text{if} \quad 0 < x_{i-1} \leq r_1 \\ \omega_2 + \alpha_2 \log x_{i-1} + \beta_2 \psi_{i-1} & \text{if} \quad r_1 < x_{i-1} \leq r_2 \\ \omega_3 + \alpha_3 \log x_{i-1} + \beta_3 \psi_{i-1} & \text{if} \quad r_2 < x_{i-1} < \infty. \end{cases} \tag{10}$$

The basic specification $x_i = \exp(\psi_i)\,\epsilon_i$ as in equation (7) still applies. The threshold parameters r_1 and r_2 determine the regime boundaries. The idea behind the specification of equation (10) is that there exist regimes where the trading or quoting process is 'slow', 'normal' and 'fast', and that a linear specification would lead to an overshooting of the expected duration after a very short duration or a long duration. ZRT assume the generalized gamma distribution for ϵ_i, but they allow the shape parameter ν to vary across regimes, thereby rendering the duration distribution to be regime-specific.

Estimation

Parameter estimation is performed by a grid search over r_1 and r_2 and by maximizing the likelihood function conditional on given threshold values. In our application the regime boundaries r_1 and r_2 are chosen to be the duration deciles. The grid search uses all possible permutations of the deciles under the restriction $r_1 < r_2$. We focus on a TACD version that employs the generalized gamma distribution and the logarithmic specification (10). Alternative TACD specifications would make use of the alternative distributions outlined in Sub-section 3.1 or employ a regime-adapted version of (6) or (9).

Density forecast

Density forecasts of the TACD are obtained in the same fashion as for ACD and Log-ACD models. However, one has to take account of the fact that equation (10) implies that the TACD density forecast depends on the regime (i.e. x_{i-1}) from which it is issued. Hence, one has to pick the correct regime parameter values to compute the predictive density.

3.4 Stochastic Conditional Duration Models

Model specification

Bauwens and Veredas (1999) (henceforth BV) assume that financial duration processes are driven by a single dynamic stochastic factor. This latent variable is interpreted as being inversely related to the information arrival process which triggers bursts of activity. BV assume that the factor follows the first order autoregressive process

$$\tilde{\psi}_i = \omega + \beta \tilde{\psi}_{i-1} + u_i, \tag{11}$$

where u_i is independently normally distributed with zero mean and variance σ^2. Following the logic that lead to the Log-ACD specification and writing

$$x_i = \exp(\tilde{\psi}_i)\, \epsilon_i \tag{12}$$

yields the SCD model. The process $\{\epsilon_i\}$ is assumed to be independent of the process $\{u_i\}$. The set of sensible distributions for the duration innovations ϵ_i is the same as for ACD models.

Estimation

Estimation of the SCD parameters is more demanding compared to standard ACD models. The factor structure complicates the exact evaluation of the likelihood since one has to integrate out the random variables u_i. Taking the natural log of x_i as defined in equation (12) the model is seen to be a linear space state model with error terms $\ln \epsilon_i$ and u_i. BV approximate the distribution of $\ln \epsilon_i$ by a

Gaussian density. Consistent parameter estimates are obtained by quasi-maximum likelihood using the Kalman filter technique and prediction error decomposition. BV reported similar results when assuming Weibull or gamma distributions for ϵ_i. In the present paper, we focus on the Weibull-SCD specification.

Density forecast

Whatever the distributional assumption for the innovation ϵ_i, the conditional distribution of x_i is a mixture of this distribution and the conditional lognormal distribution of $\exp(\tilde{\psi}_i)$. This mixture is the one-step-ahead forecast density that is implied by the SCD model. Note that although the parameter estimation procedure uses a density approximation, the probability integral transform z for the SCD model can be computed exactly. This procedure requires a bidimensional numerical integration, one for computing the forecast density (integrating out u_i), the other for computing the z that is implied by this density.

3.5 Stochastic Volatility Duration Models

Model specification

The ACD and Log-ACD models reviewed in the previous sub-sections share the implicit assumption that the dynamics of higher moments of the duration process are governed by the dynamics of the conditional mean. Ghysels, Gouriéroux, and Jasiak (1998), henceforth referred to as GGJ, have argued that this feature restricts the flexibility that is needed when modeling financial duration processes. As a more versatile tool, GGJ introduce a nonlinear two factor model that disentangles the movements of the mean and of the variance of durations. Since the second factor is responsible for the variance heterogeneity, the model is referred to as the stochastic volatility duration (SVD) model. The departure point for the SVD is a standard static duration model in which it is assumed that the durations are independently and exponentially distributed with a gamma heterogeneity,

$$x_i = \frac{U_i}{aV_i}, \tag{13}$$

where U_i and V_i are two independent variables with distributions gamma(1,1) (i.e. exponential) and gamma(b, b), respectively. Equation (13) can be rewritten to include Gaussian factors:

$$x_i = \frac{H(1, F_{1i})}{aH(b, F_{2i})} \tag{14}$$

where F_{1i} and F_{2i} are independent standard normal variables, and $H(b, F) = G(b, \varphi(F))$, where $G(b, .)$

is the quantile function of the gamma(b, b) distribution and $\varphi()$ the cdf of the standard normal. GGJ extend this static setup and propose to model duration dynamics through the two underlying Gaussian factors. We focus on a bivariate VAR representation for the process $F_i = (F_{1i}, F_{2i})'$. The marginal distribution of F_i is constrained to be $N_2(0, I)$. This ensures that the marginal distribution of x_i belongs to the class of exponential distributions with gamma heterogeneity. Restricting our attention to a VAR of order one, we have

$$F_i = \Lambda F_{i-1} + u_i, \tag{15}$$

where u_i is a Gaussian white noise random vector with covariance matrix $\Sigma(\Lambda)$ such that $var(F_i) = I$. For a VAR of order one this is achieved by setting $\Sigma = I - \Lambda\Lambda'$.

Estimation

Due to the nonlinear dynamic latent factor structure, the likelihood function of the SVD model is difficult to evaluate. On the other hand, the SVD model is easy to simulate. This suggests parameter estimation via simulation based techniques (see Gouriéroux and Monfort, 1996). The SVD model belongs to the class of nonlinear dynamic factor models for which Gouriéroux and Jasiak (2001) propose convenient, simulation based estimators. SVD estimation is performed in two steps. In the first step, we exploit that the marginal distribution of x_i is a Pareto distribution of the second type that depends only on the parameters a and b, with density given by

$$f_{P_2}(x_i) = \frac{ab^{b+1}}{(ax_i + b)^{b+1}}. \tag{16}$$

Hence, it is straightforward to estimate a and b by quasi-maximum likelihood. In the second step, the parameter matrix Λ is estimated by the method of simulated moments, adopting the procedure outlined in Gouriéroux and Jasiak (1999). For the empirical application, autocorrelations of x_i and x_i^2 up to lag order 7 are selected for estimating Λ which is assumed to be a diagonal matrix.

Density forecast

Because of analytical intractability, the SVD density forecast has to be computed by simulation. The idea to generate the density forecast and the corresponding probability integral transform series rests on the possibility to infer the corresponding states of the factors from the observable past durations. Given the model parameters (or their consistent estimates) it is possible to simulate the conditional distributions of the future states of the factors. One can then again exploit that the SVD implies a one-to-one correspondence of factors and durations, i.e. from the simulated distribution of the factors we can deduce the conditional duration distribution which is the predictive distribtion we are interested in. The simulated distribution can then be used to compute the probability integral transform z. Given a sufficiently large number of replications in each simulation, the predictive densities and the probability integral transform series can be precisely estimated. However, if a a long sequence of one-step-ahead forecasts is required this procedure can be time consuming, since it has to be carried out for each one-step-ahead forecast step.

insert table 1 about here

4 Application

4.1 Data

For the model comparison we used transaction data from New York Stock Exchange (NYSE). The NYSE trading process is organized as a hybrid mechanism that combines a market maker system and an order book system. Assigned to each stock is a specialist, who makes the market for this stock, i.e. he manages the trading and quoting processes and provides liquidity when necessary by taking the other side of the trade. Apart from an opening auction, trading is continuous from 9h30 to 16h.[4] Data were extracted from the Trade and Quote (TAQ) database supplied by the NYSE. The TAQ data consists of two parts: the first reports all trades, while the second lists the best bid and best ask (offer) prices posted by the specialists. We use data for the months of September, October, and November 1996 and focus on four actively traded Dow Jones stocks, Boeing, Coca-Cola, Disney, and Exxon. Exxon was, at that time, the largest stock in the S&P 500 index. We have also investigated IBM data, the stock that was scrutinized in the seminal application of ACD models by Engle and Russell (1998). However, for the period under investigation we found evidence for long memory in IBM trade, price and volume durations. Since models that allow for long memory in financial duration processes are not considered in the present paper we refrain from reporting the IBM results.

We investigate three types of financial duration processes. *Trade durations* are defined as the time spells between successive trades. Engle and Russell (1998) show that when a conditional duration model is fitted to a series of trade durations, its conditional hazard provides a measure of the instantaneous trading intensity. *Price durations* are defined by thinning the quote process with respect to a minimum change in the mid-price of the quotes. More precisely, a price duration is defined as the time interval

[4] Harris (2002) gives lucid insights into the workings of real world trading systems.

needed to observe a cumulative change in the mid-price not less than a threshold that we set at \$0.125 (see Giot (2000), for a more thorough discussion of price durations). As pointed out by Engle and Russell (1998), price durations provide a measure of the instantaneous volatility of the quoted mid-price process. Finally, thinning the quote process such that the duration measures the time until a cumulated volume of at least v shares has been traded one obtains so-called *volume durations.* For the purpose of this paper we set $v = 25,000$. Volume duration data have an immediate appeal for characterizing the liquidity of a stock as they indicate the time needed to trade a given amount of shares.

For all three types of duration processes, overnight durations (i.e. the durations between the first event of a day and the last recorded event of the previous day) and durations corresponding to events recorded outside the regular NYSE opening were removed. The resulting duration time series have also been scrutinized in the papers by Bauwens and Giot (2001), Bauwens and Veredas (1999), Fernandes and Grammig (2000), Giot (2000), and Grammig and Maurer (2000).

insert table 2 about here

As documented in previous empirical work, the three types of duration processes feature a strong time-of-day effect (intra-day seasonality, called diurnality), which stems from predetermined market characteristics such as opening/closing of trading or lunch time for traders. To take into account this quasi-deterministic diurnality, the time-of-day standardized durations are computed as

$$X_i = x_i \, \phi(t_i) \tag{17}$$

where X_i is the raw (trade, price or volume) duration, $\phi(t_i)$ is the time-of-day effect at time t_i, and x_i denotes the time-of-day standardized (trade, price or volume) duration. The deterministic time-of-day effect for each set of durations is defined as the expected (trade, price or volume) duration conditioned on time-of-day and on the day of the week. This implies that, for example, that the time-of-day effect of Monday can be different from the time-of-day effect of Tuesday. The expectation is computed by averaging the durations data over thirty minutes intervals for each day of the week. Cubic splines are then used on the thirty minutes intervals to smooth the time-of-day function. For brevity of notation, we will henceforth refer to time-of-day standardized durations simply as durations.

insert figure 1 about here

The resulting time series exhibit those characteristic features that challenged and fueled econometric modeling of financial duration processes. As shown in Table 2 the three types of financial duration processes imply a Q-statistic that is significant at any reasonable level. Figure 3 displays the corresponding sample autocorrelograms for Boeing which are representative for the other stocks, too. The Boeing autocorrelograms are representative in that for all stocks the autocorrelation coefficients are initially largest for volume durations and smallest for trade durations. All three duration series feature a slow decay of the autocorrelation function. Table 2 shows that both trade and price durations are characterized by overdispersion (the ratio of standard deviation to mean of the duration series is greater than one), whilst volume durations are underdispersed.

insert figure 2 about here

In Figure 2, we plot the kernel densities for the three types of durations. It can be seen that volume durations are characterized by clearly hump-shaped densities, with most of the mass around 0.5-1. This is in sharp contrast with trade and price durations which exhibit essentially decreasing density functions, but with a pronounced hump at very small durations. These salient characteristics of the data are similar to what has been documented in the literature, e.g. in Engle and Russell (1998), Giot (2000), and Bauwens and Giot (2001). The hump close to the origin is not an artefact of the kernel density estimation of the density of a positive variable. To take this into account, we employed the gamma kernel proposed by Chen (2000) which is designed for kernel density estimation when the data, like the duration sequences in the context of this paper, have positive support.

4.2 Results

4.2.1 Description

For each stock and for each type of duration process, we estimated the financial duration models summarized in Table 1 and computed the z-series (the probability integral transforms) implied by each model. We specified the autoregressive models with one lag as this is the standard specification most frequently used in the original papers. The evaluation was done out-of-sample, i.e. parameter estimation was conducted on the first two-thirds of the sample, and then the one-step-ahead forecast densities and the z series were computed based on the last third of the sample. To establish a benchmark, we also evaluate forecasts that are based on the assumption that each duration process is iid standard exponential (recall that the durations have been standardized to have mean one). That is we simply issue the density forecast e^{-x}. Although very simple, this distributional assumption is implicit in many theoretical microstructure models like the one by Easley, Kiefer, O'Hara and Paperman (1996). We refer to this specification as the Poisson model as it implicitly assumes a Poisson process for the event arrival process generating the durations.

We present the results of the comparison in tables 3 and 4 and in figures 3-6. Table 3 reports the out-of-sample $p-$values for the test of iid $U(0,1)$ of the z sequences and the $p-$values of the Ljung-Box $Q-$statistic for z and z^2 using the first 15 autocorrelations. The table contains a summary of test results for all stocks and types of financial duration. Disney out-of-sample results are not reported as the models deliver $p-$values which are so small that any model is rejected at conventional levels of significance. We therefore also present the Disney in-sample results in table 4. In-sample evaluation means that the parameter estimation and the one-step ahead density forecast evaluation are performed on the same data, namely the last one-third of the data. This in-sample exercise reveals that some models (the same that perform well out-of-sample for the other stocks) do offer satisfactory results. This indicates a structural break in the financial duration processes of the Disney stock. Within the respective columns of tables 3 and 4 models are informally ranked using the following procedures: For price and volume duration processes we first sorted those models not rejected at 5% by neither of the two $Q-$statistics (for z and z^2) by the $p-$value of Pearson's goodness-of-fit statistic. Those models rejected by either of the $Q-$statistics were ranked subsequently, applying the same sorting by $p-$value of the goodness of fit statistic. As the $p-$values of the Pearson statistic are very small for any model estimated on trade durations a different ranking was applied. Those models passing both serial correlation tests at 1% level were sorted by the average of the $p-$values of the $Q-$statistics. The rejected models were ranked subsequently sorting them in the same way.

insert table 3 about here

insert table 4 about here

To conserve space when presenting the graphical evaluation results we display autocorrelograms and histograms of the probability integral series for only one selected stock per duration process. Histograms and correlograms for the selected stock are representative for the other stocks, too. For price durations we select Exxon, for volume durations Coca Cola, and for trade durations Boeing. Because of space restrictions we also do not show the graphical results for each model, but again focus on representative specifications and the most interesting results. In some respects ACD and Log-ACD specifications based on different distributional assumptions produce similar results (e.g. serial dependence of the z sequence). Hence, we display the graphical results for one ACD specification only. This enables us to highlight and discuss the most important, general patterns.

4.2.2 Interpretation of results

Serial dependence of the probability integral transform series

The $p-$values of the Q statistics for the z and z^2 series produced by the financial duration models show that several specifications seem quite successful in capturing the dynamic structure of financial duration processes. The benchmark Poisson model, however, performs very badly. This is an important result for microstructure theory as the Poisson process is the work horse specification upon which many theoretical models are based. The SVD model performance is disappointing given the promise to offer more flexibility in modeling higher order dynamics. The model does not offer an improvement in situations when the data generating process is challenging for the other models, too (e.g. Coca Cola trade durations, Disney) and even does badly when the more simpler ACD approaches perform quite well (e.g. Exxon and Boeing price durations). Considering a richer dynamic structure than a VAR(1) for the Gaussian factors is clearly indicated. Another restriction of the SVD is the Pareto assumption for the unconditional duration distribution. This assumption can cause severe problems in the first step of SVD estimation, in the present context when trying to fit the model on volume duration data. Given the humped shape of the unconditional volume duration distribution (see Figure 2), the Pareto assumption is definitely inappropriate. As a matter of fact, it turned out that the first SVD estimation step is not feasible, and since the distribution parameters are required for the second step the SVD is not estimable on volume duration data. Hence, if the SVD is to be applied for this type of data, a modification of the marginal duration distribution is clearly indicated.

insert figure 3 about here

The second dynamic factor model included in the comparison, the SCD, offers a performance comparable to ACDs and Log-ACDs, but no improvement regarding critical cases. The Threshold ACD is also not able to outperform the simpler ACD specifications and does not excel in a situation where its flexibility would be required: We refer to the Disney price, volume and trade durations, for which all models are rejected out-of-sample, whilst the in-sample results regarding price and volume durations produced by the standard ACD models are quite good. This is a clear indication for a structural break, yet the TACD is not able to account for it. On the contrary, for some processes the TACD performs worse than simple ACD and Log-ACD specifications (see for example Boeing trade and price durations). The results in table 3 generally indicate that ACD models (log or linear) based on the generalized gamma distribution offer the best results in terms of Q-statistics $p-$values for z and z^2.

The autocorrelograms in figure 3 emphasize these results: The ACFs implied by Generalized Gamma Log-ACD and Burr ACD fluctuate nicely within the bands of the confidence intervals. We want to stress that w.r.t to the independence aspect of the probability integral transform sequence, ACD and Log-ACD models based on the exponential distribution perform quite as well as specifications based on more flexible distributions. It is quite remarkable that straightforward ACD and Log-ACD seem to capture the dynamic dependence in financial duration processes in most cases in a satisfactory way, keeping in mind that all models assume an ARMA(1,1)-type lag structure.

The autocorrelograms show that the SCD is also quite successful with only a simple AR(1) structure on the dynamic factor. They also document the failure of the Poisson model to account for dynamics in financial duration processes and underline the poor performance of the SVD, especially regarding trade durations.

Iid uniformity of the probability integral transform series

For many applications, like option pricing (Pringent, Renault and Scaillet 2001) and empirical market microstructure (Bauwens and Giot, 2000), **price durations** represent the most interesting financial duration process. The good news is that we can identify some quite successful specifications among the financial duration models considered in this paper. Interestingly, those models are not the most complex ones. Tables 3 (out-of-sample) and 4 (Disney in-sample) show that ACD and Log-ACD specifications based on the generalized gamma and the Burr distribution produce $p-$values which imply that the iid $U(0,1)$ null hypothesis for the z sequences cannot be rejected at conventional significance levels. On the other hand, ER's standard Exponential and Weibull ACD perform much worse. It seems clear that the reason for their rejection originates in the distributional assumption for the innovation distribution, since the test for serial independence of the z sequences yielded comparable (and satisfactory) results for all ACD specifications irrespective of the assumed innovation distribution. The performance of the more complex models, the Threshold ACD and the two factor models, SVD and SCD, is worse than those of the ACD and Log-ACD models (if the latter two are based on a flexible innovation distribution).

insert figure 4 about here

The histograms displayed in figure 4 (showing Exxon results which are representative for the other stocks, too) indicate the reasons for specification problems. The histograms for the Poisson model, the Exponential and the Weibull ACD as well as the SVD have a distinct non-uniform shape indicating that far too few realisations fall into the very low tails of the forecast densities. If the data were really produced by the assumed data generating process one would expect much more observations. On the other hand, small (but not very small) durations are over-represented: The frequencies associated with the third to sixth histogram bins are above the confidence interval. The SCD histogram displays peaks that lead one to doubt iid uniformity of the z sequence and indicate that the mixture distribution implied by the SCD does not offer the required flexibility.

It is worth noting that among the models considered in this paper the Weibull ACD model is the most affected. The reason for this result can be found as follows: Because the estimates of the γ parameter of the Weibull distribution are smaller than 1 for all stocks the density tends to infinity as x tends to zero. As a consequence, there are not enough very small durations compared to what is predicted by the Weibull distribution, and correspondingly there are too few z-values in the first quintile bin. We conclude that a successful model has to be built on a distribution that can put a lot of probability mass on small durations, but not too much on very small durations. This seems already suggested by the kernel densities in figure 2 (though of course these display the unconditional distributions). And indeed, the histograms for the Generalized Gamma Log-ACD, the Burr ACD and the TACD are much improved, and in accordance with what would be expected under uniformity. Another good news result, especially for empirical market microstructure analysis, is that the Log-ACD specification does not perform worse than the linear ACD. As mentioned above, the Log-ACD specification offers greater flexibility when including predetermined variables suggested by microstructure theory.

insert figure 5 about here

Regarding **trade durations** we find that all models are rejected by the test of iid uniformity behavior. Yet, looking at the $z-$histograms one will notice obvious qualitative differences among models. Some models do not perform badly, but due to the large sample size confidence bands are very narrow. Figure 5 depicts the representative histograms for Boeing trade durations. The histograms show that the same problems identified when discussing the price duration results prevail and are even aggravated. The histograms of the models based on the Exponential and Weibull distribution have a distinct non-uniform shape with, as for price durations, much too few realisations in the low tail of the forecast density compared to what would be expected if the conditional densities were correctly specified. On the other hand $z-$values in the third to sixth histogram bins are over-represented, indicating that the innovation distribution puts to little mass on small (but not very small) durations. The histograms of generalized gamma based models (recall that our TACD specification is also based on the generalized gamma distribution) are much improved and come closer to uniformity, whilst the SVD and SCD histograms, and also those for the Burr based ACD models, indicate clear departures from uniformity.

insert figure 6 about here

Tables 3, 4 and figure 6 show that the models which are successful for modeling price durations (i.e. ACD and Log-ACD models based on the Burr or generalized gamma distribution) are also those that are useful for modeling **volume duration** processes. However, the set of successful models is supplemented with the Weibull ACD and the Weibull SCD. Clearly, the problems identified when using the Weibull distribution when modeling price and trade duration processes are alleviated. The reason is that the estimated Weibull γ parameters are now greater than one which implies a hump-shaped innovation density, and also a hump-shaped conditional density forecast. The problem that the density function approaches infinity for small durations when $\gamma < 1$ vanishes. It is therefore not surprising that the Exponential ACD performs quite badly on volume durations as the exponential density (like the Pareto) cannot take on a shape that is not monotonically decreasing. In fact, this property is also responsible for the failure of the SVD when applied to modeling volume durations. Figure 6 shows that for the Poisson model and the Exponential ACD (both issue an exponential conditional density forecast) the histogram is clearly non-uniform. Both models predict much more probability mass at small durations which is not there. On the other hand, the histogram for the Weibull ACD and the Weibull SCD are much more compatible with uniformity. Again, the Burr and Generalized Gamma ACD and Log-ACD deliver quite satisfactory results whilst the TACD performance is rather disappointing.

5 Conclusion

We used density forecast evaluation techniques to compare the performance of recently proposed financial duration models developed for modeling price, trade and volume duration processes. These models have frequently been applied for empirical microstructure analysis, but so far no comparison between the competing approaches has been attempted. We investigated the performance of basic ACD and Log-ACD models. These approaches adopt elements of the more familar ARCH and Exponential ARCH models used for modeling conditional variances of asset returns. Like in the ARCH literature, the basic ACD and Log-ACD specifications have been generalized in terms of innovation distributions (Generalized Gamma and Burr ACD) and the modeling of nonlinear duration dynamics (Threshold ACD). Transferring results on stochastic volatility models, dynamic factor models, (SVD and the SCD) have been put forth as an alternative framework for modeling financial duration processes.

We compared ACD, Log-ACD (both based on a variety of innovation distributions), Threshold ACD, SCD and SVD models. We also considered density forecasts issued by the Poisson model, the work horse of market microstructure theory. This model implies density forecasts that are iid standard exponential. However, the results indicate that the Poisson model is clearly unsuitable for all types of financial duration processes. On the other hand, the results show that simpler approaches perform at least as well as more complex methods. TACD, SCD and SVD models are much more costly to estimate and to evaluate than ACD and Log-ACD models, but their performance, compared to the simpler approaches, is not superior. For the SVD model a richer dynamic structure than the VAR(1) factors structure as well as a reconsideration of the marginal duration distribution is clearly required. Recapitulating we tend to conclude that neither TACD, SCD or SVD should be considered as a starting point in an empirical research.

Applied to modeling trade duration processes, ACD and Log-ACD models successfully accounted for duration dynamics whilst none of the models provide an acceptable specification for the conditional duration distribution. This suggests that there is a need for financial duration models which imply conditional duration densities that are able to put more probability mass on small durations, but not too much on very small durations. Using the generalized gamma distribution clearly offers an improvement, but so far no overall successful specification could be identified.

In many respects, price durations represent the most interesting financial duration process due to the close links to microstructure theory and option pricing. Our results indicate that for the modeling of price duration processes, employing the basic ACD specifications based on the exponential and the Weibull distribution is not advisable. Especially in the context of empirical microstructure studies we rather recommend employing Generalized Gamma or Burr Log-ACD. We can conclude that these models provide a robust and useful modeling framework (also for the modeling of volume durations). One could summarize the results by the advise 'keep it simple but not too simple'.

An avenue to be considered for future research is to investigate the effects of the intra-day seasonality adjustment procedure. For the purpose of this paper we chose a two step approach (that is we first deseasonalized the duration data and estimated the models based on time-of-day standardized durations in the second step), since for some models a simultaneous estimation procedure is intractable. The results reported in Veredas, Rodriguez-Poo and Espasa (2001) suggest that it is worth investigating whether the problems related to the modeling of trade duration processes could be alleviated by using a more refined procedure.

Acknowledgements

We thank Stefan Frey, Roland Straub, and Sandra Vuletic for excellent research assistance, Craig Lewis and Nick Hautsch for offering insights into this work and the participants of sessions at the Econometric Society World Congress in Seattle, August 2000, and the Annual Meeting of the German Economic Association in Berlin, September 2000, for useful comments. We are grateful to the editor, associate editor and two referees whose suggestions greatly improved the exposition of this paper. This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

References

Bauwens, L. & Giot, P. (2000). The logarithmic ACD model : An application to the bid-ask quote process of three NYSE stocks, *Annales d'Economie et de Statistique*, 60, 117–149.

Bauwens, L. & Giot, P. (2001). *Econometric modelling of stock market intraday activity*, Dordrecht: Kluwer Academic Publishers.

Bauwens, L. & Veredas, D. (1999). The stochastic conditional duration model: A latent factor model for the analysis of financial durations, CORE Discussion Paper 9958, University of Louvain. Forthcoming in Journal of Econometrics.

Berkowitz, J. (2001). Testing density forecasts, with applications to risk management, *Journal of Business and Economic Statistics*, 19, 465–474.

Brock, W. A., Dechert, W. D., Scheinkman, J. A., & LeBaron, B. (1996). A test for independence based on the correlation dimension, *Econometric Reviews*, 15, 197–235.

Chen, S. X. (2000). Gamma kernel estimators for density functions, *Annals of the Institute of Statistical Mathematics,*, 52, 471–480.

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

Easley, D., Kiefer, N., O'Hara, M., & Paperman, J. P. (1996). Liquidity, information and infrequently traded stocks, *Journal of Finance*, 51, 1405–1436.

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

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987–1007.

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, 187–212.

Fernandes, M. & Grammig, J. (2000). Non-parametric specification tests for conditional duration models, european University Institute and University of Frankfurt.

Ghysels, E., Gouriéroux, C., & Jasiak, J. (1998). Stochastic volatility duration models, in: Olsen and Associates (Ed.), *High Frequency Data in Finance II: Proceedings*, Zurich. Forthcoming in Journal of Econometrics.

Giot, P. (2000). Time transformations, intraday data and volatility models, *Journal of Computational Finance*, 4, 31–62.

Goodhart, C. A. E. & O'Hara, M. (1997). High frequency data in financial markets: Issues and applications, *Journal of Empirical Finance*, 4, 73–114.

Gouriéroux, C. & Jasiak, J. (2001). Dynamic factor models, *Econometric Reviews*, 20, 385–424.

Gouriéroux, C., Jasiak, J., & Le Fol, G. (1999). Intraday market activity, *Journal of Financial Markets*, 2, 193–226.

Gouriéroux, C. & Monfort, A. (1996). *Simulation-based Econometric Methods*, Oxford: Oxford University Press.

Grammig, J. & Maurer, K.-O. (2000). Non-monotonic hazard functions and the autoregressive conditional duration model, *The Econometrics Journal*, 3, 16–38.

Harris, L. (2003). *Trading and Exchanges*, Oxford: Oxford University Press.

Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models, *Review of Economic Studies*, 65, 361–393.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data*, Cambridge: Cambridge University Press.

Lunde, A. (1999). A generalized gamma autoregressive conditional duration model, Aalborg University.

Madhavan, A. (2000). Market microstructure: A survey, *Journal of Financial Markets*, 3, 205–258.

O'Hara, M. (1995). *Market Microstructure Theory*, Oxford: Basil Blackwell.

Prigent, J.-L., Renault, E., & Scaillet, O. (2001). An autoregressive conditional binomial option pricing model, in: Geman, Madan, Pliska, & Vorst (Eds.), *Selected Papers from the First World Congress of the Bachelier Finance Society*, Heidelberg: Springer.

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

Shephard, N. (1994). Partial non-gaussian state space, *Biometrika*, 81, 115–131.

Veredas, D., Rodriguez-Poo, J., & Espasa, A. (2002). On the (intradaily) seasonality and dynamics of a financial point process: A semiparametric approach, CORE Discussion Paper 23, Université Catholique de Louvain.

Zhang, M. Y., Russell, J., & Tsay, R. S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data, *Journal of Econometrics*, 104, 179–207.

Table 1: List of Models

Short Name	Equations	Long Name
BACD	(5)-(6)	Burr ACD
GGACD	(5)-(6)	Generalized Gamma ACD
WACD	(5)-(6)	Weibull ACD
EACD	(5)-(6)	Exponential ACD
$BLACD_1$	(7)-(8)	Burr Log-ACD (type 1)
$GGLACD_1$	(7)-(8)	Generalized Gamma Log-ACD (type 1)
$WLACD_1$	(7)-(8)	Weibull Log-ACD (type 1)
$BLACD_2$	(7)-(9)	Burr Log-ACD (type 2)
$GGLACD_2$	(7)-(9)	Generalized Gamma Log-ACD (type 2)
$WLACD_2$	(7)-(9)	Weibull Log-ACD (type 2)
TACD	(7)-(10)	Generalized Gamma Threshold Log-ACD (type 1)
SVD	(14)-(15)	Stochastic Volatility Duration
WSCD	(7)-(11)	Weibull SCD

Table 2: Descriptive Statistics for the Data

	Trade durations	Price durations	Volume durations
		BOEING	
Number	23930	2620	1576
Overdispersion	1.218	1.338	0.804
$Q(10)$	1357.15	322.30	477.69
		COCA-COLA	
Number	39622	1609	3022
Overdispersion	1.181	1.171	0.876
$Q(10)$	2012.14	69.71	304.34
		DISNEY	
Number	32821	2160	1778
Overdispersion	1.226	1.209	0.720
$Q(10)$	1629.18	137.31	1136.12
		EXXON	
Number	28371	2717	2045
Overdispersion	1.194	1.196	0.658
$Q(10)$	797.21	68.18	237.85

A trade duration is given by the time interval between two consecutive trade events. A price duration is measured by the time interval between two bid-ask quotes during which a cumulative change in the mid-price of at least $0.125 is observed. A volume duration denotes the time interval between two bid-ask quotes during which the cumulative traded volume amounts to at least 25,000 shares. Overdispersion is defined as the ratio of standard deviation to mean (all means are equal to 1 to the third decimal as a result of the diurnal adjustment procedure). $Q(10)$ denotes the Ljung-Box Q-statistic of order 10 on the durations.

Table 3: Out-of sample test results.

Price durations											
BA				KO				XON			
	$U(0,1)$	$Q(15)$	$Q^2(15)$		$U(0,1)$	$Q(15)$	$Q^2(15)$		$U(0,1)$	$Q(15)$	$Q^2(15)$
GGLACD$_2$	6.64	18.58	7.54	GGLACD$_2$	71.36	80.77	90.12	GGLACD$_1$	49.86	80.40	85.18
GGLACD$_1$	2.18	15.97	10.93	GGLACD$_1$	51.44	76.09	82.96	BLACD$_2$	15.29	51.82	63.13
GGACD	0.04	24.02	10.75	GGACD	46.07	82.76	88.19	BLACD$_1$	14.73	81.07	84.98
BACD	0.01	23.05	8.54	WLACD$_2$	15.90	85.35	90.16	GGLACD$_2$	14.60	39.42	49.92
BLACD$_1$	0.01	13.84	10.44	BLACD$_2$	15.18	82.05	90.29	BACD	5.52	68.79	76.92
WLACD$_1$	0.00	18.02	13.03	WACD	11.40	83.74	89.42	EACD	0.02	84.47	88.20
BLACD$_2$	0.00	19.00	7.61	BLACD$_1$	9.18	76.79	83.43	WSCD	0.00	75.96	68.37
WSCD	0.00	11.15	8.41	TACD	9.18	69.66	91.18	WACD	0.00	83.73	88.19
WACD	0.00	12.65	4.22	WSCD	6.37	82.76	79.97	WLACD$_1$	0.00	86.45	90.65
EACD	0.00	10.15	3.17	BACD	6.15	79.18	88.34	WLACD$_2$	0.00	68.52	79.76
WLACD$_2$	0.00	11.69	4.85	SVD	4.50	8.38	5.39	TACD	10.06	4.78	14.51
TACD	19.15	0.79	0.24	EACD	1.77	84.71	89.61	GGACD	87.76	2.81	1.95
SVD	2.74	0.00	0.00	WLACD$_1$	0.33	82.30	82.40	SVD	0.01	0.08	1.61
Poisson	0.00	0.00	0.00	Poisson	0.10	35.21	10.34	Poisson	0.00	0.26	0.02
Volume durations											
BA				KO				XON			
	$U(0,1)$	$Q(15)$	$Q^2(15)$		$U(0,1)$	$Q(15)$	$Q^2(15)$		$U(0,1)$	$Q(15)$	$Q^2(15)$
WLACD$_1$	24.87	45.70	6.98	WSCD	28.82	74.41	67.19	BACD	99.75	95.84	97.92
WSCD	19.47	70.21	35.34	BACD	5.97	67.56	73.52	BLACD$_2$	98.15	92.63	96.55
BLACD$_1$	19.18	57.39	15.99	GGACD	3.22	67.00	74.80	GGLACD$_2$	97.09	92.68	96.65
GGLACD$_1$	14.76	55.18	13.10	WACD	3.19	67.50	74.19	GGACD	95.91	96.01	97.97
BLACD$_2$	11.56	53.58	41.09	GGLACD$_2$	1.68	69.02	79.48	WLACD$_2$	93.61	93.94	97.01
BACD	10.46	59.16	44.03	BLACD$_1$	1.55	72.71	73.63	WACD	91.89	95.87	97.76
GGACD	7.93	57.16	41.17	WLACD$_1$	1.38	75.86	79.58	BLACD$_1$	89.21	92.61	92.59
WACD	6.51	55.25	38.26	BLACD$_2$	1.23	68.77	79.13	WLACD$_1$	88.72	92.95	92.25
GGLACD$_2$	4.08	52.93	39.91	GGLACD$_1$	1.06	73.68	75.26	GGLACD$_1$	77.95	92.55	92.68
WLACD$_2$	2.30	48.15	34.74	WLACD$_2$	0.63	68.54	79.56	WSCD	21.12	92.74	89.20
EACD	0.00	67.18	58.89	TACD	0.16	88.78	91.97	TACD	5.81	49.94	45.66
TACD	0.00	5.10	0.08	EACD	0.00	66.59	74.51	EACD	0.00	91.24	96.65
Poisson	0.00	0.00	0.00	Poisson	0.00	0.00	0.00	Poisson	0.00	0.00	0.00
Trade durations											
BA				KO				XON			
	$U(0,1)$	$Q(15)$	$Q^2(15)$		$U(0,1)$	$Q(15)$	$Q^2(15)$		$U(0,1)$	$Q(15)$	$Q^2(15)$
GGACD	0.00	58.96	61.24	GGACD	0.00	1.00	2.53	GGLACD$_1$	0.00	37.08	2.26
GGLACD$_1$	0.00	37.41	33.72	BLACD$_2$	0.00	0.15	0.46	WLACD$_1$	0.00	35.72	1.43
BLACD$_2$	0.00	33.91	37.13	BLACD$_1$	0.00	0.09	0.42	BACD	0.00	27.13	3.12
WLACD$_2$	0.00	33.40	37.00	WSCD	0.00	0.44	0.02	BLACD$_1$	0.00	21.01	1.37
GGLACD$_2$	0.00	33.74	36.39	GGLACD$_2$	0.00	0.12	0.32	WSCD	0.00	21.31	0.16
BLACD$_1$	0.00	32.71	34.23	GGLACD$_1$	0.00	0.11	0.18	BLACD$_2$	0.00	12.09	1.46
WLACD$_1$	0.00	35.54	31.38	BACD	0.00	0.08	0.17	WACD	0.00	10.13	1.20
WSCD	0.00	49.79	16.59	WLACD$_2$	0.00	0.01	0.06	GGLACD$_2$	0.00	6.33	0.60
BACD	0.00	30.99	33.58	WLACD$_1$	0.00	0.02	0.03	EACD	0.00	4.57	0.48
WACD	0.00	30.13	33.46	WACD	0.00	0.00	0.01	WLACD$_2$	0.00	4.30	0.62
EACD	0.00	28.13	33.22	TACD	0.00	0.00	0.00	GGACD	0.00	0.66	0.03
TACD	0.00	0.05	0.00	EACD	0.00	0.00	0.00	TACD	0.00	0.03	0.02
SVD	0.00	0.00	0.00	SVD	0.00	0.00	0.00	SVD	0.00	0.00	0.00
Poisson	0.00	0.00	0.00	Poisson	0.00	0.00	0.00	Poisson	0.00	0.00	0.00

$Q(15)$ and $Q^2(15)$ represent the $p-$values (in %) of the Ljung-Box Q-statistic based on the first 15 autocorrelations of the (squared) probability integral transforms. The $U(0,1)$ columns report the $p-$value (in %) for a test of iid $U(0,1)$ of the z sequence. The table presents an informal ranking of models. W.r.t price and volume durations we first sort by the $p-$value of the iid U(0,1) test those models that are not rejected at 5% by neither of the two $Q-$statistics for z and z^2. Those models that are rejected at 5% are ranked subsequently, with the same sorting applied. W.r.t trade durations those models passing the serial correlation tests at 1% level are sorted by the average of the $p-$values of the $Q-$statistics. The models rejected at 1% are ranked subsequently, sorted in the same fashion.

Table 4: In-sample test results for Disney.

	Price				Volume		
	$U(0,1)$	$Q(15)$	$Q^2(15)$		$U(0,1)$	$Q(15)$	$Q^2(15)$
$GGLACD_1$	78.97	77.85	71.54	$WLACD_1$	73.81	9.25	6.80
$GGLACD_2$	75.68	65.26	71.49	$WLACD_2$	43.99	21.95	31.91
BACD	58.59	72.67	75.06	WACD	41.05	12.95	19.87
$BLACD_2$	19.43	69.61	76.37	$GGLACD_1$	38.21	9.30	7.34
$BLACD_1$	16.45	78.09	73.23	BACD	36.63	12.89	20.82
GGACD	0.46	76.56	82.26	$BLACD_1$	29.99	9.75	7.88
EACD	0.02	68.77	74.07	TACD	18.93	8.77	5.12
WACD	0.01	68.93	74.54	GGACD	15.65	14.11	23.30
$WLACD_1$	0.00	74.38	72.30	$GGLACD_2$	10.44	20.86	32.06
$WLACD_2$	0.00	69.39	79.22	$BLACD_2$	4.41	21.54	32.20
WSCD	0.00	87.40	88.65	EACD	0.00	6.39	14.85
TACD	22.56	0.61	0.01	WSCD	88.88	7.10	2.79
SVD	18.57	0.00	0.00	Poisson	0.00	0.00	0.00
Poisson	0.00	0.00	0.00				

$Q(15)$ and $Q^2(15)$ represent the $p-$values (in %) of the Ljung-Box Q-statistics based on the first 15 autocorrelations of the (squared) probability integral transforms. The $U(0,1)$ columns report the p-value (in %) for a test of iid $U(0,1)$ of the z sequence. The table presents an informal ranking of models. We first sort by the $p-$value of the iid U(0,1) test those models that are not rejected at 5% by neither of the two $Q-$statistics for z and z^2. Those models that are rejected at 5% are ranked subsequently, with the same sorting applied. For trade durations all $p-$values are close to zero.

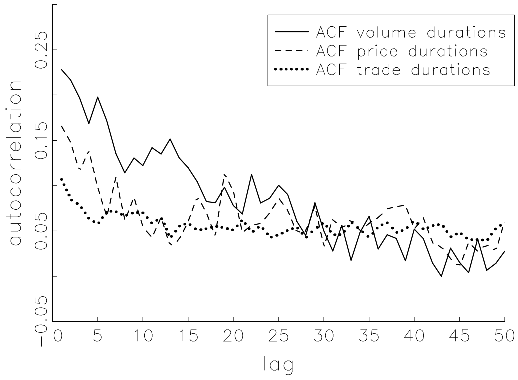

Figure 1: Sample autocorrelation functions for Boeing trade, price and volume durations.

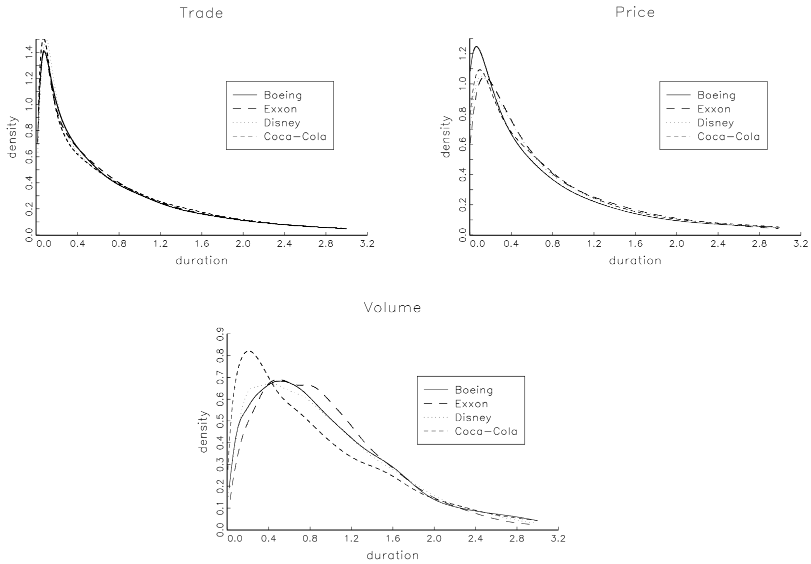

Figure 2: Kernel densities for trade, price and volume duration data using the gamma kernel proposed by Chen (2000). As suggested by Chen (2000) the bandwidth was set at $(0.9\,s\,n^{-0.2})^2$, where s is the standard deviation of the data and n the number of observations.

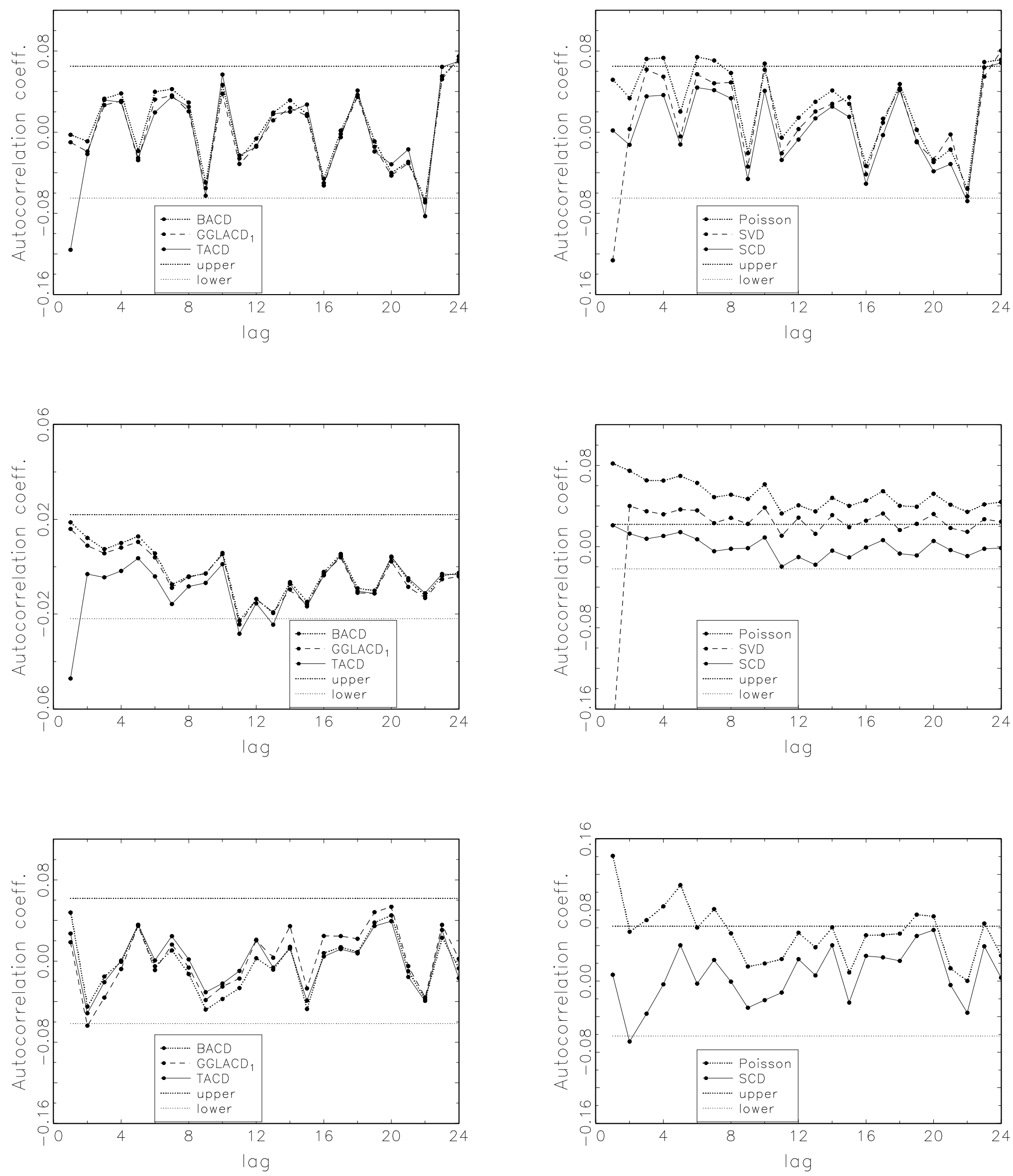

Figure 3: Estimates of the autocorrelation functions of the probability integral transform z (out-of-sample). Top panels: Exxon price durations, middle panels: Boeing trade durations, bottom panels: Coca Cola volume durations. The horizontal lines superimposed on the autocorrelograms are Bartlett's approximate 95% confidence intervals under the null that z is iid.

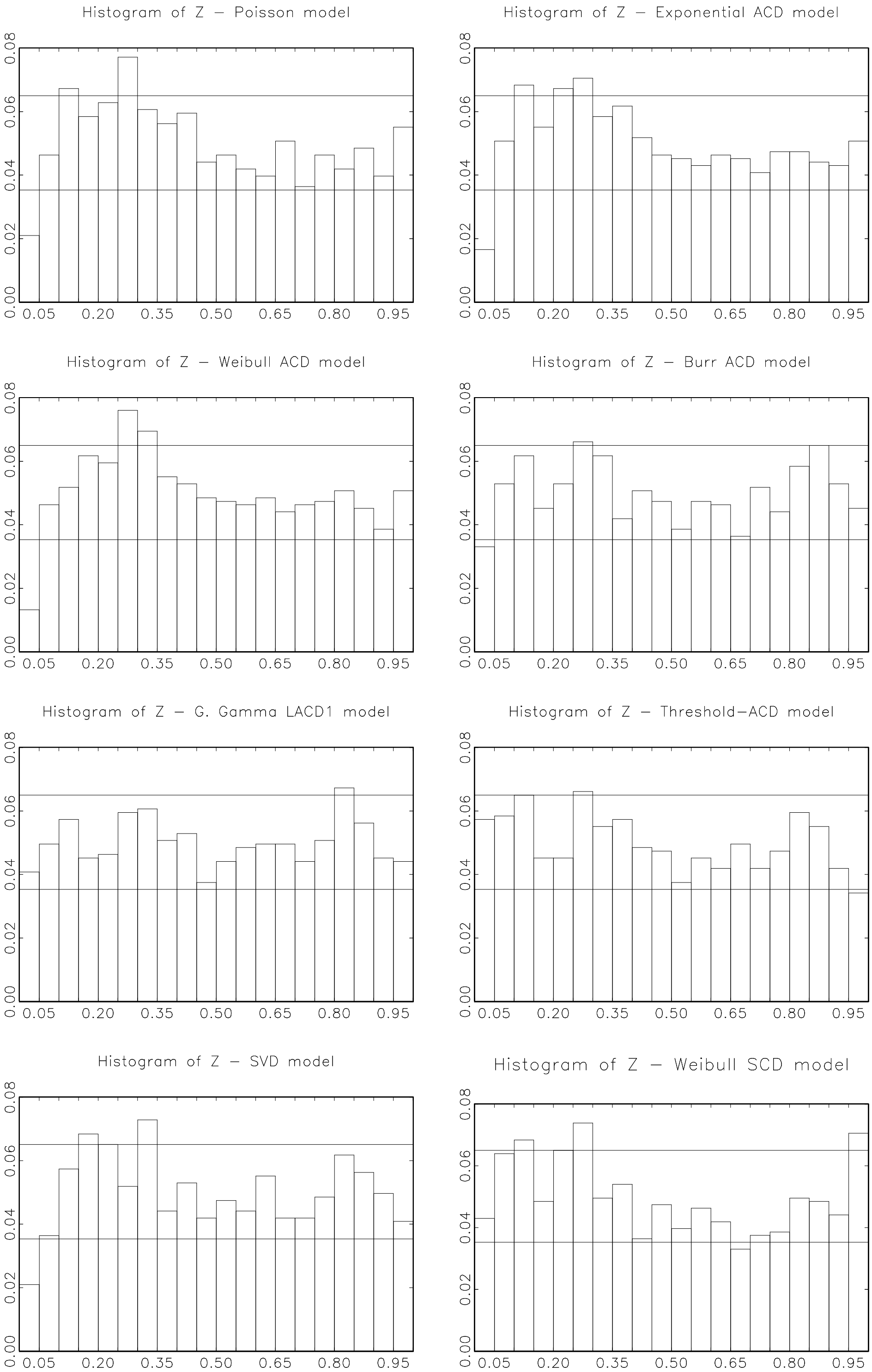

Figure 4: Histograms of the probability integral transform z for various financial duration models. Exxon price durations (out-of-sample). First two thirds of the sample were used for estimation, last one-third for out-of-sample evaluation. The horizontal lines superimposed on the histograms are approximate 95 % confidence intervals for the individual bin heights under the null that z is iid $U(0,1)$.

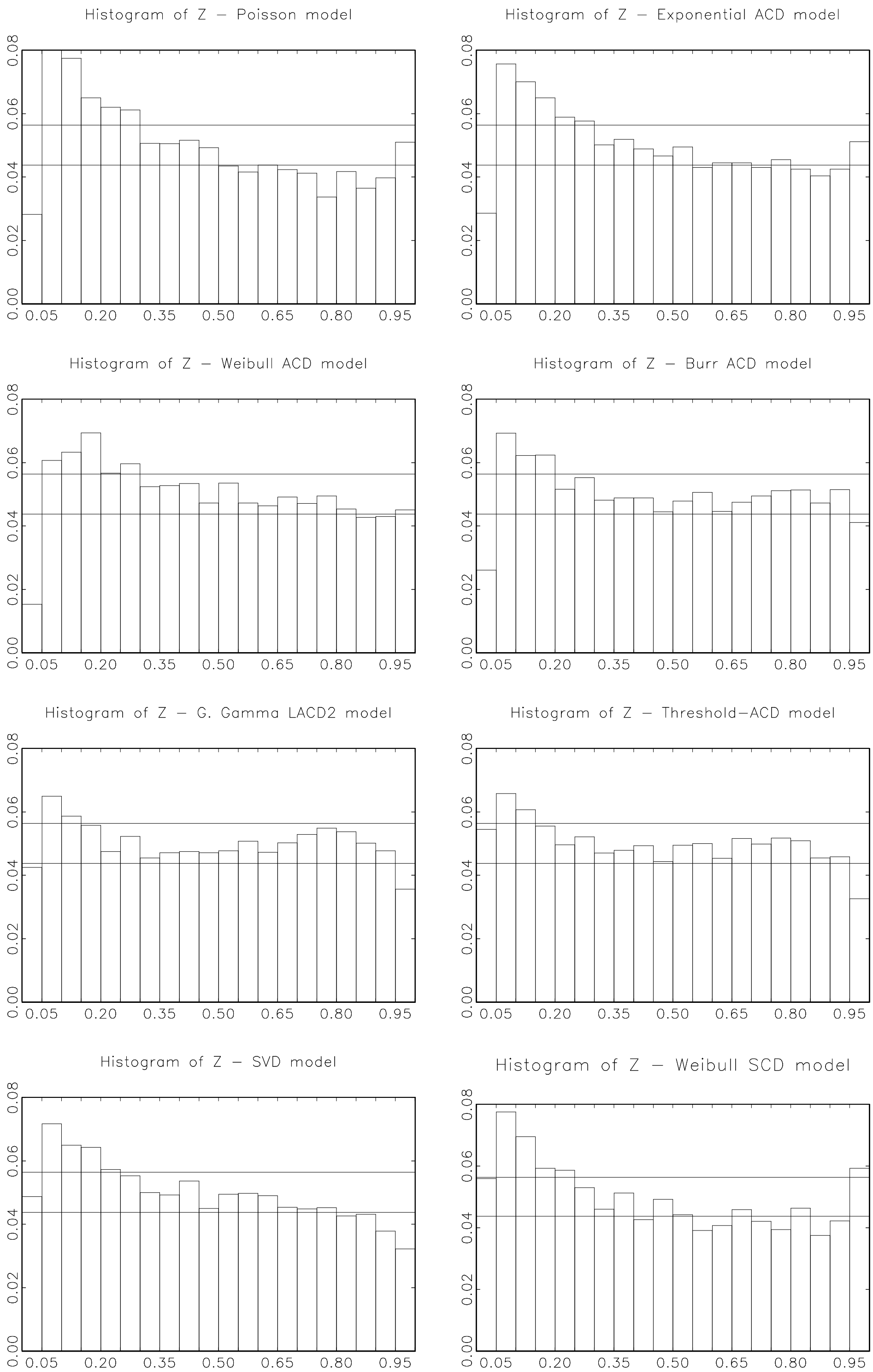

Figure 5: Histograms of the probability integral transform z for various financial duration models. Boeing trade durations (out-of-sample). First two thirds of the sample were used for estimation, last one-third for out-of-sample evaluation. The horizontal lines superimposed on the histograms are approximate 99 % confidence intervals for the individual bin heights under the null that z is iid $U(0,1)$.

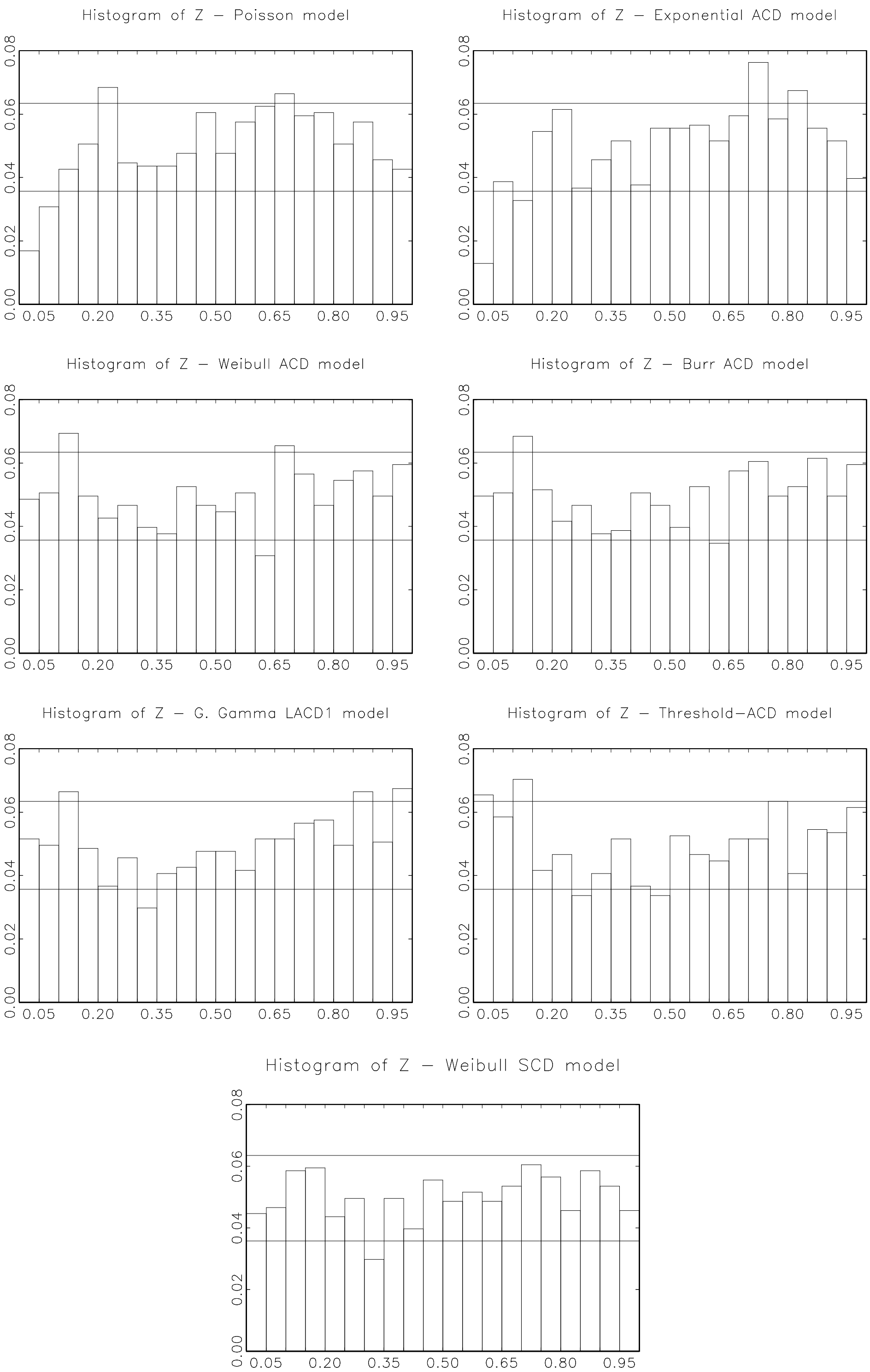

Figure 6: Histograms of the probability integral transform z for various financial duration models. Coca Cola volume durations (out-of-sample). First two thirds of the sample were used for estimation, last one-third for out-of-sample evaluation. The horizontal lines superimposed on the histograms are approximate 95 % confidence intervals for the individual bin heights under the null that z is iid $U(0,1)$.