

Comparison of alternative ACD models via density and interval forecasts: Evidence from the Australian stock market[☆]

David Allen^{a,*}, Zdravetz Lazarov^a, Michael McAleer^b, Shelton Peiris^c

^a School of Accounting, Finance and Economics, Edith Cowan University, Joondlaup Drive 100, WA 6027, Australia

^b School of Economics and Commerce, University of Western Australia, Australia

^c Department of Mathematics and Statistics, University of Sydney, Australia

Available online 24 December 2008

Abstract

In this paper a number of alternative autoregressive conditional duration (ACD) models are compared using a sample of data for three major companies traded on the Australian Stock Exchange. The comparison is performed by employing the methodology for evaluating density and interval forecasts, developed by Diebold et al. [F. Diebold, A. Gunther, S. Tay, Evaluating density forecasts with applications to financial risk management, *International Economic Review* 39 (1998) 863–883] and Christoffersen [P. Christoffersen, Evaluating interval forecasts, *International Economic Review* 39 (1998) 841–862], respectively. Our main finding is that the generalized gamma and log-normal distributions for the error terms have similar performance and perform better than the exponential and Weibull distributions. Additionally, there seems to be no substantial difference between the standard ACD specification of Engle and Russell [R. Engle, J. Russell, Autoregressive conditional duration: a new model for irregularly-spaced transaction data, *Econometrica* 66 (1998) 1127–1162] and the log-ACD specification of Bauwens and Giot [L. Bauwens, P. Giot, The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks, *Annales d'Economie et de Statistique* 60 (2000) 117–150].

© 2009 IMACS. Published by Elsevier B.V. All rights reserved.

Keywords: ACD models; Comparison; Forecasts; Australia

1. Introduction

The introduction of the autoregressive conditional duration model (ACD) by Engle and Russell [11] as well as the increased availability of high-frequency data has sparked a substantial amount of work modelling of the financial durations, both theoretical as well as empirical. Correctly specifying the dynamics of times between transactions of financial assets is important to such issues as testing market microstructure theories (see the survey by Madhavan [21]), optimizing execution costs (see Kissel and Glantz [19]) and intra-daily estimation and forecasting of the volatility (see for example, Ghysels and Jasiak [15] and Engle [10]).

In their seminal paper, Engle and Russell [11] consider a linear specification for the conditional mean equation and exponential and Weibull distributions for the error terms. Their approach has been subsequently extended by many authors. Bauwens and Giot [3] put forth a log-linear specification for the conditional mean which always guarantees the positivity of the durations without imposing restrictions on the coefficients. Zhang et al. [25] introduced a non-linear

[☆] This paper is part of the ARC linkage research project “Modelling stock market liquidity in Australia and the Asia Pacific Region”.

* Corresponding author. Tel: +61 8 6304 5471; fax: +61 8 6304 5271.

E-mail address: d.allen@ecu.edu.au (D. Allen).

version of the ACD model in the spirit of the linear autoregressive threshold models. Bauwens and Veredas [5] proposed a stochastic duration model which is analogous to the stochastic volatility models in the same way as the Engle and Russel's ACD model is analogous to the GARCH model. Ghysels et al. [14] considered a rather complicated version of the ACD model, which allows disentangling the dynamics of the mean and the variance of the duration process.

In addition to the exponential and Weibull distributions for the error terms, other distributions have been proposed. Grammig and Maurer [16], and Grammig and Wellner [17] consider the Burr distribution and Lunde [20] considers the generalized gamma distribution. The log-normal distribution, although seemingly a natural candidate, has received limited attention in the literature with the exception of the work by Allen et al. [2] and Sun et al. [24].

Despite this impressive body of work, only a limited number of papers have been devoted to testing the specifications of the alternative ACD models. Two notable exceptions are the work of Fernandes and Grammig [13] and Bauwens et al. [4]. The first paper evaluates ACD models by gauging the distance between the parametric density of the duration process and its non-parametric estimate, using the methods developed by Ajt-Sahalia [1]. Only the standard ACD specification of Engle and Russell [11] is considered, using error terms with exponential, Weibull, generalized gamma and Burr distributions. Employing only one sample of durations for Exxon, these authors find that the Burr and generalized gamma distributions perform better than the exponential and Weibull distributions.

Bauwens et al. [4], using the methodology for evaluating density forecasts by Diebold et al. [9], consider a number of alternative ACD specifications for three stocks traded on New York stock exchange. One of their main findings is that the exponential and Weibull distributions are often mis-specified, while on the other hand Burr and generalized gamma distributions perform much better. Additionally, they find that the specification for the conditional mean (e.g. the standard ACD, log-ACD as well as non-linear approaches such as the Threshold ACD model) does not seem to affect the models' performance much. In particular, their results suggest very little difference between the ACD and log-ACD specifications.

Given the very little work which has been devoted to evaluating ACD models, which additionally has been done exclusively with US data, in this paper we test a series of ACD models using data from the Australian Stock Exchange. We consider four different distributions for the residuals: exponential, Weibull, generalized gamma and the log-normal one. Two specifications for the conditional mean: standard ACD specification of Engle and Russell [11] and the log-ACD specification of Bauwens and Giot [3] are included in our study. To check for possible variation in the shape of the residual distribution, we also consider two models with GARCH-type variation of the residuals' variance.

The alternative models are evaluated on the basis of their density and interval forecasts, using the methods developed by Diebold et al. [9], and Christoffersen [7], respectively. Our results confirm the findings of Fernandes and Grammig [13] and Bauwens et al. [4] concerning the superior performance of the generalized gamma distribution in comparison to the exponential and Weibull distributions. Also, we find that the ACD and log-ACD specifications are pretty similar, which is also confirmed by Bauwens and Veredas [5]. Additionally, we find that the log-normal distribution has a surprisingly good performance similar to that of the generalized gamma distribution, although it has one free parameter less.

The rest of the paper is organized as follows. In Section 2, we provide a brief summary of the Diebold et al. [9] and Christoffersen [7] methodologies for evaluating density and interval forecasts, respectively. Section 3 contains an overview of the alternative ACD specifications used in the paper. Description of the data and ACD models estimations are discussed in Section 4. Comparison of the different models is provided in Section 5. Finally, Section 6 summarizes the main findings of the paper.

2. Evaluation of ACD models by density and interval forecasts

The literature on forecasting has traditionally focused on evaluating point forecasts. However, in the case of ACD modelling, point forecasts cannot tell us much about the suitability of the model in question, and especially about the appropriateness of the residual distribution, for several reasons. First, Engle and Russell [11] showed that if the equation for the mean is correctly specified, then a QMLE estimation of a model with exponentially distributed error terms will produce consistent and asymptotically normal coefficient estimates. Since the point forecasts do not depend on the shape of the residual distribution, this might render comparison of alternative ACD models difficult.

Second, we find that even though sometimes there is a substantial difference in coefficient estimates, there is not much difference in the precision of the point forecasts even among models with different specification for the mean, such as the standard ACD and the log-ACD model (results not reported here for the sake of brevity).

Additionally, point forecasts provide evidence of how well a model captures the dynamics of the durations around the mean. However, most empirical tests of market microstructure theories are concerned with the dynamics of the (very) short or (very) long durations. For example, periods of high volatility are associated with clustering of short durations.

Recently, Diebold et al. [9] proposed a simple and intuitive method for evaluating nested and non-nested models on the basis of evaluating density forecasts. As a motivation for their approach, these authors showed that forecast users would always prefer a model which produces the correct density function, regardless of their loss function. In addition, their method gives a broader perspective on the models ability to capture the dynamics of durations (in our case) of different sizes.

The basic idea of Diebold et al. [9] approach dates as early as Rosenblatt [23]. The latter author showed that if z is a continuous random variable with cumulative density function $F(x)$, then $F(z)$ is uniformly distributed in the interval $[0,1]$ ($U[0,1]$ random variable). Diebold et al. [9] generalized this result in the following way. Let y_0, y_1, \dots be a time series and let $\{f_i(x|\Psi_{i-1})\}_{i=1}^\infty$ be a sequence of one-step ahead density forecasts conditional on the information at time $i-1$: $\Psi_{i-1} = \{y_{i-1}, y_{i-2}, \dots, y_0\}$, produced by a particular model. The corresponding cumulative density forecasts are given by $F_i(x) = \int_{-\infty}^x f_i(v|\Psi_{i-1})dv$. If the model is correctly specified, then the sequence of probability integral transforms:

$$\{F_i(y_i|\Psi_{i-1})\}_{i=1}^\infty, \quad (1)$$

consists of IID $U[0,1]$ random variables.

In addition to formally testing the assumptions of independence of the random variables $\{F_i(y_i|\Psi_{i-1})\}_{i=1}^\infty$ and their goodness-of-fit to the uniform distribution, Diebold et al. [9] advocate the use of a visual inspection of the z -histogram of $\{F_i(y_i|\Psi_{i-1})\}_{i=1}^\infty$ to detect deviations and assess closeness to the uniform distribution. The approach of informal assessment of the histogram of the probability integral transforms has been used in financial econometrics field by Bauwens et al. [4] and Sun et al. [24] for evaluation of ACD models and by Corsi et al. [8] for evaluation of realized volatility models.

In order to directly assess the ability of alternative models to capture the dynamics of durations of different sizes, we also employ the simple techniques for evaluating interval forecasts, first formally developed by Christoffersen [7]. More specifically, let for a given ACD model and a duration time series d_1, \dots, d_N the threshold levels $dl(p)_i$, $p = 0.99, 0.95, 0.90, 0.80$ computed at time $i-1$, denote lower bounds on duration sizes at times $i, i = 2, \dots, N$, such that:

$$P[d_i \in [dl(p)_i, \infty) | \Psi_{i-1}] = p. \quad (2)$$

With each interval type $[dl(p)_i, \infty)$, we can associate a binary random variable I_i^p which takes value one if $d_i \in [dl(p)_i, \infty)$ and zero otherwise. It is simple to show (see [7]) that each set of random variables $\{I_i^p\}_{i=2}^N$, $p = 0.99, 0.95, 0.90, 0.80$ consists of IID Bernoulli random variables with parameter p , if the ACD model in question is correctly specified. Intuitively, we can assess and compare how competing ACD models capture the dynamics of the (very) short durations (e.g. durations which occur with probability less than $1-p$) by assessing how the corresponding binary variables I_i^p differ from being IID Bernoulli distributed with parameter p . Similarly, to gauge the model's ability to describe the behavior of (long) durations, we compute the threshold levels $dr(p)_i$, $p = 0.99, 0.95, 0.90, 0.80$ such that:

$$P[d_i \in [0, dr(p)_i] | \Psi_{i-1}] = p. \quad (3)$$

Additionally, we calculate two other sets of binary random variables which take value one if the corresponding duration d_i is in the interval $[dl(p)_i, dr(p)_i]$ for $p = 0.90, 0.80$. We use these two sets of random variables to assess how the medium sized durations are captured. Obviously:

$$P[d_i \in [dl(p)_i, dr(p)_i] | \Psi_{i-1}] = 2p - 1. \quad (4)$$

First, we test whether the estimated Bernoulli parameter of the sequence of the corresponding binary variables differs from its assumed value by employing the following simple likelihood-ratio test by Christoffersen [7]. Let I_t , $t = 1, \dots, s$

be a sequence of identically and independently distributed Bernoulli variables with parameter λ . Then a test of the hypothesis $\lambda = p$ versus the alternative $\lambda \neq p$ can be formulated as a standard likelihood ratio test:

$$LR = -2 \log \left(\frac{L(p)}{L(\pi)} \right) \sim \chi^2(1), \quad (5)$$

where $L(p) = (1-p)^{s_0} \cdot p^{s_1}$ is the likelihood function under the null hypothesis and $L(\pi) = (1-\pi)^{s_0} \cdot \pi^{s_1}$ is the likelihood function under the alternative. By s_0 and s_1 are denoted the number of 1's and 0's in the sequence $\{I_t\}_{t=1,\dots,s}$ and $\pi = s_1/(s_0 + s_1) = s_1/s$ is the maximum likelihood estimator of the Bernoulli parameter of $\{I_t\}_{t=1,\dots,s}$.

The test above is valid under the assumptions that the binary variables I_t , $t = 1, \dots, s$ are independent. The presence of serial autocorrelation could bias the inference about the Bernoulli parameter. Keeping this in mind, we also perform a test for serial correlation. In addition, it is desirable that an ACD model is able to capture the clustering of short (long) durations, since periods with concentrated occurrence of short (long) durations is associated with important economic phenomenon such as high (low) volatility.

We check for first-order serial correlation by testing the null hypothesis of independence versus the alternative hypothesis that the series of binary variables follows a Markov chain. Let $\{I_t\}_{t=1,\dots,s}$ be a first-order Markov chain with a transition matrix:

$$\Pi = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}, \quad (6)$$

where $\pi_{km} = P[I_t = k | I_{t-1} = m]$ are the transition probabilities. The approximate likelihood function (omitting the first observation) is

$$L(\{I_t\}_{t=1,\dots,s}; \pi_{01}, \pi_{11}) = (1 - \pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1 - \pi_{11})^{n_{10}} \pi_{11}^{n_{11}}, \quad (7)$$

where the n_{km} is the number of consecutive pairs $\{I_t = k, I_{t-1} = m\}$. The maximum-likelihood estimation for the parameters π_{01} and π_{11} can be easily computed as $\hat{\pi}_{01} = n_{01}/(n_{00} + n_{01})$ and $\hat{\pi}_{11} = n_{11}/(n_{10} + n_{11})$. Under the null hypothesis of independence, the transition matrix reduces to:

$$\Pi_0 = \begin{bmatrix} 1 - \pi & \pi \\ 1 - \pi & \pi \end{bmatrix}. \quad (8)$$

The corresponding likelihood function is

$$L(\{I_t^*\}_{t=1,\dots,s}; \pi) = (1 - \pi)^{n_{00} + n_{10}} \pi^{n_{01} + n_{11}}, \quad (9)$$

and the maximum-likelihood estimate for π is $\hat{\pi} = (n_{01} + n_{11})/s$. From standard results on Markov chains ([18]) it follows that the LR test for independence

$$LR = -2 \log(L(\{I_t\}_{t=1,\dots,s}; \hat{\pi}) / L(\{I_t\}_{t=1,\dots,s}; \hat{\pi}_{01}, \hat{\pi}_{11})), \quad (10)$$

is asymptotically distributed as a $\chi^2(1)$.

Using the above results, we perform tests for unconditional coverage and independence of all interval forecasts and report their $\chi^2(1)$ respective statistics.

3. Model overview

In this section we give a short overview of the different ACD models which are used for density and interval forecasts. In addition to the widely used ACD and log-ACD models, we also introduce two ACD models with GARCH-type time-varying variance for the log-normally distributed error terms.

The literature on ACD modelling started with the seminal paper of Engle and Russel [11]. They proposed modelling the time series of financial durations $\{d_i\}_{i=0}^\infty$ as a stochastic process of the following form:

$$d_i = \mu_i \cdot \varepsilon_i. \quad (11)$$

Here μ_i is the conditional mean value of the duration d_i , i.e. $\mu_i = E[d_i | \Psi_{i-1}]$ where Ψ_{i-1} is the information set at time $i - 1$: $\Psi_{i-1} = \{d_i, \dots, d_0\}$. The specification of the conditional mean is usually assumed to have the following autoregressive linear form:

$$\mu_i = \alpha + \beta \cdot d_i + \gamma \cdot \mu_{i-1}. \quad (12)$$

Of course, more lagged values of the duration and the conditional duration can be included in (12). However, to keep the considered models as parsimonious as possible we consider only the specification (12), which is also the one most widely used in the literature. The random variables $\{\varepsilon_i\}_{i=1}^{\infty}$ are positive, identically and independently distributed with mean value one. The above model is commonly referred in the literature as the ACD(1,1) model.

It is clear that the ACD model is very similar to the GARCH model and this similarity can be exploited to derive its asymptotic properties (see for example [11]). Engle and Russel [11] initially considered the exponential distribution for the residuals $\{\varepsilon_i\}_{i=1}^{\infty}$ with density:

$$f(x) = \exp(-x), \quad x \geq 0. \quad (13)$$

The rationale for this choice is parsimony and the fact that if the conditional mean μ_i is correctly specified, then the QMLE estimator is consistent and asymptotically normal (under some regularity conditions). As an extension, these authors also considered the standard Weibull distribution for the residuals, which nests the exponential distribution and has density of the form:

$$f(x|\nu) = \nu \cdot \left(\Gamma \left(1 + \frac{1}{\nu} \right) \right)^{\nu} \cdot x^{\nu-1} \cdot \exp \left(- \left(\Gamma \left(1 + \frac{1}{\nu} \right) \cdot x \right)^{\nu} \right), \quad x \geq 0. \quad (14)$$

Here $\nu > 0$ is the shape parameter of the Weibull distribution, which reduces to the exponential distribution when $\nu = 1$. Lunde [20] was the first one to propose the use of the standard generalized gamma distribution, which nests the standard Weibull distribution. It has density of the following form:

$$f(x|\nu, \kappa) = \frac{\nu \cdot x^{\kappa\nu-1}}{\lambda^{\kappa\nu} \cdot \Gamma(\kappa)} \cdot \exp \left(- \left(\frac{x}{\lambda} \right)^{\nu} \right), \quad x \geq 0, \quad (15)$$

where $\lambda = \Gamma(\kappa)/\Gamma(\kappa + 1/\nu)$. Here ν and κ are shape parameters and the standard generalized gamma distribution reduces to the standard Weibull distribution when $\kappa = 1$. In addition to nesting specifications (13) and (14), the specification (15) allows for non-monotonic hazard function, while the standard Weibull density has necessarily a monotonic hazard function and the exponential density has a constant one.

Besides these three distributions, we also consider a log-normal specification for the residuals. The corresponding density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2} \cdot x} \exp \left[- \frac{(\ln x + (1/2)\sigma^2)^2}{2\sigma^2} \right], \quad x \geq 0. \quad (16)$$

The density (16) is the density of the exponent of a normally distributed random variable with standard deviation σ^2 and mean $-1/2\sigma^2$. The value of the mean is chosen in order to guarantee that the log-normal error terms have expected value one.

The log-normal distribution seems to be a natural choice for positive residual terms, but surprisingly it has received little attention in the literature. To our knowledge, the only papers which perform estimations of ACD models with log-normally distributed error terms are Allen et al. [2] and Sun et al. [24]. Thus, one of the contributions of this paper is to shed some light on the empirical properties of ACD models with log-normally distributed residuals. In addition, for notational purposes, we denote the corresponding ACD(1,1) models with exponential, Weibull, generalized gamma and log-normal distribution for the residuals, by EACD(1,1), WACD(1,1), GACD(1,1) and LACD(1,1) models, respectively.

Generally, the positivity of the conditional mean μ_i given by (12) can not be guaranteed unless positivity restrictions on the coefficients are imposed, i.e. $\alpha \geq 0$, $\beta > 0$ and $\gamma \geq 0$. Although it seems natural to assume that $\beta > 0$ and $\gamma \geq 0$, it is not so for the intercept coefficient α . In addition, negative values for the conditional mean could arise, if exogenous regressors are included in (12) which affect negatively the conditional durations.

To account for this drawback of the standard ACD models, Bauwens and Giot [3] propose the so-called log-ACD model. It differs from the standard ACD model by using the following specification for the conditional mean:

$$\mu_i = \exp(\alpha + \beta \cdot \log(d_i) + \gamma \cdot \log(\mu_{i-1})). \quad (17)$$

The volatility counterpart of the log-ACD model is the Nelson's EGARCH model [22], and as in the case of the EGARCH model, the asymptotic properties of the log-ACD model are unknown (see for example, [12]). However, Bauwens et al. [6] derive models moment conditions and Allen et al. [2] consider its finite sample properties.

Given that the log-ACD model has gained popularity as a tool for modelling financial durations, we include it in our study. As in the case of the standard ACD model, four specifications for the residuals with exponential, Weibull, generalized gamma and log-normal distribution, are considered. These models are denoted as Log-EACD(1,1), Log-WACD(1,1), Log-GACD(1,1) and Log-LACD(1,1) models, respectively.

We also proceed to check the importance of higher-order moments of the durations and residuals for density and interval forecasting. An important feature of the ACD and log-ACD models is that the dynamics of the higher-order moments of the conditional durations is completely specified by the conditional mean. However, this might be too restrictive as argued by Ghysels et al. [14], who proposed a model which disentangles the movements of the mean and variance. They call their model the stochastic volatility duration (SVD) model. In this SVD setting, the dynamics of the durations is driven by a complicated non-linear two-factor latent model, which is difficult to estimate.

To be as parsimonious as possible, we propose two models with time-varying variance for the residuals, which nest and are a straightforward generalization of the LACD(1,1) and Log-LACD(1,1) models, respectively. The first one, which we call the GARCH Volatility ACD (GV-ACD(1,1)) model specifies the dynamics of the durations as in (11) and (12). However, the variance of the residuals is time-varying and it is given in the following way:

$$\log(\varepsilon_i) \sim N\left(-\frac{1}{2}\sigma_i^2, \sigma_i^2\right),$$

$$\sigma_i^2 = \alpha^1 + \beta^1 \cdot \left(\log(\varepsilon_{i-1}) + \frac{1}{2}\sigma_{i-1}^2\right)^2 + \gamma^1 \cdot \sigma_{i-1}^2. \quad (18)$$

The specification (18) is quite similar to the standard GARCH volatility specification. The difference arises in using the term $(\log(\varepsilon_{i-1}) + 1/2\sigma_{i-1}^2)^2$ instead of $\log(\varepsilon_{i-1})^2$ as a proxy for the variance of the duration at time $i - 1$. This is done to account for the fact that the conditional mean of $\log(\varepsilon_{i-1})$ is non-zero and has value $1/2\sigma_{i-1}^2$. Similarly, we consider a version of the GV-ACD(1,1) model where the conditional mean has the form (17), and call this model GV-Log-ACD(1,1) model.

4. Data description and ACD model estimation

The data used in this study comes from the Australian stock exchange (ASX). ASX is the primary stock exchange in Australia operating as a pure limit order market. On 25 July 2006 it merged with Sydney Futures Exchange, creating the ninth largest listed exchange in the world. The normal trading hours are between 10:00 a.m. and 16:00 p.m. with an initial opening period between 10:00 a.m. and 10:10 a.m., when securities open according to the first letter of their ticker symbol in groups of five. We omit all trades in the opening period and consider only transactions that occur in the normal trading period. A similar approach has been adopted by other authors (e.g. [11]). Our data consists of complete sets of transactions on three stocks randomly chosen from the Australian Stock Market index ASX 20, namely National Australia Bank (a major bank with ticker symbol NAB), Westpac Banking Corporation (another major bank with ticker symbol WBC) and Telstra (the main telecommunication company with ticker symbol TLS). To ensure robustness of the results we use different time periods for each stock.

The transactions data for NAB spans the time interval 2 January 2004 to 31 March 2004, which consists of 63 trading days and 64 561 durations. Following Engle and Russel [11] this full dataset is used to “diurnally adjust” the data in order to account for the time-of-the day effects. This is achieved by using a linear spline with knots at round hours, e.g. 10:00 a.m., 11:00 a.m., 12:00 a.m., 13:00 p.m., 14:00 p.m., 15:00 p.m. and 16:00 p.m. We also experiment with other approaches for deseasonalization. However, this seems to have little impact on coefficient estimates well

as our findings concerning ACD model forecasts. Similar results have been reported by other authors, for example ([3]). Estimations of different ACD models using all the data for whole period of three months suggests that most of the considered models are miss-specified as indicated by small but significant autocorrelation in the fitted residuals. In view of these findings, we restrict our attention to the first 10,000 durations. In the case of NAB, the time interval for the considered duration data shrinks to 2 January 2004 to 19 January 2004. The transactions data for WBC and TLS is seasonally adjusted in the same manner, and after restricting the two duration samples to the first 10,000 observations, we end up with time intervals spanning 2 April 2002 to 22 April 2002 (15 trading days) and 2 January 2003 to 23 January 2003 (16 trading days), respectively. In Tables 1–3 are presented the estimations of the ACD(1,1) models described in Section 3 for NAB, WBC and TLS, respectively. We report the coefficient estimates, together with the corresponding heteroscedasticity robust errors. All coefficients are highly significant, which is not surprising given the large number of observations.

As a robustness check, the Ljung–Box test for autocorrelation in the fitted residuals for each model is performed. The results, not reported here for the sake of brevity, indicate very small degree of autocorrelation which is almost always insignificant at any conventional statistical level. Similar results hold for the squared fitted residuals.

As it can be seen, the sum of the coefficients of the autoregressive terms is always greater than 0.9, indicating a high degree of persistence. In the case of the WACD and log-WACD models, the residual shape parameter ν is quite close to unity, the value at which the Weibull distribution collapses to the exponential one. Also, the coefficient estimates for conditional mean for the EACD and WACD models, as well as log-EACD and log-WACD, are quite close to each other. This suggests that nothing much is gained by replacing the exponential distribution with a Weibull one in the specification for the residuals. In Section 5, we will see that the two distributions also exhibit quite close forecasting performances.

For the generalized gamma distribution, the corresponding shape parameter ν is well below one. Also, in the case of NAB and WBC, the coefficient estimates for the conditional mean are substantially different from the corresponding estimates for the EACD, WACD, log-EACD and log-WACD models. Additionally, for these two companies, the (Log-)GACD and (Log-)LACD models seem to give pretty close coefficient estimates, although the generalized gamma and the log-normal distribution are non-nested, and the latter has one parameter less. Results in Section 5 indicate that the (Log-)GACD model exhibits similar forecasting performance to the (Log-LACD) model, respectively. In contrast, for TLS, the three nested distributions give estimates for the conditional mean which are close to each other, while the log-normal ACD specification produces noticeably different estimates.

The GARCH volatility models GV-LACD and GV-Log-LACD give estimates for the conditional mean pretty close to the estimates for their fixed residual variance models, the LACD and Log-LACD models. The estimates for the variance equations always give highly significant autoregressive coefficients. For the GV-LACD model, ARCH-type coefficient β^1 is quite small and falls in the range of 1–4%. There seems to be evidence for more time-variation in the residual variance for NAB and TLS, than for WBC, since for the former two companies the GARCH-type coefficient γ^1 is much smaller. Indeed, the fitted conditional residual variance shows greater variability for these two companies than for WBC (results not reported here for the sake of brevity).

Estimations of the GV-Log-LACD models suggest a much smaller degree of variation in the residual variance. The ARCH-type coefficient β^1 is about 0.4% for NAB and WBC, and about 1.7% for TLS. The GARCH-type coefficient γ^1 is quite high and very close to unity for the three companies.

5. Performance comparison

In this section we analyze the density and interval forecasts of the alternative ACD models. First, we report the z -histograms of the probability integral transforms (1) and their corresponding autocorrelations up to lag 10. Then the performance of forecasts for intervals with different sizes is discussed.

Visual inspection of the z -histograms for the probability integral transforms $\{F_i(y_i|\Psi_{i-1})\}_{i=1}^\infty$ has been advocated by Diebold et al. [9] as a simple and reliable tool for detecting deviations from uniformity (see Section 2). Figs. 1–3 contain the corresponding z -histograms with 40 bins, for NAB, WBC and TLS, respectively.

First, notice that there is no discernible substantial difference between the z -histograms of the standard ACD models and their log-ACD counterparts. On the other hand, there is a substantial difference in the histogram shape for models with different residual distribution. EACD, WACD, log-EACD and log-EACD histograms exhibit a bump on the left-hand side, which is especially pronounced for NAB and WBC. This suggests that these models might not be able to

Table 1
Autocorrelations of the probability integral transforms for NAB ^a.

Lag	EACD			WACD			GACD			LACD			GV-LACD		
	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value
1	0.050	25.334	0.000	0.047	21.906	0.000	0.030	8.8121	0.003	0.029	8.4222	0.004	0.031	9.3973	0.002
2	0.046	46.266	0.000	0.043	40.043	0.000	0.033	19.585	0.000	0.034	19.707	0.000	0.033	20.308	0.000
3	0.030	55.205	0.000	0.028	47.654	0.000	0.018	22.694	0.000	0.019	23.179	0.000	0.015	22.696	0.000
4	0.019	59.000	0.000	0.017	50.621	0.000	0.011	23.930	0.000	0.013	24.804	0.000	0.007	23.245	0.000
5	0.007	59.549	0.000	0.005	50.922	0.000	0.004	24.056	0.000	0.006	25.161	0.000	0.001	23.259	0.000
6	0.009	60.290	0.000	0.007	51.368	0.000	0.011	25.353	0.000	0.015	27.413	0.000	0.007	23.797	0.001
7	0.001	60.291	0.000	−0.001	51.377	0.000	0.001	25.355	0.001	0.003	27.524	0.000	−0.005	24.068	0.001
8	−0.005	60.567	0.000	−0.006	51.694	0.000	−0.002	25.382	0.001	0.001	27.545	0.001	−0.008	24.687	0.002
9	−0.003	60.641	0.000	−0.003	51.795	0.000	0.005	25.593	0.002	0.008	28.190	0.001	−0.003	24.771	0.003
10	−0.009	61.456	0.000	−0.009	52.564	0.000	−0.003	25.675	0.004	0.001	28.190	0.002	−0.009	25.562	0.004

Lag	Log-EACD			Log-WACD			Log-GACD			Log-LACD			GV-Log-LACD		
	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value
1	0.049	24.294	0.000	0.047	21.906	0.000	0.017	2.9519	0.086	0.014	2.0859	0.149	0.039	15.036	0.000
2	0.044	43.852	0.000	0.043	40.043	0.000	0.022	7.6160	0.022	0.020	6.1735	0.046	0.041	31.881	0.000
3	0.028	51.447	0.000	0.028	47.654	0.000	0.007	8.1158	0.044	0.006	6.5669	0.087	0.023	37.208	0.000
4	0.017	54.259	0.000	0.017	50.621	0.000	0.001	8.1202	0.087	0.001	6.5686	0.161	0.016	39.710	0.000
5	0.003	54.349	0.000	0.005	50.922	0.000	−0.006	8.5172	0.130	−0.006	6.8788	0.230	0.008	40.371	0.000
6	0.006	54.672	0.000	0.007	51.368	0.000	0.004	8.7026	0.191	0.006	7.2427	0.299	0.016	42.885	0.000
7	−0.005	54.934	0.000	0.001	51.377	0.000	−0.008	9.3446	0.229	−0.007	7.6693	0.363	0.002	42.921	0.000
8	−0.010	55.997	0.000	0.006	51.694	0.000	−0.010	10.259	0.247	−0.008	8.3112	0.404	−0.001	42.937	0.000
9	−0.010	56.945	0.000	0.003	51.795	0.000	−0.004	10.420	0.318	−0.002	8.3442	0.500	0.005	43.146	0.000
10	−0.016	59.558	0.000	0.009	52.564	0.000	−0.011	11.583	0.314	−0.009	9.1011	0.523	−0.003	43.235	0.000

^a The table presents the autocorrelation coefficients, Q-statistics and the associated p-values for the residuals of the corresponding ACD models up to lag 10.

Table 2
Autocorrelations of the probability integral transforms for WBC ^a.

Lag	EACD			WACD			GACD			LACD			GV-LACD		
	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value
1	0.071	50.662	0.000	0.068	45.636	0.000	0.064	40.519	0.000	0.064	41.241	0.000	0.022	4.7096	0.030
2	0.023	56.055	0.000	0.021	49.875	0.000	0.017	43.282	0.000	0.017	44.290	0.000	0.029	13.068	0.001
3	0.022	61.003	0.000	0.020	53.93	0.000	0.016	45.986	0.000	0.017	47.247	0.000	0.016	15.525	0.001
4	0.010	61.959	0.000	0.008	54.638	0.000	0.004	46.137	0.000	0.004	47.434	0.000	0.010	16.526	0.002
5	0.004	62.117	0.000	0.002	54.697	0.000	0.005	46.439	0.000	0.008	48.069	0.000	0.006	16.871	0.005
6	0.010	63.088	0.000	0.009	55.437	0.000	0.012	47.898	0.000	0.014	50.120	0.000	0.016	19.326	0.004
7	−0.014	65.072	0.000	0.015	57.550	0.000	0.014	49.737	0.000	0.012	51.559	0.000	0.004	19.502	0.007
8	0.010	66.014	0.000	0.009	58.352	0.000	0.015	51.975	0.000	0.018	54.673	0.000	0.003	19.581	0.012
9	−0.013	67.654	0.000	0.013	59.949	0.000	0.010	52.982	0.000	0.008	55.332	0.000	0.009	20.456	0.015
10	−0.002	67.710	0.000	0.002	59.99	0.000	0.001	52.987	0.000	0.002	55.386	0.000	0.002	20.518	0.025
Lag	Log-EACD			Log-WACD			Log-GACD			Log-LACD			GV-Log-LACD		
	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value
1	0.052	27.174	0.000	0.068	45.636	0.000	0.038	14.594	0.000	0.037	13.667	0.000	0.019	3.6065	0.058
2	0.007	27.678	0.000	0.021	49.875	0.000	0.004	14.769	0.001	0.005	13.887	0.001	0.029	12.276	0.002
3	0.013	29.449	0.000	0.02	53.930	0.000	0.003	14.849	0.002	0.002	13.929	0.003	0.020	16.430	0.001
4	0.002	29.500	0.000	0.008	54.638	0.000	0.007	15.371	0.004	0.008	14.552	0.006	0.019	19.865	0.001
5	0.001	29.500	0.000	0.002	54.697	0.000	0.001	15.387	0.009	0.001	14.552	0.012	0.016	22.309	0.000
6	0.008	30.198	0.000	0.009	55.437	0.000	0.007	15.912	0.014	0.008	15.227	0.019	0.027	29.512	0.000
7	−0.015	32.521	0.000	0.015	57.550	0.000	0.018	19.086	0.008	0.017	18.240	0.011	0.015	31.871	0.000
8	0.012	33.972	0.000	0.009	58.352	0.000	0.014	21.035	0.007	0.015	20.593	0.008	0.015	34.155	0.000
9	−0.011	35.087	0.000	0.013	59.949	0.000	0.011	22.235	0.008	0.010	21.614	0.010	0.021	38.500	0.000
10	0.001	35.094	0.000	0.002	59.990	0.000	0.001	22.235	0.014	0.001	21.615	0.017	0.016	41.109	0.000

^a The table presents the autocorrelation coefficients, Q-statistics and the associated *p*-values for the residuals of the corresponding ACD models up to lag 10.

Table 3
Autocorrelations of the probability integral transforms for TLS ^a.

Lag	EACD			WACD			GACD			LACD			GV-LACD		
	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value
1	0.061	37.350	0.000	0.060	35.784	0.000	0.0370	13.470	0.000	0.046	20.902	0.000	0.048	23.497	0.000
2	0.014	39.406	0.000	0.013	37.605	0.000	-0.013	15.050	0.001	-0.002	20.957	0.000	0.002	23.535	0.000
3	0.008	40.058	0.000	0.007	38.127	0.000	-0.014	17.045	0.001	-0.002	20.999	0.000	0.001	23.539	0.000
4	-0.009	40.822	0.000	-0.009	39.001	0.000	-0.032	27.420	0.000	-0.020	25.185	0.000	-0.016	26.043	0.000
5	-0.002	40.875	0.000	-0.003	39.082	0.000	-0.023	32.597	0.000	-0.012	26.633	0.000	-0.007	26.550	0.000
6	0.007	41.357	0.000	0.006	39.488	0.000	-0.010	33.557	0.000	-0.001	26.645	0.000	0.005	26.754	0.000
7	0.005	41.567	0.000	0.004	39.656	0.000	-0.008	34.122	0.000	-0.001	26.657	0.000	0.004	26.946	0.000
8	0.004	41.719	0.000	0.003	39.764	0.000	-0.006	34.493	0.000	0.001	26.658	0.001	0.005	27.233	0.001
9	0.021	46.224	0.000	0.021	44.127	0.000	0.012	35.902	0.000	0.017	29.466	0.001	0.023	32.417	0.000
10	0.015	48.392	0.000	0.015	46.242	0.000	0.004	36.026	0.000	0.008	30.184	0.001	0.014	34.401	0.000

Lag	Log-EACD			Log-WACD			Log-GACD			Log-LACD			GV-Log-LACD		
	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value	AC	Q-stat.	p-Value
1	0.067	45.168	0.000	0.064	41.161	0.000	0.037	13.470	0.000	0.034	11.632	0.001	0.060	36.013	0.000
2	0.017	47.902	0.000	0.016	43.681	0.000	-0.013	15.050	0.001	-0.014	13.709	0.001	0.011	37.291	0.000
3	0.006	48.293	0.000	0.006	44.098	0.000	-0.014	17.045	0.001	-0.015	15.876	0.001	0.010	38.274	0.000
4	-0.014	50.152	0.000	-0.012	45.434	0.000	-0.032	27.420	0.000	-0.032	26.311	0.000	-0.010	39.188	0.000
5	-0.009	50.979	0.000	-0.008	46.009	0.000	-0.023	32.597	0.000	-0.022	31.377	0.000	-0.002	39.240	0.000
6	-0.001	50.987	0.000	0.001	46.014	0.000	-0.010	33.557	0.000	-0.009	32.219	0.000	0.006	39.646	0.000
7	-0.003	51.067	0.000	-0.002	46.068	0.000	-0.008	34.122	0.000	-0.006	32.625	0.000	0.007	40.141	0.000
8	-0.004	51.254	0.000	-0.003	46.171	0.000	-0.006	34.493	0.000	-0.005	32.840	0.000	0.008	40.839	0.000
9	0.011	52.470	0.000	0.012	47.693	0.000	0.012	35.902	0.000	0.013	34.601	0.000	0.022	45.724	0.000
10	0.004	52.630	0.000	0.005	47.977	0.000	0.004	36.026	0.000	0.004	34.796	0.000	0.013	47.370	0.000

^a The table presents the autocorrelation coefficients, Q-statistics and the associated *p*-values for the residuals of the corresponding ACD models up to lag 10.

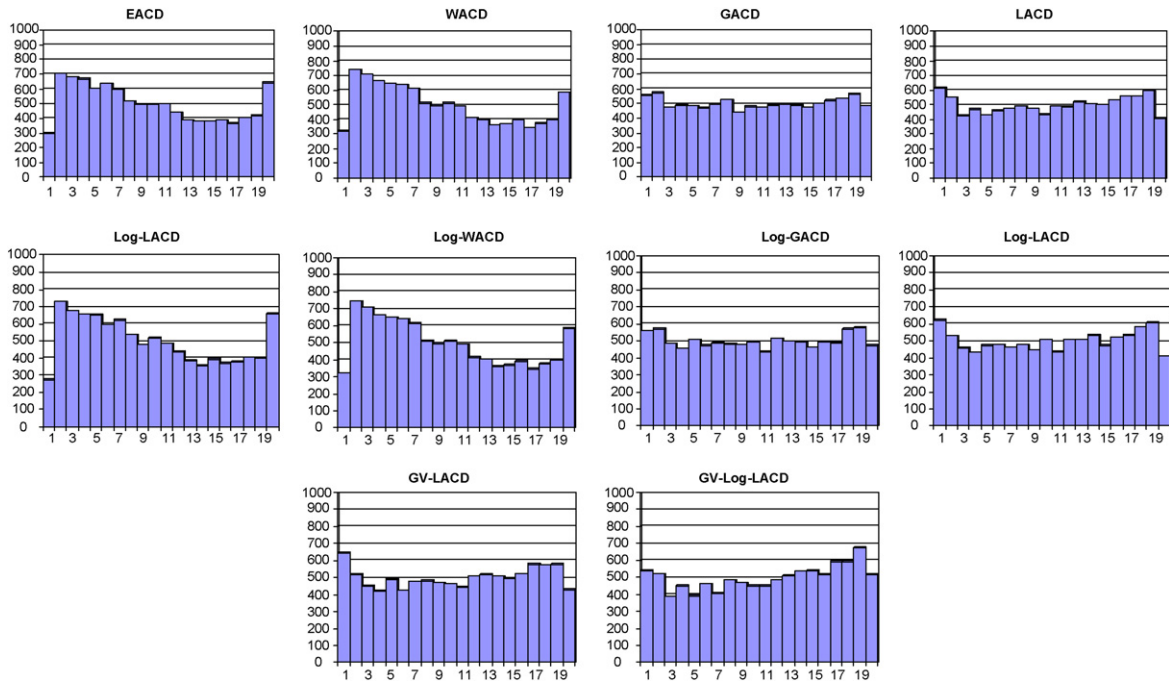


Fig. 1. Histograms of the probability integral transforms for National Australia Bank (NAB).

capture the left-tail of the durations distribution very well, which is confirmed by the analyzes of the interval forecasts below.

The z -histograms for models with generalized gamma distribution seem to match the z -histogram of the uniform distribution most closely. For NAB and WBC, the LACD and log-LACD models have z -histograms pretty similar to their GACD and log-GACD counterparts. However, for TLS, the z -histogram for the LACD and log-LACD models has a pronounced bump on the right-hand side and big spike in the first bin.

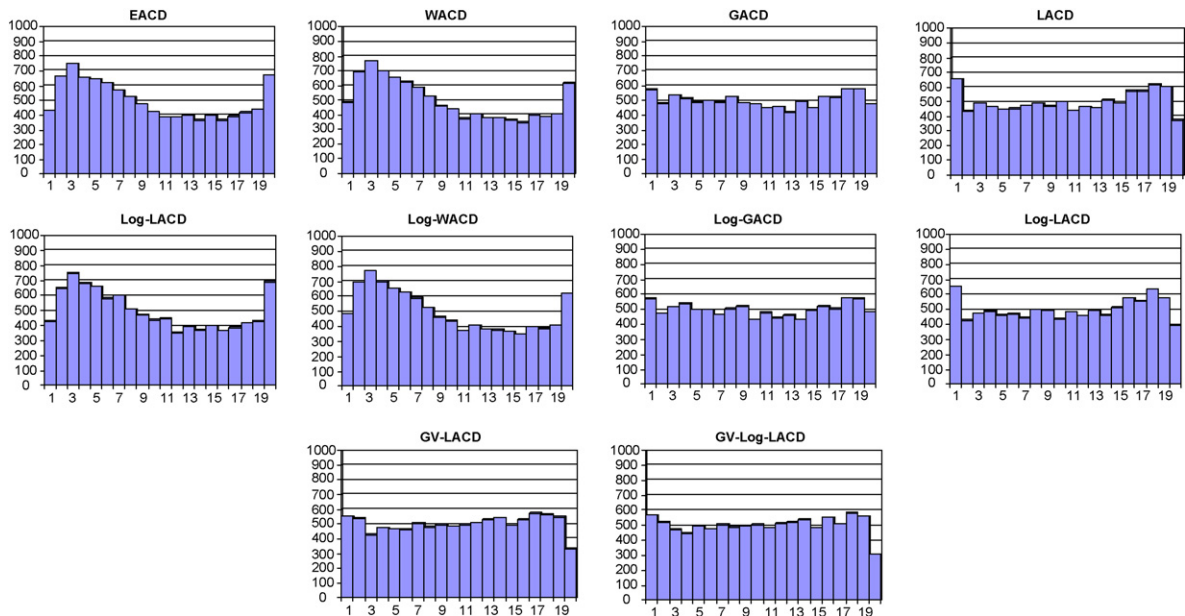


Fig. 2. Histograms of the probability integral transforms for Westpack Banking Corporation (WBC).

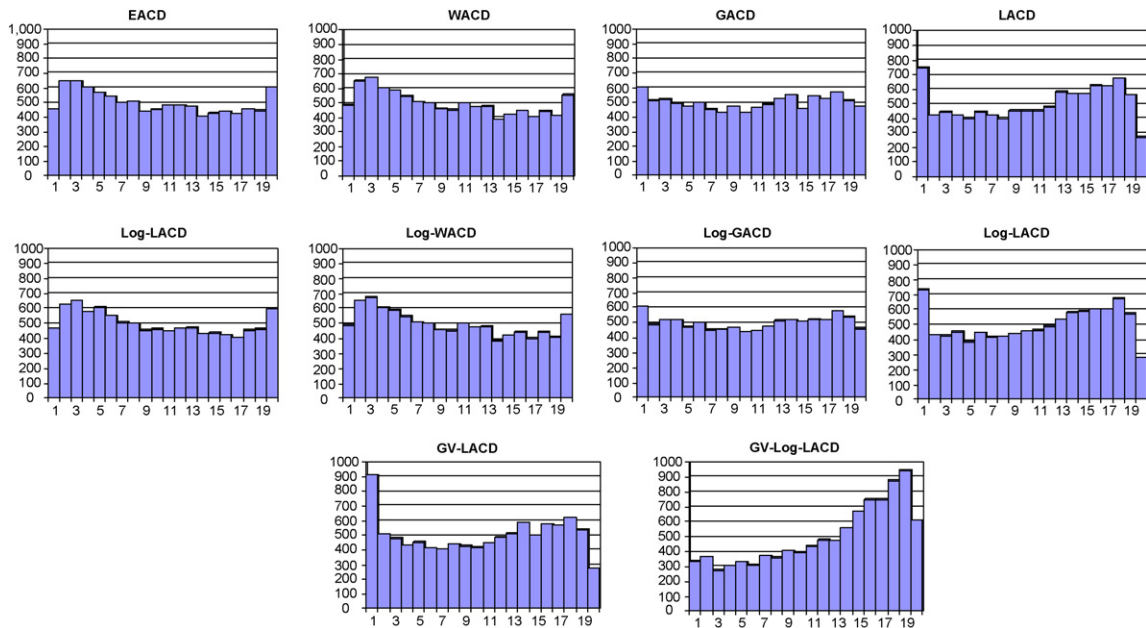


Fig. 3. Histograms of the probability integral transforms for Telstra Corporation (TLS).

The evidence for GV-LACD and GV-Log-LACD models is mixed. For NAB, there seems to be not much difference between GARCH Volatility models and their fixed variance counterparts. On the other hand, for WBC, the z -histograms for GARCH Volatility ACD models is visibly smoother, with no spikes in the first bin. Finally, for TLS, the GARCH Volatility specifications exhibit worsening in the z -histograms, esp. for the GV-Log-LACD model.

We also report the autocorrelations up to lag 10 as well as the Ljung–Box statistics of the probability integral transforms (1) in Tables 4–6. As we can see, there is always a small, but statistically significant autocorrelation, with the exception of the Log-GACD and Log-LACD models for NAB.

Generally, the autocorrelation pattern is smaller for GACD, LACD, log-GACD and log-LACD models than for the corresponding models with exponential and Weibull distributions. For example, the first-order autocorrelation usually decreases by more than a half for models with exponential and Weibull distributions for the residuals. The evidence for the GARCH Volatility models is mixed: for NAB there is no change in the autocorrelation pattern when extending the LACD model to GV-LACD model, while for GV-Log-LACD model the autocorrelation increases in comparison to the Log-LACD model. For WBC, the autocorrelation decreases and for TLS, the autocorrelation increases for both GV-LACD and GV-Log-LACD models.

The evidence from the behavior of the probability integral transforms suggests that generalized gamma and log-normal distributions might be a better choice than the exponential and Weibull distributions. To investigate this possibility more thoroughly we perform analysis of the interval forecasts below.

In Tables 7–9 are presented the descriptive statistics for different interval forecasts for NAB, WBC and TLS, respectively. We report the corresponding $\chi^2(1)$ -statistics for unconditional coverage and independence. We do not report the independence statistics for intervals of the type $[dl(0.99), \infty]$ and $[0, dr(0.99)]$ representing very short and long durations, since it is difficult to make inference about the serial dependence of the corresponding binary variables due to the small size of the sample. For example, if the binary variable $I_i^{0.99}$, $i = 2, \dots, 10000$ takes value 0 approximately every 100 observations, we can expect only a few consecutive occurrences of two zeros in the sample, even in the presence of significant temporal dependence among $\{I_i^{0.99}\}_{i=2}^{10000}$.

Let us first consider the results for NAB. It can be seen that quite often, the null hypothesis are rejected at any conventional statistical level. Keep in mind however, that the number of observations is very high. On the other hand, the $\chi^2(1)$ -statistics are lower for GACD, log-GACD, and LACD and log-LACD models for most of the cases, although there are a few exceptions. On average, it seems that the dynamics of the (very) long durations is better captured than the dynamics of the (very) short durations. The independence $\chi^2(1)$ -statistics for intervals of the type $[0, dr(p)]$ are

Table 4
ACD model estimates for NAB.

Model	Coefficient estimates					
	α	β	γ	ν	κ	σ^a
ACD conditional mean specification and error distribution coefficients estimates ^b						
EACD	0.03644 (0.0033)***	0.0743 (0.0032)***	0.9097 (0.0042)***	—	—	—
WACD	0.0443 (0.0054)***	0.0766 (0.0045)***	0.9025 (0.0058)***	0.8984 (0.0039)***	—	—
GACD	0.1216 (0.0133)***	0.1067 (0.0073)***	0.8352 (0.0111)***	0.1830 (0.0199)***	22.0644 (4.8137)***	—
LACD	0.1530 (0.0152)***	0.1155 (0.0080)***	0.8198 (0.0118)***	—	—	1.3831 (0.0218)***
GV-LACD	0.0790 (0.0111)***	0.1051 (0.0074)***	0.8644 (0.0098)***	—	—	—
Log-EACD	0.0466 (0.0020)***	0.0621 (0.0022)***	0.9267 (0.0027)***	—	—	—
Log-WACD	0.0509 (0.0029)***	0.0652 (0.0031)***	0.9190 (0.0043)***	0.8969 (0.0048)***	—	—
Log-GACD	0.0937 (0.0075)***	0.0910 (0.0057)***	0.8551 (0.0106)***	0.1476 (0.0213)***	33.9771 (9.8641)***	—
Log-LACD	0.1067 (0.0085)***	0.0937 (0.0060)***	0.8450 (0.0115)***	—	—	1.3704 (0.0217)***
Log-GV-LACD	0.0989 (0.0078)***	0.0915 (0.0059)***	0.8544 (0.0107)***	—	—	—
Model	Coefficient estimates					
	α^b				β^b	γ^b
Coefficient estimates of the variance equation ^b						
GV-LACD		0.0411 (0.0100)***		0.0223 (0.0034)***		0.9476 (0.010)***
GV-Log-LACD		0.0180 (0.0092)***		0.0040 (0.0020)***		0.9828 (0.0078)***

As usual, coefficients significant at 10%, 5% and 1% significance level are marked with *, ** and ***, respectively.

^a This table contains the coefficient estimates for the autoregressive specifications for the time-varying variance of the error terms of the corresponding ACD models.

^b This table contains the coefficients estimates and standard errors for the autoregressive specifications for the corresponding ACD models, as well as the coefficients estimates for the error term distributions with the exception of the Garch volatility ACD models.

Table 5
ACD model estimates for WBC.

Model	Coefficient estimates					
	α	β	γ	ν	κ	σ^a
ACD conditional mean specification and error distribution coefficients estimates ^b						
EACD	0.1182 (0.0103)***	0.0874 (0.0047)***	0.8571 (0.0042)***	–	–	–
WACD	0.1225 (0.0137)***	0.0903 (0.0062)***	0.8514 (0.0107)***	0.8862 (0.0074)***	–	–
GACD	0.1632 (0.0207)***	0.1094 (0.0085)***	0.8174 (0.0150)***	0.2405 (0.0230)***	11.782 (2.8004)***	–
LACD	0.1928 (0.0245)***	0.1199 (0.010)***	0.8084 (0.0160)***	–	–	1.526 (0.0245)***
GV-LACD	0.1823 (0.0234)***	0.1287 (0.0102)***	0.8063 (0.0160)***	–	–	–
Log-EACD	0.12023 (0.0067)***	0.0893 (0.0040)***	0.8245 (0.0094)***	–	–	–
Log-WACD	0.1221 (0.0087)***	0.0921 (0.0053)***	0.8200 (0.0123)***	0.8866 (0.0074)***	–	–
Log-GACD	0.1431 (0.0114)***	0.1050 (0.0068)***	0.7958 (0.0106)***	0.2301 (0.02901)***	12.9123 (3.1941)***	–
Log-LACD	0.1627 (0.0137)***	0.1067 (0.0070)***	0.7900 (0.0167)***	–	–	1.5163 (0.0243)***
Log-GV-LACD	0.1576 (0.0131)***	0.1052 (0.0068)***	0.7963 (0.0160)***	–	–	–
Coefficient estimates of the variance equation ^a						
Model	Coefficient estimates					
	α^b	β^b			γ^b	
GV-LACD		0.2429 (0.0682)***		0.0352 (0.0070)***		0.8059 (0.0489)***
GV-Log-LACD		–		0.004 (0.0006)***		0.9961 (0.0006)***

As usual, coefficients significant at 10%, 5% and 1% significance level are marked with *, ** and ***, respectively.

^a This table contains the coefficient estimates for the autoregressive specifications for the time-varying variance of the error terms of the corresponding ACD models. Estimating the GV-Log-ACD model shows that the intercept coefficient in the conditional variance specification is not significant and the GV-Log-ACD model is re-estimated without an intercept in the conditional variance specification.

^b This table contains the coefficients estimates and standard errors for the autoregressive specifications for the corresponding ACD models, as well as the coefficients estimates for the error term distributions with the exception of the Garch volatility ACD models.

Table 6
ACD model estimates for TLS.

Model	Coefficient estimates					
	α	β	γ	ν	κ	σ^a
ACD conditional mean specification and error distribution coefficients estimates ^b						
EACD	0.1411 (0.0144)***	0.0823 (0.0048)***	0.8651 (0.0087)***	—	—	—
WACD	0.1408 (0.0170)***	0.0830 (0.0058)***	0.8642 (0.0103)***	0.9295 (0.0071)***	—	—
GACD	0.1538 (0.0213)***	0.0917 (0.0074)***	0.8522 (0.0128)***	0.4455 (0.0294)***	3.7945 (0.4777)***	—
LACD	0.2138 (0.0302)***	0.1115 (0.0094)***	0.8315 (0.0160)***	—	—	1.5000 (0.0244)***
GV-LACD	0.1967 (0.0287)***	0.1108 (0.010)***	0.8381 (0.0150)***	—	—	—
Log-EACD	0.0981 (0.0068)***	0.0659 (0.0034)***	0.8739 (0.0080)***	—	—	—
Log-WACD	0.0992 (0.0082)***	0.0671 (0.0041)***	0.8719 (0.010)***	0.9276 (0.0068)***	—	—
Log-GACD	0.1134 (0.0108)***	0.0765 (0.0055)***	0.8549 (0.0125)***	0.4280 (0.0282)***	4.0940 (0.5190)***	—
Log-LACD	0.1538 (0.0154)***	0.0850 (0.0063)***	0.8303 (0.0154)***	—	—	1.4934 (0.0242)***
Log-GV-LACD	0.1564 (0.0158)***	0.0836 (0.0062)***	0.8278 (0.0158)***	—	—	—
Model	Coefficient estimates					
	α^b				β^b	γ^b
Coefficient estimates of the variance equation ^a						
GV-LACD		0.3097 (0.0950)***		0.0300 (0.0064)***		0.7633 (0.0671)***
GV-Log-LACD		—		0.0017 (0.0003)***		0.9984 (0.0003)***

As usual, coefficients significant at 10%, 5% and 1% significance level are marked with *, ** and ***, respectively.

^a This table contains the coefficient estimates for the autoregressive specifications for the time-varying variance of the error terms of the corresponding ACD models. Estimating the GV-Log-ACD model shows that the intercept coefficient in the conditional variance specification is not significant and the GV-Log-ACD model is re-estimated without an intercept in the conditional variance specification.

^b This table contains the coefficients estimates and standard errors for the autoregressive specifications for the corresponding ACD models, as well as the coefficients estimates for the error term distributions with the exception of the Garch volatility ACD models.

Table 7
Statistics for the interval forecasts for NAB ^a.

Interval type	EACD			WACD			GACD			LACD			GV-LACD		
	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.
$(dl(0.99), \infty)$	0.9998	181.29	–	0.9998	181.29	–	0.9957	41.730	–	0.9922	5.2800	–	0.9934	13.261	
$(dl(0.95), \infty)$	0.9706	104.15	27.236	0.9908	521.71	44.454	0.9445	6.1703	9.4282	0.9381	27.831	14.410	0.9356	40.205	0.1717
$(dl(0.90), \infty)$	0.9001	0.0001	25.762	0.9368	170.76	21.738	0.8873	17.316	9.7089	0.8828	31.367	5.9234	0.8836	28.579	2.2045
$(dl(0.80), \infty)$	0.7654	71.950	31.119	0.8143	12.983	27.702	0.7912	4.8101	10.140	0.7931	2.9678	9.1187	0.7960	1.0051	8.1292
$(0, dr(0.80))$	0.8166	17.555	4.1464	0.8142	12.800	3.6403	0.7897	4.8101	0.6166	0.7873	9.9579	0.6879	0.7841	15.540	0.8877
$(0, dr(0.90))$	0.8939	4.0756	7.8800	0.8994	0.0413	5.5504	0.8952	2.5352	4.5077	0.8994	0.0413	4.2075	0.8993	0.0559	4.5044
$(0, dr(0.95))$	0.9359	38.609	3.0788	0.9439	7.5620	0.0119	0.9516	0.5411	0.5655	0.9588	17.280	4.8167	0.9571	11.112	1.2479
$(0, dr(0.99))$	0.9793	88.388	–	0.985	21.903	–	0.9949	29.551	–	0.9976	84.065	–	0.9970	68.241	–
$(dl(0.90), dr(0.90))$	0.7940	2.2485	14.718	0.8363	86.416	12.422	0.7826	18.573	8.7339	0.7823	19.212	6.6983	0.7829	17.946	1.4866
$(dl(0.80), dr(0.80))$	0.5820	13.497	1.5030	0.6286	34.283	5.1616	0.5810	15.031	0.8830	0.5805	15.829	0.7929	0.5801	16.481	0.1717
Interval type	Log-EACD			Log-WACD			Log-GACD			Log-LACD			GV-Log-LACD		
	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.
$(dl(0.99), \infty)$	0.9995	160.93	–	0.9995	160.93	–	0.9956	40.050	–	0.9942	20.980	–	0.9356	40.205	
$(dl(0.95), \infty)$	0.9589	17.688	14.976	0.9881	435.36	42.021	0.9489	0.2553	10.410	0.9475	1.3008	9.4770	0.9459	3.4596	5.5259
$(dl(0.90), \infty)$	0.8764	58.064	24.986	0.9249	74.675	24.007	0.8955	2.2307	11.514	0.8983	0.3233	9.8496	0.8938	4.2089	6.1700
$(dl(0.80), \infty)$	0.7317	270.69	27.201	0.7910	5.0296	28.860	0.8048	1.4368	13.310	0.8152	14.689	7.1061	0.8097	5.9298	7.5435
$(0, dr(0.80))$	0.8504	170.25	7.9390	0.8419	116.12	5.7408	0.7689	58.362	0.9416	0.7563	113.63	1.6130	0.7620	86.457	1.9923
$(0, dr(0.90))$	0.9210	52.375	8.1357	0.9206	50.327	9.0747	0.8809	38.483	0.7846	0.8774	53.386	0.6768	0.8807	39.272	1.0315
$(0, dr(0.95))$	0.9533	2.3351	0.7582	0.9567	9.8656	0.4490	0.9442	6.8490	0.0001	0.9463	2.8250	0.0562	0.9481	0.7551	0.0457
$(0, dr(0.99))$	0.9871	7.7886	–	0.9899	0.0103	–	0.9932	11.647	–	0.996	47.048	–	0.9963	52.812	
$(dl(0.90), dr(0.90))$	0.7975	0.3957	12.615	0.8456	138.32	17.374	0.7765	33.622	5.2010	0.7758	35.625	5.1444	0.7745	39.493	1.8158
$(dl(0.80), dr(0.80))$	0.5826	12.616	0.2541	0.6320	42.979	4.2920	0.5821	13.348	2.8510	0.5827	12.472	2.2026	0.5717	33.225	0.0315

^a The table presents the estimations for the corresponding Bernoulli parameters and the associated chi-squared statistics for unconditional coverage and independence. The critical values for the chi-squared distribution are 2.706, 3.841, 5.024, 6.635 and 10.828 with the associated p -values being 0.10, 0.05, 0.025, 0.01 and 0.001.

Table 8
Statistics for the interval forecasts for WBC ^a.

Interval type	EACD			WACD			GACD			LACD			GV-LACD		
	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.
$(dl(0.99), \infty)$	0.9998	181.30	–	0.9999	189.76	–	0.9957	41.735	–	0.9902	0.0403	–	0.9977	86.976	–
$(dl(0.95), \infty)$	0.9569	10.479	2.2051	0.9857	369.17	0.4160	0.9426	11.042	1.2348	0.9343	47.464	0.2283	0.9355	40.743	2.1422
$(dl(0.90), \infty)$	0.8907	9.3788	11.180	0.9256	79.135	0.9410	0.8949	2.8588	7.0881	0.8909	8.9842	9.0896	0.8918	7.3151	0.6782
$(dl(0.80), \infty)$	0.7501	147.19	42.383	0.8049	1.4978	23.627	0.7900	6.1991	20.586	0.7949	1.6282	20.369	0.7950	1.5654	7.7749
$(0, dr(0.80))$	0.8080	4.0210	4.3129	0.8071	3.1615	23.627	0.7852	13.484	4.1626	0.7831	17.533	1.1999	0.7825	18.785	0.9989
$(0, dr(0.90))$	0.8890	13.054	1.5504	0.8954	2.3301	3.5796	0.8948	2.9709	4.5524	0.9021	0.4884	7.1282	0.9024	0.6393	8.1288
$(0, dr(0.95))$	0.9328	56.524	0.4401	0.9412	15.488	1.4763	0.9525	1.3319	0.6372	0.9627	37.066	0.0612	0.9625	35.850	0.3316
$(0, dr(0.99))$	0.9750	160.46	–	0.9835	35.697	–	0.9934	13.262	–	0.9979	93.066	–	0.9977	86.976	–
$dl(0.90dr(0.90))$	0.7797	25.187	0.0139	0.8211	28.552	1.6521	0.7898	6.4475	0.0229	0.7931	2.9678	0.0114	0.8162	2.1021	5.2740
$dl(0.80dr(0.80))$	0.5581	72.544	3.9049	0.6121	6.0818	2.4017	0.5753	25.342	2.6512	0.5781	19.944	2.3623	0.5775	21.047	0.0726
Interval type	Log-EACD			Log-WACD			Log-GACD			Log-LACD			GV-Log-LACD		
	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.
$(dl(0.99), \infty)$	0.9997	173.89	–	0.9999	189.76	–	0.9961	48.917	–	0.9900	0.0001	–	0.9975	81.237	–
$(dl(0.95), \infty)$	0.9571	11.111	0.7144	0.9867	395.67	0.0334	0.9429	10.183	2.1241	0.9346	45.740	1.6867	0.9349	44.045	1.9909
$(dl(0.90), \infty)$	0.8924	6.2956	2.1886	0.9268	87.113	0.6658	0.8954	2.3301	0.5160	0.8918	7.3151	0.5371	0.8908	9.1800	0.6392
$(dl(0.80), \infty)$	0.7492	152.41	14.210	0.8057	2.0311	4.8585	0.7898	6.4475	3.3352	0.7954	1.3265	2.1032	0.7935	2.6359	3.9300
$(0, dr(0.80))$	0.8083	4.3307	5.7885	0.8071	3.1615	5.2578	0.7867	10.912	3.3352	0.7840	15.734	2.4373	0.7832	17.328	4.2211
$(0, dr(0.90))$	0.8883	14.738	0.0010	0.8971	0.9330	1.1571	0.8946	3.2016	2.2861	0.9032	1.1417	0.7782	0.9030	1.0024	0.4797
$(0, dr(0.95))$	0.9310	68.341	0.0589	0.9396	21.436	0.0758	0.9516	0.5411	0.5915	0.9608	26.417	0.8890	0.9607	25.911	0.4497
$(0, dr(0.99))$	0.9753	154.91	–	0.9831	39.854	–	0.9943	22.096	–	0.9975	81.237	–	0.9975	81.237	–
$dl(0.90dr(0.90))$	0.7808	22.563	0.6526	0.8240	37.097	0.5253	0.7901	6.0767	0.0057	0.7951	1.5039	0.0083	0.7938	2.3998	0.0049
$dl(0.80dr(0.80))$	0.5576	74.276	5.0278	0.6129	6.9171	1.5990	0.5766	22.756	1.3927	0.5795	17.486	1.2536	0.5767	22.563	2.2540

^a The table presents the estimations for the corresponding Bernoulli parameters and the associated chi-squared statistics for unconditional coverage and independence. The critical values for the chi-squared distribution are 2.706, 3.841, 5.024, 6.635 and 10.828 with the associated p -values being 0.10, 0.05, 0.025, 0.01 and 0.001.

Table 9
Statistics for the interval forecasts for TLS ^a.

Interval type	EACD			WACD			GACD			LACD			GV-LACD		
	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.
$(dl(0.99), \infty)$	1.0000	–	–	0.9999	189.76	–	0.9984	110.05	–	0.9861	13.708	–	0.9729	201.35	
$(dl(0.95), \infty)$	0.9542	3.8077	9.9781	0.9718	117.92	20.576	0.9397	21.038	6.1241	0.9252	113.17	16.772	0.9086	293.10	1.0294
$(dl(0.90), \infty)$	0.8897	11.469	22.298	0.9110	13.885	19.639	0.8885	14.247	22.333	0.8831	30.307	18.875	0.8576	179.08	3.0211
$(dl(0.80), \infty)$	0.7651	73.179	24.272	0.7952	1.4435	22.753	0.7872	10.114	18.756	0.7967	0.6861	18.681	0.7667	66.743	13.807
$(0, dr(0.80))$	0.8074	3.4364	1.4750	0.8071	3.1615	1.3270	0.7917	4.2828	3.1062	0.7873	9.9579	2.5188	0.7995	0.0169	2.0226
$(0, dr(0.90))$	0.8954	2.3301	2.1270	0.9012	0.1579	2.4040	0.9010	0.1092	3.7863	0.9169	33.438	3.5835	0.9182	38.955	3.1369
$(0, dr(0.95))$	0.9397	21.0374	3.6746	0.9450	5.1154	2.0771	0.9524	3.1738	0.6372	0.9730	132.74	0.2464	0.9722	122.74	0.4422
$(0, dr(0.99))$	0.9811	63.447	–	0.9850	21.903	–	0.9927	8.1204	–	0.9986	117.68	–	0.9985	113.80	
$(dl(0.90), dr(0.90))$	0.7851	13.665	0.7150	0.8123	9.5755	0.6150	0.7896	6.7007	0.3363	0.8001	0.0001	0.2323	0.7758	35.625	1.2412
$(dl(0.80), dr(0.80))$	0.5725	31.383	2.4010	0.6024	0.2322	1.8540	0.5790	18.345	1.3266	0.5841	10.543	1.8179	0.5775	0.5662	0.8517
Interval type	Log-EACD			Log-WACD			Log-GACD			Log-LACD			GV-Log-LACD		
	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point stat.	Ind stat.	Est. value	Point Stat.	Ind stat.	Est. value	Point stat.	Ind stat.
$(dl(0.99), \infty)$	0.9999	189.76	–	0.9999	189.76	–	0.9981	99.540	–	0.9856	17.222	–	0.9975	81.237	
$(dl(0.95), \infty)$	0.9533	11.1107	0.7144	0.9733	136.61	5.5789	0.9388	24.750	0.3960	0.9267	100.62	2.2350	0.9349	44.045	1.9909
$(dl(0.90), \infty)$	0.8908	9.1800	5.3223	0.9127	18.615	6.4670	0.8901	10.608	4.1878	0.8835	28.921	4.2017	0.8908	9.1800	0.6392
$(dl(0.80), \infty)$	0.7680	61.725	16.957	0.7951	1.5039	10.550	0.7865	11.239	10.716	0.7951	1.5039	5.2150	0.7935	2.6359	3.9300
$(0, dr(0.80))$	0.8081	4.1230	6.4789	0.8081	4.1230	6.8048	0.7910	5.0296	6.1452	0.7874	15.734	5.1167	0.7832	17.328	4.2211
$(0, dr(0.90))$	0.8941	3.8153	2.4093	0.8999	0.0013	2.2949	0.8999	0.0013	2.2949	0.9148	25.460	0.0038	0.9030	1.0024	0.4797
$(0, dr(0.95))$	0.9404	18.348	0.0738	0.9466	2.3903	0.0906	0.9535	2.6307	0.0173	0.9718	117.92	0.0001	0.9607	25.911	0.4497
$(0, dr(0.99))$	0.9815	58.368	–	0.9856	17.222	–	0.9927	8.1204	–	0.9982	102.93	–	0.9975	81.237	
$(dl(0.90), dr(0.90))$	0.7850	13.847	2.5517	0.8127	10.215	1.9715	0.7901	6.0767	2.5761	0.7984	0.1637	1.5154	0.7938	2.3998	0.0049
$(dl(0.80), dr(0.80))$	0.5762	23.537	6.7427	0.6033	0.4433	6.4401	0.5776	20.861	4.7653	0.5826	12.616	4.7942	0.5767	22.563	2.2540

^a The table presents the estimations for the corresponding Bernoulli parameters and the associated chi-squared statistics for unconditional coverage and independence. The critical values for the chi-squared distribution are 2.706, 3.841, 5.024, 6.635 and 10.828 with the associated p -values being 0.10, 0.05, 0.025, 0.01 and 0.001.

substantially smaller than the corresponding statistics for intervals of the type $[dl(p), \infty]$, $p = 0.95, 0.90, 0.80$, i.e. the long durations are better captured by the long durations. Also, the corresponding $\chi^2(1)$ -statistics for unconditional coverage exhibit similar pattern, although there are three exceptions: Log-GACD, Log-LACD, and GV-Log-LACD models. Note, that for the exponential and Weibull distribution, this was suggested by the left-hand bump of the z -histogram of the probability integral transform.

Comparison of the LACD and GV-LACD models as well as the Log-LACD and GV-Log-LACD models, reveals that the GARCH-type specification can capture some of the autocorrelation in the interval forecasts as evidenced by the much lower values for the independence statistics, which quite often indicate insignificance at 5–10% confidence levels. This observation is contrary to the unchanged and even increased autocorrelation of the probability integral transform for the GV-type of models. Finally, the results suggest that there is no advantage of choosing an Log-ACD model vs. an ACD model with standard specification of the conditional mean.

Analysis of the interval forecasts for WBC shows that they are qualitatively similar to the interval forecasts for NAB. The models with exponential and Weibull distribution have the worst performance. The ACD models with generalized gamma and log-normal distribution as well as models with the Log-ACD and the standard ACD specification for the mean, have similar performance. With regards to the GV-LACD and GV-Log-ACD models, we don't find evidence for improvement over the LACD and Log-LACD models, on contrary they show slightly worse performance, indicated by the higher values for the corresponding statistics. Interestingly, as in the case of NAB, this is in contrary with the evidence from the autocorrelations of the probability integral transforms, which decrease after we assume a time-varying form for the variance of the residuals.

For TLS, as for the case of NAB and WBD, the models with exponential and Weibull distribution have the worst performance and the models with generalized gamma and log-normal seem to perform much better. Interestingly, by looking at Fig. 3, the probability integral transform for the LACD and Log-LACD models seems as “bad” an approximation to the uniform distribution as the probability integral transforms for the corresponding models with exponential and Weibull distribution. However, the models with log-normal distribution for the residuals on average have worse $\chi^2(1)$ -statistics for unconditional coverage than their counterparts with generalized gamma distribution. However, the $\chi^2(1)$ -statistics for independence are pretty similar for both types of models.

Allowing the variance of the residuals to be time-varying shows that substantial decrease in the $\chi^2(1)$ -statistics for independence, which is most pronounced for the GV-Log-LACD model. Indeed, the GV-LACD and GV-Log-LACD models have the lowest independence $\chi^2(1)$ -statistics. Also, the statistics for the unconditional average decrease on average. Again, this is in contradiction with the evidence implied by the statistics for the probability integral transforms. As we mentioned before, the z -histograms for the GV-LACD and GV-Log-LACD seem to be worse than their LACD and Log-LACD counterparts, and the difference is very clearly seen for the histogram of the GV-Log-LACD model. Additionally, the first order autocorrelation for the probability integral transform either remains the same (GV-LACD model) or increases (GV-Log-LACD model) when we specify a Garch-type variance for the residuals.

In a nutshell, the results of this section are as follows. Evidence from analyzing both probability integral transforms and interval forecasts suggests that models with generalized gamma and log-normal residual distribution are a better than models with exponential and Weibull distribution. Using Log-ACD specification for the conditional mean does not make much difference. The choice of log-normal residual distribution gives pretty close performance to the generalized gamma distribution, even though it has one free parameter less. Finally, there seems to be mixed evidence for time-variation in the variance of the residuals.

6. Conclusion

In this paper, we analyze different approaches to modelling the durations of three major Australian companies traded on the Australian stock exchange. More specifically, we consider four different types of distributions for the residuals, namely the exponential, Weibull, generalized gamma and log-normal distributions, as well as GARCH-type time-varying specification for the variance of the log-normally distributed residuals. Additionally, we assume alternative specification for the conditional mean (the Log-ACD approach of Bauwens and Giot [3]).

Estimations of the alternative models show that the coefficient estimates of models with exponential and the Weibull distribution for the residuals are close to each other. Additionally, the estimated shape parameter of the Weibull distribution κ is close to unity, the value at which the Weibull distribution collapses to the exponential one. For NAB and WBC, the models with generalized gamma and the log-normal distribution are close to each other, but for TLS they seem to be substantially different. These preliminary findings suggest that the models with generalized gamma and log-normal distribution for the residuals perform similarly and they jointly differ from the models with exponential and the Weibull distribution, which is confirmed later in the paper.

To compare the suitability of the alternative ACD specifications, we employ the method of evaluating density forecasts proposed by Diebold et al. [9] and the methodology for evaluating interval forecasts by Christoffersen [7]. The results strongly suggest that the generalized gamma and the log-normal distribution perform better than the exponential and the Weibull distribution. However, given the fact that the log-normal distribution has only one free parameter compared to the two free parameters of the generalized gamma distribution, it might very well be the preferred choice. Additionally, a promising unexplored topic would be the investigation of different generalizations of the log-normal distribution such as the Inverse Gaussian and Student t -distribution.

There seems to be no gain or loss in choosing between a standard ACD specification for the conditional mean and the Log-ACD specification of Bauwens and Giot [3]. Finally, the results show mixed evidence for time-variation in the shape of the residual distributions as evidenced by the result for the GARCH Volatility ACD models.

Acknowledgments

We are grateful to Australian Research Council for financial support. The financial data has been graciously provided by Securities Research Institute (SIRCA) which is our industry partner.

References

- [1] Y. Ajt-Sahalia, Testing continuous-time models of the spot interest rate, *Review of Financial Studies* 9 (1996) 385–426.
- [2] D. Allen, F. Chan, M. McAleer, S. Peiris, Finite sample properties of the QMLE for the Log-ACD Model: Application to Australian Stocks Working Paper, Edith Cowan University, Australia, 2005.
- [3] L. Bauwens, P. Giot, The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks, *Annales d'Economie et de Statistique* 60 (2000) 117–150.
- [4] L. Bauwens, P. Giot, J. Grammig, D. Veredas, A comparison of financial duration models via density forecasts, *International Journal of Forecasting* 20 (2004) 589–609.
- [5] L. Bauwens, D. Veredas, The stochastic conditional duration model: a latent factor model for the analysis of financial durations, *Journal of Econometrics* 119 (1999) 381–412.
- [6] L. Bauwens, F. Galli, P. Giot, The Moments of ACD Models, Core Discussion Paper, (2003).
- [7] P. Christoffersen, Evaluating interval forecasts, *International Economic Review* 39 (1998) 841–862.
- [8] F. Corsi, S. Mittnik, C. Pigorsch, U. Pigorsch, The volatility of realized volatility, *Econometric Reviews* 27 (2008) 46–78.
- [9] F. Diebold, A. Gunther, S. Tay, Evaluating density forecasts with applications to financial risk management, *International Economic Review* 39 (1998) 863–883.
- [10] R. Engle, The econometrics of ultra-high frequency data, *Econometrica* 68 (2000) 1–22.
- [11] R. Engle, J. Russell, Autoregressive conditional duration: a new model for irregularly-spaced transaction data, *Econometrica* 66 (1998) 1127–1162.
- [12] D. Feng, G. Jiang, P. Song, Stochastic conditional duration models with “leverage effect for financial transaction data, *Journal of financial econometrics* 3 (2004) 390–421.
- [13] M. Fernandes, J. Grammig, Non-Parametric specification tests for conditional duration models, *Journal of Econometrics* 127 (2005) 35–68.
- [14] E. Ghysels, C. Gouriéroux, J. Jasiak, Stochastic volatility duration models, CREST Working Paper 9746, INSEE, 1997.
- [15] E. Ghysels, J. Jasiak, GARCH for irregularly spaced financial data: the ACD-GARCH model, *Studies in Nonlinear Dynamics and Econometrics* 2 (1998) 133–149.
- [16] J. Grammig, K. Maurer, Non-monotonic hazard functions and the autoregressive conditional duration model, *Econometrics Journal* 3 (2000) 16–38.
- [17] J. Grammig, M. Wellner, Modeling the interdependence of volatility and inter-transaction duration processes, *Journal of Econometrics* 106 (2002) 369–400.
- [18] P.G. Hoel, A test for markov chains, *Biometrika* 41 (1954) 430–433.
- [19] R. Kissel, M. Glantz, Optimal Trading Strategies Quantitative Approaches for Managing Market Impact and Trading Risk, Amacom, USA, 2003.
- [20] A. Lunde, A generalized gamma autoregressive conditional duration model, Working Paper, Aalborg University, 2000.
- [21] A. Madhavan, Market microstructure: a survey, *Journal of Financial Markets* 3 (2000) 205–258.

- [22] D. Nelson, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* 59 (1991) 347–370.
- [23] M. Rosenblatt, Remarks on a multivariate transformation, *Annals of Mathematical Statistics* 23 (1952) 470–472.
- [24] W. Sun, S. Rachev, F. Fabozzi, P. Kalem, Long-range dependence and heavy tailedness in modelling trade duration, Working Paper, University of Karlsruhe, 2005.
- [25] M.Y. Zhang, J. Russell, R.S. Tsay, A Non-linear autoregressive-conditional duration model with applications to financial transactions data, *Journal of Econometrics* 104 (2001) 179–207.