Birnbaum-Saunders autoregressive conditional duration models applied to high-frequency financial data

Helton Saulo[1,2], **Jeremias Leão**[3], **Víctor Leiva**[4,5]*, **Robert G. Aykroyd**[6]

[1]Institute of Mathematics and Statistics, Universidade Federal de Goiás, Brazil
[2]Department of Statistics, Universidade de Brasília, Brazil
[3]Department of Statistics, Universidade Federal do Amazonas, Brazil
[4]Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Chile
[5]Faculty of Administration, Accounting and Economics, Universidade Federal de Goiás, Brazil
[6]Department of Statistics, University of Leeds, UK

Abstract

Modern financial markets now record the precise time of each stock trade, along with price and volume, with the aim of analysing the structure of the times between trading events – leading to a big data problem. In this paper, we propose and compare two Birnbaum-Saunders autoregressive conditional duration models specified in terms of time-varying conditional median and mean durations. These models provide a novel alternative to the existing autoregressive conditional duration models due to their flexibility and ease of estimation. Influence diagnostic tools are developed to allow goodness-of-fit assessment and to detect departures from assumptions, including the presence of outliers and influential cases. Both global and local influence tools are considered based on the parameter estimates under different perturbation schemes. A thorough Monte Carlo study is presented to evaluate the performance of the maximum likelihood estimators, and the forecasting ability of the models is assessed using the traditional and density forecast evaluation techniques. The Monte Carlo study suggests that the parameter estimators are asymptotically unbiased, consistent and normally distributed. Finally, a full analysis of a real-world financial transaction data set, from the German DAX in 2016, is presented to illustrate the proposed approach and to compare the fitting and forecasting performances with existing models in the literature. One case related to the duration time is identified as potentially influential, but its removal does not change resulting inferences demonstrating the robustness of the proposed approach. Fitting and forecasting performances favor the proposed models and, in particular, the median-based approach gives additional protection against outliers, as expected.

Keywords: Big data; Birnbaum-Saunders distribution; forecasting ability; influence diagnostics; likelihood-based methods; Monte Carlo simulation; R software.

*Corresponding author. Victor Leiva. E-mail address: victorleivasanchez@gmail.com. URL: www.victorleiva.cl.

1 Introduction

The family of autoregressive conditional duration (ACD) models proposed by Engle and Russell (1998) has been the primary tool to deal with high frequency financial data on transactions, leading to a big data problem. The ACD model is the counterpart of GARCH models for dealing with trade duration (TD) data and it is used to capture the clustering structure, which conveys meaningful information, observed in high frequency financial data; see Duchesne and Pacurar (2008), Liu and Heyde (2008) and Pacurar (2008). TD data possess a number of unique characteristics such as: an irregular nature in which they are collected; a large number of observations or cases; a diurnal pattern, where activity is higher at the beginning and closing than in the middle of the trading day; asymmetry and an inverse bathtub shaped hazard rate (HR); see Bhatti (2010) and Leiva et al. (2014b).

Some extensions of the original ACD model have been proposed in the literature; see, for example, Grammig and Maurer (2000), Bauwens and Giot (2000), Meitz and Terasvirta (2006), Chiang (2007), Pacurar (2008), Bhatti (2010), and Leiva et al. (2014b). These versions take into account the following aspects: (A1) the shape of the HR of TD data; (A2) the conditional dynamics established in terms of mean or median; (A3) the linear form of the conditional mean or median dynamics; and (A4) the time series properties. Some recent applications of ACD models are discussed in Diana (2015) and Dionne et al. (2015).

This paper focuses on the Birnbaum-Saunders ACD (BSACD) model; see Bhatti (2010). It is based on a skew distribution, which has an HR with inverse bathtub shape; see Birnbaum and Saunders (1969) and Kundu et al. (2008). The BS distribution originates from material fatigue and has interesting properties, doing it widely studied. Some of its recent applications range across fields different to engineering, such as business, environment, finance, industry and medicine, which have been conducted by an international, transdisciplinary group of researchers; see Jin and Kawczak (2003), Bhatti (2010), Lio et al. (2010), Castillo et al. (2011), Saulo et al. (2013), Leiva et al. (2015, 2016b, 2017), Wanke and Leiva (2015), Garcia-Papani et al. (2016), Marchant et al. (2016b), and Leao et al. (2017). The BSACD model proposed by Bhatti (2010) is constructed in terms of a conditional median duration, rather than an ACD model in the sense of Engle and Russell (1998) based on the mean. However, the BSACD model provides: (B1) a realistic assumption for ACD data in terms of the shape of both the probability density function (PDF) and HR of the BS distribution; (B2) a natural parameterization in terms of a conditional median duration instead of the mean, since the scale parameter of the BS distribution is also its median, see (A3); and (B3) an easy parameter estimation due to fast convergence and obtainment of initial values from the modified method of moments.

Bhatti (2010) suggested that (B2) might possibly: (C1) improve the model fit, since for asymmetric, heavy-tailed distributions, as occurs with TD data, the median is often considered as a better measure of central tendency than the mean; and (C2) increase the forecasting ability due to the fact that the mean is greater than the median for skew distributions. In this context, we consider two models: A first new mean-based model (BSACD1 in short) specified in terms of a time-varying conditional mean duration, as usual in ACD models, using a reparameterized version of the BS distribution (see Leiva et al., 2014a; Santos-Neto et al., 2016); and a second median-based model (BSACD2 in short) specified in terms of a time-varying conditional median duration. Thus, the primary objective of this paper is to compare both BSACD1 and BSACD2 models. The secondary objectives are: (i) to obtain the maximum likelihood (ML) estimators of the BSACD1 and BSACD2 model parameters and to evaluate their performance by a Monte Carlo (MC) simulation study; (ii) to derive influence diagnostic tools for the BSACD1 and BSACD2 models and to assess the robustness of each model to atypical cases (see Liu, 2000); (iii) to fit the BSACD1 and BSACD2 models to a real-world data set for evaluating (C1); (iv) to establish the forecasting ability of the BSACD1 and BSACD2 models for detecting (C2); and (v) to compare the fitting and forecasting performances of the BSACD1 and BSACD2 models with existing models in the literature. Note that the BSACD2 model focusses on the median of TDs, which is not the typical interested parameter considered in the literature. The median duration is able to communicate to investors and play an important role in the stock market, because it provides a more robust alternative to the usual mean, and hence it can be interpreted and used in exactly the same way. However, this robustness property means that the median is not affected by extremes or outliers – this is particularly important for skew data such as TDs; see Section 4.1 for a discussion about outlier detection framework in times series and ACD models. This means that future predictions will not depend significantly on previous freak events – a sharp jump in the market last week does not mean a similar jump every week.

The rest of the paper proceeds as follows. In Section 2, we describe the BS distribution and its mean-based reparameterized version. In Section 3, we introduce the BSACD1 and BSACD2 models and derive ML-based estimation and inference for their parameters. In Section 4, we derive global and local influence tools and calculate the normal curvatures of local influence under three different perturbation schemes. Moreover, we consider two types of residuals for the BSACD models. In Section 5, we carry out an MC simulation study to evaluate the behavior of ML-based inference. In Section 6, we apply the ACD models and derived tools to a real-world financial data set. Finally, in Section 7, we discuss conclusions and future research on the topic of this work.

2 BS distributions

2.1 The BS distribution

A random variable X is BS distributed if it can be represented by the transformation of a standard normal random variable, $Z \sim \mathrm{N}(0,1)$ say, given by

$$X = \sigma \left[\kappa Z/2 + \left\{(\kappa Z/2)^2 + 1\right\}^{1/2}\right]^2, \tag{1}$$

where $\kappa > 0$ and $\sigma > 0$ are shape and scale parameters, respectively. In this case, the notation $X \sim \mathrm{BS}(\kappa, \sigma)$ is used and the corresponding PDF is obtained as

$$f(x;\kappa,\sigma) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\kappa^2}\left[\frac{x}{\sigma} + \frac{\sigma}{x} - 2\right]\right) \frac{x^{-3/2}[x+\sigma]}{2\kappa\sigma^{1/2}}, \quad x > 0. \tag{2}$$

Then, it can be shown that $\mathrm{E}[X] = \sigma[1 + \kappa^2/2]$, $\mathrm{Var}[X] = [\kappa\sigma]^2[1 + 5\kappa^2/4]$ and the median of the distribution is the scale parameter σ. Note that the BS distribution is closed under scale and reciprocal transformations, that is, $b\,X \sim \mathrm{BS}(\kappa, b\,\sigma)$, with $b > 0$, and $1/X \sim \mathrm{BS}(\kappa, 1/\sigma)$, respectively. The survival function (SF) and HR of X are expressed as

$$S(x;\kappa,\sigma) = \Phi\left(-\frac{1}{\kappa}\left[\sqrt{\frac{x}{\sigma}} - \sqrt{\frac{\sigma}{x}}\right]\right), \quad h(x;\kappa,\sigma) = \frac{\phi\left(\frac{1}{\kappa}\left[\sqrt{\frac{x}{\sigma}} - \sqrt{\frac{\sigma}{x}}\right]\right) x^{-3/2}[x+\sigma]}{2\kappa\sigma^{1/2}\Phi\left(-\frac{1}{\kappa}\left[\sqrt{\frac{x}{\sigma}} - \sqrt{\frac{\sigma}{x}}\right]\right)}, \quad x > 0,$$

where ϕ and Φ are the standard normal PDF and cumulative distribution function (CDF), respectively.

2.2 A reparameterized BS distribution

Consider a reparameterized version of the BS (RBS) distribution by setting $\kappa = \sqrt{2/\tau}$ and $\sigma = \tau\mu/[\tau+1]$, such that $\tau = 2/\kappa^2$ and $\mu = \sigma[1+\kappa^2/2]$, where $\tau > 0$ is a shape and precision parameter and $\mu > 0$ is a scale parameter and the mean of the distribution; see Leiva et al. (2014a) and Santos-Neto et al. (2016). In this case, the PDF of $X \sim \mathrm{RBS}(\mu,\tau)$ is given by

$$f(x;\mu,\tau) = \frac{\exp(\tau/2)\sqrt{\tau+1}}{4\sqrt{\pi\mu}x^{3/2}}\left[x+\frac{\tau\mu}{\tau+1}\right]\exp\left(-\frac{\tau}{4}\left[\frac{x\{\tau+1\}}{\tau\mu}+\frac{\tau\mu}{x\{\tau+1\}}\right]\right),\quad x>0, \tag{3}$$

and the RBS and standard normal random variables are related by the transformation

$$X = \left[\frac{\tau\mu}{\tau+1}\right]\left[Z/\sqrt{2\tau}+\sqrt{\left\{Z/\sqrt{2\tau}\right\}^2+1}\right]^2. \tag{4}$$

Note from equation (3) that the mean and variance of X are now $\mathrm{E}[X] = \mu$ and $\mathrm{Var}[X] = \mu^2[2\tau+5]/[\tau+1]^2$, respectively. In addition, the SF and HR of X are obtained as

$$\begin{aligned}
S(x;\mu,\tau) &= \Phi\left(-\sqrt{\frac{\tau}{2}}\left[\sqrt{\frac{\{\tau+1\}x}{\mu\tau}}-\sqrt{\frac{\mu\tau}{\{\tau+1\}x}}\right]\right),\quad x>0,\\
h(x;\mu,\tau) &= \frac{\sqrt{\tau+1}}{4\sqrt{\pi\mu}x^{3/2}}\left[x+\frac{\tau\mu}{\tau+1}\right]\frac{\exp\left(-\frac{\tau}{4}\left[\frac{x\{\tau+1\}}{\tau\mu}+\frac{\tau\mu}{x\{\tau+1\}}-2\right]\right)}{\Phi\left(-\sqrt{\frac{\tau}{2}}\left[\sqrt{\frac{\{\tau+1\}x}{\mu\tau}}-\sqrt{\frac{\mu\tau}{\{\tau+1\}x}}\right]\right)},\quad x>0.
\end{aligned}$$

3 Birnbaum-Saunders ACD models

3.1 Mean-based BSACD model (BSACD1)

Suppose that data collection starts at a time T_0 and a sequence of successive times $T_1,\ldots,T_n$ at which market events, or trades, occur is recorded. Then, let $X_i = T_i - T_{i-1}$ be a duration time, that is, the time elapsed between two successive occurrence times, T_{i-1} and T_i, for $i = 1,\ldots,n$.

The ACD model is then specified, in terms of these duration times, by a conditional mean duration $\mathrm{E}[X_i|\Omega_{i-1}] = \mu_i$, where μ_i is the RBS mean introduced in Section 2.2. Note that Ω_{i-1} is a set which includes all information available until time T_{i-1}. In this case, the BSACD1 model can be defined by

$$X_i = \mu_i\,\varepsilon_i,\quad i=1,\ldots,n, \tag{5}$$

where ε_i are independent and identically distributed (IID) random variables following the RBS distribution with mean equal to one and precision τ, denoted by $\varepsilon_i \overset{\mathrm{IID}}{\sim} \mathrm{RBS}(1,\tau)$, and then the X_is are independent (IND) not identically distributed, that is, $X_i \overset{\mathrm{IND}}{\sim} \mathrm{RBS}(\mu_i,\tau)$. In addition, autoregressive (AR) and moving average (MA) processes, of order p_1 and q_1 respectively, can be defined for the model given in equation (5) through the representation

$$\log(\mu_i) = \varpi + \sum_{j=1}^{p_1} \alpha_j \log(\mu_{i-j}) + \sum_{j=1}^{q_1} \frac{\beta_j X_{i-j}}{\mu_{i-j}},$$

leading to the notation BSACD1(p_1, q_1), as usual in ARMA models.

By using the model given in equation (5) and the PDF expressed in equation (3), the log-likelihood function for $\boldsymbol{\theta} = [\varpi, \alpha_1, \ldots, \alpha_{p_1}, \beta_1, \ldots, \beta_{q_1}, \tau]^\top$ is obtained as

$$\ell(\boldsymbol{\theta}) = \frac{n\tau}{2} - \frac{n\log(16\pi)}{2} - \frac{1}{2}\sum_{i=1}^{n} \log\left(\frac{[\tau+1]x_i^3\mu_i}{[\tau x_i + x_i + \tau\mu_i]^2}\right) - \sum_{i=1}^{n}\frac{x_i[\tau+1]}{4\mu_i} - \sum_{i=1}^{n}\frac{\tau^2\mu_i}{4[\tau+1]x_i}. \quad (6)$$

Estimation, inference and local influence for $\boldsymbol{\theta}$ can be based on the log-likelihood function given in equation (6). As usual, to obtain the ML estimates of the model parameter $\boldsymbol{\theta}$, one must maximize expression defined in (6) by equating the score vector $\dot{\boldsymbol{\ell}}(\boldsymbol{\theta})$, which contains the first derivatives of $\ell(\boldsymbol{\theta})$, to zero, providing the likelihood equations. They must be solved by an iterative procedure for non-linear optimization, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method, which is regarded as the best-performing algorithm; see Mittelhammer et al. (2000, p. 199) and Leiva et al. (2014b). The BFGS method is implemented in the `R` software (see R-Team, 2016) available at `http://cran.r-project.org`, by the functions `optim` and `optimx`. Inference for $\boldsymbol{\theta}$ of the BSACD1 model can be based on the asymptotic distribution of the ML estimator $\widehat{\boldsymbol{\theta}}$. This estimator is consistent and has an asymptotic multivariate normal joint distribution with mean $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$, which may be obtained from the corresponding expected Fisher information matrix $\mathcal{I}(\boldsymbol{\theta})$. Thus, we have that

$$\sqrt{n}\,[\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}] \xrightarrow{\mathcal{D}} \mathrm{N}_{p_1+q_1+2}(\mathbf{0}, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} = \mathcal{J}(\boldsymbol{\theta})^{-1}),$$

as $n \to \infty$, where $\xrightarrow{\mathcal{D}}$ means "convergence in distribution" and $\mathcal{J}(\boldsymbol{\theta}) = \lim_{n\to\infty}[1/n]\mathcal{I}(\boldsymbol{\theta})$. Note that $\widehat{\mathcal{I}}(\boldsymbol{\theta})^{-1}$ is a consistent estimator of the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$. In practice, one may approximate the expected Fisher information matrix by its observed version obtained from the Hessian matrix $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$, which contains the second derivatives of $\ell(\boldsymbol{\theta})$ given in equation (6), whereas the diagonal elements of its inverse matrix can be used to approximate the corresponding standard errors (SEs); see Efron and Hinkley (1978) for details about the use of observed versus expected Fisher information matrices. In addition, local influence diagnostics can also be considered also using the log-likelihood function; see, for example, Cook (1987) and Liu (2000).

3.2 Median-based BSACD model (BSACD2)

Now, we specify an ACD model in terms of a conditional median duration, $\sigma_i = F_{\text{BS}}^{-1}(0.5|\Omega_{i-1})$ say, where F_{BS}^{-1} is the inverse function of the CDF, or quantile function (QF), of the BS distribution introduced in Section 2.1, with Ω_{i-1} analogously defined as in Section 3.1. In this case, the BSACD2 model can be formulated as

$$X_i = \sigma_i\, \varrho_i, \quad i = 1, \ldots, n, \quad (7)$$

where ϱ_i are IID random variables following the BS distribution with shape κ and scale (median) equal to one, denoted by $\varrho_i \stackrel{\text{IID}}{\sim} \text{BS}(\kappa, 1)$, and then $X_i \stackrel{\text{IND}}{\sim} \text{BS}(\kappa, \sigma_i)$. In addition, an ARMA$(p_2, q_2)$ process can be defined for the model given in equation (7) by

$$\log(\sigma_i) = \varsigma + \sum_{j=1}^{p_2} \gamma_j \log(\sigma_{i-j}) + \sum_{j=1}^{q_2} \frac{\delta_j X_{i-j}}{\sigma_{i-j}},$$

inducting the notation BSACD2(p_2, q_2). Upon the model defined in equation (7) and considering the PDF given in equation (2), the log-likelihood function for $\boldsymbol{\xi} = [\varsigma, \gamma_1, \ldots, \gamma_{p_2}, \delta_1, \ldots, \delta_{q_2}, \kappa]^\top$ is expressed as

$$\begin{aligned} \ell(\boldsymbol{\xi}) \quad = \quad & -n\log(\sqrt{2\pi}) - n\log(2) - \frac{3}{2}\sum_{i=1}^{n}\log(x_i) - \frac{1}{2\kappa^2}\sum_{i=1}^{n}\left[\frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2\right] \\ & + \sum_{i=1}^{n}\log(x_i + \sigma_i) - n\log(\kappa) - \frac{1}{2}\sum_{i=1}^{n}\log(\sigma_i). \end{aligned} \tag{8}$$

Estimation, inference and local influence for $\boldsymbol{\xi}$ can be based on the log-likelihood function given in equation (8) analogously as in the case of the BSACD1 model.

4 Influence diagnostics and residual analysis

4.1 Outlier detection framework in ACD models

We may record four types of outliers in time series: (a) additive outliers; (b) innovative outliers; (c) level shift outliers; and (d) transitory change outliers. Types (a) and (b) are mostly considered in time series. An additive outlier affects a single case, whereas an innovative outlier affects not only a particular observation but also the subsequent cases. Indeed, there are similarities between the influence statistics and tests for outlier detection. For instance, in the AR(1) model studied by Zevallos et al. (2012), the statistics for computing the influential points under the innovative perturbation scheme are the same as the innovative outlier detection test discussed in Fox (1972). A similar relationship can be obtained between the test for detecting additive outliers and the statistics for determining influential points under the data perturbation scheme; see Zevallos et al. (2012).

In the case of multiple outlier detection in time series modeling, Chen and Liu (1993) introduced detection procedures for avoiding masking and spurious effects. They have considered the four above-mentioned types of outliers. In the ACD literature, Chiang and Wang (2012) constructed a procedure to detect additive outliers for the logarithmic ACD model discussed by Bauwens and Giot (2000).

In the Cook distance, which is the deletion measure of influence of individual cases used in our paper, the masking and swamping issues pose some limitations on the assessment of the mutual influence and interactions of groups of cases, since this distance is concerned with individual cases; see Lawrence (1995). Nevertheless, a deeper study of outlier detection taking into account the joint and conditional deletion influence measures based on the Cook distance is beyond the scope of our paper. We leave it to be investigated in a further research.

4.2 Global influence

Global influence is related to case-deletion, that is, an approach to assess the effect of dropping the case i from the data set. Let a quantity with subscript "(i)" be that calculated with the case i deleted. Then, $\ell_{(i)}$ is the log-likelihood function, which is defined in equation (6) (or (8)), but evaluated at

$$\widehat{\boldsymbol{\theta}}_{(i)} = [\widehat{\varpi}_{(i)}, \widehat{\alpha}_{1(i)}, \ldots, \widehat{\alpha}_{p_1(i)}, \widehat{\beta}_{1(i)}, \ldots, \widehat{\beta}_{q_1(i)}, \widehat{\tau}_{(i)}]^\top$$

or at

$$\widehat{\boldsymbol{\xi}}_{(i)} = [\widehat{\varsigma}_{(i)}, \widehat{\gamma}_{1(i)}, \ldots, \widehat{\gamma}_{p_2(i)}, \widehat{\delta}_{1(i)}, \ldots, \widehat{\delta}_{q_2(i)}, \widehat{\kappa}_{(i)}]^\top,$$

according to which of the two models, BSACD1 or BSACD2 respectively, is being considered.

A first measure of global influence can be defined as the standardized norm of $\widehat{\boldsymbol{\theta}}_{(i)} - \widehat{\boldsymbol{\theta}}$, known as the generalized Cook distance (GCD), which is given by

$$\mathrm{GCD}_i(\boldsymbol{\theta}) = [\widehat{\boldsymbol{\theta}}_{(i)} - \widehat{\boldsymbol{\theta}}]^\top[-\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})][\widehat{\boldsymbol{\theta}}_{(i)} - \widehat{\boldsymbol{\theta}}], \quad i = 1, \ldots, n, \tag{9}$$

where $\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})$ is the Hessian matrix of $\ell(\boldsymbol{\theta})$ evaluated at $\widehat{\boldsymbol{\theta}}$. Alternatively, one can compute $\mathrm{GCD}_i(\varpi)$, $\mathrm{GCD}_i(\alpha_m)$, $\mathrm{GCD}_i(\beta_s)$ and $\mathrm{GCD}_i(\tau)$, with $m = 1, \ldots, p_1$ and $s = 1, \ldots, q_1$, whose values reveal the impact of the case i on the estimates of ϖ, α_m, β_s and τ, respectively. Moreover, $\widehat{\boldsymbol{\theta}}_{(i)}$ and $\widehat{\boldsymbol{\theta}}$ can be compared by their likelihood distance (LD) as $\mathrm{LD}_i(\boldsymbol{\theta}) = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{(i)})]$, for $i = 1, \ldots, n$. Similarly for the parameter $\boldsymbol{\xi}$ of the BSACD2 model, a GCD as that given in equation (9) may be employed.

4.3 Local influence

Local influence relies on the curvature of the surface of the log-likelihood function. Consider the likelihood displacement (LL) given by $\mathrm{LL}_i(\boldsymbol{\theta}) = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})]$, for $i = 1, \ldots, n$, where $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ corresponds to the ML estimate of $\boldsymbol{\theta}$ for a perturbed model, $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_n]^\top$ is a perturbation vector, and $\ell(\boldsymbol{\theta}_{\boldsymbol{\omega}})$ is the log-likelihood function of the model perturbed by $\boldsymbol{\omega}$. The local behavior of $\mathrm{LL}_i(\boldsymbol{\theta})$ around $\boldsymbol{\omega}_0$, the non-perturbation vector, was studied by Cook (1987), who showed that the normal curvature for $\boldsymbol{\theta}$ in the direction $\boldsymbol{d}$, with $||\boldsymbol{d}|| = 1$, is expressed as

$$C_d(\boldsymbol{\theta}) = 2|\boldsymbol{d}^\top \boldsymbol{\nabla}^\top \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})^{-1} \boldsymbol{\nabla}\, \boldsymbol{d}|,$$

where $\boldsymbol{\nabla}$ is a $[1 + p_1 + q_1] \times n$ matrix of perturbations with elements

$$\nabla_{ji} = \left.\frac{\partial^2 \ell(\boldsymbol{\theta}_{\boldsymbol{\omega}})}{\partial \boldsymbol{\theta}_j \partial \boldsymbol{\omega}_i}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}, \quad j = 1, \ldots, p_1 + q_1, \quad i = 1, \ldots, n.$$

Judgement of the local influence is usually based on index plots. We consider the index graph of the eigenvector $\boldsymbol{d}_{\max}$ corresponding to the largest eigenvalue of

$$\ddot{\boldsymbol{F}} = -\boldsymbol{\nabla}^\top \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})^{-1} \boldsymbol{\nabla}, \tag{10}$$

$C_{d_{\max}}(\boldsymbol{\theta})$ say, which can detect those cases that are potentially influential on $\widehat{\boldsymbol{\theta}}$. However, we may be interested only in $\widehat{\varpi}$, $\widehat{\boldsymbol{\alpha}} = [\widehat{\alpha}_1, \ldots, \widehat{\alpha}_{p_1}]^\top$, $\widehat{\boldsymbol{\beta}} = [\widehat{\beta}_1, \ldots, \widehat{\beta}_{q_1}]^\top$ or $\widehat{\tau}$, such that the normal curvature in the direction $\boldsymbol{d}$ is $C_l(\widehat{v}) = 2|\boldsymbol{d}^\top \boldsymbol{\nabla}^\top [\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})^{-1} - \ddot{\boldsymbol{\ell}}(\widehat{v})] \boldsymbol{\nabla}\, \boldsymbol{d}|$, with $v = \varpi, \boldsymbol{\alpha}, \boldsymbol{\beta}$ or τ and

$$\ddot{\boldsymbol{\ell}}(\widehat{\varpi}) = \begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\alpha}})^{-1} & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\beta}})^{-1} & \mathbf{0} \\ \cdot & \cdot & \cdot & \ddot{\boldsymbol{\ell}}(\widehat{\tau})^{-1} \end{bmatrix}, \quad \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\alpha}}) = \begin{bmatrix} \ddot{\boldsymbol{\ell}}(\widehat{\varpi})^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\beta}})^{-1} & \mathbf{0} \\ \cdot & \cdot & \cdot & \ddot{\boldsymbol{\ell}}(\widehat{\tau})^{-1} \end{bmatrix},$$
$$\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\beta}}) = \begin{bmatrix} \ddot{\boldsymbol{\ell}}(\widehat{\varpi})^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\alpha}})^{-1} & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \cdot & \ddot{\boldsymbol{\ell}}(\widehat{\tau})^{-1} \end{bmatrix}, \quad \ddot{\boldsymbol{\ell}}(\widehat{\tau}) = \begin{bmatrix} \ddot{\boldsymbol{\ell}}(\widehat{\varpi})^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\alpha}})^{-1} & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\beta}})^{-1} & \mathbf{0} \\ \cdot & \cdot & \cdot & 0 \end{bmatrix}.$$

In addition, we can consider the direction $\boldsymbol{d} = \boldsymbol{e}_{in}$, with $\boldsymbol{e}_{in}$ being an $n \times 1$ vector of zeros with one at the ith position, that is, the canonical basis of $\mathbb{R}^n$, $\{\boldsymbol{e}_{in}, 1 \leq i \leq n\}$ say. Here, the normal curvature is given by

$$C_i(\boldsymbol{\theta}) = 2\boldsymbol{e}_{in}|\ddot{\boldsymbol{F}}|\boldsymbol{e}_{in} = 2|\ddot{F}_{ii}|, \quad i = 1, \ldots, n, \tag{11}$$

where $\ddot{F}_{ii}$ is the ith diagonal element of $\ddot{\boldsymbol{F}}$ defined in equation (10). Therefore, if $C_i(\widehat{\boldsymbol{\theta}}) > 2\overline{C}(\widehat{\boldsymbol{\theta}})$, where $\overline{C}(\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} C_i(\widehat{\boldsymbol{\theta}})/n$, then the case i is considered as potentially influential. This procedure is named the total local influence method; see Lesaffre and Verbeke (1998).

Next, considering the model defined in equation (5) and its log-likelihood function given by equation (6), we derive the different perturbation matrices for each scheme. Similarly for the parameter $\boldsymbol{\xi}$ of the BSACD2 model, the total local influence method based on C_i given in equation (11) can be derived.

Case-weight perturbation Under this perturbation scheme, one is interested in evaluating whether the contributions of the cases with different weights affect the ML estimate of $\boldsymbol{\theta}$. Consider a weight vector $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_n]^\top$. Then, the perturbed log-likelihood is given by

$$\ell(\boldsymbol{\theta}_{\boldsymbol{\omega}}) = \sum_{i=1}^{n} \omega_i \ell_i(\boldsymbol{\theta}),$$

with $0 \leq \omega_i \leq 1$, for $i = 1, \ldots, n$, and $\boldsymbol{\omega}_0 = [1, \ldots, 1]^\top$, where

$$\ell_i(\boldsymbol{\theta}) \;=\; \frac{\tau}{2} - \frac{\log(16\pi)}{2} - \frac{1}{2}\log\left(\frac{[\tau+1]x_i^3\mu_i}{[\tau x_i + x_i + \tau\mu_i]^2}\right) - \frac{x_i[\tau+1]}{4\mu_i} - \frac{\tau^2\mu_i}{4[\tau+1]x_i}, \quad x_i > 0.$$

Data perturbation We here assume an additive perturbation for each case of the model given in equation (5), namely,

$$X_i(\omega_i) = X_i + \omega_i\, s(X_i) = \mu_i(\omega_i)\, \varepsilon_i, \quad \varepsilon_i \sim \text{RBS}(1, \tau),$$

with

$$\log(\mu_i(\omega_i)) = \varpi + \sum_{j=1}^{p_1} \alpha_j \log(\mu_{i-j}(\omega_{i-j})) + \sum_{j=1}^{q_1} \frac{\beta_j\, X_{i-j}(\omega_{i-j})}{\mu_{i-j}(\omega_{i-j})},$$

where $s(X_i)$ is a scale factor and $\omega_i \in \mathbb{R}$, for $i = 1, \ldots, n$. Note that $X_i(\omega_i) \sim \text{RBS}(\mu_i(\omega_i), \tau)$. Then, the corresponding log-likelihood function is given by $\ell(\boldsymbol{\theta}_{\boldsymbol{\omega}}) = \sum_{i=1}^{n} \ell_{\omega_i}(\boldsymbol{\theta})$, with

$$\ell_{\omega_i}(\boldsymbol{\theta}) = \frac{\tau}{2} - \frac{\log(16\pi)}{2} - \frac{1}{2}\log\left(\frac{[\tau+1]x_i^3(\omega_i)\mu_i(\omega_i)}{[\tau x_i(\omega_i) + x_i(\omega_i) + \tau\mu_i(\omega_i)]^2}\right) - \frac{x_i(\omega_i)[\tau+1]}{4\mu_i(\omega_i)} - \frac{\tau^2\mu_i(\omega_i)}{4[\tau+1]x_i(\omega_i)},$$

for $x_i(\omega_i) > 0$ and $\boldsymbol{\omega}_0 = [0, \ldots, 0]^\top$.

Innovative perturbation Note that the log-likelihood function given in equation (6) relies on the assumption that ε_is are IID random variables following the RBS distribution. Nevertheless, the model fitting may be strongly affected by the existence of influential cases. We assume that a perturbation vector $\boldsymbol{\omega}$ is introduced to equation (5) through the conditional mean (or median) duration as

$$Z_i = X_i = \mu_i\,\varepsilon_i, \quad \varepsilon_i \sim \text{RBS}(1/\omega_i, \tau),$$

with

$$\log(\mu_i) = \varpi + \sum_{j=1}^{p_1} \alpha_j \log(\mu_{i-j}) + \sum_{j=1}^{q_1} \frac{\beta_j\, X_{i-j}}{\mu_{i-j}}.$$

Then, the corresponding log-likelihood function is given by $\ell(\boldsymbol{\theta}_{\boldsymbol{\omega}}) = \sum_{i=1}^{n} \ell_{\omega_i}(\boldsymbol{\theta})$, where

$$\ell_{\omega_i}(\boldsymbol{\theta}) \;=\; \frac{\tau}{2} - \frac{\log(16\pi)}{2} - \frac{1}{2}\log\left(\frac{[\tau+1]x_i^3\mu_i}{\omega_i[\tau x_i + x_i + \tau\omega_i^{-1}\mu_i]^2}\right) - \frac{x_i[\tau+1]\omega_i}{4\mu_i} - \frac{\tau^2\mu_i}{4[\tau+1]x_i\omega_i},$$

for $x_i > 0$ and $\boldsymbol{\omega}_0 = [1, \ldots, 1]^\top$.

4.4 Residual analysis

Goodness-of-fit and departures from the assumptions of the model can be assessed by means of residual analysis. In particular, two types of residuals are considered in this paper. The first is a generalized Cox-Snell (GCS) residual given by

$$r_i^{\text{GCS}} = -\log(\widehat{S}(x_i|\Omega_{i-1})), \quad i = 1, \ldots, n, \tag{12}$$

where $\widehat{S}$ is the SF fitted to the ACD data. The SF for the mean-based BSACD model (BSACD1) is given by

$$S(x_i; \mu_i, \tau) \;=\; \Phi\left(-\sqrt{\frac{\tau}{2}}\left[\sqrt{\frac{\{\tau+1\}x_i}{\mu_i\tau}} - \sqrt{\frac{\mu_i\tau}{\{\tau+1\}x_i}}\right]\right), \quad x_i, \mu_i, \tau > 0,$$

whereas for the median-based BSACD model (BSACD2) this is given by

$$S(x_i; \kappa, \sigma_i) \;=\; \Phi\left(-\frac{1}{\kappa}\left[\sqrt{\frac{x_i}{\sigma_i}} - \sqrt{\frac{\sigma_i}{x_i}}\right]\right), \quad x_i, \kappa, \sigma_i > 0.$$

If the model is correctly specified, then the GCS residual is unit exponential, EXP(1) in short, distributed whatever the ACD model specification; see Bhatti (2010).

The randomized quantile (RQ) residual is the second type of residual to be considered. It is usually applied to generalized additive models for location, scale and shape; see Dunn and Smyth (1996). The RQ residual is given by

$$r_i^{\mathrm{RQ}} = \Phi^{-1}(\widehat{S}(x_i|\Omega_{i-1})), \quad i = 1, \ldots, n,$$

where Φ^{-1} is the inverse function of the standard normal CDF and $\widehat{S}$ is the fitted SF as in equation (12). The RQ residual follows a standard normal distribution when the model is correctly specified, again regardless of the ACD model.

5 Monte Carlo simulation

Two MC simulation studies were carried out to evaluate the performance of the ML estimators for the BSACD1($p_1 = 1, q_1 = 1$) and BSACD2($p_2 = 1, q_2 = 1$) models. The order of the lags for these models are set as $p_l = 1$ and $q_l = 1$, for $l = 1, 2$, because a higher order for BSACD models does not improve the model fit; see Bhatti (2010). Thereby, in the following, any BSACD1($p_1 = 1, q_1 = 1$) or BSACD2($p_2 = 1, q_2 = 1$) model is simply denoted as BSACD1 or BSACD2 (the same applies to other models considered in this work). The first study considers the simulated TDs generated from the BSACD1 and BSACD2 models, whereas the second one has as its data generating process the logarithmic ACD model (see Bauwens and Giot, 2000) with generalized gamma errors (GGACD) . All numerical evaluations were done in the R software; see R-Team (2016).

5.1 Simulation study 1

The first simulation scenario considers: sample size $n \in \{50, 100, 500, 1000, 2000\}$, vector of true parameters $[\varpi, \alpha, \beta, \tau] = [0.1, 0.9, 0.1, 1.65]$ (BSACD1) and $[\varsigma, \gamma, \delta, \kappa] = [0.1, 0.9, 0.1, 1.1]$ (BSACD2), and a number of $10,000$ MC replications for each sample size. The BSACD1 and BSACD2 samples were generated using the transformations defined in equations (1) and (4), respectively. The ML estimation results are presented in Table 1. The following sample statistics for the ML estimates are reported: empirical mean, coefficients of skewness (CS) and of kurtosis (CK), relative bias (RB), and root mean squared error (RMSE) defined as the square root of the mean squared error. A look at the results in Table 1 allows us to conclude that, as the sample size increases, the RB and RMSE of all the estimators decrease, indicating that they are asymptotically unbiased, as expected. Moreover, $\widehat{\beta}$, $\widehat{\delta}$, $\widehat{\tau}$ and $\widehat{\kappa}$ seem to be consistent and marginally asymptotic normal distributed. However, $\widehat{\varpi}$, $\widehat{\varsigma}$, $\widehat{\alpha}$ and $\widehat{\gamma}$ are somewhat skewed and with high kurtosis, but they tend to the normal case, as the sample size increases.

5.2 Simulation study 2

The second simulation scenario considers: sample size $n \in \{50, 100, 500, 1000, 2000\}$, vector of true parameters $[a, b, c, \vartheta, \varkappa] = [0.1, 0.9, 0.1, 0.75, 1.5]$ (GGACD) and $10,000$ MC replications for each sample size. The values of ϑ and $\varkappa$ have been chosen so as to simulate data with a unimodal HR ($\vartheta < 1$ and $\vartheta, \varkappa - 1 > 0$), which is a characteristic present in financial durations; see Bhatti (2010)

Table 1: Summary statistics from simulated ACD data for the indicated model, estimator and n.

	BSACD1					BSACD2				
	n					n				
Statistic	50	100	500	1000	2000	50	100	500	1000	2000
	$\widehat{\varpi}$					$\widehat{\varsigma}$				
True value	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
Mean	0.1992	0.1518	0.1480	0.1101	0.1068	0.0182	0.1657	0.1583	0.0973	0.1002
CS	2.9381	4.6335	8.4225	1.2638	0.9232	9.2936	4.8474	8.3807	0.8559	0.5952
CK	12.4462	26.2824	87.7129	5.6169	4.7492	121.8675	26.8107	73.4060	4.3755	3.7650
RB	0.9922	0.5184	0.4800	0.1012	0.0676	1.8183	0.3430	0.5838	0.0265	0.0027
RMSE	0.5636	0.5295	0.2909	0.0672	0.0459	0.2550	0.7303	0.5506	0.0609	0.0434
	$\widehat{\alpha}$					$\widehat{\gamma}$				
True value	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000
Mean	0.6513	0.7580	0.8634	0.8894	0.8938	0.1803	0.7950	0.8659	0.8946	0.8966
CS	−1.4732	−3.0760	−8.4600	−1.0580	−0.8020	1.4751	−3.8060	−8.2852	−0.6918	−0.5089
CK	4.3876	12.713	85.732	4.8490	4.4175	3.4129	16.817	71.8364	3.6231	3.5347
RB	0.2763	0.1577	0.0407	0.0117	0.0069	1.7996	0.1166	0.0379	0.0059	0.0037
RMSE	0.4965	0.3820	0.1600	0.0376	0.0255	0.8051	0.3782	0.2096	0.0254	0.0181
	$\widehat{\beta}$					$\widehat{\delta}$				
True value	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
Mean	0.1950	0.1844	0.1220	0.1121	0.1061	0.0395	0.1556	0.1164	0.1096	0.1049
CS	0.0873	0.2552	−0.0363	0.1138	0.0868	2.0847	−0.0316	−0.7590	0.0512	0.0338
CK	3.3799	3.4704	4.1325	3.0158	2.7908	6.5525	3.8752	6.7512	3.1480	2.8253
RB	0.9503	0.8445	0.2204	0.1214	0.0611	1.6041	0.5562	0.1638	0.0958	0.0487
RMSE	0.4416	0.1234	0.0392	0.0242	0.0155	0.1037	0.0799	0.0286	0.0170	0.0107
	$\widehat{\tau}$					$\widehat{\kappa}$				
True value	1.6529	1.6529	1.6529	1.6529	1.6529	1.1000	1.1000	1.1000	1.1000	1.1000
Mean	1.3199	1.5136	1.5924	1.6197	1.6349	0.2547	1.1684	1.1276	1.1144	1.1077
CS	−0.7339	0.5348	0.1600	0.1152	0.1263	1.4431	0.2468	0.3117	0.1930	0.0496
CK	3.7721	3.3853	3.3622	3.2757	3.1453	3.1905	3.3262	3.6865	3.4992	3.1465
RB	0.2015	0.0843	0.0366	0.0201	0.0109	1.7684	0.0622	0.0251	0.0131	0.0070
RMSE	0.6895	0.2994	0.1265	0.0861	0.0570	0.9765	0.1231	0.0499	0.0317	0.0202

and Leiva et al. (2014b). In the GGACD model, the PDF of X_i can be written as

$$f(x_i;\psi_i,\vartheta,\varkappa)=\frac{\vartheta}{\varphi(\vartheta,\varkappa)\psi_i\Gamma(\varkappa)}\left(\frac{x_i}{\varphi(\vartheta,\varkappa)\psi_i}\right)^{\varkappa\vartheta-1}\exp\left(-\left(\frac{x_i}{\varphi(\vartheta,\varkappa)\psi_i}\right)^{\vartheta}\right), \tag{13}$$

where $\varphi(\vartheta,\varkappa)=\Gamma(\varkappa)/\Gamma(\varkappa+\vartheta^{-1})$ and $\log(\psi_i)=a+b\,\log(\psi_{i-1})+c\,X_i/\psi_i$. The GGACD samples were generated by considering the PDF given in equation (13). For both BSACD1 and BSACD2 models, Table 2 presents the empirical mean, CS and CK of the ML estimators. Note that the RB and RMSE are not computed as their true values because they do not apply in the true GGACD model. Hence, we consider only the ACD parameters as comparison. From Table 2, we observe that, in general, $\widehat{\varpi}$, $\widehat{\alpha}$, $\widehat{\delta}$, $\widehat{\varsigma}$, $\widehat{\gamma}$ and $\widehat{\delta}$ are persistently skewed with high kurtosis. Nonetheless, these results are expected since the simulated durations were generated using the GGACD model. Overall, the results associated with the BSACD1 model are closer to the simulation model ones. Now, a quick glance at the estimates in Table 3 reveals that $\widehat{a}$ and $\widehat{b}$ are highly skewed with a great degree of kurtosis. Moreover, $\widehat{b}$ remains close to the normal distribution in terms of skewness and kurtosis values.

Table 2: Summary statistics from simulated ACD data for the indicated model, estimator and n.

	BSACD1					BSACD2				
	n					n				
Statistic	50	100	500	1000	2000	50	100	500	1000	2000
	$\widehat{\varpi}$					$\widehat{\varsigma}$				
a (GGACD)	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
Mean	0.1640	0.1160	0.1361	0.2261	0.0972	−0.1057	−0.1022	−0.1026	−0.1030	−0.0982
CS	1.5929	2.2436	2.7077	2.0290	3.4038	−1.1457	−0.9851	−1.0081	−2.2631	−2.3702
CK	5.9769	9.5881	11.5575	6.9964	17.0474	13.0222	21.3692	10.5784	15.5903	13.9089
	$\widehat{\alpha}$					$\widehat{\gamma}$				
b (GGACD)	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000
Mean	0.6207	0.7342	0.7482	0.6517	0.8037	0.8023	0.7239	0.7462	0.7972	0.8098
CS	−1.4016	−2.0746	−2.4127	−1.7828	−3.0732	−2.3469	−2.0830	−2.3519	−2.9810	−3.3607
CK	4.2612	7.4591	9.1998	5.7076	14.0063	9.4004	7.2681	8.8088	13.8514	16.3301
	$\widehat{\beta}$					$\widehat{\delta}$				
c (GGACD)	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
Mean	0.1270	0.1094	0.0840	0.0802	0.0696	0.0445	0.0438	0.0321	0.0278	0.0250
CS	−0.1549	0.2365	0.2107	0.0701	0.4838	−0.0320	−0.2496	−0.1600	−0.0105	0.0145
CK	4.0667	5.8606	4.8427	4.7059	5.8006	4.4591	4.5015	4.2197	4.1956	4.9335

6 Analysis of high-frequency financial data

6.1 Exploratory data analysis

The BSACD1 and BSACD2 models are now used to analyse a real high-frequency financial data set, corresponding to price durations of BASF-SE stock on 19th April 2016 downloaded from the Dukascopy site (`www.dukascopy.com`). A data adjustment was applied, using the `R` package `ACDm` (see Belfrage, 2015), to allow for the fact that this type of data has an active trading pattern in the opening and closing hours and a dormant trading pattern around noon; see Engle and Russell (1998). Table 4 provides descriptive statistics for the BASF-SE data set, including central tendency statistics, standard deviation (SD), coefficient of variation (CV), CS and CK. From this table, note the right skewed nature and high kurtosis level of the data distribution. The skewness is confirmed by the histogram shown in Figure 1(left).

Table 3: Summary statistics from simulated ACD data in the GGACD model for the indicated estimator and n.

Statistic	n = 50	100	500	1000	2000
			$\widehat{a}$		
True value	0.1000	0.1000	0.1000	0.1000	0.1000
Mean	0.1668	0.1397	0.0869	0.0971	0.0987
CS	2.8514	4.0351	6.2753	6.1861	3.1767
CK	21.2327	23.5028	48.2212	83.6398	24.9667
RB	0.6680	0.3979	0.1300	0.0290	0.0121
RMSE	0.6983	0.4595	0.2342	0.2332	0.1846
			$\widehat{b}$		
True value	0.9000	0.9000	0.9000	0.9000	0.9000
Mean	0.8496	0.8343	0.8629	0.8902	0.8946
CS	−3.6517	−3.7168	−5.7847	−5.4554	−3.1589
CK	18.8634	18.9606	41.7511	69.2005	27.4577
RB	0.0559	0.0729	0.0411	0.0108	0.0059
RMSE	0.2683	0.2628	0.1984	0.0702	0.0457
			$\widehat{c}$		
True value	0.1000	0.1000	0.1000	0.1000	0.1000
Mean	−0.0426	0.0189	0.0544	0.0570	0.0579
CS	0.9863	0.5387	0.1293	0.2406	0.4003
CK	4.2246	3.5639	4.1752	3.3597	3.6020
RB	1.4265	0.8101	0.4557	0.4295	0.4202
RMSE	1.4360	0.8173	0.4570	0.4301	0.4205

Table 4: Summary statistics for the BASF-SE data.

n	Min	Median	Mean	Max	SD	CV	CS	CK
2194	0.061	0.682	1.067	9.776	1.167	109.35%	2.521	8.902

The shape of an HR, which is defined by $h(x) = f(x)/[1 - F(x)]$, with f and F being the PDF and CDF of X, respectively, is a relevant characteristic to decide whether a particular distribution is suitable or not for a data set. The scaled total time on test (TTT) function (see Aarset, 1987) is usually a good tool to characterize the HR and is given by $W(u) = H^{-1}(u)/H^{-1}(1)$, for $0 \leq u \leq 1$, where $H^{-1}(u) = \int_0^{F^{-1}(u)}[1 - F(y)]\mathrm{d}y$, with F^{-1} being the inverse function of the CDF of X. A plot of the points $[k/n, W_n(k/n)]$, with

$$W_n(k/n) = \frac{\sum_{i=1}^{k} x_{(i)} + [n-k]x_k}{\sum_{i=1}^{n} x_{(i)}}, \quad k = 1, \dots, n,$$

and $x_{(i)}$ being the ith observed order statistic, provides an approximation for W. Figure 1(centre) suggests an inverse bathtub HR for the BASF-SE data set, as expected; see Bhatti (2010) and Leiva et al. (2014b).

Figure 1(right) shows the usual and adjusted boxplots, with the latter being useful in cases where the data follow a skew distribution; see Hubert and Vanderveeken (2008). From this figure, we note that potential outliers considered by the usual boxplot are not outliers when the adjusted boxplot is observed.

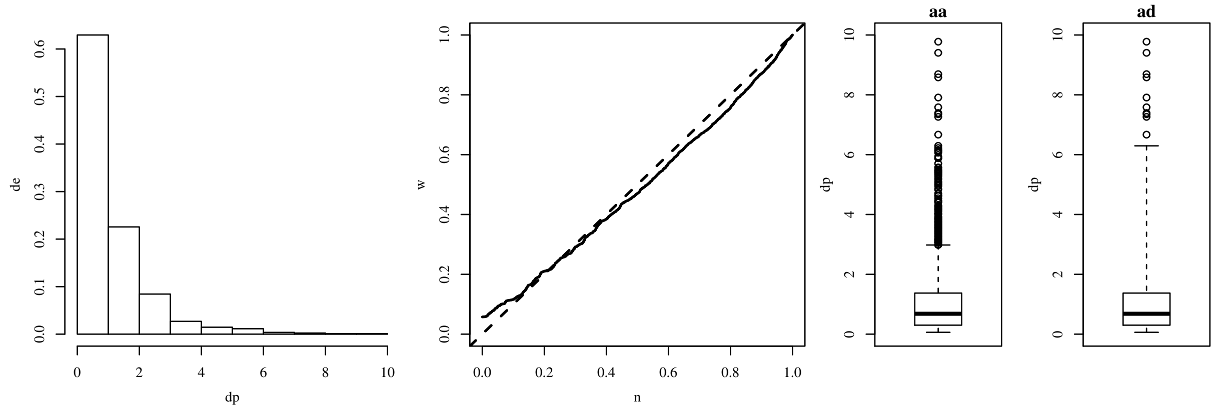


Figure 1: Histogram (left), TTT plot (centre) and boxplots (right) for the BASF-SE data.

6.2 Estimation and model validation

Table 5 reports the ML estimates, computed by the BFGS method, SEs and p-values of the t-test for BSACD1(1,1) and BSACD2(1,1) model parameters. In addition, we report the Akaike (AIC) and Bayesian information (BIC) criteria and evaluate the absence of autocorrelation in the residuals, providing the p-values of the Ljung-Box (LB) statistic, $Q(\gamma)$ say, for up to γth order serial correlation. For comparison, the results of the GGACD model, in addition to ACD models based on the lognormal (LNACD) and log-Student-t (LtACD) distributions, are given as well; see Bhatti (2010). From Table 5, note that, at a 1% significance level, the LB statistics provide no evidence of serial correlation in the residuals. We also note that the ACD models based on the BS (BSACD1 and BSACD2) models provide better adjustments compared to the other models based on the values of AIC and BIC. Moreover, the AIC and BIC values of the BSACD1 model do not substantially differ from the values of the BSACD2 model, suggesting no adjustment improvement.

Hypothesis testing of H_0: $\theta = \theta_0$ versus H_1: $\theta \neq \theta_0$ can be performed using the Wald statistic defined by $W = [\widehat{\theta} - \theta_0]/\mathrm{SE}(\widehat{\theta})$, which is approximately N(0,1) distributed under H_0, where $\widehat{\theta}$ and θ_0 are the corresponding estimator and its proposed value under H_0, respectively. From Table 5, note that all the ACD parameters are statistically significant at the 5% level. Figure 2 displays the QQ plots with simulated envelope of the GCS and RQ residuals for the BSACD1, BSACD2, GGACD, LNACD and LtACD models. From this figure, observe that the GCS residuals show good agreement with the EXP(1) distribution and the RQ residuals with the $\mathrm{N}(0, 1)$ distribution in the BSACD1 and BSACD2 models.

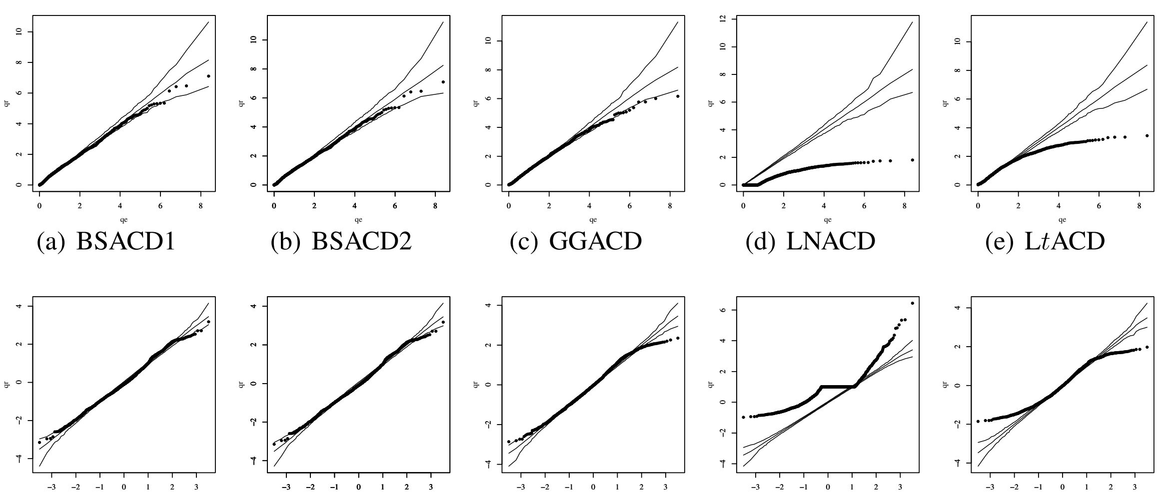
6.3 Diagnostic analysis

Global influence Figure 3 presents the case-deletion measures $\mathrm{GCD}_i(\boldsymbol{\theta})$ and $\mathrm{LD}_i(\boldsymbol{\theta})$ presented in Section 4.2. We note that both the $\mathrm{GCD}_i(\boldsymbol{\theta})$ and $\mathrm{LD}_i(\boldsymbol{\theta})$ statistics indicate that the case #1118 is a possible influential observation.

Table 5: ML estimates (with SE in parentheses) and model selection measures for fit to the BASF-SE data.

ACD model	Parameter	ML estimate	p-value	$Q(4)$	$Q(16)$	AIC	BIC
BSACD1	ϖ	$-0.0485(0.0189)$	0.0101	0.6227	0.2400	4461.66	4484.43
	α	0.5799(0.1763)	0.0010				
	β	0.0692(0.0183)	0.0001				
	τ	1.3946(0.0421)					
BSACD2	ς	$-0.2756(0.0975)$	0.0047	0.6238	0.2403	4461.66	4484.43
	γ	0.5800(0.1758)	0.0009				
	δ	0.0403(0.0106)	0.0001				
	κ	1.1974(0.0180)					
GGACD	a	0.0724(0.0183)	<0.0001	0.7964	0.4344	4488.384	4516.851
	b	0.7204(0.0936)	<0.0001				
	c	0.0494(0.0117)	<0.0001				
	ϑ	15.0013(6.5247)					
	$\varkappa$	0.2467(0.0547)					
LNACD	a_1	$-0.2666(0.0799)$	0.0008	0.9472	0.7028	6505.187	6482.413
	b_2	0.5951(0.1502)	<0.0001				
	c_2	0.0494(0.0117)	<0.0001				
	ξ	1.0675(0.0161)					
LtACD	a_3	$-0.2278(0.0592)$	0.0001	0.8662	0.6887	6782.066	6753.599
	b_3	0.6498(0.1193)	<0.0001				
	c_3	0.0561(0.0124)	<0.0001				
	ζ	0.8786(0.0176)					
	ν	3					

where a_1, b_1, c_1 are the ARMA parameters of the LNACD model and ξ its shape parameter, whereas a_2, b_2, c_2 are the ARMA parameters of the LtACD model, ξ its scale parameter and ν its degrees of freedom.



(a) BSACD1 (b) BSACD2 (c) GGACD (d) LNACD (e) LtACD

(f) BSACD1 (g) BSACD2 (h) GGACD (i) LNACD (j) LtACD

Figure 2: QQ plot and its envelope for the GCS (top) and RQ (bottom) residuals in the indicated model with the BASF-SE data.

Case-weight perturbation Index plots of C_i under case perturbation are shown in Figure 4, detecting the case #1119 as a potential influential observation for both of the BSACD1 and BSACD2 models.

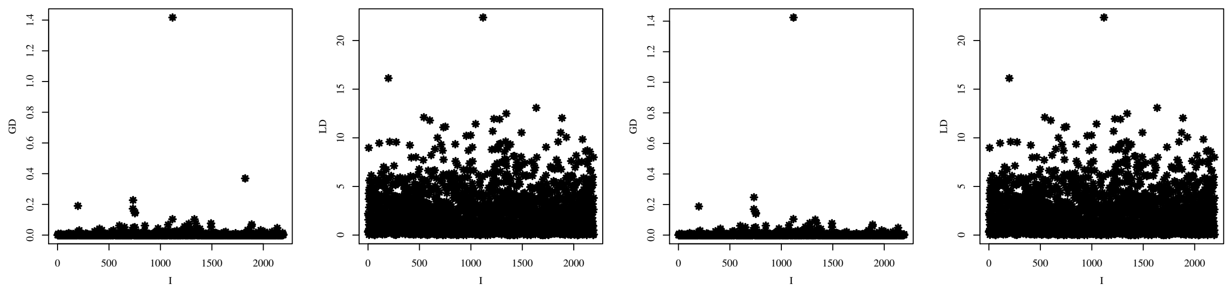


Figure 3: GCD and LD for the BSACD1 (left) and BSACD2 (right) models with the BASF-SE data.

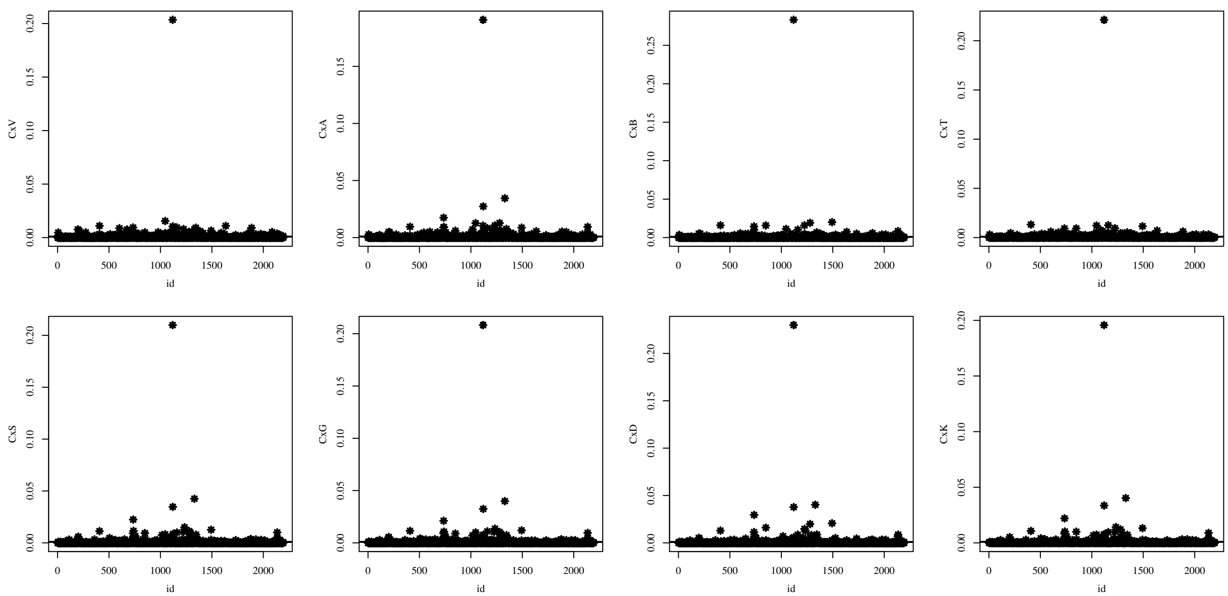


Figure 4: Index plots of C_i for the indicated parameter under case-weight perturbation in the BSACD1 (top) and BSACD2 (bottom) models with the BASF-SE data.

Data perturbation Index plots of C_i under data perturbation are displayed in Figure 5, where the case #1119 is again detected as a potential influential observation for both models.

Innovative perturbation Figures 6(a)-(d) show index plots of C_i under innovative perturbation, where the case #1119 is once again detected as a potential influential observation for the BSACD1 and BSACD2 models.

Relative change Here, the impact of the detected influential cases on the model inference is checked. We compute the relative change (RC), which is obtained by removing influential cases and re-estimating the parameters and the corresponding SEs as $\mathrm{RC}_{\theta_{j(i)}} = |[\widehat{\theta}_j - \widehat{\theta}_{j(i)}]/\widehat{\theta}_j| \times 100\%$ and $\mathrm{RC}_{\mathrm{SE}(\theta_{j(i)})} = |[\widehat{\mathrm{SE}}(\widehat{\theta}_j) - \widehat{\mathrm{SE}}(\widehat{\theta}_j)_{(i)}]/\widehat{\mathrm{SE}}(\widehat{\theta}_j)| \times 100\%$, where $\widehat{\theta}_{j(i)}$ and $\widehat{\mathrm{SE}}(\widehat{\theta}_j)$ are the ML estimate of θ_j and its corresponding SE, respectively, after removing the case i, for $j = 1, \ldots, 4$ and $i = 1, \ldots, n$, with $\theta_1 = \varpi, \varsigma$; $\theta_2 = \alpha, \gamma$; $\theta_3 = \beta, \delta$; and $\theta_4 = \tau, \kappa$. Table 6 reports the RCs in the parameter estimates and SEs, as well as the p-values of the corresponding t-test obtained by considering the data with dropped cases. From this table, note that the largest RCs are related to the removal of the case #1118.

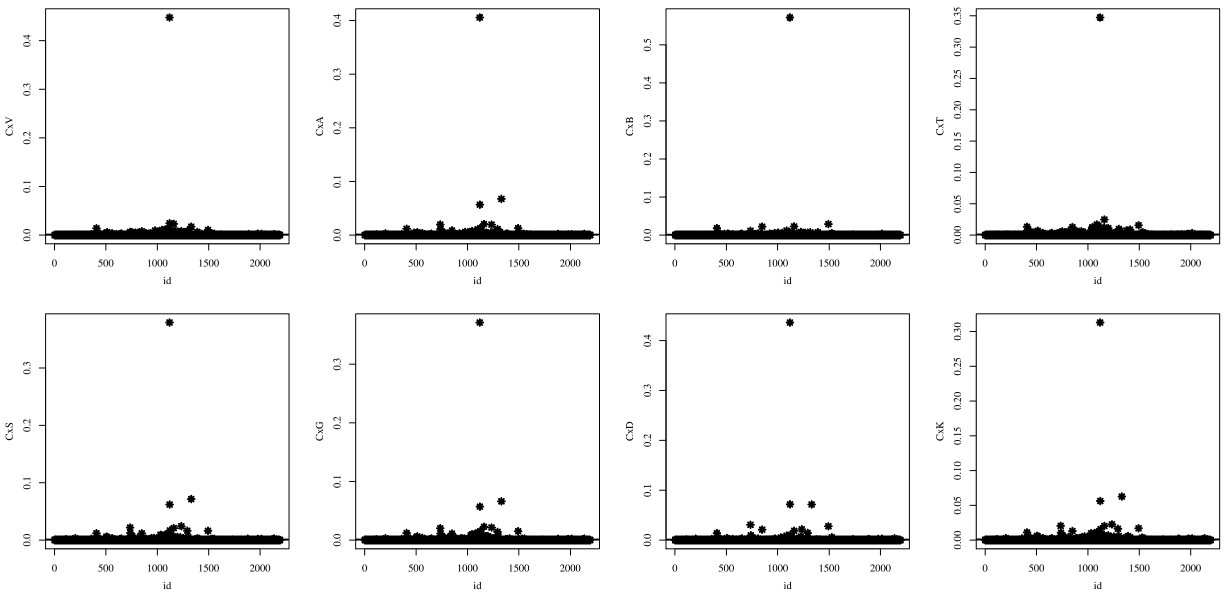


Figure 5: Index plots of C_i for the indicated parameter under data perturbation for the BSACD1 (top) and BSACD2 (bottom) models with the BASF-SE data.

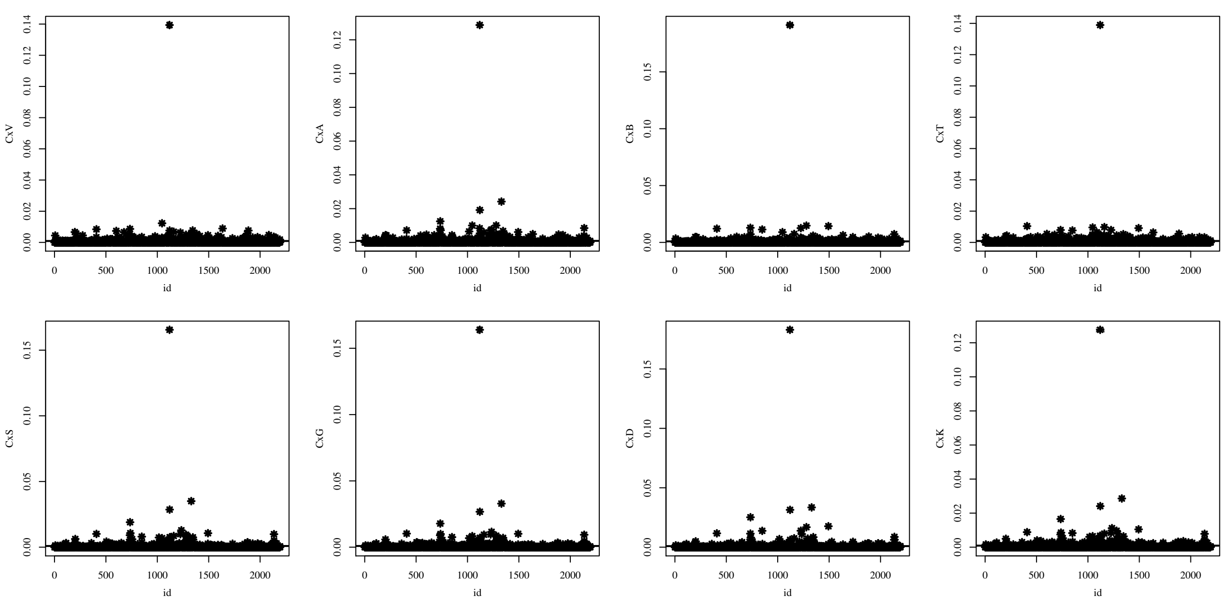


Figure 6: Index plots of C_i for the indicated parameter under innovative perturbation for the BSACD1 (top) and BSACD2 (bottom) models with the BASF-SE data.

In terms of model sensitivity, we observe a more pronounced influence of the removed cases on the ML estimates of the BSACD1 model parameters than on the ML estimates of the BSACD2 model parameters. This result can be interpreted as a robustness of the BSACD2 model to atypical cases. Thus, no inferential changes are found for either model; namely, the diagnostic measures identify potentially influential cases, but these do not alter the inference of the models.

Table 6: RCs (in %) in ML estimates and their corresponding SEs for the indicated parameter and removed cases, and respective p-values with the BASF-SE data.

Removed case(s)		BSACD1				BSACD2			
		$\widehat{\varpi}$	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\tau}$	$\widehat{\varsigma}$	$\widehat{\gamma}$	$\widehat{\delta}$	$\widehat{\kappa}$
{1118}	$\mathrm{RC}_{\theta_{j(i)}}$	40.38	12.62	37.30	23.29	8.12	2.80	3.91	1.73
	$\mathrm{RC}_{\mathrm{SE}(\theta_{j(i)})}$	(6.44)	(15.37)	(7.84)	(0.83)	(12.22)	(15.22)	(8.05)	(0.37)
	p-value	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
{1119}	$\mathrm{RC}_{\theta_{j(i)}}$	1.41	2.56	0.6	6.58	2.26	4.39	0.08	0.49
	$\mathrm{RC}_{\mathrm{SE}(\theta_{j(i)})}$	(1.11)	(1.50)	(0.6)	(0.25)	(1.34)	(1.52)	(0.65)	(0.09)
	p-value	[0.008]	[<0.001]	[<0.001]	[<0.001]	[0.005]	[<0.001]	[<0.001]	[<0.001]
{1118,1119}	$\mathrm{RC}_{\theta_{j(i)}}$	37.67	18.22	35.36	28.25	7.97	3.62	3.71	2.12
	$\mathrm{RC}_{\mathrm{SE}(\theta_{j(i)})}$	(6.53)	(14.04)	(7.85)	(1.03)	(12.22)	(13.89)	(8.15)	(0.44)
	p-value	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]

6.4 Forecasting performance

We evaluate the forecasting ability of the proposed models using traditional forecasting and density forecast evaluation.

Traditional forecasting technique Forecasts from an ACD model can be carried out in an analogous way to that used for GARCH models; see Tsay (2009). Consider the BSACD1 and BSACD2 models and suppose that the forecast origin is $i = h$. For a 1-step-ahead forecast, the models state that for BSACD1 $x_{h+1} = \mu_{h+1}\,\varepsilon_{h+1}$ with $\mu_{h+1} = \exp\left(\varpi + \alpha\log(\mu_h) + \beta\, x_h/\mu_h\right)$ and for BSACD2 $x_{h+1} = \sigma_{h+1}\,\varrho_{h+1}$ with $\sigma_h = \exp\left(\varsigma + \gamma\log(\sigma_h) + \delta\, x_h/\sigma_h\right)$. Let $x_h(1)$ be the 1-step-ahead forecast of x_{h+1} at the origin h. Then, $x_h(1) = \mathrm{E}[\mu_{h+1}\,\varepsilon_{h+1}] = \mu_{h+1}$ (BSACD1) and $x_h(1) = F_{\mathrm{BS}}^{-1}(0.5|\Omega_h) = \sigma_{h+1}$ (BSACD2). For multi-step-ahead forecasts, we use $x_{h+j} = \mu_{h+j}\,\varepsilon_{h+j}$ (BSACD1). Then, its j-step-ahead ($j > 1$) forecast under the BSACD1 model is

$$\begin{aligned}
x_h(j) = \mathrm{E}[\mu_{h+j}\,\varepsilon_{h+j}] &= \mathrm{E}[\exp(g(\varepsilon_{h+j-1}))]\mathrm{E}[\exp(\varpi + \alpha\log(\mu_{h+j-1}))] \\
&= \Xi_1\mathrm{E}[\exp\left(\varpi + \alpha\log(\mu_{h+j-1})\right)] \\
&= \Xi_1\mathrm{E}[\exp\left(\varpi + \alpha(\varpi + \alpha\log(\mu_{h+j-2}) + g(\varepsilon_{h+j-2}))\right)] \\
&\cdots \\
&= \Xi_1\Xi_2\cdots\Xi_{j-1}\left(\exp\left(\frac{\varpi[1-\alpha^{j-1}]}{1-\alpha} + \alpha^{j-1}\log(\mu_{h+1})\right)\right),
\end{aligned}$$

where $\Xi_m = \mathrm{E}[\exp\left(\alpha^{m-1}g(\varepsilon_{h+j-m})\right)]$, with $g(\varepsilon_{h+j-m}) = \beta\varepsilon_{h+j-m} = \beta\, x_{h+j-m}/\mu_{h+j-m}$. An estimate of Ξ_m can be computed by $\widehat{\Xi}_m = [1/N]\sum_{h=1}^{N}\exp\left(\alpha^{m-1}g(\varepsilon_h)\right)$, with $m = 1,\ldots,j-1$. Similarly, j-step-ahead ($j > 1$) forecasts may be obtained for the BSACD2 model; see Dufour and Engle (2000). Now, we compare the forecast ability of the BSACD1, BSACD2, GGACD, LNACD and LtACD models. Table 7 presents the estimated mean square error (MSE) for 6-step-ahead forecasts from these models. The results indicate the good performance of the ACD models based on BS distributions.

Table 7: Forecasts 6-step-ahead and MSE from the indicated model with the BASF-SE data.

j	1	2	3	4	5	6
BSACD1	0.054	0.040	0.002	1.263	0.086	0.089
BSACD2	0.011	0.004	0.034	1.617	0.201	0.017
GGACD	0.075	0.087	0.010	0.883	0.005	0.311
LNACD	0.012	0.006	0.031	1.596	0.194	0.020
LtACD	0.016	0.008	0.029	1.599	0.181	0.024

Density forecast evaluation technique Next, the density forecast (DF) evaluation technique, proposed by Diebold et al. (1998), is used to compare the forecasting ability of the BSACD1, BSACD2, GGACD, LNACD and LtACD models. In our case, this technique consists of checking whether a sequence of one-step-ahead DFs generated by an ACD model, $\{f_i(x_i|\Omega_{i-1})\}$ say, and a sequence of PDFs defining the data generating process, $\{p_i(x_i|\Omega_{i-1})\}$ say, are such that

$$\{f_i(x_i|\Omega_{i-1})\} = \{p_i(x_i|\Omega_{i-1})\}, \tag{14}$$

or not. Due to the fact that $\{p_i(x_i|\Omega_{i-1})\}$ is never observed, we compute the probability integral transform given by

$$z_i = \int_{-\infty}^{x_i} f_i(u)\mathrm{d}u.$$

Under a null hypothesis based on equation (14), the sequence $\{z_i\}$ of $\{x_i\}$ with respect to $\{f_i(x_i|\Omega_{i-1})\}$ are IID random variables following the U(0,1) distribution; see Diebold et al. (1998) and Bauwens et al. (2004). Therefore, graphical analysis, such as histogram, autocorrelation function (ACF), and partial ACF (PACF) plots, can be used to verify if independence and uniformity are met. Also, the Kolmogorov-Smirnov (KS) and LB tests may be employed to corroborate goodness-of-fit and independence, respectively.

The sequence $\{z_i\}$ for each model is computed out-of-sample; namely, parameter estimates are calculated on the first part of the sample, and then the sequence $\{z_i\}$ is calculated based on the other part. Table 8 reports the p-values of the KS test for the sequence $\{z_i\}$ and the p-values of the LB test for the sequences $\{z_i\}$ and $\{z_i^2\}$ over 4 and 16 lags, where $\{z_i^2\}$ is used to test a possible non-linear serial correlation. From this table, note that the hypothesis of uniformity for the sequence $\{z_i\}$ is confirmed by the results of the KS test for the BSACD1 and BSACD2 models. This confirms the superiority of these models in terms of out-of-sample forecasting ability over the GGACD, LNACD and LtACD models; see Table 8. Moreover, by means of the ACF and PACF plots displayed in Figure 7, independence of the sequence $\{z_i\}$ for the BSACD1 and BSACD2 models is observed, which is confirmed by the p-values of the LB test; see Table 8.

From these results, we can also conclude that the BSACD1 and BSACD2 models have a quite similar out-of-sample forecasting performance and better than the GGACD, LNACD and LtACD models for the BASF-SE data.

Table 8: Out-of-sample test results (KS and LB p-values) for the indicated model with the BASF-SE data.

	p-values				
Model	KS	$Q(4)$	$Q^2(4)$	$Q(16)$	$Q^2(16)$
BSACD1	0.123	0.660	0.676	0.913	0.895
BSACD2	0.124	0.514	0.497	0.519	0.539
GGACD	0.016	0.494	0.363	0.931	0.796
LNACD	<0.001	0.278	0.080	0.780	0.239
LtACD	<0.001	0.227	0.070	0.703	0.230

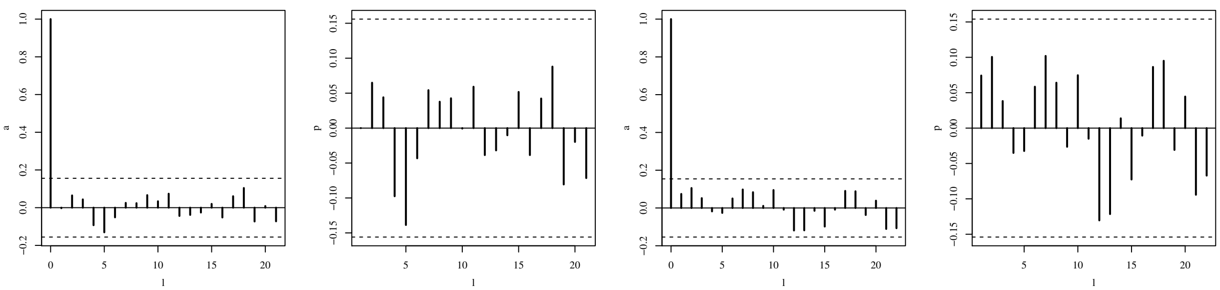


Figure 7: ACF and PACF plots of the probability integral transform $\{z_i\}$ for the BSACD1 (left) and BSACD2 (right) models with the BASF-SE data.

7 Concluding remarks

We have compared and analyzed two Birnbaum-Saunders autoregressive conditional duration models based on the mean and median durations, the former one being a new model to be proposed in the present research. We have considered inference about the model parameters, influence diagnostics and two types of residuals for these models. A Monte Carlo simulation study was carried out to evaluate the behaviour of the maximum likelihood estimators of the corresponding parameters. We have applied the proposed and existing models to a recent real-world data set of financial transactions from the German DAX stock exchange. We have evaluated the global and local influence of atypical cases based on the proposed models for these data. The influence diagnostic study suggested the BSACD2 model is more robust to atypical cases than the BSACD1 model hence making it a more reliable choice in highly unpredictable market conditions. In addition, the forecasting ability of the proposed and existing models based on the traditional and density forecast evaluation techniques has been assessed. In general, the results have shown that the two Birnbaum-Saunders autoregressive conditional duration models have similar performances in terms of model fitting and forecasting ability, and that they outperform the existing models in the literature. As part of future research, it is of interest to study outlier detection taking into account the joint and conditional deletion influence measures based on the Cook distance, as well as propose an outlier detection procedure to evaluate and estimate their effects in ACD models; see Chiang and Wang (2012). In addition, influence diagnostic tools can be extended to more general Birnbaum-Saunders ACD models, such as those based on scale mixture of normals or versions of extreme value; see Leiva et al. (2016a). Furthermore, multivariate models can also be explored; see Marchant et al. (2016a). Work on some of these issues is currently in progress and we hope to report some findings in future papers.

Acknowledgement The authors thank the Editors and reviewers for their constructive comments on an earlier version of this manuscript. The research was partially supported by CNPq and CAPES grants from the Brazilian government and by FONDECYT 1160868 grant from the Chilean government.

References

Aarset, M. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, 36:106–108.

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

Bauwens, L., Giot, P., Joachim, G., and David, V. (2004). A comparison of financial duration models via density forecasts. *International Journal of Forecasting*, 20:589–609.

Belfrage, M. (2015). R package ACDm: Tools for autoregressive conditional duration model. `https://cran.r-project.org/web/packages/ACDm`.

Bhatti, C. (2010). The Birnbaum-Saunders autoregressive conditional duration model. *Mathematics and Computers in Simulation*, 80:2062–2078.

Birnbaum, Z. W. and Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied Probability*, 6:319–327.

Castillo, N., Gómez, H., and Bolfarine, H. (2011). Epsilon Birnbaum-Saunders distribution family: Properties and inference. *Statistical Papers*, 52:871–883.

Chen, C. and Liu, L-M. (1993). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, 88:284–297.

Chiang, M-H. (2007). A smooth transition autoregressive conditional duration model. *Studies in Nonlinear Dynamics and Econometrics*, 11:108–144.

Chiang, M-H. and Wang, L-M. (2012). Additive outlier detection and estimation for the logarithmic autoregressive conditional duration model. *Communications in Statistics: Simulation and Computation*, 41:287–301.

Cook, R. D. (1987). Influence assessment. *Journal of Applied Statistics*, 14:117–131.

Diana, T. (2015). Measuring the impact of traffic flow management on interarrival duration: An application of autoregressive conditional duration. *Journal of Air Transport Management*, 42:219–225.

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

Dionne, G., Pacurar, M., and Zhou, X. (2015). Liquidity-adjusted intraday value at risk modeling and risk management: An application to data from Deutsche Börse. *Journal of Banking and Finance*, 59:202–219.

Duchesne, P. and Pacurar, M. (2008). Evaluating financial time series models for irregularly spaced data: A spectral density approach. *Computers and Operations Research*, 35:130–155.

Dufour, A. and Engle, R. F. (2000). The ACD model: Predictability of the time between consecutive trades. Technical Report 2000-05, University of Reading, Reading, UK.

Dunn, P. and Smyth, G. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, 5:236–244.

Efron, B. and Hinkley, D. (1978). Assessing the accuracy of the maximum likelihood estimator: Observed vs. expected Fisher information. *Biometrika*, 65:457–487.

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

Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society B*, pages 350–363.

Garcia, F., Uribe, M., Leiva, V., and Aykroyd, R. (2016). Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. *Stochastic Environmental Research and Risk Assessment*, in press .

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

Hubert, M., Vanderveeken, S. (2008). Outlier detection for skewed data. *Journal of Chemometrics*, 22:235–246.

Jin, X. and Kawczak, J. (2003). Birnbaum-Saunders and lognormal kernel estimators for modelling durations in high frequency financial data. *Annals of Economics and Finance*, 4:103–124.

Kundu, D., Kannan, N., and Balakrishnan, N. (2008). On the hazard function of Birnbaum-Saunders distribution and associated inference. *Computational Statistics and Data Analysis*, 52:2692–2702.

Lawrence, A. J. (1995). Deletion influence and masking in regression. *Journal of the Royal Statistical Society B*, 57:181–189.

Leao, J., Leiva, V., Saulo, H., and Tomazella, V. (2017). Birnbaum-Saunders frailty regression models: Diagnostics and application to medical data. *Biometrical Journal*, page in press.

Leiva, V., Ferreira, M., Gomes, M., Lillo, C. (2016a). Extreme value Birnbaum-Saunders regression models applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, 30:1045–1058.

Leiva, V., Marchant, C., Ruggeri, F., and Saulo, H. (2015). A criterion for environmental assessment using Birnbaum-Saunders attribute control charts. *Environmetrics*, 26:463–476.

Leiva, V., Ruggeri, F., Saulo, H., and Vivanco, J. F. (2017). A methodology based on the Birnbaum-Saunders distribution for reliability analysis applied to nano-materials. *Reliability Engineering and System Safety*, 157:192–201.

Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., and Barros, M. (2014a). Birnbaum-Saunders statistical modelling: A new approach. *Statistical Modelling*, 14:21–48.

Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., and Barros, M. (2016b). A methodology for stochastic inventory models based on a zero-adjusted Birnbaum-Saunders distribution. *Applied Stochastic Models in Business and Industry*, 32:74–89.

Leiva, V., Saulo, H., Leão, J., and Marchant, C. (2014b). A family of autoregressive conditional duration models applied to financial data. *Computational Statistics and Data Analysis*, 79:175–191.

Lesaffre, E. and Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, 54:570–582.

Lio, Y. L., Tsai, T. R., and Wu, S. J. (2010). Acceptance sampling plans from truncated life tests based on the Birnbaum-Saunders distribution for percentiles. *Communications in Statistics: Simulation and Computation*, 39:119–136.

Liu, S. (2000). On local influence in elliptical linear regression models. *Statistical Papers*, 41:211–224.

Liu, S. and Heyde, C. C. (2008). On estimation in conditional heteroskedastic time series models under non-normal distributions. *Statistical Papers*, 49:455–469.

Marchant, C., Leiva, V., and Cysneiros, F. J. A. (2016a). A multivariate log-linear model for Birnbaum-Saunders distributions. *IEEE Transactions on Reliability*, 65:816–827.

Marchant, C., Leiva, V., Cysneiros, F. J. A., and Vivanco, J. F. (2016b). Diagnostics in multivariate generalized Birnbaum-Saunders regression models. *Journal of Applied Statistics*, 43:2829–2849.

Meitz, M. and Terasvirta, T. (2006). Evaluating models of autoregressive conditional duration. *Journal of Business and Economic Statistics*, 24:104–12.

Mittelhammer, R. C., Judge, G. G., and Miller, D. J. (2000). *Econometric Foundations*. Cambridge University Press, New York, US.

Pacurar, M. (2008). Autoregressive conditional durations models in finance: A survey of the theoretical and empirical literature. *Journal of Economic Surveys*, 22:711–751.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Santos-Neto, M., Cysneiros, F. J. A., Leiva, V., and Barros, M. (2016). Reparameterized Birnbaum-Saunders regression models with varying precision. *Electronic Journal of Statistics*, 10:2825–2855.

Saulo, H., Leiva, V., Ziegelmann, F. A., and Marchant, C. (2013). A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, 27:1479–1491.

Tsay, R. (2009). Autoregressive conditional duration models. In Mills, T.C. and Patterson, K., editors, *Handbook of Econometrics*, volume 2 (Applied Econometrics), pages 1004–1024. Palgrave MacMillan, UK.

Wanke, P. and Leiva, V. (2015). Exploring the potential use of the Birnbaum-Saunders distribution in inventory management. *Mathematical Problems in Engineering*, Article ID 827246:1–9.

Zevallos, M., Santos, B., and Hotta, L. K. (2012). A note on influence diagnostics in AR (1) time series models. *Journal of Statistical Planning and Inference*, 142:2999–3007.