



UNIVERSITY OF LEEDS

This is a repository copy of *Periodic autoregressive conditional duration*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/172378/>

Version: Accepted Version

Article:

Aknouche, A, Almohaimeed, B and Dimitrakopoulos, S orcid.org/0000-0002-0043-180X
(2021) Periodic autoregressive conditional duration. *Journal of Time Series Analysis*. ISSN 0143-9782

<https://doi.org/10.1111/jtsa.12588>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Periodic autoregressive conditional duration

Abdelhakim Aknouche^{*†}, Bader Almohaimed^{*}, Stefanos Dimitrakopoulos^{‡ §}

Abstract

We propose an autoregressive conditional duration (ACD) model with periodic time-varying parameters and multiplicative error form. We name this model periodic autoregressive conditional duration (PACD). First, we study the stability properties and the moment structures of it. Second, we estimate the model parameters, using (profile and two-stage) Gamma quasi-maximum likelihood estimates (QMLEs), the asymptotic properties of which are examined under general regularity conditions. Our estimation method encompasses the exponential QMLE, as a particular case. The proposed methodology is illustrated with simulated data and two empirical applications on forecasting Bitcoin trading volume and realized volatility. We found that the PACD produces better in-sample and out-of-sample forecasts than the standard ACD.

Keywords: Positive time series, autoregressive conditional duration, periodic time-varying models, multiplicative error models, exponential QMLE, two-stage Gamma QMLE.

1 Introduction

Recent research in time series analysis tends to avoid transforming original data prior to modeling and prefers to represent them directly through models that take into account the actual support of their distributions. Such an approach parallels to that of *generalized linear models* (GLM) for independent data (McCullagh and Nelder, 1989). In this way, numerous time series models with “*specific values*” have, recently, received great interest, such as integer-valued models, including count and binary specifications, and positive-valued models.

^{*}College of Science, Qassim University, Saudi Arabia

[†]Faculty of Mathematics, University of Science and Technology Houari Boumediene, Algeria

[‡]Economics Division, Leeds University, UK

[§]Correspondence to: Stefanos Dimitrakopoulos (s.dimitrakopoulos@leeds.ac.uk)

A well-known model for positive-valued time series data is the *autoregressive conditional duration* (ACD), introduced by Engle and Russell (1998). Originally designed to model durations between financial events in high-frequency microstructure markets, the ACD model is also useful for modeling a broad range of positive data, such as regularly-spaced return range series (Chou, 2005), daily realized volatility (Lanne, 2006; Zheng et al, 2015; Aknouche and Francq, 2019) and trading volume (Li, 2020; Aknouche and Francq, 2020). Various generalizations of the ACD model have been proposed to take into account additional facts of positive time series data (Pacurar, 2008; Hautsch, 2012; Bhogal and Variyam, 2019).

As in the case of GARCH models, it has been documented that the high persistence observed in empirical studies utilizing the standard ACD specification, is in fact artificial and can be avoided by considering ACD models with time-varying parameters (Diebold, 1986; Andersen and Bollerslev, 1997; Mikosch and Starica, 2004; Hujer and Vuletic, 2007; Caporin et al, 2017; Gallo and Ortanto, 2018). In this paper, we are dealing with time-varying ACD models, by proposing an ACD model, the parameters of which are allowed to evolve periodically over time. We name this model periodic autoregressive conditional duration (PACD).

Such a model aims to represent seasonally varying positive-valued series. The observed process is defined as the product of a unit mean *independent and periodically distributed* innovation process with the conditional mean of the model having a GARCH-type specification with periodic time-varying parameters. We first study the stability properties of the PACD model, such as the existence of periodically stationary and ergodic solutions with finite moments or log-moments. Such properties are needed in the estimation stage, which is the second contribution of this paper.

To estimate the model parameters, the exponential quasi-maximum likelihood (EQMLE) is used, since it is well-adapted to the support of the distribution of the data, and it does not require specifying a distribution for the periodically distributed innovation sequence. However, because of the periodicity of that sequence, the EQMLE may be less efficient than the Gamma QMLE (GQMLE) which, in fact, can account for the periodicity of the model innovation.

Consequently, we propose a two-stage Gamma QMLE (2S-GQMLE) which i) utilizes the EQMLE (or a profile GQMLE) in the first stage, ii) estimates the variance innovations, and then iii) uses the latter as a by-product in the second stage of the computation of the GQMLE.

Consistency and asymptotic normality (CAN) of the proposed QMLEs are established and the relative efficiency of the 2S-GQMLE is studied for some specific conditional distributions.

The PACD can be used to model various seasonal positive-valued phenomena (realized volatility, trading volumes and transaction rates). The day-of-the-week pattern may be present in all these phenomena, which means that each day of the week may have its own distribution (Franses and Paap, 2000; Boynton et al, 2009; Tsiakas, 2006; Charles, 2010). In that sense, a time-invariant ACD model for daily data is just an average model that does not take into account the specificities of the underlying measures across days. Other examples of non-financial series that may be characterized by periodicity are wind power and wind speed series (Ambach and Croonenbroeck, 2015; Ambach and Schmid, 2015; Ziel et al, 2016).

Our empirical applications concern Bitcoin trading volume data and the UN realized volatility. Both series are likely to be characterized by the day-of-the-week effect and we show that the PACD produces better in-sample and out-of-sample forecasts than the benchmark ACD.

The rest of this paper is outlined as follows. In Section 2 we define the PACD and some special cases of it, and describe the link/relationship between the PACD and the periodic GARCH of Bollerslev and Ghysels (1996). In Section 3 we derive the stability conditions of our model. In Section 4, various Gamma QMLEs are proposed and their asymptotic properties are studied. In section 5 we conduct a simulation study and in section 6 we present the empirical results from two series (Bitcoin trading volume and UN realized volatility). An online supplementary material accompanies this paper that contains all the proofs and some additional data results.

2 Periodic Autoregressive Conditional Duration model

All random variables and processes in this paper are defined on a probability space (Ω, \mathcal{F}, P) and valued in the set of positive real numbers $\mathbb{R}_+ = (0, \infty)$, which is endowed with the Borel field $\mathcal{B}(\mathbb{R}_+)$. Let $S \geq 1$ be a positive integer called the period, and $\omega_t^0, \alpha_{t1}^0, \dots, \alpha_{tq}^0, \beta_{t1}^0, \dots, \beta_{tp}^0$ ($p, q \in \mathbb{N} = \{0, 1, \dots\}$) be positive real parameters S -periodic over time, i.e. $\omega_t^0 = \omega_{t+kS}^0, \alpha_{ti}^0 = \alpha_{t+kS,i}^0$ ($i = 1, \dots, q$) and $\beta_{tj}^0 = \beta_{t+kS,j}^0$ ($j = 1, \dots, p$) for all integers k and t . Let also $\{\xi_t, t \in \mathbb{Z}\}$ be a sequence of positive random variables with $E\xi_t = 1$ for all t , and a finite $Var(\xi_t) = \sigma_{0t}^2 > 0$. Assume that $\{\xi_t, t \in \mathbb{Z}\}$ is independent and S -periodically distributed in the sense

that $\{\xi_t, t \in \mathbb{Z}\}$ is an independent sequence and $\xi_t \stackrel{D}{=} \xi_{t+S}$ for all t , where $\stackrel{D}{=}$ denotes equality in distribution.

A positive-valued stochastic process $\{Y_t, t \in \mathbb{Z}\}$ is said to be an MEM (multiplicative error model; Engle, 2002) *periodic autoregressive conditional duration* with orders p and q (henceforth PACD(p, q)) if Y_t is given for all $t \in \mathbb{Z}$ by

$$Y_t = \psi_t \xi_t \quad (2.1a)$$

and

$$\psi_t = \omega_t^0 + \sum_{i=1}^q \alpha_{ti}^0 Y_{t-i} + \sum_{j=1}^p \beta_{tj}^0 \psi_{t-j}, \quad (2.1b)$$

where the innovation term ξ_t is independent of ψ_{t-j} for all $j \geq 1$. To ensure the almost sure (a.s.) positivity of ψ_t , it is assumed that $\omega_t^0 > 0$, $\alpha_{ti}^0 \geq 0$, and $\beta_{tj}^0 \geq 0$, for all $t \in \mathbb{Z}$, $i = 1, \dots, q$ and $j = 1, \dots, p$. To emphasize the periodicity of the model, let $t = nS + v$ for $n \in \mathbb{Z}$ and $0 \leq v \leq S - 1$. Then, equation (2.1b) can be written as follows

$$\psi_{nS+v} = \omega_v^0 + \sum_{i=1}^q \alpha_{vi}^0 Y_{nS+v-i} + \sum_{j=1}^p \beta_{vj}^0 \psi_{nS+v-j}, \quad n \in \mathbb{Z}, \quad 0 \leq v \leq S - 1, \quad (2.1c)$$

where by season or channel v ($0 \leq v \leq S - 1$) (sometimes in this paper we rather write $1 \leq v \leq S$) we denote the set $\{\dots, v - S, v, v + S, v + 2S, \dots\}$ with corresponding parameters ω_v^0 , α_{vi}^0 , β_{vj}^0 and $\sigma_{0v}^2 = \text{Var}(\xi_{nS+v})$. Let \mathcal{F}_t be the σ -Algebra generated by $\{Y_{t-i}, i \geq 0\}$. The conditional mean and conditional variance of the model (2.1) are given respectively by

$$E(Y_t | \mathcal{F}_{t-1}) = \psi_t \quad (2.2a)$$

and

$$\text{Var}(Y_t | \mathcal{F}_{t-1}) = \sigma_{0t}^2 \psi_t^2. \quad (2.2b)$$

The PACD model, thus, follows the quadratic variance-to-mean relationship (i.e. the conditional variance is proportional to the squared conditional mean), where $\sigma_{0t}^2 > 0$ is the variance of ξ_t and is S -periodic by construction (from the *ipd_S* property of the innovation sequence $\{\xi_t, t \in \mathbb{Z}\}$). The specification (2.1) is an MEM in the sense of Engle (2002), but the condi-

tional mean equation (2.1b) has rather periodic time-varying coefficients. For $S = 1$, model (2.1) reduces to the standard autoregressive conditional duration (ACD in short) of Engle and Russell (1998). No specification for the distribution of $\{\xi_t, t \in \mathbb{Z}\}$ is imposed apart the semi-parametric quadratic variance-to-mean function (2.2b). However, a useful family of conditional distributions satisfying (2.2b) is the Gamma distribution with shape $\frac{1}{\sigma_{0t}^2}$ and scale $\frac{1}{\sigma_{0t}^2 \psi_t}$, that is

$$Y_t | \mathcal{F}_{t-1} \sim \Gamma \left(\frac{1}{\sigma_{0t}^2}, \frac{1}{\sigma_{0t}^2 \psi_t} \right), \quad (2.3)$$

where ψ_t satisfies (2.1b). In the latter case, the innovation term ξ_t in (2.1) will be marginally Gamma distributed

$$\xi_t \sim \Gamma \left(\frac{1}{\sigma_{0t}^2}, \frac{1}{\sigma_{0t}^2} \right), \quad (2.4)$$

and the process defined by (2.3) is called Gamma PACD(p, q). A notable particular case of model (2.3) appears when the variance $\sigma_{0t}^2 \equiv 1$ is constant, so $\xi_t \sim \Gamma(1, 1)$, which corresponds to the *exponential* PACD. Another useful and versatile family of conditional distributions we use in this paper is the Beta prime (BP) distribution. Recall that a nonnegative random variable Y is said to have a BP distribution with shape parameters $a > 0$ and $b > 0$, and scale parameter $c > 0$ ($Y \sim BP(a, b, c)$) if its probability density function is given by

$$f(y; a, b, c) = \frac{1}{c \text{Beta}(a, b)} \left(\frac{y}{c} \right)^{a-1} \left(1 + \frac{y}{c} \right)^{-a-b}, \quad y \geq 0,$$

where $\text{Beta}(a, b)$ is the classical Beta function (cf. Johnson et al, 1995). In contrast with the Gamma distribution, the BP distribution allows dealing with zero values and can be fat-tailed (Moghaddam et al, 2019). The mean and variance of Y are respectively $EY = \frac{ac}{b-1}$ ($b > 1$) and $\text{Var}(Y) = \frac{c^2 a(a+b-1)}{(b-2)(b-1)^2}$ ($b > 2$). When $c = 1$, the $BP(a, b, 1)$ is called the standard BP distribution and simply writes $BP(a, b)$. Naturally if $X \sim BP(a, b)$ then for any $c > 0$, $Y = cX \sim BP(a, b, c)$. Thus if the innovation ξ_t is BP distributed with the following parametrization

$$\xi_t \sim BP(2\sigma_{0t}^{-2} + 1, 2\sigma_{0t}^{-2} + 2),$$

then $E\xi_t = 1$ and $\text{Var}(\xi_t) = \sigma_{0t}^2$ as required by (2.1). Moreover, the conditional distribution of

$Y_t = \psi_t \xi_t$ would be

$$Y_t | \mathcal{F}_{t-1} \sim BP(2\sigma_{0t}^{-2} + 1, 2\sigma_{0t}^{-2} + 2, \psi_t).$$

Note also that $E\xi_t^2 < \infty$ when $\sigma_{0t}^2 > 0$, $E\xi_t^3 < \infty$ when $0 < \sigma_{0t}^2 < 2$, and $E\xi_t^4 < \infty$ when $0 < \sigma_{0t}^2 < 1$.

As in the time-invariant case, the periodic ACD model can be seen as a squared periodic GARCH (PGARCH) model as proposed by Bollerslev and Ghysels (1996). Indeed, consider the following real-valued PGARCH(p, q) process given by

$$X_t = \sqrt{h_t} \eta_t \tag{2.5a}$$

and

$$h_t = \omega_t^0 + \sum_{i=1}^q \alpha_{ti}^0 X_{t-i}^2 + \sum_{j=1}^p \beta_{tj}^0 h_{t-j}, \tag{2.5b}$$

where $\{\eta_t, t \in \mathbb{Z}\}$ is an ipd_S sequence with mean zero and unit variance, and the parameters ω_t^0 , α_{ti}^0 and β_{tj}^0 are defined as above. It is clear that the squared PGARCH process defined by $Y_t = X_t^2$ ($t \in \mathbb{Z}$) satisfies the PACD equation (2.1) with $\xi_t = \eta_t^2$ and $\psi_t = h_t$. Conversely, let $\{Y_t, t \in \mathbb{Z}\}$ be a PACD model given by (2.1), and assume $\{z_t, t \in \mathbb{Z}\}$ is an independent and identically distributed (iid) sequence uniformly distributed in $\{-1, 1\}$ (see also Francq and Zakoian, 2019 for the non-periodic case $S = 1$). Assume $\{z_t, t \in \mathbb{Z}\}$ and $\{\xi_t, t \in \mathbb{Z}\}$ are independent and define the process $\{X_t, t \in \mathbb{Z}\}$ by

$$X_t = z_t \sqrt{Y_t} = \sqrt{h_t} \eta_t,$$

where $h_t = \psi_t$ satisfies (2.5b) and $\eta_t = z_t \sqrt{\xi_t}$ is a term of an ipd_S sequence. Hence $\{X_t, t \in \mathbb{Z}\}$ is a PGARCH model in the sense (2.5). Note finally that a PACD model admits a weak periodic ARMA (PARMA) (Lund and Basawa, 2000; Francq et al, 2011). Setting $Y_t = \psi_t + \varepsilon_t$, the process $\{Y_t, t \in \mathbb{Z}\}$ may be written in the following PARMA

$$Y_t = \omega_t^0 + \sum_{i=1}^{\max(p,q)} (\alpha_{ti}^0 + \beta_{ti}^0) Y_{t-i} + \varepsilon_t - \sum_{j=1}^p \beta_{tj}^0 \varepsilon_{t-j},$$

where

$$\varepsilon_t = Y_t - E(Y_t | \mathcal{F}_{t-1}) = \psi_t (\xi_t - 1) \quad (2.6)$$

is a zero-mean term of a martingale difference sequence with a finite periodic variance $E\varepsilon_t^2 = E(\xi_t - 1)^2 E\psi_t^2 = \sigma_{0t}^2 E\psi_t^2$.

A more general PACD, which is not necessarily MEM is defined through a conditional distribution of the form

$$Y_t | \mathcal{F}_{t-1} \sim F_{\psi_t}, \quad (2.7)$$

where F_ψ is a cumulative probability distribution (with positive support) with mean ψ , and ψ_t is given by (2.1b). Notice finally that our PACD is a particular case of the more general time-dependent ACD (tdACD) in which the conditional mean and variance parameters are time-varying, but not necessarily function of the sample size. Time dependence here is understood in the sense of Miller (1968) and Hallin (1984). In a different perspective, Bortoluzzo et al (2010) proposed a class of *slowly varying* tdACD models (denoted by tvACD) in which the coefficients are time-varying but depend on the sample size similarly to Dahlhaus and Subba Rao (2006). The tvACD process is locally stationary in the sense of Dahlhaus (1997). Mishra and Ramanathan (2010) established some probabilistic and asymptotic results for the tvACD model. Our PACD model, however, is not nested within the tvACD model.

3 Periodic ergodicity and finite moment conditions

We now give necessary and/or sufficient conditions for model (2.1) to be strictly periodically stationary and *periodically ergodic*. Such properties are recalled in the Supplementary material. We also consider conditions for the existence of finite moments. Combining (2.1a) and (2.1b) we obtain the following stochastic recurrence equation (SRE)

$$\underline{Y}_{nS+v} = A_{nS+v}^0 \underline{Y}_{nS+v-1} + B_{nS+v}^0, \quad (3.1)$$

driven by the ipd_S sequence $\{(A_{nS+v}^0, B_{nS+v}^0), n \in \mathbb{Z}, 0 \leq v \leq S-1\}$, where $\underline{Y}_{nS+v} = (Y_{nS+v}, \dots, Y_{nS+v-q+1}, \psi_{nS+v}, \dots, \psi_{nS+v-p+1})'$, $B_{nS+v}^0 = (\omega_v^0 \xi_{nS+v}, 0_{(q-1) \times 1}, \omega_v^0, 0_{(p-1) \times 1})'$, and

$$A_{nS+v}^0 = \begin{pmatrix} \alpha_{v1}^0 \xi_{nS+v} & \cdots & \alpha_{v,q-1}^0 \xi_{nS+v} & \alpha_{vq}^0 \xi_{nS+v} & \beta_{v1}^0 \xi_{nS+v} & \cdots & \beta_{v,p-1}^0 \xi_{nS+v} & \beta_{vp}^0 \xi_{nS+v} \\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_{v1}^0 & \cdots & \alpha_{v,q-1}^0 & \alpha_{vq}^0 & \beta_{v1}^0 & \cdots & \beta_{v,p-1}^0 & \beta_{vp}^0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

$0_{m \times n}$ being the null matrix of dimension $m \times n$. Let

$$\gamma^S = \inf \left\{ \frac{1}{n} E \log \|A_{nS}^0 \dots A_2^0 A_1^0\|, n \geq 1 \right\}$$

be the top Lyapunov exponent associated with the ipd_S -driven SRE (3.1) (Aknouche et al, 2020). Let also

$$\beta_{nS+v}^0 = \begin{pmatrix} \beta_{v1}^0 & \cdots & \beta_{v,p-1}^0 & \beta_{vp}^0 \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix},$$

and denote by $\rho(A^0)$ the spectral radius of the squared matrix A^0 , i.e. the maximum modulus of the eigenvalues of A^0 . The following result gives the conditions for equation (3.1) to have a unique strictly periodically stationary and periodically ergodic solution.

Theorem 3.1 i) *Assume $E(\log(\xi_v)) < \infty$ for all $1 \leq v \leq S$. A necessary and sufficient condition for model (2.1) to have a unique nonanticipative strictly periodically stationary and periodically ergodic solution is that*

$$\gamma^S < 0. \tag{3.2}$$

Such a solution is given for all $n \in \mathbb{Z}$ and $0 \leq v \leq S - 1$ by

$$\underline{Y}_{nS+v} = \sum_{j=0}^{\infty} \prod_{i=0}^{j-1} A_{nS+v-i}^0 B_{nS+v-j}^0, \quad (3.3)$$

where the series in the right hand side of (3.3) converges absolutely almost surely.

ii) If (2.1) admits a strictly periodically stationary solution then

$$\rho \left(\prod_{v=0}^{S-1} \beta_{S-v}^0 \right) < 1. \quad (3.4)$$

In the special case where $p = q = 1$, the periodic stationarity condition (3.2) is simplified as follows

$$\sum_{v=1}^S E (\log (\alpha_v^0 \xi_{v-1} + \beta_v^0)) < 0,$$

while (3.4) reduces to $\prod_{v=1}^S \beta_v^0 < 1$.

Conditions for the existence of moments of the $PACD(p, q)$ process are given as follows.

Theorem 3.2 Assume $E\xi_v < \infty$ for all $1 \leq v \leq S$. A sufficient condition for the process given by (2.1) to be strictly periodically stationary and periodically ergodic with $EY_v < \infty$ ($1 \leq v \leq S$) is that

$$\rho \left(\prod_{v=0}^{S-1} E (A_{S-v}^0) \right) < 1. \quad (3.5)$$

Some remarks are in order:

- In the case, where $S = 1$, the conditional mean coefficients are time-invariant, that is $\omega_v^0 = \omega^0$, $\alpha_{vi}^0 = \alpha_i^0$ and $\beta_{vj}^0 = \beta_j^0$. Therefore, using a similar device by Chen and An (1998), (3.5) reduces to the following stationarity in mean condition

$$\sum_{i=1}^q \alpha_i^0 + \sum_{j=1}^p \beta_j^0 < 1$$

as provided by Engle and Russell (1998).

- When $p = q = 1$, the periodic stationarity-in-mean condition (3.5) is equivalent to the following condition

$$\prod_{v=1}^S (\alpha_v^0 + \beta_v^0) < 1. \quad (3.6)$$

As many time-varying models, the PACD(1, 1) model may be locally unstable, i.e. $\alpha_v^0 + \beta_v^0 \geq 1$ for some season v , but even though stable in the sense (3.6).

Theorem 3.3 i) Under (3.2) there exists $\kappa > 0$ such that for all $1 \leq v \leq S$

$$E\psi_v^\kappa < \infty \quad \text{and} \quad EY_v^\kappa < \infty. \quad (3.7)$$

ii) Let $\{Y_t, t \in \mathbb{Z}\}$ be a strictly periodically stationary solution of (2.1) and assume that $E\xi_v^m$ ($m \in \mathbb{N}^*$) is finite for all $1 \leq v \leq S$. A sufficient condition for EY_v^m to be finite (for all $1 \leq v \leq S$) is that

$$\rho \left(\prod_{v=0}^{S-1} E \left(A_{S-v}^{0 \otimes m} \right) \right) < 1, \quad (3.8)$$

where $A^{0 \otimes m}$ is the Kronecker product: $A^0 \otimes A^0 \otimes \cdots \otimes A^0$ with m factors.

In the special case of Gamma PACD with $p = q = 1$, explicit conditions equivalent to (3.8) can be given. These conditions are also necessary for the existence of finite moments. For the next Proposition and its proof, to simplify the notation, we will denote the parameters ω_v^0 , α_v^0 and β_v^0 by ω_{0v} , α_{0v} and β_{0v} , respectively.

Proposition 3.1 The Gamma PACD (1, 1) model (2.3) admits a unique nonanticipative periodically ergodic solution $\{Y_t, t \in \mathbb{Z}\}$ such that:

i) $EY_v < \infty$ ($1 \leq v \leq S$) if and only if (3.6) holds.

ii) $EY_v^2 < \infty$ ($1 \leq v \leq S$) if and only if $E\xi_v^2 < \infty$ ($1 \leq v \leq S$), (3.6) and

$$\prod_{v=1}^S (\alpha_{0v}^2 E\xi_{v-1}^2 + 2\alpha_{0v}\beta_{0v} + \beta_{0v}^2) < 1. \quad (3.9)$$

iii) $EY_v^3 < \infty$ ($1 \leq v \leq S$) if and only if $E\xi_v^3 < \infty$ ($1 \leq v \leq S$), (3.6), (3.9) and

$$\prod_{v=1}^S (\alpha_{0v}^3 E\xi_{v-1}^3 + 3(\sigma_{0,v-1}^2 + 1)\alpha_{0v}^2\beta_{0v} + 3\alpha_{0v}\beta_{0v}^2 + \beta_{0v}^3) < 1. \quad (3.10)$$

iv) $EY_v^4 < \infty$ ($1 \leq v \leq S$) if and only if $E\xi_v^4 < \infty$ ($1 \leq v \leq S$), (3.6), (3.9), (3.10) and the following hold

$$\prod_{v=1}^S (\alpha_{0v}^4 E\xi_{v-1}^4 + 4(1 + \sigma_{0,v-1}^2)(1 + 2\sigma_{0,v-1}^2)\alpha_{0v}^3\beta_{0v} + 6(1 + \sigma_{0,v-1}^2)\alpha_{0v}^2\beta_{0v}^2 + 4\alpha_{0v}\beta_{0v}^3 + \beta_{0v}^4) < 1. \quad (3.11)$$

For the particular exponential PACD(1, 1) model, $Y_{nS+v} | \mathcal{F}_{nS+v-1} \sim \Gamma\left(1, \frac{1}{\psi_{nS+v}}\right)$, just replace in Proposition 3.1 the moments $E\xi_v^2$, $E\xi_v^3$ and $E\xi_v^4$ by 1, 6 and 24 respectively, and σ_{0v}^2 by 1 for all $1 \leq v \leq S$.

4 Gamma quasi-maximum likelihood estimates

Let Y_1, Y_2, \dots, Y_T be a series generated from the PACD(p, q) model, which we can rewrite in the following form

$$\begin{aligned} Y_{nS+v} &= \psi_{nS+v} \xi_{nS+v} \\ \psi_{nS+v} &= \psi_{nS+v}(\theta_0) = \omega_v^0 + \sum_{i=1}^q \alpha_{vi}^0 Y_{nS+v-i} + \sum_{j=1}^p \beta_{vj}^0 \psi_{nS+v-j} \end{aligned} \quad n \in \mathbb{Z}, 0 \leq v \leq S-1, \quad (4.1)$$

where the true parameter $\theta_0 = (\theta_1^{0'}, \theta_2^{0'}, \dots, \theta_S^{0'})'$ with $\theta_v^0 = (\omega_v^0, \alpha_{v1}^0, \dots, \alpha_{vq}^0, \beta_{v1}^0, \dots, \beta_{vp}^0)'$ ($1 \leq v \leq S$) belongs to a parameter space $\Theta \subset \left((0, \infty) \times [0, \infty)^{(p+q)}\right)^S$. The true innovation variance parameter, which is denoted by $\boldsymbol{\sigma}_0^2 = (\sigma_{01}^2, \dots, \sigma_{0S}^2)'$, belongs to a parametric space $\Delta \subset (0, \infty)^S$. The sample size $T = NS$ ($N \geq 1$) is assumed without loss of generality a multiple of S . Given initial values $Y_0, \dots, Y_{1-q}, \tilde{\psi}_0, \dots, \tilde{\psi}_{1-p}$ and a generic parameter $\theta \in \Theta$ define

$$\tilde{\psi}_{nS+v}(\theta) = \omega_v + \sum_{i=1}^q \alpha_{vi} Y_{nS+v-i} + \sum_{j=1}^p \beta_{vj} \tilde{\psi}_{nS+v-j}(\theta), \quad 1 \leq v \leq S, n \geq 0, \quad (4.2a)$$

as an observable proxy for $\psi_{nS+v}(\theta)$. The latter is defined as a periodically stationary solution of the following generic model ($\theta \in \Theta$)

$$\psi_{nS+v}(\theta) = \omega_v + \sum_{i=1}^q \alpha_{vi} Y_{nS+v-i} + \sum_{j=1}^p \beta_{vj} \psi_{nS+v-j}(\theta), \quad 0 \leq v \leq S-1, n \in \mathbb{Z}. \quad (4.2b)$$

4.1 Exponential and profile Gamma QMLEs

The true conditional distribution of (4.1) is unknown due to the unspecification of the law of ξ_v ($1 \leq v \leq S$). Thus, a quasi-maximum likelihood estimate (QMLE) which does not require any precise knowledge of the conditional distribution is suitable for estimating the parameter θ_0 involved in the conditional mean. Among many possible QMLEs, the one computed on the basis of the exponential distribution (EQMLE in short) is especially useful for positive

duration data because it reduces to the maximum likelihood estimate when ξ_v is exponentially distributed (Aknouche and Francq, 2020). A more general QMLE, which can be more efficient than the EQMLE in the periodic time-varying innovation context, is the one computed on the basis of the Gamma distribution with arbitrary fixed variance parameters. Let $(\sigma_t^2)_t$ be fixed known positive numbers, S -periodic over t , i.e. $\sigma_{nS+v}^2 = \sigma_v^2$, for all $n \in \{0, \dots, N-1\}$. (This does not be confused with the unknown true σ_{0t}^2). The profile Gamma likelihood associated with $\boldsymbol{\sigma}^2 = (\sigma_1^2, \dots, \sigma_S^2)' > \mathbf{0}$ is given, ignoring constants, by

$$\tilde{L}_T(\theta, \boldsymbol{\sigma}^2) = \frac{1}{T} \sum_{t=1}^T \tilde{l}_t(\theta, \sigma_t^2) = \frac{1}{NS} \sum_{v=1}^S \sum_{n=0}^{N-1} \tilde{l}_{nS+v}(\theta, \sigma_v^2), \quad (4.3a)$$

$$\tilde{l}_{nS+v}(\theta, \sigma_v^2) = \frac{1}{\sigma_v^2} \left(\frac{Y_{nS+v}}{\tilde{\psi}_{nS+v}(\theta)} + \log \tilde{\psi}_{nS+v}(\theta) \right). \quad (4.3b)$$

The profile Gamma QMLE (GQMLE) $\hat{\theta}_G$ of θ_0 is, then, the minimizer of $\tilde{L}_T(\theta, \boldsymbol{\sigma}^2)$ over Θ ,

$$\hat{\theta}_G = \arg \min_{\theta \in \Theta} \tilde{L}_T(\theta, \boldsymbol{\sigma}^2), \quad (4.4)$$

for some arbitrarily fixed and known $\boldsymbol{\sigma}^2$.

When $\boldsymbol{\sigma}^2 = (1, \dots, 1)'$, the GQMLE defined by (4.4) reduces to the EQMLE denoted by $\hat{\theta}_E = \arg \min_{\theta \in \Theta} \tilde{L}_T(\theta, \mathbf{1})$ (Aknouche and Francq, 2020), that is

$$\hat{\theta}_E = \arg \min_{\theta \in \Theta} \frac{1}{NS} \sum_{v=1}^S \sum_{n=0}^{N-1} \frac{Y_{nS+v}}{\tilde{\psi}_{nS+v}(\theta)} + \log \tilde{\psi}_{nS+v}(\theta).$$

Let $\gamma^S(A^0)$ be the top Lyapunov exponent associated with $(A_{nS+v}^0)_{n,v}$ in (3.1). Let also $\boldsymbol{\beta}_v$ defined as in (3.1) with β_{vj} in place of β_{vj}^0 . To establish the strong consistency of $\hat{\theta}_G$ we need the following assumptions.

A1 $\gamma^S(A^0) < 0$ and $\forall \theta \in \Theta$, $\rho \left(\prod_{v=0}^{S-1} \boldsymbol{\beta}_{S-v} \right) < 1$.

A2 $\theta_0 \in \Theta$ and Θ is compact.

A3 The polynomials $\alpha_v^0(z) = \sum_{i=1}^q \alpha_{vi}^0 z^i$ and $\beta_v^0(z) = 1 - \sum_{j=1}^p \beta_{vj}^0 z^j$ have no common root, $\alpha_v^0(1) \neq 0$, and $\alpha_{vq}^0 + \beta_{vp}^0 \neq 0$ for all $1 \leq v \leq S$.

A4 ξ_v is non-degenerate and $E\xi_v = 1$ for all $1 \leq v \leq S$.

As seen in Section 3, $\gamma^S(A^0) < 0$ in **A1** ensures periodic stationarity and periodic ergodicity of the PACD model (4.1). The condition $\rho \left(\prod_{v=0}^{S-1} \boldsymbol{\beta}_{S-v} \right) < 1$ is imposed for the invertibility of

equation (4.2b) for any $\theta \in \Theta$. The compactness assumption **A2** is standard while **A3** and **A4** are made to guarantee the identifiability of the model.

Theorem 4.1 Let $(\widehat{\theta}_G)$ be a sequence of GQMLEs defined by (4.4). Under **A1-A4**,

$$\widehat{\theta}_G \rightarrow \theta_0 \text{ a.s. as } N \rightarrow \infty \text{ for all } \boldsymbol{\sigma}^2 > \mathbf{0}.$$

Turn now to the asymptotic normality property of $\widehat{\theta}_G$. The following assumptions are to be considered.

A5 θ_0 belongs to the interior of Θ .

A6 $E\xi_v^2 = \sigma_{0v}^2 \in (0, \infty)$ for all $1 \leq v \leq S$, the matrices

$$I(\theta_0, \boldsymbol{\sigma}_0^2, \boldsymbol{\sigma}^2) = \sum_{v=1}^S \frac{\sigma_{0v}^2}{\sigma_v^4} E \left(\frac{1}{\psi_v^2(\theta_0)} \frac{\partial \psi_v(\theta_0)}{\partial \theta} \frac{\partial \psi_v(\theta_0)}{\partial \theta'} \right) \text{ and } J(\theta_0, \boldsymbol{\sigma}^2) = \sum_{v=1}^S \frac{1}{\sigma_v^2} E \left(\frac{1}{\psi_v^2(\theta_0)} \frac{\partial \psi_v(\theta_0)}{\partial \theta} \frac{\partial \psi_v(\theta_0)}{\partial \theta'} \right) \quad (4.5)$$

are finite, and $J(\theta_0, \boldsymbol{\sigma}^2)$ is nonsingular for all $\boldsymbol{\sigma}^2 > \mathbf{0}$.

Theorem 4.2 Under **A1-A6** we have

$$\sqrt{N} \left(\widehat{\theta}_G - \theta_0 \right) \xrightarrow{D} \mathcal{N} \left(0, \Sigma(\theta_0, \boldsymbol{\sigma}_0^2, \boldsymbol{\sigma}^2) \right) \text{ as } N \rightarrow \infty \text{ for all } \boldsymbol{\sigma}^2 > \mathbf{0}, \quad (4.6a)$$

where

$$\Sigma(\theta_0, \boldsymbol{\sigma}_0^2, \boldsymbol{\sigma}^2) = J(\theta_0, \boldsymbol{\sigma}^2)^{-1} I(\theta_0, \boldsymbol{\sigma}_0^2, \boldsymbol{\sigma}^2) J(\theta_0, \boldsymbol{\sigma}^2)^{-1} \quad (4.6b)$$

is block-diagonal and \xrightarrow{D} stands for convergence in distribution.

Remark 4.1

i) When $\boldsymbol{\sigma}^2 = (1, \dots, 1)' := \mathbf{1}$, the EQMLE has a covariance matrix in a "sandwich" form and is, in general, not asymptotically efficient unless $\boldsymbol{\sigma}_0^2 = \mathbf{1}$ and the conditional distribution is exponential.

ii) For the special exponential PACD(p, q) model corresponding to $Var(\xi_v) = 1$ for all $1 \leq v \leq S$, if we set $\boldsymbol{\sigma}^2 = \mathbf{1}$ then $J(\theta_0, \mathbf{1}) = I(\theta_0, \mathbf{1}, \mathbf{1})$ and the asymptotic covariance matrix of the EQMLE reduces to $\Sigma = J(\theta_0, \mathbf{1})^{-1}$. The EQMLE is thus asymptotically efficient.

iii) If ξ_v has a constant variance, i.e. $\sigma_{0v}^2 = Var(\xi_v) = \sigma_0^2$ for all $1 \leq v \leq S$, then it suffices to take $\boldsymbol{\sigma}^2 = (1, \dots, 1)'$ and apply the EQMLE. We would have $I(\theta_0, \boldsymbol{\sigma}_0^2, \mathbf{1}) = \sigma_0^2 J(\theta_0, \mathbf{1})$ and the covariance matrix would be equal to $\Sigma = \sigma_0^2 J(\theta_0, \mathbf{1})^{-1}$. In this case, the EQMLE is the

best QMLE among all QMLEs belonging to the linear exponential family.

iv) For the non-periodic ACD corresponding to $S = 1$ and then $\sigma_{0v}^2 = \sigma_0^2$ for all $1 \leq v \leq S$, it is natural to take $\sigma_v^2 = \sigma^2$ for all $1 \leq v \leq S$. In this case, the profile likelihood (4.3) would be given by $\tilde{L}_T(\theta, \boldsymbol{\sigma}^2) = \frac{1}{\sigma^2} \frac{1}{T} \sum_{t=1}^T \frac{Y_t}{\tilde{\psi}_t(\theta)} + \log \tilde{\psi}_t(\theta)$ and the resulting GQMLE then amounts to maximizing $\frac{1}{T} \sum_{t=1}^T \frac{Y_t}{\tilde{\psi}_t(\theta)} + \log \tilde{\psi}_t(\theta)$ which is nothing else but the EQMLE criterion. This is why, in general, the EQMLE is the most used QMLE for non-periodic ACD even when the latter is strictly (conditionally) Gamma distributed.

v) Although $\hat{\theta}_G$ is consistent and asymptotically Gaussian for all fixed known $\boldsymbol{\sigma}^2$, and thus can be used in applications, it is generally inefficient unless $\boldsymbol{\sigma}^2$ is fixed to the true variance parameter $\boldsymbol{\sigma}_0^2 = (\sigma_{01}^2, \dots, \sigma_{0S}^2)'$ which is generally unknown.

vi) When the profile variance $\boldsymbol{\sigma}^2$ coincides with the true variance $\boldsymbol{\sigma}_0^2$ we would have $J(\theta_0, \boldsymbol{\sigma}_0^2) = I(\theta_0, \boldsymbol{\sigma}_0^2, \boldsymbol{\sigma}_0^2)$ and $\Sigma = J(\theta_0, \boldsymbol{\sigma}_0^2)^{-1}$, where the GQMLE is the most efficient among all QMLEs belonging to the exponential family. As $\boldsymbol{\sigma}_0^2$ is generally unknown, a crucial step is to get a consistent estimate $\hat{\boldsymbol{\sigma}}^2$ and construct with it an estimated (profile) log-likelihood from which a new Gamma QMLE, called the two-stage Gamma QMLE (2S-GQMLE), is computed. The resulting estimate would have the aforementioned efficiency property.

vii) As for the Gaussian QMLE of GARCH models, the GQMLEs do not require any moment condition on the observed process (Y_t) , but only on the innovation process (ξ_v) , which is a desirable property for any QMLE.

viii) Theorem 4.1 also holds for the non-MEM PACD (2.7). It suffices to replace the assumptions **A1-A4** by the following:

A1' The process $\{Y_t, t \in \mathbb{Z}\}$ is strictly periodically stationary and periodically ergodic.

A2' $EY_t^{1+\epsilon} < \infty$ for some $\epsilon > 0$.

A3' $\psi_t(\theta) = \psi_t(\theta_0)$ a.s. $\Rightarrow \theta = \theta_0$.

4.2 Estimating the innovation variances

To estimate the unknown variances $\boldsymbol{\sigma}_0^2 = (\sigma_{01}^2, \dots, \sigma_{0S}^2)'$ under the MEM constraint recall (2.1)-(2.2) and let

$$u_t = (Y_t - \psi_t)^2 - \text{Var}(Y_t | \mathcal{F}_{t-1}) = \psi_t^2 ((\xi_t - 1)^2 - \sigma_{0t}^2).$$

Then

$$\frac{(Y_t - \psi_t)^2}{\psi_t^2} = \sigma_{0t}^2 + v_t, \quad (4.7a)$$

where $v_t = \frac{u_t}{\psi_t^2} = (\xi_t - 1)^2 - \sigma_{0t}^2$. The sequence (v_t) is thus zero-mean *ipd*_S with variance $E((\xi_t - 1)^2 - \sigma_{0t}^2)^2$, which is finite under the following assumption.

A7 $E\xi_v^4 < \infty$ for all $v = 1, \dots, S$.

Since $\psi_t = \psi_t(\theta_0)$ depends on the unknown parameter θ_0 , the regressand in (4.7a) is unobservable. If we replace θ_0 by a consistent estimate, say the GQMLE in (4.4), then we get the following approximate regression but with observable regressand

$$\frac{(Y_t - \hat{\psi}_t)^2}{\hat{\psi}_t^2} = \sigma_{0t}^2 + \hat{v}_t, \quad (4.7b)$$

where $\hat{\psi}_t = \psi_t(\hat{\theta}_G)$. From (4.7b) a feasible OLS estimate (OLSE) of σ_0^2 is given by

$$\hat{\sigma}_v^2 = \frac{1}{N} \sum_{n=0}^{N-1} \frac{(Y_{nS+v} - \hat{\psi}_{nS+v})^2}{\hat{\psi}_{nS+v}^2}, \text{ for all } v = 1, \dots, S. \quad (4.8)$$

The following result shows that the OLSE $\hat{\sigma}_v^2$ ($1 \leq v \leq S$) is consistent and asymptotically Gaussian.

Theorem 4.3 *Under A1-A4*

$$\hat{\sigma}_v^2 \rightarrow \sigma_{0v}^2 \text{ a.s. as } N \rightarrow \infty, \text{ for all } v = 1, \dots, S. \quad (4.9a)$$

If in addition A7 holds then for all $v = 1, \dots, S$

$$\sqrt{N} (\hat{\sigma}_v^2 - \sigma_{0v}^2) \xrightarrow{D} \mathcal{N}(0, \Lambda_v) \text{ as } N \rightarrow \infty, \quad (4.9b)$$

where $\Lambda_v = E((\xi_v - 1)^2 - \sigma_{0v}^2)^2$.

An immediate stronger consequence of the latter Theorem is that

$$\sqrt{N} (\hat{\sigma}^2 - \sigma_0^2) \xrightarrow{D} \mathcal{N}(0, \Lambda),$$

where $\hat{\sigma}^2 = (\hat{\sigma}_1^2, \dots, \hat{\sigma}_S^2)'$ and $\Lambda = \text{diag}(\Lambda_1, \dots, \Lambda_S)$. This means that $\hat{\sigma}_v^2$ is asymptotically

independent of $\widehat{\sigma}_s^2$ for all $v \neq s$. A consistent estimate of the limiting variance Λ_v in (4.9b) is given by

$$\widehat{\Lambda}_v = \frac{1}{N} \sum_{n=0}^{N-1} \left(\left(\widehat{\xi}_{nS+v} - 1 \right)^2 - \widehat{\sigma}_v^2 \right)^2, \quad v = 1, \dots, S, \quad (4.10)$$

where $\widehat{\xi}_{nS+v} = \frac{Y_{nS+v}}{\widehat{\psi}_{nS+v}}$ is the residual of model (2.1). With (4.9b) and (4.10), the asymptotic matrices in (4.5) may also be estimated. A consistent estimate of Σ , while replacing the true parameters θ_0 and $\boldsymbol{\sigma}_0^2 = (\sigma_{01}^2, \dots, \sigma_{0S}^2)'$ by their respective estimates $\widehat{\theta}_G$ and $\widehat{\boldsymbol{\sigma}}^2 = (\widehat{\sigma}_1^2, \dots, \widehat{\sigma}_S^2)'$, is

$$\widehat{\Sigma}(\boldsymbol{\sigma}^2) := \widehat{\Sigma}(\boldsymbol{\sigma}^2, \widehat{\theta}_G) = \widehat{J}^{-1}(\boldsymbol{\sigma}^2) \widehat{I}^{-1}(\boldsymbol{\sigma}^2) \widehat{J}^{-1}(\boldsymbol{\sigma}^2), \quad (4.11)$$

where

$$\begin{aligned} \widehat{J}(\boldsymbol{\sigma}^2) &: = \widehat{J}(\widehat{\theta}_G, \boldsymbol{\sigma}^2) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{v=1}^S \frac{1}{\sigma_v^2 \psi_{nS+v}^2(\widehat{\theta}_G)} \frac{\partial \psi_{nS+v}(\widehat{\theta}_G)}{\partial \theta} \frac{\partial \psi_{nS+v}(\widehat{\theta}_G)}{\partial \theta'}, \\ \widehat{I}(\boldsymbol{\sigma}^2) &: = \widehat{I}(\widehat{\theta}_G, \boldsymbol{\sigma}^2) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{v=1}^S \frac{\widehat{\sigma}_v^2}{\sigma_v^4 \psi_{nS+v}^2(\widehat{\theta}_G)} \frac{\partial \psi_{nS+v}(\widehat{\theta}_G)}{\partial \theta} \frac{\partial \psi_{nS+v}(\widehat{\theta}_G)}{\partial \theta'}. \end{aligned}$$

4.3 Two-stage Gamma QMLE

We have seen above that the asymptotic distribution and then the asymptotic efficiency of the profile GQMLE depend on the choice of the profile variance $\boldsymbol{\sigma}^2$. To improve the asymptotic efficiency of the GQMLE, we can replace in (4.4) the profile variances $\boldsymbol{\sigma}^2$ by the OLS estimates $\widehat{\boldsymbol{\sigma}}^2 = (\widehat{\sigma}_1^2, \dots, \widehat{\sigma}_S^2)'$ given by (4.8). The resulting estimate is denoted by 2S-GQMLE and is given by the following steps.

Algorithm 4.1 Two-stage GQMLE

- i) Fix an arbitrarily known $\boldsymbol{\sigma}^2 > \mathbf{0}$, for example $\boldsymbol{\sigma}^2 = (1, \dots, 1)'$.
- ii) Get the profile GQMLE $\widehat{\theta}_G$ from (4.4).
- iii) Estimate the innovation variance $\boldsymbol{\sigma}_0^2$ using $\widehat{\boldsymbol{\sigma}}^2$ in (4.8).
- iv) Consider the 2S-GQMLE as a solution of the following problem

$$\widehat{\theta}_G^* = \arg \min_{\theta \in \Theta} \widetilde{L}_T(\theta, \widehat{\boldsymbol{\sigma}}^2) = \arg \min_{\theta \in \Theta} \sum_{n=0}^{N-1} \sum_{v=1}^S \frac{Y_{nS+v}}{\widehat{\sigma}_v^2 \psi_{nS+v}(\theta)} + \frac{1}{\widehat{\sigma}_v^2} \log \widetilde{\psi}_{nS+v}(\theta). \quad (4.12)$$

Consistency of asymptotic normality of $\widehat{\theta}_G^*$ are a by-product of Theorems 4.1-4.2.

Corollary 4.1 *Under A1-A4*

$$\widehat{\theta}_G^* \rightarrow \theta_0 \text{ a.s. as } N \rightarrow \infty.$$

If in addition A5-A6 hold then

$$\sqrt{N} \left(\widehat{\theta}_G^* - \theta_0 \right) \xrightarrow{D} \mathcal{N} \left(0, \Sigma \left(\theta_0, \boldsymbol{\sigma}_0^2, \boldsymbol{\sigma}_0^2 \right) \right) \text{ as } N \rightarrow \infty, \quad (4.13)$$

where

$$\Sigma \left(\theta_0, \boldsymbol{\sigma}_0^2, \boldsymbol{\sigma}_0^2 \right) = J \left(\theta_0, \boldsymbol{\sigma}_0^2 \right)^{-1} = \left(\sum_{v=1}^S \frac{1}{\sigma_{0v}^2} E \left(\frac{1}{\psi_v^2(\theta_0)} \frac{\partial \psi_v(\theta_0)}{\partial \theta} \frac{\partial \psi_v(\theta_0)}{\partial \theta'} \right) \right)^{-1}.$$

The latter result shows that whatever the distribution of $(\xi_v)_v$ is, the 2S-GQMLE $\widehat{\theta}_G^*$ is asymptotically the most efficient one among all QMLEs belonging to the linear exponential family (cf. Gouriéroux et al, 1984; Wooldridge, 1999). In particular, $\widehat{\theta}_G^*$ is never asymptotically less efficient than the profile GQMLE $\widehat{\theta}_G$ and therefore than the EQMLE $\widehat{\theta}_E$. Similarly to (4.11), a consistent estimate of $J \left(\theta_0, \boldsymbol{\sigma}_0^2 \right)$ in Corollary 4.1 is

$$\widehat{J} := \widehat{J} \left(\widehat{\theta}_G^*, \widehat{\boldsymbol{\sigma}}^2 \right) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{v=1}^S \frac{1}{\widehat{\sigma}_v^2 \psi_{nS+v}^2 \left(\widehat{\theta}_G^* \right)} \frac{\partial \psi_{nS+v} \left(\widehat{\theta}_G^* \right)}{\partial \theta} \frac{\partial \psi_{nS+v} \left(\widehat{\theta}_G^* \right)}{\partial \theta'}. \quad (4.14)$$

4.4 Testing for periodic parameter variation

4.4.1 Conditional mean parameters

As for any periodic time-varying model, an important issue in building a PACD model is to test for the periodic variation of its parameters. This may validate the periodic structure of the model and also reduce the complexity of the model and thus simplify its estimation procedure. For specified (Gaussian) periodic ARMA and (conditionally Gaussian) GARCH models, Franses and co-workers (e.g. Boswijk and Franses, 1995,1996; Franses and Paap, 2000,2004) among others (see also Aknouche 2015) employed likelihood ratio (LR), Lagrange Multiplier (LM), Fisher (F), or Wald tests for the null hypothesis of no periodic parameter variation against its opposite as alternative. For our PACD model, such a null hypothesis writes for the conditional

mean parameter as follows

$$H_0^\theta : \theta_1^0 = \theta_2^0 = \dots = \theta_S^0, \quad i.e. \quad M\theta_0 = \mathbf{0}, \quad (4.15)$$

where

$$M = \begin{pmatrix} I_k & -I_k & 0_k & \cdots & 0_k \\ 0_k & I_k & -I_k & \cdots & 0_k \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_k & 0_k & \cdots & I_k & -I_k \end{pmatrix}$$

is of dimension $(S - 1)k \times Sk$ with I_k and 0_k being respectively the identity matrix and the null square matrix of dimension $k = p + q + 1$. Note that the matrix M is not unique and can be replaced by any appropriate block-permutation of it. Since we do not specify a distribution for the PACD model, quasi- versions of the LR and the LM statistics might be derived to test (4.15), but they would depend on the nuisance parameter σ_0^2 , which is a complicated task. We thus alternatively use the Wald test which is more simple since it does not rely on the conditional distribution of the PACD model. A feasible Wald statistic for (4.15), based on the asymptotic distribution of the 2S-GQMLE as given by (4.13), is defined by

$$W^\theta = \left(M\widehat{\theta}_G^* \right)' \left(M \frac{1}{N} \widehat{J}^{-1} M' \right)^{-1} M\widehat{\theta}_G^*, \quad (4.16)$$

where \widehat{J} is given by (4.14). The use of the 2S-GQMLE rather than a profile GQMLE is made because the latter is never asymptotically more efficient than the former. Under H_0^θ in (4.15) and the assumptions of Corollary 4.1 it can be seen that

$$W^\theta \xrightarrow{D} \chi_{((S-1)(p+q+1))}^2 \quad \text{as } N \rightarrow \infty.$$

Given a size $\alpha \in (0, 1)$, let χ_α be the critical value such that $P(W^\theta > \chi_\alpha) \rightarrow \alpha$ as $N \rightarrow \infty$. The null H_0^θ is thus rejected if $W^\theta > \chi_\alpha$.

Note that if the null (4.15) is rejected, this does not imply that all parameters across season are different, and it may be possible that at least two seasonal parameters are significantly equal. In lieu of a global hypothesis over all seasons as in (4.15), we can also consider the

following pairwise null hypothesis

$$H_0^{\theta,vs} : \theta_v^0 - \theta_s^0 = 0 \text{ for some } v \neq s \quad (4.17)$$

for which we use again the Wald test. As the asymptotic variance in (4.13) is block diagonal, the limiting result (4.13) can be rewritten in the following reduced form

$$\sqrt{N} \left(\widehat{\theta}_v^* - \theta_v^0 \right) \xrightarrow{D} \mathcal{N} \left(0, J_v \left(\theta_v^0, \sigma_{0v}^2 \right)^{-1} \right) \text{ as } N \rightarrow \infty \text{ for all } 1 \leq v \leq S,$$

where $J_v \left(\theta_v^0, \sigma_{0v}^2 \right) = \frac{1}{\sigma_{0v}^2} E \left(\frac{1}{\psi_v^2(\theta_0)} \frac{\partial \psi_v(\theta_0)}{\partial \theta} \frac{\partial \psi_v(\theta_0)}{\partial \theta'} \right)$ and $\widehat{\theta}_v^*$ is the subvector of $\widehat{\theta}_G^*$ corresponding to the parameter θ_v^0 over the v th season. For all $v \neq s$, the estimates $\widehat{\theta}_v^*$ and $\widehat{\theta}_s^*$ are asymptotically independent, implying that

$$\sqrt{N} \left(\widehat{\theta}_v^* - \widehat{\theta}_s^* \right) \xrightarrow{D} \mathcal{N} \left(0, J_{vs} \right),$$

where $J_{vs} = J_v \left(\theta_v^0, \sigma_{0v}^2 \right)^{-1} + J_s \left(\theta_s^0, \sigma_{0s}^2 \right)^{-1}$. A Wald statistic for testing (4.17) is given by

$$W_{vs}^\theta = \left(\widehat{\theta}_v^* - \widehat{\theta}_s^* \right)' \left(\frac{1}{N} \widehat{J}_{vs} \right)^{-1} \left(\widehat{\theta}_v^* - \widehat{\theta}_s^* \right), \quad (4.18)$$

where from (4.14)

$$\widehat{J}_{vs} = \left(\frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\widehat{\sigma}_v^2 \psi_{nS+v}^2(\widehat{\theta}_G^*)} \frac{\partial \psi_{nS+v}(\widehat{\theta}_G^*)}{\partial \theta} \frac{\partial \psi_{nS+v}(\widehat{\theta}_G^*)}{\partial \theta'} \right)^{-1} + \left(\frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\widehat{\sigma}_s^2 \psi_{nS+s}^2(\widehat{\theta}_G^*)} \frac{\partial \psi_{nS+s}(\widehat{\theta}_G^*)}{\partial \theta} \frac{\partial \psi_{nS+s}(\widehat{\theta}_G^*)}{\partial \theta'} \right)^{-1}$$

is a consistent estimate of J_{vs} . Under the assumptions of Corollary 4.1 and $H_0^{\theta,vs}$ in (4.17) we have

$$W_{vs}^\theta \xrightarrow{D} \chi_{(p+q+1)}^2 \text{ as } N \rightarrow \infty.$$

4.4.2 Innovation variances

We can also test for the periodicity of the innovation sequence (ξ_t) . For this we consider a similar null hypothesis to (4.15)

$$H_0^\sigma : \sigma_{01}^2 = \sigma_{02}^2 = \dots = \sigma_{0S}^2 \quad i.e. \quad L\sigma_0^2 = \mathbf{0}, \quad (4.19)$$

where the $(S - 1) \times S$ matrix L is equal to M in (4.15) when $k = 1$. The Wald statistic for testing (4.19) is given by

$$W^\sigma = (L\widehat{\sigma}^2)' \left(L \frac{1}{N} \widehat{\Lambda} L' \right)^{-1} L\widehat{\sigma}^2, \quad (4.20)$$

where $\widehat{\Lambda} = \text{diag}(\widehat{\Lambda}_1, \dots, \widehat{\Lambda}_1)'$ is the diagonal matrix formed by $\widehat{\Lambda}_v$ given by (4.10). Under the assumptions of Theorem 4.3 we have

$$W^\sigma \xrightarrow{D} \chi_{(S-1)}^2 \text{ as } N \rightarrow \infty.$$

A pairwise null hypothesis for the innovation variances can also be considered

$$H_0^{\sigma,vs} : \sigma_{0v}^2 - \sigma_{0s}^2 = 0 \text{ for some } v \neq s.$$

The corresponding Wald statistic,

$$W_{vs}^\sigma = N \widehat{\Delta}_{vs}^{-1} (\widehat{\sigma}_{0v}^2 - \widehat{\sigma}_{0s}^2)^2, \quad (4.21)$$

satisfies

$$W_{vs}^\sigma \xrightarrow{D} \chi_{(1)}^2 \text{ as } N \rightarrow \infty,$$

where $\widehat{\Delta}_{vs} = \widehat{\Delta}_v + \widehat{\Delta}_s$ follows from the asymptotic independence between $\widehat{\sigma}_v^2$ and $\widehat{\sigma}_s^2$ ($v \neq s$).

5 Simulation study

We examine the finite-sample behavior of the Gamma QMLEs, as defined above, using many simulated PACD(1,1) series with sample size $T = 2000$. We consider three distributions for the innovation (ξ_t) in (2.1), namely i) the exponential distribution ($\xi_t \sim \mathcal{E}(1) \equiv \Gamma(1, 1)$) so that $Y_t | \mathcal{F}_{t-1} \sim \Gamma(1, 1/\psi_t)$ (cf. Table 5.1), ii) the Gamma distribution ($\xi_t \sim \Gamma(\sigma_{0t}^{-2}, \sigma_{0t}^{-2})$) which entails $Y_t | \mathcal{F}_{t-1} \sim \Gamma(\sigma_{0t}^{-2}, \sigma_{0t}^{-2}/\psi_t)$, where σ_{0v}^{-2} ($1 \leq v \leq S$) are given in Tables 5.2, and iii) the beta prime distribution, i.e. $\xi_t \sim BP(2\sigma_{0t}^{-2} + 1, 2\sigma_{0t}^{-2} + 2)$ (cf. Tables 5.3-5.4) so that $Y_t | \mathcal{F}_{t-1} \sim GBP(2\sigma_{0t}^{-2} + 1, 2\sigma_{0t}^{-2} + 2, \psi_t)$.

We take $S = 5$ for the two Gamma cases (Tables 5.1-5.2) and $S = 4$ for the Beta prime distribution (Tables 5.3-5.4). Such choices are representative of numerous real daily trading

measurements for $S = 5$, such as trading volumes and realized volatilities, and quarterly data when $S = 4$. The true conditional mean parameters, which are reported in Tables 5.1-5.4, are chosen so that the PACD model is stable in the sense of Section 3, while implying fairly persistent series that are in accordance with the empirical evidence. For each case and each series we compute the profile GQMLE and the two-stage GQMLE (2S-GQMLE), using 1000 Monte Carlo replications. In Tables 5.1-5.3 we take the profile variance $\sigma^2 = \mathbf{1}$ so that the profile GQMLE reduces to the EQMLE. In Table 5.4 the profile variance of the first stage profile GQMLE is $\sigma^2 = (1, 1.5, 2, 0.8)'$.

The starting parameter value in the nonlinear optimization routines (4.4) and (4.12) is set to the true value, while the unobservable starting values Y_0 and $\psi_0(\theta)$ of the PACD(1, 1) equation are set to the intercept ω_0^0 . The two-stage GQMLE is calculated, with the mentioned profile GQMLE computed in the first stage.

Means, standard deviations and (estimated) asymptotic standard errors (ASE) of the estimates $\hat{\theta}_G$ and $\hat{\theta}_G^*$ over the 1000 replications are reported in Table 5.1 for the exponential PACD(1,1) model, in Table 5.2 for the homolog Gamma PACD(1,1) model, and in Tables 5.3-5.4 for the Beta prime model. It can be observed from Tables 5.1-5.4 that the results are consistent with asymptotic theory. They are indeed almost identical in the exponential case with a slight superiority of the EQMLE over the 2S-GQMLE (cf. Table 5.1). The two estimates are, in fact, asymptotically efficient in this case but the EQMLE is much simpler to compute. For the Gamma PACD model in Table 5.2, $\hat{\theta}_G^*$ outperforms $\hat{\theta}_E$ in terms of bias and variability, as expected. In all cases, the 2S-GQMLE is the least risky one in the misspecification case. In the Beta prime case, the 2S-GQMLE $\hat{\theta}_G^*$ still dominates its competing estimates $\hat{\theta}_E$ (Table 5.3) and $\hat{\theta}_G$ (Table 5.4), even with some σ_{0v}^2 exceeding 1 which implies that $E(\xi_v^4) = \infty$. In fact, we have seen from Theorem 4.3 that for the consistency of $\hat{\sigma}_v^2$ (and thus for that the asymptotic normality of $\hat{\theta}_G^*$) we only need $E(\xi_v^2) < \infty$, i.e. $\sigma_{0v}^2 > 0$. However, when $\sigma_{0v}^2 > 1$ the ASE of $\hat{\sigma}_v^2$ is arbitrary since $E(\xi_v^4) < \infty$, a necessary condition for asymptotic normality of $\hat{\sigma}_v^2$, is no longer satisfied.

On the other hand, the ASEs estimated from (4.10)-(4.11) and (4.13) are close to the standard deviations of the estimates and are consistent with asymptotic theory. In particular, it can be seen that the ASEs of $\hat{\theta}_G^*$ are in general (slightly) smaller than $\hat{\theta}_E$ and $\hat{\theta}_G$ (Tables

5.2-5.4).

		EQMLE			2S-GQMLE			
v	θ_v^0	ω_v^0	α_v^0	β_v^0	σ_{0v}^2	ω_v^0	α_v^0	β_v^0
1	True	0.5	0.6	0.35	1	0.5	0.6	0.35
	Mean	0.5126	0.5976	0.3497	0.9849	0.5127	0.5976	0.3497
	StD	0.3284	0.0693	0.0695	0.0948	0.3284	0.0695	0.0695
	ASE	0.2059	0.0681	0.0579	0.0186	0.2059	0.0682	0.0579
2	True	0.9	0.4	0.5	1	0.9	0.4	0.5
	Mean	0.8953	0.3984	0.5030	0.9884	0.8955	0.3984	0.5029
	StD	0.3589	0.0678	0.0900	0.1018	0.3590	0.0677	0.0899
	ASE	0.1992	0.0658	0.0605	0.0188	0.1993	0.0658	0.0605
3	True	1.5	0.5	0.5	1	1.5	0.5	0.5
	Mean	1.4735	0.4961	0.5113	0.9795	1.4731	0.4962	0.5112
	StD	0.4820	0.0797	0.1055	0.0951	0.4811	0.0800	0.1054
	ASE	0.2424	0.0759	0.0697	0.0179	0.2424	0.0759	0.0697
4	True	0.45	0.45	0.45	1	0.45	0.45	0.45
	Mean	0.4662	0.4458	0.4479	0.9798	0.4664	0.4458	0.4479
	StD	0.4095	0.0633	0.0799	0.09617	0.4095	0.0633	0.0799
	ASE	0.2298	0.0627	0.0591	0.0181	0.2298	0.0627	0.0591
5	True	0.7	0.55	0.4	1	0.7	0.55	0.4
	Mean	0.6865	0.5493	0.4060	0.9813	0.6867	0.5493	0.4060
	StD	0.3776	0.0723	0.0785	0.0934	0.3773	0.0722	0.0785
	ASE	0.2264	0.0690	0.0617	0.0182	0.2264	0.0690	0.0617

Table 5.1. EQMLE and 2S-GQMLE results for 1000 PACD₅(1,1) series with $n = 2000$ generated from the exponential $\Gamma(1, 1/\psi_{nS+v})$ distribution; $\sigma^2 = (1, 1, 1, 1, 1)'$.

		EQMLE			2S-GQMLE			
v	θ_v^0	ω_v^0	α_v^0	β_v^0	σ_{0v}^2	ω_v^0	α_v^0	β_v^0
1	True	0.2	0.4	0.5	0.5	0.2	0.4	0.5
	Mean	0.2036	0.3990	0.5026	0.4982	0.1957	0.3992	0.5039
	StD	0.1600	0.0412	0.0766	0.0441	0.1571	0.0398	0.0743
	ASE	0.0748	0.0402	0.0395	0.0031	0.0744	0.0402	0.0395
2	True	0.9	0.3	0.6	0.3	0.9	0.3	0.6
	Mean	0.8855	0.3040	0.6043	0.2983	0.8942	0.3012	0.6023
	StD	0.1464	0.0669	0.0884	0.0240	0.1439	0.0560	0.0759
	ASE	0.0704	0.0404	0.0443	0.0008	0.0702	0.0404	0.0442
3	True	0.3	0.5	0.4	1.5	0.3	0.5	0.4
	Mean	0.3328	0.5012	0.3878	1.4728	0.3389	0.5024	0.3859
	StD	0.2825	0.1088	0.1262	0.1643	0.2802	0.1073	0.1237
	ASE	0.1893	0.1076	0.1160	0.0573	0.1898	0.1080	0.1164
4	True	0.4	0.45	0.45	1	0.4	0.45	0.45
	Mean	0.4127	0.4495	0.4462	0.9872	0.4048	0.4483	0.4488
	StD	0.2586	0.0680	0.0988	0.0979	0.2548	0.0659	0.0963
	ASE	0.1359	0.0649	0.0659	0.0189	0.1358	0.0648	0.0659
5	True	0.5	0.55	0.35	2	0.5	0.55	0.35
	Mean	0.4838	0.5491	0.3602	1.9465	0.4799	0.5515	0.3612
	StD	0.2645	0.0926	0.0968	0.2253	0.2454	0.0828	0.0838
	ASE	0.1945	0.1006	0.0980	0.1203	0.1945	0.1009	0.0985

Table 5.2. EQMLE and 2S-GQMLE results for 1000 PACD₅(1,1) series with $T = 2000$ generated from the Gamma $\Gamma(1/\sigma_{0v}^2, 1/\sigma_{0v}^2\psi_{nS+v})$ distribution; $\boldsymbol{\sigma}^2 = (1, 1, 1, 1, 1)'$.

		EQMLE			2S-GQMLE			
v	θ_v^0	ω_v^0	α_v^0	β_v^0	σ_{0v}^2	ω_v^0	α_v^0	β_v^0
1	True	1	0.6	0.3	1	1	0.6	0.3
	Mean	0.9786	0.6009	0.3065	0.9691	0.9790	0.6008	0.3065
	Std	0.4374	0.0728	0.0797	0.2841	0.4351	0.0728	0.0788
	ASE	0.3498	0.0305	0.0578	0.1098	0.3498	0.0305	0.0579
2	True	0.9	0.3	0.6	1	0.9	0.3	0.6
	Mean	0.9053	0.2990	0.6037	0.9602	0.9065	0.2990	0.6036
	Std	0.5243	0.0659	0.1061	0.3356	0.5277	0.0662	0.1057
	ASE	0.2228	0.0312	0.0455	0.1323	0.2231	0.0313	0.0456
3	True	1.5	0.5	0.5	1	1.5	0.5	0.5
	Mean	1.4776	0.4980	0.5049	0.9528	1.4759	0.4977	0.5054
	Std	0.6110	0.0800	0.1135	0.2767	0.6105	0.0799	0.1135
	ASE	0.2893	0.0367	0.0561	0.1046	0.2889	0.0367	0.0560
4	True	0.7	0.4	0.4	1	0.7	0.4	0.4
	Mean	0.7723	0.4002	0.3889	0.9546	0.7700	0.3998	0.3895
	Std	0.4958	0.0655	0.0829	0.2952	0.4944	0.0648	0.0824
	ASE	0.3069	0.0347	0.0585	0.1151	0.3068	0.0347	0.0585

Table 5.3. EQMLE and 2S-GQMLE results for 1000 PACD₄(1,1) series with $T = 2000$ from the Beta Prime $BP(2\sigma_{0v}^{-2} + 1, 2\sigma_{0v}^{-2} + 2)$ innovation distribution; $\boldsymbol{\sigma}^2 = (1, 1, 1, 1)'$.

		GQMLE			2S-GQMLE			
v	θ_v^0	ω_v^0	α_v^0	β_v^0	σ_{0v}^2	ω_v^0	α_v^0	β_v^0
1	True	1	0.6	0.3	1	1	0.6	0.3
	Mean	1.0043	0.5976	0.3000	0.9428	0.9950	0.5978	0.3012
	StD	0.4851	0.0734	0.0816	0.2533	0.4721	0.0700	0.0760
	ASE	0.3157	0.0703	0.0710	0.0863	0.3160	0.0703	0.0712
2	True	0.9	0.3	0.6	0.5	0.9	0.3	0.6
	Mean	0.9004	0.3000	0.6033	0.4910	0.8863	0.2981	0.6065
	StD	0.4331	0.0512	0.0893	0.0758	0.4313	0.0475	0.0860
	ASE	0.1964	0.0444	0.0502	0.0083	0.1959	0.0443	0.0502
3	True	1.5	0.5	0.5	0.8	1.5	0.5	0.5
	Mean	1.4905	0.5020	0.5025	0.7798	1.4919	0.4970	0.5057
	StD	0.6040	0.1010	0.1345	0.1888	0.5877	0.0899	0.1182
	ASE	0.2999	0.0761	0.0837	0.0474	0.2978	0.0758	0.0836
4	True	0.7	0.4	0.4	1.2	0.7	0.4	0.4
	Mean	0.7316	0.4026	0.3951	1.1273	0.7492	0.4030	0.3933
	StD	0.5698	0.0725	0.0923	0.3723	0.5704	0.0702	0.0902
	ASE	0.3415	0.0695	0.0770	0.1732	0.3413	0.0696	0.0771

Table 5.4. GQMLE and 2S-GQMLE results for 1000 PACD₄(1,1) series with $T = 2000$ from the Beta Prime $BP(2\sigma_{0v}^{-2} + 1, 2\sigma_{0v}^{-2} + 2)$ innovation distribution; $\sigma^2 = (1, 1.5, 2, 0.8)'$.

6 Empirical applications

6.1 Application to Bitcoin trading volume data

In our application, we fit the PACD(1,1) model to the daily Bitcoin trading volume (BTV). The choice of $p = q = 1$ is made for parsimony reasons as the number of parameters highly increases with p and q . Moreover, we found that for larger orders, the estimated parameters are not significant, and the information criteria AIC and BIC, computed for specific conditional distributions such as the Gamma and BP distributions, are not quite smaller. The dataset was

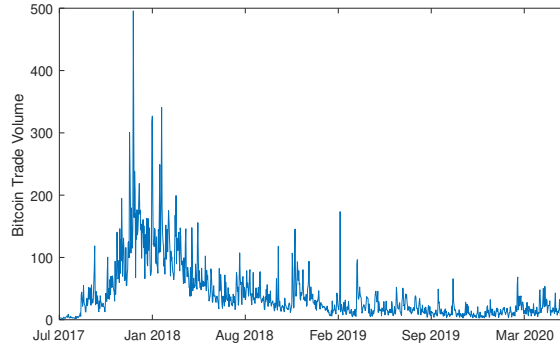


Figure 6.1: Daily Bitcoin trading volume (BTV).

obtained from the webpage www.blockchain.com. This series spans from July, 3, 2017 to June, 26, 2020, with a total of $T = 1092 = 7 \times 156$ observations. Figure 6.1 displays the time series plot of the data.

Periodicity in cryptocurrencies is a recently studied topic. For example, Mbangwa (2019) found evidence of the presence of the day-of-the-week pattern in Bitcoin prices. In our paper we investigate time-dependent anomalies in Bitcoin trading volume. Baur *et al* (2019) and Wang *et al* (2020) identified trading volume patterns over an average week with lower trading activity on weekends compared with weekdays. This finding, which is consistent with that reported for currency markets, suggests the presence of institutional investors who play an active and important role in the trading of the Bitcoin, and that retail traders do not dominate the Bitcoin market. In other words, the overall lower trading activity during weekend is driven by the lower institutional trades.

Thus our aim here is to show that the Bitcoin volume data are also characterized by the day-of-the-week effect, which implies a period of $S = 7$. Such a case is different from the data usually encountered in non-cryptocurrency returns (such as stocks, exchange rates), which are characterized by a periodicity of $S = 5$, due to the existence of non-trading days at each week (Franses and Paap, 2000; Tsiakas, 2006).

Table 6.1 provides some descriptive statistics for the full sample and for each day of the week separately. The mean of BTV series is clearly different from one day to another. The difference is more pronounced for the Kurtosis and skewness across the days. Also, the estimated kernel densities of the data across the days are visually different (see Supplementary material). In that regard, we suspect that the day-of-the week effect may characterize the Bitcoin trading volume series.

Day	Full series	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Mean	40.8394	33.3621	41.3257	42.9304	46.3669	46.7979	44.2866	30.8063
StD	47.2340	40.9090	41.2095	49.5528	52.3069	57.7127	48.9096	34.3056
Kurtosis	17.2929	9.1734	6.6888	16.2964	11.3676	25.8981	10.0401	5.8247
Skewness	2.9736	2.3813	1.9404	3.1176	2.5808	3.8368	2.3894	1.8163

Table 6.1. Day-of-the-week pattern in the BTV series.

We first estimate a standard ACD(1,1) model (i.e. PACD with $S = 1$), using the EQMLE as recommended in Remark 4.1, (iv). Parameter estimates of higher-order ACD(p, q) are not significantly different from zero. This model is used as a competitor to our PACD(1,1). The initial parameter values are set to $\theta^{(0)} = (\omega^{(0)}, \alpha^{(0)}, \beta^{(0)})' = (0.1, 0.3, 0.5)'$ and the starting values of the conditional mean equation are fixed to $Y_0 = \psi_0 = \omega^{(0)}$. The estimated parameters and their asymptotic standard errors (ASE) in parentheses, obtained from Theorem 4.2-4.3 with $S = 1$, are reported in Table 6.2. In particular, the ASE of $\hat{\sigma}^2$ is computed from (4.10). The persistence parameter estimate, $\hat{\alpha} + \hat{\beta} = 0.9865$, indicates a strong persistence in the series, as expected.

$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}^2$	$\hat{\alpha} + \hat{\beta}$	IMSFE	IMAFAE
0.8293 (0.2062)	0.4615 (0.0270)	0.5250 (0.0296)	0.3351 (0.0019)	0.9865	701.9662	14.01243

Table 6.2. EQML estimates for the ACD(1,1); BTV series.

Table 6.2 also displays the in-sample mean square (one-step ahead) forecast error (IMSFE) and the in-sample mean absolute forecast error (IMAFAE) given by $\text{IMSFE} = \frac{1}{T} \sum_{t=1}^T (Y_t - \hat{\psi}_t)^2$ and $\text{IMAFAE} = \frac{1}{T} \sum_{t=1}^T |Y_t - \hat{\psi}_t|$, respectively. Unreported sample autocorrelations of the residuals consolidate the validity of the estimated ACD(1,1). Since this model does not take into account the day-of-the-week effect, we fit a 7-periodic PACD(1,1) to the BTV series. To this end, we utilize the 2S-GQMLE by starting from the EQMLE in the first stage with the following initial parameter values for the optimization routine: $\omega^{(0)} = (0.6, 0.25, 0.35, 3, 2, 4, 1.5)'$, $\alpha^{(0)} = (0.25, 0.15, 0.1, 0.3, 0.3, 0.4, 0.35)'$ and $\beta^{(0)} = (0.7, 0.8, 0.85, 0.4, 0.5, 0.2, 0.45)'$. These ini-

tial values are arbitrary but we checked that the estimates are robust even with other initial values.

Day	v	$\hat{\sigma}_v^2$	$\hat{\omega}_v$	$\hat{\alpha}_v$	$\hat{\beta}_v$	$\hat{\alpha}_v + \hat{\beta}_v$	IMSE	IMAFE
Mon	1	0.2526 (0.0050)	0.0165 (0.3537)	0.5205 (0.0421)	0.5645 (0.0481)	1.0850		
Tue	2	0.2716 (0.0102)	2.9945 (0.4361)	0.5423 (0.0379)	0.7173 (0.0430)	1.2597		
Wed	3	0.3682 (0.0180)	0.3109 (0.2705)	0.1544 (0.0414)	0.8710 (0.0428)	1.0244		
Thu	4	0.2839 (0.0075)	0.0125 (0.6255)	0.4951 (0.0690)	0.5641 (0.0772)	1.0591		
Fri	5	0.3299 (0.0092)	0.2889 (0.5360)	0.3663 (0.0655)	0.6358 (0.0714)	1.0021		
Sat	6	0.2129 (0.0014)	1.3040 (0.5645)	0.4422 (0.0724)	0.4786 (0.0810)	0.9207		
Sun	7	0.2372 (0.0043)	0.3513 (0.4721)	0.4339 (0.0781)	0.2256 (0.0914)	0.6596		
All	$\hat{\phi}$					0.9025	651.9418	13.1164
W^θ		1182.382						
W^σ		18.5338						

Table 6.3. 2S-GQML estimates for the PACD(1,1); BTV series.

The 2S-GQML estimates and their ASEs in parentheses are reported in Table 6.3. We observe that the estimates are quite different across the days. To comfort this fact we test the hypothesis of no conditional mean parameter variation by computing the global and pairwise Wald statistics given by (4.16) and (4.18) respectively. We also do the same for the variance parameters using (4.20) and (4.21). The global Wald statistics $W_\theta = 1182.382$ and $W^\sigma = 18.5338$ indicate that the null hypotheses H_0^θ and H_0^σ of no periodic (conditional mean and variance) parameter variation cannot be accepted at any reasonable level. Table 6.4 displays the pairwise Wald statistics for the conditional mean parameters (lower triangular part of table in bold) and the variances parameters (upper triangular part of the table). For the conditional mean variation, only the statistics W_{41}^θ , W_{45}^θ and W_{65}^θ are significant at the level $\alpha = 0.05$. For the pairwise variance hypothesis, only six statistics among the 21 ones are significant at 0.05 (cf. Table 6.4). Overall, the time-invariant hypothesis cannot be accepted reasonably.

$v \setminus s$	1	2	3	4	5	6	7
1	0	0.1674	4.0703	0.5489	2.9525	1.7111	0.1788
2	298.9884	0	2.3179	0.0597	1.2296	2.0815	0.5754
3	42.4191	806.5083	0	1.9502	0.3770	8.7020	5.4034
4	1.1300	270.0954	19.9804	0	0.8883	3.9490	1.2984
5	11.0859	395.3212	12.7764	5.1796	0	9.0424	4.4821
6	21.0663	387.5937	27.1226	14.0177	5.1002	0	0.7264
7	144.2691	541.4114	150.8982	95.7066	112.2777	77.5127	0

Table 6.4. Pairwise Wald statistics for BTV. W_{vs}^θ in bold for $v > s$.

W_{vs}^σ in the upper triangular part ($v < s$).

The persistence parameters over the days show locally explosive behaviors except for Saturday and Sunday. However, the whole persistence parameter, $\hat{\phi} := \prod_{v=1}^7 (\hat{\alpha}_v + \hat{\beta}_v) = 0.9025$ (also called the monodromy estimate) is, as expected, considerably smaller than the one given by the estimated standard ACD(1,1). All results have been obtained irrespective of any distributional specification of the models.

Note that the ASE of estimates for the PACD are larger than those obtained for the ACD. This is due to the fact that for the PACD the ASEs are computed for lower channel series with sample size $\frac{T}{S} = 156$. To get the same precision as with the ACD we should consider larger series with the sample size multiplied at least by 7. Nevertheless, in term of in-sample forecast ability (IMSFE and IMASE), the PACD model outperforms the standard ACD.

Figure 6.2 shows the probability integral transform (PIT) of the residuals $\hat{\xi}_{nS+v}$ with respect to four distributions of the innovation in the PACD framework: The exponential distribution $\xi_v \sim \Gamma(1, 1)$ (exponential PACD), the Gamma distribution with constant variance $\xi_v \sim \Gamma(\sigma_0^{-2}, \sigma_0^{-2})$ (Gamma PACD), the Gamma distribution with periodic variance $\xi_v \sim \Gamma(\sigma_{0v}^{-2}, \sigma_{0v}^{-2})$ (Gamma PACD*), and the Beta prime distribution $\xi_v \sim BP(2\sigma_{0v}^{-2} + 1, 2\sigma_{0v}^{-2} + 2)$ (Beta prime PACD). It can be seen from Figure 6.2 that the residuals fit better with the Beta prime distribution, followed by the Gamma distribution. The exponential distribution seems to be a bad model. The Gamma distribution with periodic variance is slightly preferred to the Gamma

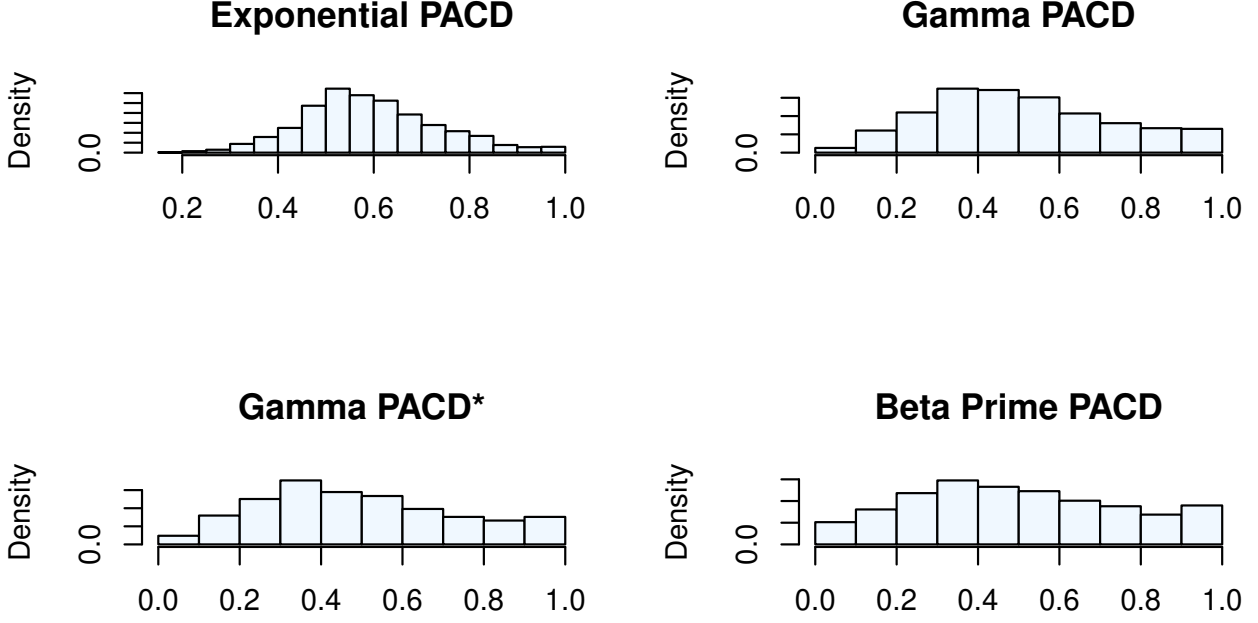


Figure 6.2: PITs of the residuals $\hat{\xi}_t$ (BTV series) for four PACD innovation distributions.

distribution with a constant variance.

Finally, we compare the out-of-sample forecasting performance of the two models. We estimate the two competing models on the basis of the first T_f observations of the series, where $1 < T_f < T$. For the PACD, we use two estimates, the EQMLE, which is mainly recommended for the PACD with constant variance, and the 2S-GQMLE. Then, we compute the one-step ahead forecast on the period $(T_f + 1, \dots, T)$ based on

$$\hat{\psi}_t = \hat{\omega}_t + \hat{\alpha}_t Y_{t-1} + \hat{\beta}_t \hat{\psi}_{t-1} \text{ for } t = T_f + 1, \dots, T.$$

We finally calculate for each model the following three criteria: i) the mean square forecast error $\text{MSFE} = \frac{1}{T-T_f} \sum_{t=T_f+1}^T (Y_t - \hat{\psi}_t)^2$, ii) the mean absolute forecast error given by $\text{MAFE} = \frac{1}{T-T_f} \sum_{t=T_f+1}^T |Y_t - \hat{\psi}_t|$, and iii) the mean QLIKE (cf. Patton, 2011; Aknouche and Francq, 2019) $\text{MQLI} = \frac{1}{T-T_f} \sum_{t=T_f+1}^T (\log \hat{\psi}_t + \frac{Y_t}{\hat{\psi}_t})$.

Table 6.5 shows the computed values of these criteria for the two models and for various truncated series with sample size T_f . For the PACD model, the forecasts based on the EQMLE are denoted by PACD_E while those computed using the 2S-GQMLE are denoted by PACD_{G^*}

(cf. Table 6.5). It can be observed that irrespective of the chosen time-cut T_f , the PACD yields better out-of-sample forecasts with regard to the aforementioned criteria. Moreover, for almost all horizons the PACD forecasts provided by 2S-GQMLE are (slightly) better than those obtained by the EQMLE. For a significant comparison between the three model forecasts in Table 6.5, we employ the Model Confidence Set method of Hansen *et al* (2011) using the R package MCS of Bernardi and Catania (2014). For all time-cut T_f , we found that the forecasts obtained by the PACD_{G*} and PACD_E models generally constitute the Superior Set Models, the ACD model being eliminated.

Overall, the PACD(1, 1) outperforms the ACD(1, 1), both in terms of in-sample and out-of-sample forecasting power.

T_f		500	600	700	800	900	1000
ACD	MSFE	207.8341	165.8957	89.6771	88.3917	94.4021	104.8815
	MAFE	8.6659	7.8318	6.9127	6.6050	6.9080	7.6159
	MQLI	3.8983	3.7735	3.6879	3.6052	3.6959	3.8487
PACD _E	MSFE	193.4029	153.9036	79.1289	80.8411	80.4482	93.4090
	MAFE	8.0311	7.1461	6.1789	5.9196	6.0210	6.6975
	MQLI	3.8756	3.7492	3.6605	3.5757	3.6635	3.8146
PACD _{G*}	MSFE	193.1508	151.0290	79.2446	80.8776	80.4284	93.9384
	MAFE	8.0187	7.1074	6.1546	5.9031	6.02684	6.6908
	MQLI	3.8759	3.7495	3.6605	3.5754	3.6635	3.8144

Table 6.5. Out-of-sample forecasting performance of the PACD and ACD; BTV series.

6.2 Application to the UN realized volatility

The second dataset is the daily UN realized volatility (RV) that covers the sample period from January 04, 1999 to December, 31, 2008 with a total of $T = 2489$ observations. Realized volatility is defined to be the integrated variability of high frequency intraday asset returns (see e.g. Barndorff-Nielsen and Shephard, 2002). It has been extensively studied in the last two decades. Periodicity in realized volatility has been well documented (Martens *et al*, 2009, Yang and Chen, 2014) and is usually caused by macroeconomic news announcements on specific

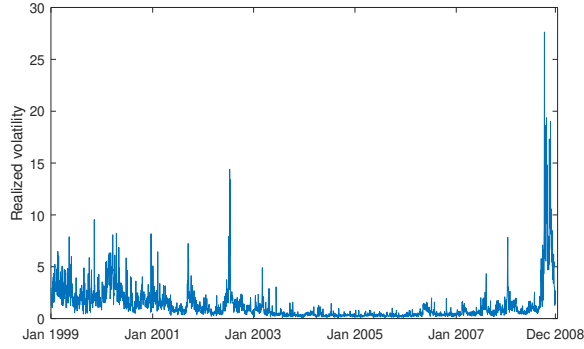


Figure 6.3: Daily realized volatility UN (UN-RV).

weekdays. The plot of the index series is displayed in Figure 6.3.

Table 6.6 reports some descriptive statistics concerning the whole series and the subseries corresponding to the five trading days. It can be easily seen that these statistics strongly indicate that the distributions of realized volatility are significantly different across the trading days. This is also confirmed by the estimated kernel density of each trading day (Supplementary material). These facts suggest using a 5-periodic PACD(1, 1) model for these data.

Day	Full series	Mon	Tue	Wed	Thu	Fri
Sample size	2489	469	511	514	504	491
Mean	1.3085	1.1674	1.2528	1.3028	1.3631	1.4511
StD	1.7699	1.4919	1.4628	1.5890	1.9073	2.2648
Kurtosis	47.1196	26.3686	13.9103	23.4364	37.6800	55.2210
Skewness	5.1689	3.9488	2.9500	3.7675	5.0477	6.0455

Table 6.6. Day-of-the-week pattern in the UN-RV series.

As a reference model, we first estimate a standard ACD(1, 1) for the data. Table 6.7 shows the EQML estimates and their asymptotic standard errors in parenthesis. The results signal a high persistence near to instability.

$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}^2$	$\hat{\alpha} + \hat{\beta}$	IMSFE	IMAFE
0.0109 (0.0031)	0.2849 (0.0157)	0.7084 (0.0162)	0.2841 (0.0005)	0.9933	1.1122	0.4842

Table 6.7. EQML estimates for the ACD(1,1); UN-RV series.

Table 6.8 displays the 2S-GQML estimates of the PACD(1,1) based on the UN-RV series.

These estimates are all significant and are quite different across the days. This is confirmed by the global Wald statistic $W^\theta = 193.9721$ which suggests that the hypothesis H_0^θ of no periodic variation cannot be accepted at, at least, the level 0.005. Moreover, the pairwise Wald statistics in Table 6.9 are all significant at the level 0.01 except $W_{24}^\theta = 2.4490$, which only significant at the level 0.1.

Day	v	$\hat{\sigma}_v^2$	$\hat{\omega}_v$	$\hat{\alpha}_v$	$\hat{\beta}_v$	$\hat{\alpha}_v + \hat{\beta}_v$	IMSFE	IMAFE
Mon	1	0.3805 (0.0082)	0.0164 (0.0098)	0.2374 (0.0390)	0.6702 (0.0417)	0.9076		
Tue	2	0.2597 (0.0017)	0.0154 (0.0084)	0.3320 (0.0291)	0.6903 (0.0348)	1.0222		
Wed	3	0.2552 (0.0012)	0.0015 (0.0088)	0.4023 (0.0327)	0.6521 (0.0354)	1.0544		
Thu	4	0.2670 (0.0019)	0.0126 (0.0078)	0.3052 (0.0359)	0.7065 (0.0388)	1.0117		
Fri	5	0.2682 (0.0023)	0.0884 (0.0132)	0.4138 (0.0390)	0.4851 (0.0442)	0.8989		
All	$\hat{\phi}$					0.8897	1.0570	0.4757
W^θ		193.9721						
W^σ		9.0576						

Table 6.8. 2S-GQML estimates for the PACD(1,1); UN-RV series.

For the variance parameter, the global Wald statistic $W^\sigma = 193.9721$ is significant at the level 0.06 whereas six pairwise statistics W_{vs}^σ from the tenth entail accepting $H_0^{\sigma,vs}$. On the other hand, the persistence parameter, given by $\prod_{v=1}^7(\hat{\alpha}_v + \hat{\beta}_v) = 0.8897$, is significantly smaller than that obtained from the the ACD. The ASEs of the estimates for the PACD are smaller than those in the first application, since the series here is quite longer. Moreover, the PACD model outperforms the standard ACD, according to the IMSFE and IMASE criteria.

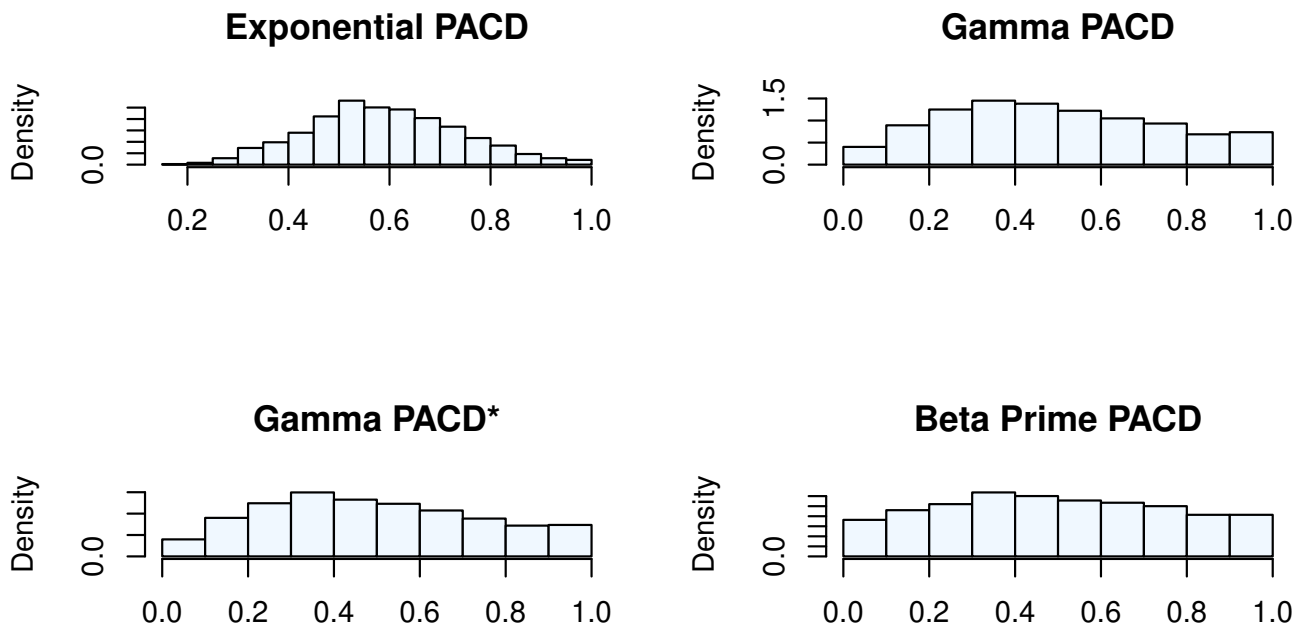


Figure 6.4: PITs of the residuals $\hat{\xi}_t$ (UN-RV series) for four PACD innovation distributions.

$v \setminus s$	1	2	3	4	5
1	0	7.8531	8.8510	6.7729	6.4023
2	123.0932	0	0.0330	0.0729	0.0889
3	119.6519	6.0346	0	0.2139	0.2315
4	87.7977	2.4490	9.5620	0	0.0016
5	55.5838	28.3102	38.6483	28.2441	0

Table 6.9. Pairwise Wald statistics for UN-RV.

W_{vs}^θ for $v > s$; W_{vs}^σ in the upper triangular part ($v < s$).

In an attempt to identify the distribution of the model, Figure 6.4 displays the PIT of the residuals with respect to the four above models. It can be seen that the residuals are strongly consistent with the BP distribution as the corresponding PIT (Beta prime PACD) is almost a straight line. The residuals fit less better to the Gamma distribution with constant or periodic variance as well. As in the previous application, the exponential distribution is the worst one.

We finally compare the out-of-sample forecasting performance of the two models, using the same devices as before. Since the model with constant variance is also plausible as the variance

Wald statistics show, we also run the EQMLE for the data. From Table 6.10 it can be concluded that for all truncated series (with sample size T_f), the PACD computed from both the EQMLE and 2S-GQMLE gives the best accurate forecasts, in terms of the MSFE and MAFE values. Regarding the mean QLIKE criterion, the PACD is clearly the best one (except for $T_f = 1100$ and $T_f = 1200$, where the models are almost comparable). The forecasts obtained by the EQMLE and 2S-GQMLE are in general comparable, where those stemming from $\hat{\theta}_G^*$ exhibit a slight superiority. As in the previous application, using again the Model Confidence Set test of Hansen et al (2011), the forecasts obtained by the PACD $_{G^*}$ and PACD $_E$ models still constitute the Superior Set of Models for almost all time-cut T_f .

T_f		1100	1200	1500	1600	1800	2000
ACD	MSFE	1.0803	1.1501	1.4973	1.6645	2.1461	2.9891
	MAFE	0.3662	0.3644	0.4219	0.4548	0.5509	0.6966
	MQLI	0.4485	0.4534	0.5903	0.6690	0.9126	1.1470
PACD $_E$	MSFE	0.9723	1.0230	1.3563	1.4962	1.9310	2.7311
	MAFE	0.3652	0.3563	0.4092	0.4351	0.5185	0.6541
	MQLI	0.4562	0.4552	0.5892	0.6685	0.9074	1.1406
PACD $_{G^*}$	MSFE	0.9729	1.0299	1.3543	1.4932	1.9275	2.7234
	MAFE	0.3656	0.3564	0.4094	0.4353	0.5189	0.6536
	MQLI	0.4564	0.4547	0.5893	0.6686	0.9074	1.1404

Table 6.10. Out-of-sample forecasting values for the PACD and ACD (UN-RV series).

7 Conclusion and future research

A GARCH-like model for positive-valued data with seasonal behavior has been proposed. The model consists of an ACD model with parameters evolving periodically over time. In our methodology for studying and building such a model, we considered QML estimates that are distribution free and are consistent and asymptotically Gaussian under general conditions. In particular, our proposed two-stage Gamma QMLE takes into account the periodicity of the innovation sequence, producing more accurate results compared to the exponential QMLE. The proposed estimates are also consistent for more general non-MEM forms.

The model was applied to two daily financial series with different periods ($S = 7$ and $S = 5$) in an attempt to capture the day-of-the-week effect. A third application to daily S&P 500 volumes (with $S = 5$) is displayed in the Supplementary material and leads to the same conclusions. The PACD model was used for forecasting in a semi-parametric way without assuming a specific distribution for the innovation. However, the PIT suggests that the beta prime distribution could be a good model for the data.

Our model can be applied to other data frequencies, such as monthly data with $S = 12$ and quarterly data with $S = 4$. Moreover, it may also be utilized as an approximate model for count time series data with large values, such as the daily number of transactions in a market. While we applied our model to specific financial return data, other seasonal positive data coming from various fields such as hydrology (e.g. river flows, rainfall data) and meteorology (e.g. wind speed, wind power) might be represented by the PACD model.

Although our model is named periodic ACD in reference to the ACD proposed by Engle and Russell (1998), it is not recommended to model intraday durations, which are rather characterized by a (stochastic) time-varying period, due to the irregularly-spaced nature of durations. Furthermore, modeling intraday positive-valued series generally requires very large periods for which the estimation of the parameters becomes very challenging.

For models with large periods, some basis functions for reducing the number of parameters, such as Fourier approximation (Bollerslev et al, 2000; Rossi and Fantazani, 2015; Bracher and Held, 2017), periodic B-splines (Ziel et al, 2015) or periodic wavelets (Ziel et al, 2016) could be adapted to our model. As suggested by a referee it would be useful to give an effective method for detecting periodicity in PACD. These issues are left for future research.

Acknowledgements The authors are deeply grateful to Prof. Shiqing Ling the Co-Editor and two referees for their careful reading, very useful comments and constructive suggestions that have led to considerable improvements in the revision. The authors are also grateful to Prof. Christian Francq for providing the UN realized volatility series.

References

- Aknouche A. 2015. Explosive strong periodic autoregression with multiplicity one. *Journal of Statistical Planning and Inference* 161: 50-72.
- Aknouche A, Francq C. 2019. Two-stage weighted least squares estimator of the conditional mean of observation-driven time series models. *MPRA paper 97382*.
- Aknouche A, Francq C. 2020. Count and duration time series with equal conditional stochastic and mean orders. *Econometric Theory*. forthcoming, DOI: <https://doi.org/10.1017/S0266466620000134>.
- Aknouche A, Demmouche N, Dimitrakopoulos S, Touche N. 2020. Bayesian analysis of periodic asymmetric power GARCH models. *Studies in Nonlinear Dynamics and Econometrics* 24: 1–24.
- Ambach D., Croonenbroeck C. 2015. Obtaining Superior Wind Power Predictions from a Periodic and Heteroscedastic Wind Power Prediction Tool. In: Steland A., Rafajkowicz E., Szajowski K. (eds) *Stochastic Models, Statistics and Their Applications* (pp. 225-232). Springer Proceedings in Mathematics & Statistics, vol 122. Springer, Cham.
- Ambach D, Schmid W. 2015. Periodic and long range dependent models for high frequency wind speed data. *Energy* 82: 277–293.
- Andersen T, Bollerslev T. 1997. Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance* 4: 115–158.
- Barndorff-Nielsen OE, Shephard N. 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society* B64: 253-280.
- Baur D, Cahill D, Godfrey K, Liu Z. 2019. Bitcoin time-of-day, day-of-week and month-of-year effects in returns and trading volume. *Finance Research Letters* 31: 78–92.
- Bernardi M, Catania L. 2014. The model confidence set package for R. *ArXiv:1410.8504v1*.
- Bhogal SK, Variyam RT. 2019. Conditional duration models for high-frequency data: A review on recent developments. *Journal of Economic Surveys* 33: 252–273.
- Bollerslev T, Cai J, Song FM. 2000. Intraday periodicity, long memory volatility, and macroeconomic announcement effects in the US Treasury bond market. *Journal of Empirical Finance* 7: 37–55.

- Bollerslev T, Ghysels E. 1996. Periodic autoregressive conditional heteroskedasticity. *Journal of Business & Economic Statistics* 14: 139–152.
- Bortoluzzo AB, Morettin PA, Toloï CM. 2010. Time-varying autoregressive conditional duration model. *Journal of Applied Statistics* 37: 847–864.
- Boswijk HP, Franses PH. 1995. Testing for periodic integration. *Economics Letters* 48: 241–248.
- Boswijk HP, Franses PH. 1996. Unit roots in periodic autoregressions. *Journal of Time Series Analysis* 17: 221–245.
- Boynton W, Oppenheimer HR, Reid SF. 2009. Japanese day-of-the-week return patterns: New results. *Global Finance Journal* 20: 1–12.
- Bracher J, Held L. 2017. Periodically stationary multivariate autoregressive models. *arXiv preprint, arXiv:1707.04635*.
- Caporin M, Rossi E, Santucci de Magistris P. 2017. Chasing volatility: a persistent multiplicative error model with jumps. *Journal of Econometrics* 198: 122–145.
- Charles A. 2010. The day-of-the-week effects on the volatility: The role of the asymmetry. *European Journal of Operational Research* 202: 143–152.
- Chen M, An HZ. 1998. A note on the stationarity and the existence of moments of the GARCH model. *Statistica Sinica* 8: 505–510.
- Chou RY. 2005. Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) Model. *Journal of Money, Credit, and Banking* 37: 561–582.
- Dahlhaus, R. 1997. Fitting time series models to nonstationary Processes. *Annals of Statistics* 25: 1–37.
- Dahlhaus R, Subba Rao S. 2006. Statistical Inference for Time-Varying ARCH Processes. *Annals of Statistics* 34: 1075–1114.
- Diebold F. 1986. Modeling the persistence of conditional variances: A comment. *Econometric Reviews* 5: 51–56.
- Engle R. 2002. New frontiers for Arch models. *Journal of Applied Econometrics* 17: 425–446.
- Engle R, Russell J. 1998. Autoregressive conditional duration: A new model for irregular spaced transaction data. *Econometrica* 66: 1127–1162.

- Francq C, Zakoian J.-M. 2019. *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons, 2nd ed.
- Francq C, Roy R, Saidi A. 2011. Asymptotic Properties of Weighted Least Squares Estimation in Weak PARMA Models. *Journal of Time Series Analysis* 32: 699–723
- Franses PH, Paap R. 2000. Modeling day-of-the-week seasonality in the S&P 500 Index. *Applied Financial Economics* 10: 483–488.
- Franses PH, Paap R. 2004. *Periodic time series models*. Oxford University Press.
- Gallo GM, Otranto E. 2018. Combining sharp and smooth transitions in volatility dynamics: a fuzzy regime approach. *Journal of the Royal Statistical Society C67*: 549–573.
- Gouriéroux C, Monfort A, Trognon A. 1984. Pseudo maximum likelihood methods: Theory. *Econometrica* 52: 681–700.
- Hallin M. 1984. Spectral Factorization of Nonstationary Moving Average Processes. *Annals of Statistics* 12: 172–192.
- Hautsch N. 2012. *Econometrics of Financial High-Frequency Data*. Berlin, Heidelberg: Springer.
- Hujer R, Vuletic S. 2007. Econometric analysis of financial trade processes by discrete mixture duration models. *Journal of Economic Dynamics and Control* 31: 635–667.
- Johnson NL, Kotz S, Balakrishnan N. 1995. *Continuous univariate distributions*. Wiley, 2nd Edition.
- Lanne M. 2006. A mixture multiplicative error model for realized volatility. *Journal of Financial Econometrics* 4: 594–616.
- Li Q. 2020. Location multiplicative error models with quasi maximum likelihood estimation. *Journal of Time Series Analysis* 41: 387–405.
- Lund R, Basawa IV. 2000. Recursive prediction and likelihood evaluation for periodic ARMA models. *Journal of Time Series Analysis* 21: 75–93.
- Martens M, Van Dijk D, De Pooter M. 2009. Forecasting S&P 500 volatility: Long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements. *International Journal of Forecasting* 25: 282–303.
- Mbanga CL. 2019. The day-of-the-week pattern of price clustering in Bitcoin. *Applied Economics Letters* 26: 1–6.

- McCullagh P, Nelder JA. 1989. Generalized Linear Models, 2nd edn. Chapman and Hall, London.
- Mikosch T, Starica C. 2004. Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects. *Review of Economics and Statistics* 86: 378–390.
- Miller KS. 1968. *Linear Difference Equations*. Benjamin.
- Mishra A, Ramanathan TV. 2017. Nonstationary autoregressive conditional duration models. *Studies in Nonlinear Dynamics and Econometrics* 21: 1–22.
- Moghaddam MD, Liu J, Serota RA. 2019. Implied and realized volatility: A study of distributions and the distribution of difference. *ArXiv:1906.02306*.
- Pacurar M. 2008. Autoregressive Conditional Duration (ACD) Models in Finance: A Survey of the Theoretical and Empirical Literature. *Journal of Economic Surveys* 22: 711–751.
- Patton AJ. 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160:, 246–256.
- Rossi E, Fantazzini D. 2015. Long memory and periodicity in intraday volatility. *Journal of Financial Econometrics* 13: 922–961.
- Tsiakas I. 2006. Periodic stochastic volatility and fat tails. *Journal of Financial Econometrics* 4: 90–135.
- Wang JN, Liu HC, Hsu YT. 2020. Time-of-day periodicities of trading volume and volatility in Bitcoin exchange: Does the stock market matter?. *Finance Research Letters* 34: 1–8.
- Wooldridge JM. 1999. Quasi-likelihood methods for count data. In M.H. Pesaran and P. Schmidt (ed.), *Handbook of Applied Econometrics, Volume 2: Microeconomics*, (pp. 35–406). Oxford: Blackwell.
- Yang K, Chen L. 2014. Realized volatility forecast: structural breaks, long memory, asymmetry, and day-of-the-week effect. *International Review of Finance* 14: 345–392.
- Ziel F, Croonenbroeck C, Ambach D. 2016. Forecasting wind power- Modeling periodic and non-linear effects under conditional heteroscedasticity. *Applied Energy* 177: 285–297.
- Ziel F, Steinert R, Husmann S. 2015. Efficient modeling and forecasting of electricity spot prices. *Energy Economics* 47: 98–111.