# Periodic autoregressive conditional duration

Aknouche, Abdelhakim and Almohaimeed, Bader and Dimitrakopoulos, Stefanos

University of Science and Technology Houari Boumediene, Qassim University, Leeds University

8 July 2020

# Periodic autoregressive conditional duration

Abdelhakim Aknouche*†, Bader Almohaimeed*, 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). All the proofs are given in the Appendix. A Supplementary material accompanies this paper.

# 2 Periodic Autoregressive Conditional Duration model

All random variables and processes in this paper are defined on a probability space $(\Omega, \mathscr{F}, P)$ and valued in the set of positive real numbers $\mathbb{R}_+ = (0, \infty)$, which is endowed with the Borel field $\mathscr{B}(\mathbb{R}_+)$. Let $S \geq 1$ be a positive integer called the period, and $\omega_t^0$, $\alpha_{t1}^0, ..., \alpha_{tq}^0$, $\beta_{t1}^0, ..., \beta_{tp}^0$ $(p, q \in \mathbb{N} = \{0, 1, ...\})$ be positive real parameters $S$-periodic over time, i.e. $\omega_t^0 = \omega_{t+kS}^0$, $\alpha_{ti}^0 = \alpha_{t+kS,i}^0$ $(i = 1, ..., q)$ and $\beta_{tj}^0 = \beta_{t+kS,j}^0$ $(j = 1, ..., p)$ 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 $\xi_t \overset{D}{=} \xi_{t+S}$ for all $t$, where $\overset{D}{=}$ 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

$$Y_t = \psi_t \xi_t \tag{2.1a}$$

and

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

where the innovation term $\xi_t$ is independent of $\psi_{t-j}$ for all $j \geq 1$. To ensure the almost sure (a.s.) positivity of $\psi_t$, it is assumed that $\omega_t^0 > 0$, $\alpha_{ti}^0 \geq 0$, and $\beta_{tj}^0 \geq 0$, for all $t \in \mathbb{Z}$, $i = 1, .., q$ and $j = 1, ..., p$. 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

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

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 $\{..., v-S, v, v+S, v+2S, ...\}$ with corresponding parameters $\omega_v^0$, $\alpha_{vi}^0$, $\beta_{vj}^0$ and $\sigma_{0v}^2 = Var\left(\xi_{nS+v}\right)$. Let $\mathscr{F}_t$ be the $\sigma$-Algebra generated by $\{Y_{t-i}, i \geq 0\}$. The conditional mean and conditional variance of the model (2.1) are given respectively by

$$E\left(Y_t | \mathscr{F}_{t-1}\right) = \psi_t \tag{2.2a}$$

and

$$Var\left(Y_t | \mathscr{F}_{t-1}\right) = \sigma_{0t}^2 \psi_t^2. \tag{2.2b}$$

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 $\xi_t$ 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 conditional 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 (2.2$b$). However, a useful family of conditional distributions satisfying (2.2$b$) is the Gamma distribution with shape $\frac{1}{\sigma_{0t}^2}$ and scale $\frac{1}{\sigma_{0t}^2 \psi_t}$, that is

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

where $\psi_t$ satisfies (2.1$b$). In the latter case, the innovation term $\xi_t$ in (2.1) will be marginally Gamma distributed

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

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

$$f(y; a, b, c) = \tfrac{1}{cBeta(a,b)} \left(\tfrac{y}{c}\right)^{a-1} \left(1 + \tfrac{y}{c}\right)^{-a-b}, \ y \geq 0,$$

where $Beta(a, b)$ 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 $Var(Y) = \frac{c^2 a(a+b-1)}{(b-2)(b-1)^2}$ $(b > 2)$. 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 $\xi_t$ is BP distributed with the following parametrization

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

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

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

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

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

and

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

where $\{\eta_t, t \in \mathbb{Z}\}$ is an $ipd_S$ sequence with mean zero and unit variance, and the parameters $\omega_t^0$, $\alpha_{ti}^0$ and $\beta_{tj}^0$ 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 (2.5$b$) 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

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

where

$$\varepsilon_t = Y_t - E\left(Y_t|\mathscr{F}_{t-1}\right) = \psi_t\left(\xi_t - 1\right) \tag{2.6}$$

is a zero-mean term of a martingale difference sequence with a finite periodic variance $E\varepsilon_t^2 =$

$E\left(\xi_t-1\right)^2 E\psi_t^2=\sigma_{0t}^2 E\psi_t^2$.

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

$$Y_t|\mathscr{F}_{t-1} \sim F_{\psi_t}, \tag{2.7}$$

where $F_\psi$ is a cumulative probability distribution (with positive support) with mean $\psi$, and $\psi_t$ is given by (2.1$b$). 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 (2.1$a$) and (2.1$b$) we obtain the following stochastic recurrence equation (SRE)

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

driven by the $ipd_S$ sequence $\left\{(A^0_{nS+v}, B^0_{nS+v}), n \in \mathbb{Z}, 0 \leq v \leq S-1\right\}$, where $\underline{Y}_{nS+v} = (Y_{nS+v}, ...$ $, Y_{nS+v-q+1}, \psi_{nS+v}, ..., \psi_{nS+v-p+1})'$, $B^0_{nS+v} = \left(\omega^0_v \xi_{nS+v}, 0_{(q-1)\times 1}, \omega^0_v, 0_{(p-1)\times 1}\right)'$, and

$$A^0_{nS+v} = \begin{pmatrix} \alpha^0_{v1}\xi_{nS+v} & \cdots & \alpha^0_{v,q-1}\xi_{nS+v} & \alpha^0_{vq}\xi_{nS+v} & \beta^0_{v1}\xi_{nS+v} & \cdots & \beta^0_{v,p-1}\xi_{nS+v} & \beta^0_{vp}\xi_{nS+v} \\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha^0_{v1} & \cdots & \alpha^0_{v,q-1} & \alpha^0_{vq} & \beta^0_{v1} & \cdots & \beta^0_{v,p-1} & \beta^0_{vp} \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

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

$$\gamma^S = \inf\left\{\frac{1}{n} E \log \left\|A^0_{nS}...A^0_2 A^0_1\right\|, \ n \geq 1\right\}$$

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

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

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*

$$\underline{Y}_{nS+v} = \sum_{j=0}^{\infty} \prod_{i=0}^{j-1} A^0_{nS+v-i} B^0_{nS+v-j}, \tag{3.3}$$

*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*

$$\rho \left( \prod_{v=0}^{S-1} \boldsymbol{\beta}^0_{S-v} \right) < 1. \tag{3.4}$$

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

$$\sum_{v=1}^{S} E\left(\log\left(\alpha^0_v \xi_{v-1} + \beta^0_v\right)\right) < 0,$$

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

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*

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

Some remarks are in order:

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

$$\sum_{i=1}^{q} \alpha^0_i + \sum_{j=1}^{p} \beta^0_j < 1$$

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

$$\prod_{v=1}^{S} \left(\alpha^0_v + \beta^0_v\right) < 1. \tag{3.6}$$

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

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

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

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*

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

*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 $\omega_v^0$, $\alpha_v^0$ and $\beta_v^0$ by $\omega_{0v}$, $\alpha_{0v}$ and $\beta_{0v}$, 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*

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

*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*

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

*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*

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

For the particular exponential PACD(1,1) model, $Y_{nS+v}|\mathscr{F}_{nS+v-1} \sim \Gamma\left(1, \frac{1}{\psi_{nS+v}}\right)$, 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 $\sigma_{0v}^2$ by 1 for all $1 \leq v \leq S$.

# 4 Gamma quasi-maximum likelihood estimates

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

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

where the true parameter $\theta_0 = (\theta_1^{0\prime}, \theta_2^{0\prime}, ..., \theta_S^{0\prime})'$ with $\theta_v^0 = (\omega_v^0, \alpha_{v1}^0, ..., \alpha_{vq}^0, \beta_{v1}^0, ..., \beta_{vp}^0)'$ $(1 \leq v \leq S)$ 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 $\boldsymbol{\sigma}_0^2 = (\sigma_{01}^2, ..., \sigma_{0S}^2)'$, where $\sigma_{0v}^2 = Var(\xi_{nS+v})$ $(1 \leq v \leq S)$, 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 $Y_0, ..., Y_{1-q}, \widetilde{\psi}_0, ..., \widetilde{\psi}_{1-p}$ and a generic parameter $\theta \in \Theta$ define

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

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

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

## 4.1 Exponential and profile Gamma QMLEs

The true conditional distribution of (4.1) is unknown due to the unspecification of the law of $\xi_v$ $(1 \leq v \leq S)$. Thus, a quasi-maximum likelihood estimate (QMLE) which does not require any precise knowledge of the conditional distribution is suitable for estimating the parameter $\theta_0$ 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 $\xi_v$ 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 $n \in \{0, ..., N-1\}$. (This does not be confused with the unknown true $\sigma_{0t}^2$). The profile Gamma likelihood associated with $\boldsymbol{\sigma}^2 = (\sigma_1^2, ..., \sigma_S^2)' > \mathbf{0}$ is given, ignoring constants, by

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

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

The profile Gamma QMLE (GQMLE) $\widehat{\theta}_G$ of $\theta_0$ is, then, the minimizer of $\widetilde{L}_T\left(\theta, \boldsymbol{\sigma}^2\right)$ over $\Theta$,

$$\widehat{\theta}_G = \arg\min_{\theta\in\Theta} \widetilde{L}_T\left(\theta, \boldsymbol{\sigma}^2\right), \tag{4.4}$$

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

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

$$\widehat{\theta}_E = \arg\min_{\theta\in\Theta} \tfrac{1}{NS}\sum_{v=1}^{S}\sum_{n=0}^{N-1} \tfrac{Y_{nS+v}}{\widetilde{\psi}_{nS+v}(\theta)} + \log \widetilde{\psi}_{nS+v}\left(\theta\right).$$

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

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

**A2** $\theta_0 \in \Theta$ *and* $\Theta$ *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** $\xi_v$ *is non-degenerate and* $E\xi_v = 1$ *for all* $1 \leq v \leq S$.

As seen in Section 3, $\gamma^S\left(A^0\right) < 0$ 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 (4.2$b$) 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* $\left(\widehat{\theta}_G\right)$ *be a sequence of GQMLEs defined by* (4.4). *Under* ***A1-A4***,

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

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

**A5** $\theta_0$ *belongs to the interior of* $\Theta$.

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

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

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

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

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

*where*

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

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

**Remark 4.1**

i) When $\boldsymbol{\sigma}^2=(1,...,1)':=\mathbf{1}$, 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 $Var\left(\xi_v\right)=1$ for all $1\leq v\leq S$, if we set $\boldsymbol{\sigma}^2=\mathbf{1}$ then $J\left(\theta_0,\mathbf{1}\right)=I\left(\theta_0,\mathbf{1},\mathbf{1}\right)$ and the asymptotic covariance matrix of the EQMLE reduces to $\Sigma=J\left(\theta_0,\mathbf{1}\right)^{-1}$. The EQMLE is thus asymptotically efficient.

iii) If $\xi_v$ has a constant variance, i.e. $\sigma_{0v}^2=Var\left(\xi_v\right)=\sigma_0^2$ for all $1\leq v\leq S$, then it suffices to take $\boldsymbol{\sigma}^2=(1,...,1)'$ and apply the EQMLE. We would have $I\left(\theta_0,\boldsymbol{\sigma}_0^2,\mathbf{1}\right)=\sigma_0^2J\left(\theta_0,\mathbf{1}\right)$ 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 $\widetilde{L}_T(\theta, \boldsymbol{\sigma}^2) = \frac{1}{\sigma^2}\frac{1}{T}\sum_{t=1}^{T}\frac{Y_t}{\widetilde{\psi}_t(\theta)} + \log\widetilde{\psi}_t(\theta)$ and the resulting GQMLE then amounts to maximizing $\frac{1}{T}\sum_{t=1}^{T}\frac{Y_t}{\widetilde{\psi}_t(\theta)} + \log\widetilde{\psi}_t(\theta)$ 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 $\widehat{\theta}_G$ 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 $\boldsymbol{\sigma}_0^2 = (\sigma_{01}^2, ..., \sigma_{0S}^2)'$ 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 $\widehat{\boldsymbol{\sigma}}^2$ 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 $(\xi_v)$, 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 $\boldsymbol{\sigma}_0^2 = (\sigma_{01}^2, ..., \sigma_{0S}^2)'$ under the MEM constraint recall (2.1)-(2.2) and let

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

Then

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

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 $E\left((\xi_t - 1)^2 - \sigma_{0t}^2\right)^2$, which is finite under the following assumption.

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

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

$$\frac{\left(Y_t-\widehat{\psi}_t\right)^2}{\widehat{\psi}_t^2} = \sigma_{0t}^2 + \widehat{v}_t, \tag{4.7b}$$

where $\widehat{\psi}_t = \psi_t\left(\widehat{\theta}_G\right)$. From (4.7$b$) a feasible OLS estimate (OLSE) of $\boldsymbol{\sigma}_0^2$ is given by

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

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

**Theorem 4.3** *Under **A1**-**A4***

$$\widehat{\sigma}_v^2 \to \sigma_{0v}^2 \ \ a.s. \ as \ N \to \infty, \ for \ all \ v = 1, ..., S. \tag{4.9a}$$

*If in addition **A7** holds then for all* $v = 1, ..., S$

$$\sqrt{N}\left(\widehat{\sigma}_v^2 - \sigma_{0v}^2\right) \overset{D}{\to} \mathscr{N}(0, \Lambda_v) \ \ as \ N \to \infty, \tag{4.9b}$$

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

An immediate stronger consequence of the latter Theorem is that

$$\sqrt{N}\left(\widehat{\boldsymbol{\sigma}}^2 - \boldsymbol{\sigma}_0^2\right) \overset{D}{\to} \mathscr{N}(0, \Lambda),$$

where $\widehat{\boldsymbol{\sigma}}^2 = (\widehat{\sigma}_1^2, ..., \widehat{\sigma}_S^2)'$ and $\Lambda = diag(\Lambda_1, ..., \Lambda_S)$. This means that $\widehat{\sigma}_v^2$ is asymptotically independent of $\widehat{\sigma}_s^2$ for all $v \neq s$. A consistent estimate of the limiting variance $\Lambda_v$ in (4.9$b$) is given by

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

where $\widehat{\xi}_{nS+v} = \frac{Y_{nS+v}}{\widehat{\psi}_{nS+v}}$ is the residual of model (2.1). With (4.9$b$) and (4.10), the asymptotic matrices in (4.5) may also be estimated. A consistent estimate of $\Sigma$, while replacing the true parameters $\theta_0$ and $\boldsymbol{\sigma}_0^2 = (\sigma_{01}^2, ..., \sigma_{0S}^2)'$ by their respective estimates $\widehat{\theta}_G$ and $\widehat{\boldsymbol{\sigma}}^2 = (\widehat{\sigma}_1^2, ..., \widehat{\sigma}_S^2)'$, is

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

where

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

## 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 efficiency of the GQMLE, we can replace in (4.4) the profile variances $\boldsymbol{\sigma}^2$ by the OLS estimates $\widehat{\boldsymbol{\sigma}}^2 = (\widehat{\sigma}_1^2, ..., \widehat{\sigma}_S^2)'$ 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 $\boldsymbol{\sigma}^2 = (1, ..., 1)'$.

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

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

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

**Corollary 4.1** *Under* ***A1-A4***

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

*If in addition* ***A5-A6*** *hold then*

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

*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. Gourieroux et *al*, 1984; Wooldridge, 1999). In particular, $\widehat{\theta}_G^*$ in 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 $J(\theta_0, \boldsymbol{\sigma}_0^2)$ in Corollary 4.1 is

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

## 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

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

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 $\boldsymbol{\sigma}_0^2$, 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

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

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 efficient than the former. Under $H_0^\theta$ in (4.15) and the assumptions of Corollary 4.1 it can be seen that

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

Given a size $\alpha \in (0,1)$, let $\chi_\alpha$ be the critical value such that $P\left(W^\theta > \chi_\alpha\right) \to \alpha$ as $N \to \infty$. The null $H_0^\theta$ 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

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

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

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

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 $\theta_v^0$ 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

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

where from (4.14)

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

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

$$W_{vs}^{\theta} \xrightarrow{D} \chi^2_{(p+q+1)} \quad \text{as} \quad N \to \infty.$$

### 4.4.2 Innovation variances

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

$$H_0^{\sigma} : \sigma_{01}^2 = \sigma_{02}^2 = .... = \sigma_{0S}^2 \quad i.e. \quad L\boldsymbol{\sigma}_0^2 = \mathbf{0}, \tag{4.19}$$

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

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

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

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

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

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

The corresponding Wald statistic,

$$W^{\sigma}_{vs}=N\widehat{\Delta}^{-1}_{vs}\left(\widehat{\sigma}^2_{0v}-\widehat{\sigma}^2_{0s}\right)^2, \tag{4.21}$$

satisfies

$$W^{\sigma}_{vs}\xrightarrow{D}\chi^2_{(1)}\ \text{ as }\ N\to\infty,$$

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

# 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 $(\xi_t)$ in (2.1), namely i) the exponential distribution ($\xi_t\sim\mathscr{E}(1)\equiv\Gamma(1,1)$) so that $Y_t|\mathscr{F}_{t-1}\sim\Gamma(1,1/\psi_t)$ (cf. Table 5.1), ii) the Gamma distribution ($\xi_t\sim\Gamma\left(\sigma_{0t}^{-2},\sigma_{0t}^{-2}\right)$) which entails $Y_t|\mathscr{F}_{t-1}\sim\Gamma\left(\sigma_{0t}^{-2},\sigma_{0t}^{-2}/\psi_t\right)$, where $\sigma_{0v}^{-2}$ $(1\leq v\leq S)$ are given in Tables 5.2, and iii) the beta prime distribution, i.e. $\xi_t\sim BP\left(2\sigma_{0t}^{-2}+1,2\sigma_{0t}^{-2}+2\right)$ (cf. Tables 5.3-5.4) so that $Y_t|\mathscr{F}_{t-1}\sim GBP\left(2\sigma_{0t}^{-2}+1,2\sigma_{0t}^{-2}+2,\psi_t\right)$.

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 $\boldsymbol{\sigma}^2 = \mathbf{1}$ so that the profile GQMLE reduces to the EQMLE. In Table 5.4 the profile variance of the first stage profile GQMLE is $\boldsymbol{\sigma}^2 = (1, 1.5, 2, 0.8)'$.

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 $\omega_0^0$. 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 $\widehat{\theta}_G$ and $\widehat{\theta}_G^*$ 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, $\widehat{\theta}_G^*$ outperforms $\widehat{\theta}_E$ 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 $\widehat{\theta}_G^*$ still dominates its competing estimates $\widehat{\theta}_E$ (Table 5.3) and $\widehat{\theta}_G$ (Table 5.4), even with some $\sigma_{0v}^2$ exceeding 1 which implies that $E(\xi_v^4) = \infty$. In fact, we have seen from Theorem 4.3 that for the consistency of $\widehat{\sigma}_v^2$ (and thus for that the asymptotic normality of $\widehat{\theta}_G^*$) we only need $E(\xi_v^2) < \infty$, i.e. $\sigma_{0v}^2 > 0$. However, when $\sigma_{0v}^2 > 1$ the ASE of $\widehat{\sigma}_v^2$ is arbitrary since $E(\xi_v^4) < \infty$, a necessary condition for asymptotic normality of $\widehat{\sigma}_v^2$, 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 $\widehat{\theta}_G^*$ are in general (slightly) smaller than $\widehat{\theta}_E$ and $\widehat{\theta}_G$ (Tables

5.2-5.4).

| $v$ | $\theta_v^0$ | EQMLE | | | 2S-GQMLE | | | |
|---|---|---|---|---|---|---|---|---|
| $v$ | $\theta_v^0$ | $\omega_v^0$ | $\alpha_v^0$ | $\beta_v^0$ | $\sigma_{0v}^2$ | $\omega_v^0$ | $\alpha_v^0$ | $\beta_v^0$ |
| 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 $\text{PACD}_5(1,1)$ series with $n = 2000$ generated from the exponential $\Gamma(1, 1/\psi_{nS+v})$ distribution; $\boldsymbol{\sigma}^2 = (1, 1, 1, 1, 1)'$.

| | | EQMLE | | | 2S-GQMLE | | | |
|---|---|---|---|---|---|---|---|---|
| $v$ | $\theta_v^0$ | $\omega_v^0$ | $\alpha_v^0$ | $\beta_v^0$ | $\sigma_{0v}^2$ | $\omega_v^0$ | $\alpha_v^0$ | $\beta_v^0$ |
| 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 $\text{PACD}_5(1,1)$ series with $T = 2000$ generated from the Gamma $\Gamma\left(1/\sigma_{0v}^2, 1/\sigma_{0v}^2\psi_{nS+v}\right)$ distribution; $\boldsymbol{\sigma}^2 = (1,1,1,1,1)'$.

| | | EQMLE | | | 2S-GQMLE | | | |
|---|---|---|---|---|---|---|---|---|
| $v$ | $\theta_v^0$ | $\omega_v^0$ | $\alpha_v^0$ | $\beta_v^0$ | $\sigma_{0v}^2$ | $\omega_v^0$ | $\alpha_v^0$ | $\beta_v^0$ |
| 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 $\text{PACD}_4(1{,}1)$ 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 | | | |
|---|---|---|---|---|---|---|---|---|
| $v$ | $\theta_v^0$ | $\omega_v^0$ | $\alpha_v^0$ | $\beta_v^0$ | $\sigma_{0v}^2$ | $\omega_v^0$ | $\alpha_v^0$ | $\beta_v^0$ |
| 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 $\text{PACD}_4(1,1)$ series with $T = 2000$ from the Beta Prime $BP(2\sigma_{0v}^{-2}+1, 2\sigma_{0v}^{-2}+2)$ innovation distribution; $\boldsymbol{\sigma}^2 = (1, 1.5, 2, 0.8)'$.

# 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

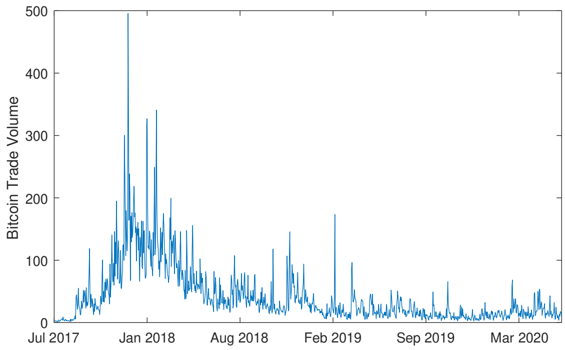



Figure 6.1: Daily Bitcoin trading volume (BTV).

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, Mbanga (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 $\theta^{(0)} = \left(\omega^{(0)}, \alpha^{(0)}, \beta^{(0)}\right)' = (0.1, 0.3, 0.5)'$ 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 $\widehat{\sigma}^2$ is computed from (4.10). The persistence parameter estimate, $\widehat{\alpha}+\widehat{\beta} = 0.9865$, indicates a strong persistence in the series, as expected.

| $\widehat{\omega}$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\sigma}^2$ | $\widehat{\alpha}+\widehat{\beta}$ | IMSFE | IMAFE |
|---|---|---|---|---|---|---|
| 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 (IMAFE) given by IMSFE$= \frac{1}{T}\sum_{t=1}^{T}(Y_t - \widehat{\psi}_t)^2$ and IMAFE$= \frac{1}{T}\sum_{t=1}^{T}\left|Y_t - \widehat{\psi}_t\right|$, 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 initial values are arbitrary but we checked that the estimates are robust even with other initial values.

| Day | $v$ | $\widehat{\sigma}_v^2$ | $\widehat{\omega}_v$ | $\widehat{\alpha}_v$ | $\widehat{\beta}_v$ | $\widehat{\alpha}_v + \widehat{\beta}_v$ | IMSFE | 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 | $\widehat{\phi}$ | | | | | 0.9025 | 651.9418 | 13.1164 |
| $W^{\theta}$ | | 1182.382 | | | | | | |
| $W^{\sigma}$ | | 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 $H_0^{\theta}$ and $H_0^{\sigma}$ 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 $W_{41}^{\theta}$, $W_{45}^{\theta}$ $W_{65}^{\theta}$ 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.

| $v \backslash s$ | 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. $W^{\theta}_{vs}$ in bold for $v > s$. $W^{\sigma}_{vs}$ in the uper triangular part ($v < s$).

The persistence parameters over the days show locally explosive behaviors except for Saturday and Sunday. However, the whole persistence parameter, $\widehat{\phi} := \prod_{v=1}^{7}(\widehat{\alpha}_v + \widehat{\beta}_v) = 0.9025$ (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 $\widehat{\xi}_{nS+v}$ 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 $\xi_v \sim \Gamma\left(\sigma_0^{-2}, \sigma_0^{-2}\right)$ (Gamma PACD), the Gamma distribution with periodic variance $\xi_v \sim \Gamma\left(\sigma_{0v}^{-2}, \sigma_{0v}^{-2}\right)$ (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

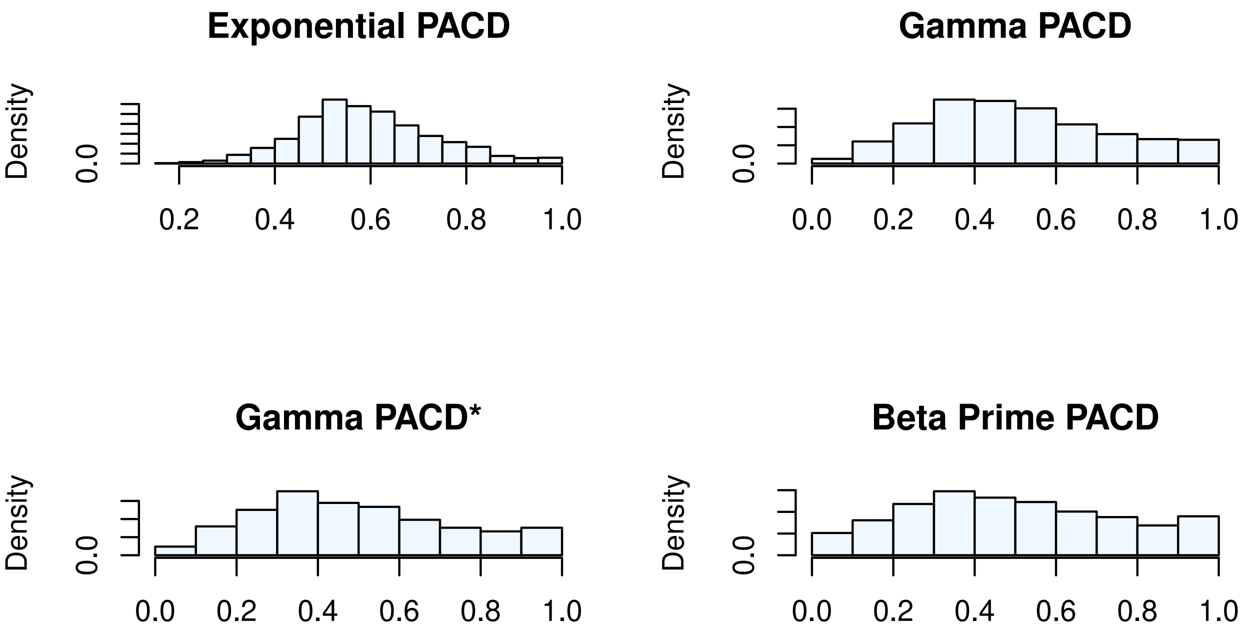



Figure 6.2: PITs of the residuals $\widehat{\xi}_t$ (BTV series) for four PACD innovation distributions.

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 $(T_f + 1, ..., T)$ based on

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

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

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 $\text{PACD}_{\text{E}}$ while those computed using the 2S-GQMLE are denoted by $\text{PACD}_{G^*}$

(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 $\text{PACD}_{\text{G}^*}$ and $\text{PACD}_{\text{E}}$ 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 |
|---|---|---|---|---|---|---|---|
| | MSFE | 207.8341 | 165.8957 | 89.6771 | 88.3917 | 94.4021 | 104.8815 |
| ACD | 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 |
| | MSFE | 193.4029 | 153.9036 | 79.1289 | 80.8411 | 80.4482 | 93.4090 |
| $\text{PACD}_{\text{E}}$ | 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 |
| | MSFE | 193.1508 | 151.0290 | 79.2446 | 80.8776 | 80.4284 | 93.9384 |
| $\text{PACD}_{\text{G}^*}$ | 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

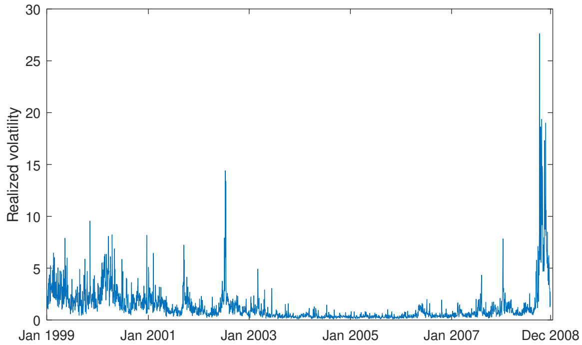



Figure 6.3: Daily realized volatility UN (UN-RV).

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.

| $\widehat{\omega}$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\sigma}^2$ | $\widehat{\alpha}+\widehat{\beta}$ | 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 $H_0^{\theta}$ 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 | $v$ | $\widehat{\sigma}_v^2$ | $\widehat{\omega}_v$ | $\widehat{\alpha}_v$ | $\widehat{\beta}_v$ | $\widehat{\alpha}_v + \widehat{\beta}_v$ | 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 | $\widehat{\phi}$ | | | | | 0.8897 | 1.0570 | 0.4757 |
| $\mathrm{W}^{\theta}$ | | 193.9721 | | | | | | |
| $\mathrm{W}^{\sigma}$ | | 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 $W_{vs}^{\sigma}$ from the tenth entail accepting $H_0^{\sigma,vs}$. On the other hand, the persistence parameter, given by $\prod_{v=1}^{7}(\widehat{\alpha}_v + \widehat{\beta}_v) = 0.8897$, 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.

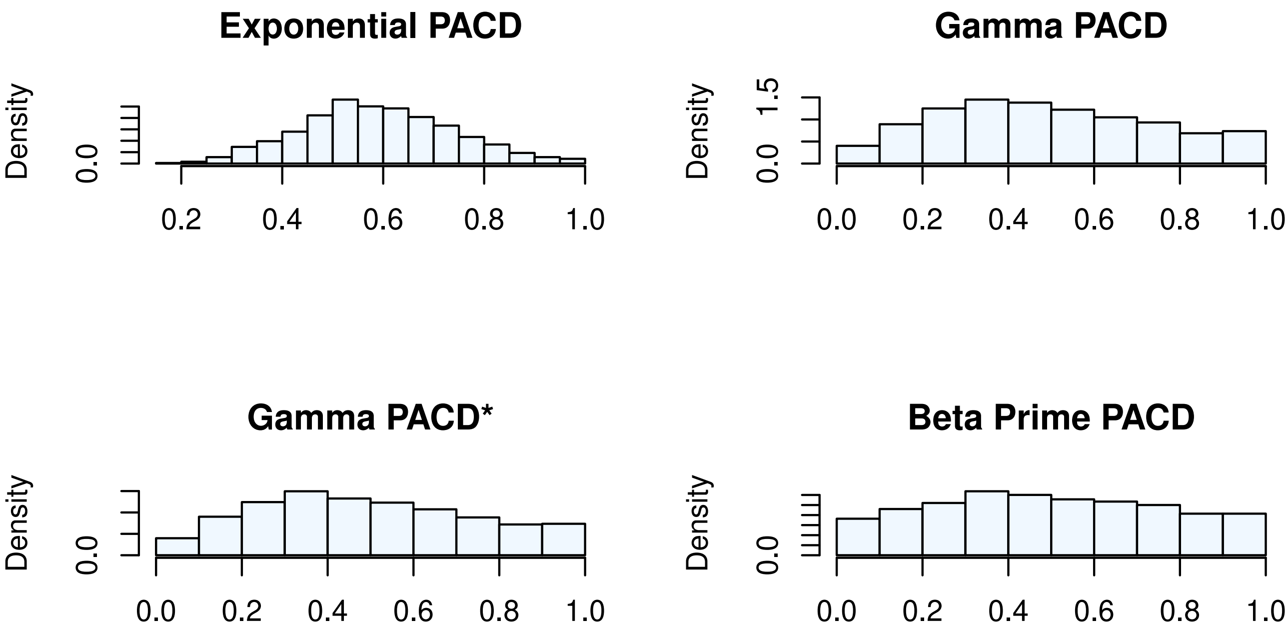



Figure 6.4: PITs of the residuals $\widehat{\xi}_t$ (UN-RV series) for four PACD innovation distributions.

| $v \setminus s$ | 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.

$W_{vs}^{\theta}$ for $v > s$; $W_{vs}^{\sigma}$ in the uper 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 $\widehat{\theta}_G^*$ 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 $\mathrm{PACD_{G^*}}$ and $\mathrm{PACD_E}$ 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 |
| $\mathrm{PACD_E}$ | 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 |
| $\mathrm{PACD_{G^*}}$ | 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.

# Appendix: Proofs

**Proof of Theorem 3.1** i) The $ipd_S$-driven SRE (3.1) can be embedded in the following system of $S$ SREs

$$\underline{Y}_{nS+v} = \mathbb{A}^0_{nS+v}\underline{Y}_{(n-1)S+v} + \mathbb{B}^0_{nS+v},\ n \in \mathbb{Z},\quad v \in \{0,...,S-1\}, \tag{A.1}$$

where $\mathbb{A}_{nS+v} = \prod_{i=0}^{S-1} A^0_{nS+v-i}$ and $\mathbb{B}^0_{nS+v} = \sum_{j=0}^{S-1}\prod_{i=0}^{j-1} A^0_{nS+v-i}B^0_{nS+v-j}$ are such that the sequence $\left\{(\mathbb{A}^0_{nS+v}, \mathbb{B}^0_{nS+v}), n \in \mathbb{Z}\right\}$ is *iid* for all $v \in \{0,...,S-1\}$. The standard top Lyapunov exponent $\gamma_v^S$ associated with the *iid*-driven SRE $(A.1)$ is given for all $v \in \{0,...,S-1\}$ by (cf. Bougerol and Picard, 1992$a$)

$$\begin{aligned}\gamma_v^S &= \inf\left\{\tfrac{1}{n}E\log\left\|\mathbb{A}^0_{nS+v}\mathbb{A}^0_{(n-1)S+v}...\mathbb{A}^0_{S+v}\right\|,\ n \geq 1\right\}\\ &= \lim_{n\to\infty}\tfrac{1}{n}\log\left\|A^0_{nS+v}A^0_{nS+v-1}...A^0_{v+1}\right\| \quad a.s.\end{aligned}$$

Since $E\xi_v < \infty$ it follows that $E\log^+\|\mathbb{A}^0_v\| < \infty$ and $E\log^+\|\mathbb{B}^0_v\| < \infty$ for all $1 \leq v \leq S$. Therefore, by Theorem 2.5 of Bougerol and Picard (1992$a$), equation $(A.1)$ admits a unique nonanticipative strictly stationary and ergodic solution $\left\{\underline{Y}_{nS+v}, n \in \mathbb{Z}, 0 \leq v \leq S-1\right\}$ under $\gamma_v^S < 0$. That solution is given for all $v \in \{0,...,S-1\}$ and $n \in \mathbb{Z}$ by

$$\underline{Y}_{nS+v} = \sum_{j=0}^{\infty}\prod_{i=0}^{j-1}\mathbb{A}^0_{(n-i)S+v}\mathbb{B}^0_{(n-j)S+v} \tag{A.2}$$

which is exactly (3.3), where the series in equality $(A.2)$ converges absolutely $a.s.$ This shows that $\left\{\underline{Y}_{nS+v}, n \in \mathbb{Z},\ 0 \leq v \leq S-1\right\}$ is the unique causal strictly periodically stationary and periodically ergodic solution of (3.1). By a sandwiching argument, it is easily seen that for all $v \in \{0,...,S-1\}$

$$\gamma_v^S = \lim_{n\to\infty}\tfrac{1}{n}\log\left\|A^0_{nS+v}A^0_{nS+v-1}...A^0_{v+1}\right\| = \lim_{n\to\infty}\tfrac{1}{n}\log\left\|A^0_{nS}A^0_{nS-1}...A^0_1\right\| := \gamma^S.$$

ii) If model (3.1) admits a nonanticipative strictly periodically stationary solution $\left\{\underline{Y}_{nS+v}, n \in \mathbb{Z}, 0 \leq v \leq S-1\right\}$ then from the non-negativity of the coefficients of $A^0_{nS+v}$ in

(3.1) it follows that for all $k > 1$,

$$\underline{Y}_v \geq \sum_{j=0}^{k} \prod_{i=0}^{j-1} A_{v-i}^0 B_{v-j}^0, \quad a.s.,$$

so the series $\sum_{j=0}^{\infty} \prod_{i=0}^{j-1} A_{v-i}^0 B_{v-j}^0$ converges $a.s.$ Therefore $\prod_{i=0}^{j-1} A_{v-i}^0 B_{v-j}^0 \to 0$ $a.s.$ as $j \to \infty$, from which we have to show that $\prod_{i=0}^{j-1} A_{v-i}^0 \to 0$ $a.s.$ as $j \to \infty$. This holds whenever

$$\lim_{j\to\infty} \prod_{i=0}^{j-1} A_{v-i}^0 e_m = 0, a.s. \text{ for all } 1 \leq m \leq r, \tag{A.3}$$

where $r = p+q$ and $(e_m)_{1\leq m\leq r}$ is the canonical basis of $\mathbb{R}^r$. Since $A_t^0$ has the same zero-structure as the matrix $A_t$ in Bougerol and Picard (1992$b$), then $(A.3)$ follows from their results using the same argument.

ii) By the nonnegativity of $\left(A_{nS+v}^0\right)_{n,v}$ we have

$$\gamma^S\left(A^0\right) \geq \gamma^S\left(\boldsymbol{\beta}^0\right) := \log \rho \left(\prod_{v=0}^{S-1} \boldsymbol{\beta}_{S-v}^0\right). \tag{A.4}$$

If (3.1) admits a strictly periodically stationary solution then $\gamma^S\left(A^0\right) < 0$. In view of $(A.4)$, it follows that

$$\gamma^S\left(\boldsymbol{\beta}^0\right) < 0 \tag{A.5}$$

which in turn implies (3.4). □

**Proof of Theorem 3.2** Theorem 3.2 is a particular case of Theorem 3.3, ii). □

**Proof of Theorem 3.3** i) The proof is similar to the one of Lemma 2.3 of Berkes et *al* (2003). See Supplementary material.

ii) Define $\left\{\widetilde{Y}_{nS+v}, n \in \mathbb{N},\, 0 \leq v \leq S-1\right\}$ by

$$\begin{cases} \widetilde{Y}_{nS+v} = A_{nS+v}^0 \widetilde{Y}_{nS+v-1} + B_{nS+v}^0 & n \geq 0,\, v > 0 \\ \widetilde{Y}_{nS+v} = 0 \qquad n \leq -1,\; 0 \leq v \leq S \end{cases} \tag{A.6}$$

and let $\underline{Y}^{(v)}$ $(0 \leq v \leq S-1)$ be a random variable having the same distribution as the term $\underline{Y}_{nS+v}$ of the unique periodically stationary solution given by (3.1). By construction $\widetilde{Y}_{nS+v} \overset{\mathscr{L}}{\to}$

$\underline{Y}^{(v)}$ as $n \to \infty$. Let $m = 2$. From the weak convergence theory (Billingsley, 1999), to show that $E\left(vec\left(\underline{Y}^{(v)}\underline{Y}^{(v)\prime}\right)\right)$ is finite for all $v$, it is sufficient to show that $\lim\inf_{n\to\infty} E\left(vec\left(\widetilde{Y}'_{nS+v}\widetilde{Y}_{nS+v}\right)\right) < \infty$ for all $v$. Set $V_{nS+v} = E\left(vec\left(\widetilde{Y}'_{nS+v}\widetilde{Y}_{nS+v}\right)\right)$. From $(A.6)$ we get the following first-order $S$-periodic difference equation

$$V_{nS+v} = E\left(A_v^{0\otimes 2}\right) V_{nS+v-1} + \left[E(A_v^0 \otimes B_v^0) + E(B_v^0 \otimes A_v^0)\right] E\left(\widetilde{Y}_{nS+v}\right) + vec(E\left(B_v^0 B_v^{0\prime}\right)). \quad (A.7)$$

The latter result has been obtained while using the fact that $E\left(A_t^{0\otimes 2}\right)$, $E(A_t^0 \otimes B_t^0)$ and $vec(E\left(B_t^0 B_t^{0\prime}\right))$ are finite $S$-periodic matrices over $t$. Since, the last two terms of the right-hand side of $(A.7)$ are bounded, it follows that $\lim_{n\to\infty} V_{nS+v}$ exists for every $1 \le v \le S$ as long as (3.8) holds, from which follows the proof for $m = 2$. For general $m$, the proof is similar. □

Before giving the proof of Proposition 3.1, we need to state the following well-known result on linear ordinary periodic difference equations. Let

$$y_{nS+v} = a_v y_{nS+v-1} + b_v, \qquad n \in \mathbb{Z},\ 0 \le v \le S-1 \qquad (A.8)$$

be an ordinary difference equation with *$S$-periodic positive coefficients $a_v = a_{nS+v} > 0$ and* $b_v = b_{nS+v} > 0$ for all $n \in \mathbb{Z}$ *and* $0 \le v \le S-1$.

**Lemma 1** *The real-valued ordinary difference equation $(A.8)$ admits a unique solution $\{y_{nS+v}, n \in \mathbb{Z}, 0 \le v \le S-1\}$ if and only if*

$$\prod_{v=1}^{S} a_v < 1. \qquad (A.9)$$

**Proof of Proposition 3.1** It is well-known that if $Y_{nS+v}|\mathscr{F}_{nS+v-1} \sim \Gamma\left(\frac{1}{\sigma_{0v}^2}, \frac{1}{\sigma_{0v}^2\psi_{nS+v}}\right)$ then the conditional moments up to the fourth order are given by

$$E\left(Y_{nS+v}|\mathscr{F}_{nS+v-1}\right) = \psi_{nS+v} \qquad (A.10a)$$

$$E\left(Y_{nS+v}^2|\mathscr{F}_{nS+v-1}\right) = \left(1+\sigma_{0v}^2\right)\psi_{nS+v}^2 \qquad (A.10b)$$

$$E\left(Y_{nS+v}^3|\mathscr{F}_{nS+v-1}\right) = \left(1+\sigma_{0v}^2\right)\left(1+2\sigma_v^2\right)\psi_{nS+v}^3 \qquad (A.10c)$$

$$E\left(Y_{nS+v}^4|\mathscr{F}_{nS+v-1}\right) = \left(1+\sigma_{0v}^2\right)\left(1+2\sigma_{0v}^2\right)\left(1+3\sigma_{0v}^2\right)\psi_{nS+v}^4. \qquad (A.10d)$$

In view of $(A.10)$ it turns out that the conditional moment of $Y_v$ of order $i$ is a polynomial in $\psi_{nS+v}$ with degree $i$ $(i = 1, 2, 3, 4)$. Hence $EY_v^i < \infty$ if and only if $E\psi_v^i < \infty$ $(i = 1, 2, 3, 4)$, $(1 \leq v \leq S)$, conditions for which are given as follows.

i) Expanding $E\psi_{nS+v}$ using $(2.1b)$ with $p = q = 1$ and $(A.10a)$, we find the following linear $S$-periodic difference equation

$$E\psi_{nS+v} = (\alpha_{0v} + \beta_{0v})\, E\psi_{nS+v-1} + \omega_{0v}, \quad n \in \mathbb{Z},\ 0 \leq v \leq S-1. \qquad (A.11)$$

By Lemma 1, there is a unique solution of $(A.11)$ if and only if (3.6) holds.

ii) For the existence of the second moments $EY_v^2$ $(1 \leq v \leq S)$, expanding $E\psi_{nS+v}^2$ using $(2.1b)$ and $(3.10a)$-$(3.10b)$, we find the following linear periodic difference equation

$$E\psi_{nS+v}^2 = \left(\alpha_{0v}^2 E\xi_{nS+v-1}^2 + 2\alpha_{0v}\beta_{0v} + \beta_{0v}^2\right) E\psi_{nS+v-1}^2 + K_v^{(1)},\ n \in \mathbb{Z},\ 0 \leq v \leq S-1, \quad (A.12)$$

where

$$K_v^{(1)} = (2\alpha_{0v}\omega_{0v} + 2\beta_{0v}\omega_{0v})\, E\psi_{v-1} + \omega_{0v}^2$$

is finite if and only if $E\psi_{v-1} < \infty$, and thus if and only if (3.6) holds. By Lemma 1, there exists a unique solution to $(A.12)$ if and only if (3.6) and (3.9) are satisfied.

iii) Expanding $E\psi_{nS+v}^3$ using $(2.1b)$ and $(A.10a)$-$(A.10c)$ we get the following linear periodic difference equation

$$E\psi_{nS+v}^3 = \left(\alpha_{0v}^3 E\xi_{nS+v-1}^3 + 3\alpha_{0v}^2\beta_{0v}\left(\sigma_{0,v-1}^2 + 1\right) + \beta_{0v}^3 + 3\alpha_{0v}\beta_{0v}^2\right) E\psi_{nS+v-1}^3 + K_v^{(2)}, \quad (A.13)$$

where

$$K_v^{(2)} = \left(3\alpha_{0v}^2\omega_{0v}E\xi_{v-1}^2 + 6\alpha_{0v}\beta_{0v}\omega_{0v} + 3\beta_{0v}^2\omega_{0v}\right) E\psi_{v-1}^2 + \left(3\beta_{0v}\omega_{0v}^2 + 3\alpha_{0v}\omega_{0v}^2\right) E\psi_{v-1} + \omega_{0v}^3$$

is finite if and only if $E\psi_{v-1}^2 < \infty$ and $E\psi_{v-1} < \infty$. By Lemma 1, equation $(A.13)$ admits a unique solution if and only if (3.6), (3.9) and (3.10) hold.

iv) Expanding $E\psi_{nS+v-1}^4$ using $(2.1b)$ and $(A.10a)$-$(A.10d)$ we get the following linear periodic difference equation

$$E\psi_{nS+v}^4 = \left[\alpha_{0v}^4 E\xi_{v-1}^4 + 4\alpha_{0v}^3\beta_{0v}\left(1+\sigma_{0,v-1}^2\right)\left(1+2\sigma_{v-1}^2\right) + 6\alpha_{0v}^2\beta_{0v}^2\left(1+\sigma_{0,v-1}^2\right) + \right.$$
$$\left.\left(4\alpha_{0v}\beta_{0v}^3 + \beta_{0v}^4\right)\right] E\psi_{nS+v-1}^4 + K_v^{(3)}, \qquad (A.14)$$

where

$$K_v^{(3)} = 4\alpha_{0v}^3\omega_{0v}EY_{v-1}^3 + 12\alpha_{0v}^2\beta_{0v}\omega_{0v}E\left(Y_{v-1}^2\psi_{v-1}\right) + 12\alpha_{0v}\beta_{0v}^2\omega_{0v}E\left(Y_{v-1}\psi_{v-1}^2\right)$$
$$+4\beta_{0v}^3\omega_{0v}E\psi_{v-1}^3 + 6\alpha_{0v}^2\omega_{0v}^2EY_{v-1}^2 + 12\alpha_{0v}\beta_{0v}\omega_{0v}^2E\left(Y_{v-1}\psi_{v-1}\right) + 6\beta_{0v}^2\omega_{0v}^2E\psi_{v-1}^2$$
$$+4\alpha_{0v}\omega_{0v}^3EY_{v-1} + 4\beta_{0v}\omega_{0v}^3E\psi_{v-1} + \omega_{0v}^4$$

is finite under (3.6), (3.9) and (3.10). By Lemma 1, equation ($A.14$) admits a unique positive solution if and only if (3.6), (3.9), (3.10) and (3.11) hold.

**Proof of Theorem 4.1** Theorem 4.1 will be proved by showing several lemmas below. In what follows $M > 0$ and $\rho \in (0,1)$ stand for constants that are not necessarily the same when appearing in different terms. Let $L_T\left(\theta, \boldsymbol{\sigma}^2\right)$ and $l_{nS+v}\left(\theta, \sigma_v^2\right)$ be defined in the same way as $\widetilde{L}_T\left(\theta, \boldsymbol{\sigma}^2\right)$ and $\widetilde{l}_{nS+v}\left(\theta, \sigma_v^2\right)$ in ($4.3a$) and ($4.3b$), respectively, with $\psi_{nS+v}(\theta)$ in place of $\widetilde{\psi}_{nS+v}(\theta)$.

**Lemma 1** *Under **A1** and **A2** we have*

$$\sup_{\theta\in\Theta}\left|L_T\left(\theta, \boldsymbol{\sigma}^2\right) - \widetilde{L}_T\left(\theta, \boldsymbol{\sigma}^2\right)\right| \to 0 \ \ a.s. \text{ as } T \to \infty.$$

**Proof** Rewrite ($4.2b$) in a vector form as follows

$$\underline{\psi}_{nS+v} = \boldsymbol{\beta}_v \underline{\psi}_{nS+v-1} + \underline{\boldsymbol{\alpha}}_{nS+v}, \ n \in \mathbb{Z}, \ 0 \le v \le S-1, \qquad (A.15)$$

where $\underline{\psi}_{nS+v} = \left(\psi_{nS+v}(\theta), ..., \psi_{nS+v-p+1}(\theta)\right)'$ and $\underline{\boldsymbol{\alpha}}_{nS+v} = \left(\omega_v + \sum_{i=1}^{q}\alpha_{vi}Y_{nS+v-i}, 0, ..., 0\right)'_{1\times p}$. By **A1** and the assumption **A2** of compactness of $\Theta$ it follows that

$$\sup_{\theta\in\Theta}\rho\left(\prod_{v=0}^{S-1}\boldsymbol{\beta}_{S-v}\right) < 1. \qquad (A.16)$$

Iterating $(A.15)$ gives

$$\underline{\psi}_{nS+v} = \sum_{k=0}^{nS+v-1} \prod_{i=0}^{k-1} \boldsymbol{\beta}_{nS+v-i} \underline{\boldsymbol{\alpha}}_{nS+v-k} + \prod_{i=0}^{nS+v} \boldsymbol{\beta}_{nS+v-i} \underline{\psi}_0, \ n \in \mathbb{Z}, \ 0 \le v \le S-1.$$

Denote by $\widetilde{\underline{\psi}}_{nS+v}$ and $\widetilde{\underline{\boldsymbol{\alpha}}}_{nS+v}$ the vectors obtained from $\underline{\psi}_{nS+v}$ and $\underline{\boldsymbol{\alpha}}_{nS+v}$, respectively, while replacing $\psi_{nS+v-j}(\theta)$ by $\widetilde{\psi}_{nS+v-j}$ with fixed initial values. From $(A.15)$ and $(A.16)$ we thus get

$$\begin{aligned}
& \sup_{\theta \in \Theta} \left\| \underline{\psi}_{nS+v} - \widetilde{\underline{\psi}}_{nS+v} \right\| \\
= & \sup_{\theta \in \Theta} \left\| \sum_{k=nS+v-q}^{nS+v-1} \prod_{i=0}^{k-1} \boldsymbol{\beta}_{nS+v-i} \left( \underline{\boldsymbol{\alpha}}_{nS+v-k} - \widetilde{\underline{\boldsymbol{\alpha}}}_{nS+v-k} \right) + \prod_{i=0}^{nS+v-1} \boldsymbol{\beta}_{nS+v-i} \left( \underline{\psi}_0 - \widetilde{\underline{\psi}}_0 \right) \right\| \\
\le & \ M\rho^{nS+v}. \qquad (A.17)
\end{aligned}$$

Using the inequality $\left|\log \frac{y}{x}\right| \le \frac{|y-x|}{\min(y,x)}$ for positive $x$ and $y$ (cf. Francq and Zakoian, 2019) it follows that

$$\begin{aligned}
& \sup_{\theta \in \Theta} \left| L_T\left(\theta, \boldsymbol{\sigma}^2\right) - \widetilde{L}_T\left(\theta, \boldsymbol{\sigma}^2\right) \right| \le \frac{1}{NS} \sum_{n=0}^{N-1} \sum_{v=1}^{S} \sup_{\theta \in \Theta} \frac{1}{\sigma_v^2} \left[ \left| \frac{\widetilde{\psi}_{nS+v} - \psi_{nS+v}(\theta)}{\widetilde{\psi}_{nS+v} \psi_{nS+v}(\theta)} \right| Y_{nS+v} + \left| \log\left( \frac{\psi_{nS+v}(\theta)}{\widetilde{\psi}_{nS+v}(\theta)} \right) \right| \right] \\
\le & \max_{1 \le v \le S} \sup_{\theta \in \Theta} \left( \frac{\omega_v^{-2}}{\sigma_v^2} \right) \frac{M}{NS} \sum_{n=0}^{N-1} \sum_{v=1}^{S} \rho^{nS+v} Y_{nS+v} + \max_{1 \le v \le S} \sup_{\theta \in \Theta} \left( \frac{\omega_v^{-2}}{\sigma_v^2} \right) \frac{M}{NS} \sum_{n=0}^{N-1} \sum_{v=1}^{S} \rho^{nS+v}.
\end{aligned}$$

The existence of $EY_v^{\delta}$ (cf, Theorem 3.3, i)) implies, by the Borel-Cantelli lemma, that $\rho^{nS+v} Y_{nS+v} \to 0$ $a.s.$ as $n \to 0$ and the conclusion follows by Césaro's lemma. $\square$

**Lemma 2** *Under **A1-A4** there is $nS+v \in \mathbb{Z}$ such that $\psi_{nS+v}(\theta) = \psi_{nS+v}(\theta_0)$ a.s. if and only if $\theta = \theta_0$.*

**Proof** From the assumption $\rho\left(\prod_{v=0}^{S-1} \boldsymbol{\beta}_{S-v}\right) < 1$ in **A1**, the polynomials $(\beta_v(L))_v$ are invertible for all $1 \le v \le S$ and all $\theta \in \Theta$. Assume $\psi_{nS+v}(\theta) = \psi_{nS+v}(\theta_0)$ $a.s.$ for some $n \in \mathbb{Z}$, $0 \le v \le S-1$. Using the second equality in (4.1) and (4.2b) we have

$$\left( \frac{\alpha_v(L)}{\beta_v(L)} - \frac{\alpha_v^0(L)}{\beta_v^0(L)} \right) Y_{nS+v} = \left( \frac{\omega_v^0}{\beta_v^0(1)} - \frac{\omega_v}{\beta_v(1)} \right) \text{ for all } 1 \le v \le S.$$

If $\frac{\alpha_v(L)}{\beta_v(L)} \neq \frac{\alpha_v^0(L)}{\beta_v^0(L)}$ for some $0 \le v \le S-1$ then there exists a deterministic periodic time-varying combination of $Y_{nS+v-j}$, $j \ge 1$. This contradicts **A4** which assumes $(\xi_{nS+v})_{n,v}$ non-degenerate, since by (2.6) we have $Y_{nS+v} - E\left(Y_{nS+v} | \mathscr{F}_{nS+v-1}\right) = \psi_{nS+v}(\xi_{nS+v} - 1) \neq 0$ with positive probability. Therefore,

$$\frac{\alpha_v(z)}{\beta_v(z)} = \frac{\alpha_v^0(z)}{\beta_v^0(z)}\ \forall\,|z| \leq 1 \text{ and } \frac{\omega_v^0}{\beta_v^0(1)} - \frac{\omega_v}{\beta_v(1)} \text{ for all } 1 \leq v \leq S,$$

and, by the assumption **A3** of no common roots between $\alpha_v^0(z)$ and $\beta_v^0(z)$, it follows that $\alpha_v(z) = \alpha_v^0(z)$, $\beta_v(z) = \beta_v^0(z)$ and $\omega_v = \omega_v^0$ for all $1 \leq v \leq S$. □

**Lemma 3** *Under* ***A1*** *for all* $\boldsymbol{\sigma}^2 > \mathbf{0}$

$$\sum_{v=1}^{S} E\left(l_v\left(\theta_0, \sigma_v^2\right)\right) < \infty,$$

*and* $\sum_{v=1}^{S} E\left(l_v\left(\theta, \sigma_v^2\right)\right)$ *is minimized at* $\theta = \theta_0$.

**Proof** By Jensen's inequality and Theorem 3.3, i) we have

$$\sum_{v=1}^{S} E\left(\log \psi_v\left(\theta_0\right)\right) = \tfrac{1}{\delta}\sum_{v=1}^{S} E\left(\log \psi_v\left(\theta_0\right)^{\delta}\right) \leq \tfrac{1}{\delta}\sum_{v=1}^{S} \log E\left(\psi_v\left(\theta_0\right)^{\delta}\right) < \infty.$$

Hence

$$\sum_{v=1}^{S} E\left(l_v\left(\theta_0, \sigma_v^2\right)\right) = \sum_{v=1}^{S} \tfrac{1}{\sigma_v^2} E\left[\xi_v + \log \psi_v\left(\theta_0\right)\right] = \sum_{v=1}^{S} \tfrac{1}{\sigma_v^2} + \sum_{v=1}^{S} \tfrac{1}{\sigma_v^2} E\left(\log \psi_v\left(\theta_0\right)\right) < \infty.$$

Using the inequality $\log(x) \leq x - 1$ we have for all $\theta \in \Theta$

$$\begin{aligned} \sum_{v=1}^{S} E\left(l_v\left(\theta, \sigma_v^2\right)\right) - \sum_{v=1}^{S} E\left(l_v\left(\theta_0, \sigma_v^2\right)\right) = & \sum_{v=1}^{S} \tfrac{1}{\sigma_v^2} E\left[\log\left(\tfrac{\psi_v(\theta)}{\psi_v(\theta_0)}\right) + \tfrac{\psi_v(\theta_0)}{\psi_v(\theta)} - 1\right] \\ \geq & \sum_{v=1}^{S} \tfrac{1}{\sigma_v^2} E\left[\log \tfrac{\psi_v(\theta)}{\psi_v(\theta_0)} + \log \tfrac{\psi_v(\theta_0)}{\psi_v(\theta)}\right] = 0, \qquad (A.18) \end{aligned}$$

showing that $\sum_{v=1}^{S} E\left(l_v\left(\theta, \sigma_v^2\right)\right)$ is minimized at $\theta_0$ for all $\boldsymbol{\sigma}^2 > \mathbf{0}$.

**Lemma 4** *For any* $\theta \neq \theta_0$ *there exists a neighborhood* $\mathscr{V}(\theta)$ *such that*

$$\lim_{N\to\infty} \inf \inf_{\overline{\theta}\in\mathscr{V}(\theta)} \widetilde{L}_{NS}\left(\overline{\theta}, \boldsymbol{\sigma}^2\right) > \tfrac{1}{S}\sum_{v=1}^{S} E\left(l_v\left(\theta_0, \sigma_v^2\right)\right).$$

**Proof** For all $\theta \in \Theta$ and any positive integer $k$, let $\mathscr{V}_k(\theta)$ be the open ball of center $\theta$ and radius $1/k$. In view of Lemma 1 we have

$$\begin{aligned}\liminf_{N\to\infty}\inf_{\overline{\theta}\in\mathscr{V}_k(\theta)\cap\Theta}\widetilde{L}_T\left(\overline{\theta},\boldsymbol{\sigma}^2\right) &\geq \liminf_{N\to\infty}\inf_{\overline{\theta}\in\mathscr{V}_k(\theta)\cap\Theta}L_T\left(\overline{\theta},\boldsymbol{\sigma}^2\right)-\liminf_{N\to\infty}\inf_{\theta\in\Theta}\left|L_T\left(\theta,\boldsymbol{\sigma}^2\right)-\widetilde{L}_T\left(\theta,\boldsymbol{\sigma}^2\right)\right|\\ &\geq \liminf_{N\to\infty}\tfrac{1}{N}\sum_{n=0}^{N-1}\tfrac{1}{S}\sum_{v=1}^{S}\inf_{\overline{\theta}\in\mathscr{V}_k(\theta)\cap\Theta}l_{nS+v}\left(\overline{\theta},\sigma_v^2\right).\end{aligned}$$

By the ergodic theorem for the stationary sequence $\left\{\sum_{v=1}^{S}l_{nS+v}\left(\overline{\theta},\sigma_v^2\right)\right\}_n$ with $E\left(\sum_{v=1}^{S}l_{nS+v}\left(\overline{\theta},\sigma_v^2\right)\right)\in\mathbb{R}\cup\{\infty\}$ (cf, Billingsley 1995, p. 495), it follows that

$$\liminf_{N\to\infty}\tfrac{1}{N}\sum_{n=0}^{N-1}\tfrac{1}{S}\sum_{v=1}^{S}\inf_{\overline{\theta}\in\mathscr{V}_k(\theta)\cap\Theta}l_{nS+v}\left(\overline{\theta},\sigma_v^2\right)=\tfrac{1}{S}\sum_{v=1}^{S}E\left(\inf_{\overline{\theta}\in\mathscr{V}_k(\theta)\cap\Theta}l_v\left(\overline{\theta},\sigma_v^2\right)\right).$$

Beppo-Levi's theorem (e.g. Billingsley, 1995, p. 219) yields

$$\tfrac{1}{S}\sum_{v=1}^{S}E\left(\inf_{\overline{\theta}\in\mathscr{V}_k(\theta)\cap\Theta}l_{nS+v}\left(\overline{\theta},\sigma_v^2\right)\right)\to\tfrac{1}{S}\sum_{v=1}^{S}E\left(l_v\left(\theta,\sigma_v^2\right)\right)\text{ as }k\to\infty,$$

and by $(A.18)$ the result follows. □

Finally, the proof of Theorem 4.1 is completed by standard compactness arguments using Lemmas 2-4. □

**Proof of Theorem 4.2** The proof of Theorem 4.2 is based on a Taylor expansion of $\frac{\partial\widetilde{L}_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta}$ at $\theta_0$ which, by **A5** and the strong consistency of $\widehat{\theta}_G$, yields

$$0=\sqrt{N}\frac{\partial\widetilde{L}_T\left(\widehat{\theta}_G,\boldsymbol{\sigma}^2\right)}{\partial\theta}=\sqrt{N}\frac{\partial L_T\left(\theta_0,\boldsymbol{\sigma}^2\right)}{\partial\theta}+\sqrt{N}\frac{\partial^2L_T\left(\theta^*,\boldsymbol{\sigma}^2\right)}{\partial\theta\partial\theta'}\left(\widehat{\theta}_G-\theta_0\right)+\sqrt{N}\left(\frac{\partial\widetilde{L}_T\left(\widehat{\theta}_G,\boldsymbol{\sigma}^2\right)}{\partial\theta}-\frac{\partial L_T\left(\widehat{\theta}_G,\boldsymbol{\sigma}^2\right)}{\partial\theta}\right)$$

where $\theta^*$ is between $\widehat{\theta}_G$ and $\theta_0$. The derivatives $\frac{\partial L_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta}$ and $\frac{\partial^2L_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta\partial\theta'}$ are given by

$$\frac{\partial L_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta}=\frac{1}{NS}\sum_{n=0}^{N-1}\sum_{v=1}^{S}\frac{\partial l_{nS+v}\left(\theta,\sigma_v^2\right)}{\partial\theta}=\frac{1}{T}\sum_{n=0}^{N-1}\sum_{v=1}^{S}\left(1-\frac{Y_{nS+v}}{\psi_{nS+v}(\theta)}\right)\frac{1}{\sigma_v^2\psi_{nS+v}(\theta)}\frac{\partial\psi_{nS+v}(\theta)}{\partial\theta}$$

$$\frac{\partial^2L_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta\partial\theta'}=$$

$$\frac{1}{NS}\sum_{n=0}^{N-1}\sum_{v=1}^{S}\left[\left(1-\frac{Y_{nS+v}}{\psi_{nS+v}(\theta)}\right)\frac{1}{\sigma_v^2\psi_{nS+v}(\theta)}\frac{\partial^2\psi_{nS+v}(\theta)}{\partial\theta\partial\theta'}+\left(\frac{2Y_{nS+v}}{\psi_{nS+v}(\theta)}-1\right)\frac{1}{\sigma_v^2\psi_{nS+v}^2(\theta)}\frac{\partial\psi_{nS+v}(\theta)}{\partial\theta}\frac{\partial\psi_{nS+v}(\theta)}{\partial\theta'}\right].$$

Therefore, the asymptotic normality result (4.6) follows whenever the following lemmas are established.

**Lemma 5** *Under **A1-A2***

$$i) \sup_{\theta\in\mathscr{V}(\theta_0)} \sqrt{N}\left\|\frac{\partial L_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta}-\frac{\partial\widetilde{L}_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta}\right\| \underset{N\to\infty}{\overset{p}{\to}} 0,\quad ii) \sup_{\theta\in\mathscr{V}(\theta_0)} \sqrt{N}\left\|\frac{\partial^2 L_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta\partial\theta'}-\frac{\partial^2\widetilde{L}_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta\partial\theta'}\right\| \underset{N\to\infty}{\overset{p}{\to}} 0$$

*for some neighborhood* $\mathscr{V}(\theta_0)$ *of* $\theta_0$.

**Proof** Following the same lines of Francq and Zakoian (2019, Section 7.4), it is easily seen that under **A1**,

$$E_{\theta_0}\left\|\frac{1}{\sigma_v^2\psi_v(\theta)}\frac{\partial\psi_v(\theta_0)}{\partial\theta}\right\|<1,\ E_{\theta_0}\left\|\frac{1}{\sigma_v^2\psi_v^2(\theta)}\frac{\partial^2\psi_v(\theta_0)}{\partial\theta\partial\theta'}\right\|<1,\ E_{\theta_0}\left\|\frac{1}{\sigma_v^2\psi_v^2(\theta)}\frac{\partial\psi_v(\theta_0)}{\partial\theta}\frac{\partial\psi_v(\theta_0)}{\partial\theta'}\right\|<1. \tag{A.19}$$

By $(A.17)$, the compactness of $\Theta$ (cf. **A2**) and the fact that $\rho\left(\prod_{v=0}^{S-1}\boldsymbol{\beta}_{S-v}\right)<1$ (cf. **A1**) we have

$$\begin{aligned}
\left|\frac{1}{\widetilde{\psi}_{nS+v}(\theta)}-\frac{1}{\psi_{nS+v}(\theta)}\right| &\leq \frac{M\rho^{nS+v}}{\psi_{nS+v}(\theta)}\frac{\psi_{nS+v}(\theta)}{\widetilde{\psi}_{nS+v}(\theta)}(1+M)\rho^{nS+v},\\
\sup_{\theta\in\Theta}\left|\frac{\partial\psi_{nS+v}(\theta)}{\partial\theta}-\frac{\partial\widetilde{\psi}_{nS+v}(\theta)}{\partial\theta}\right| &\leq M\rho^{nS+v},\ \sup_{\theta\in\Theta}\left|\frac{\partial^2\psi_{nS+v}(\theta)}{\partial\theta\partial\theta'}-\frac{\partial^2\widetilde{\psi}_{nS+v}(\theta)}{\partial\theta\partial\theta'}\right|\leq M\rho^{nS+v},
\end{aligned}$$

and

$$\begin{aligned}
&\sup_{\theta\in\mathscr{V}(\theta_0)}\sqrt{N}\left\|\frac{\partial L_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta}-\frac{\partial\widetilde{L}_T\left(\theta,\boldsymbol{\sigma}^2\right)}{\partial\theta}\right\|=\\
&\sup_{\theta\in\mathscr{V}(\theta_0)}\frac{1}{S\sqrt{N}}\left\|\sum_{n=0}^{N-1}\sum_{v=1}^{S}\left[\left(\frac{1}{\widetilde{\psi}_{nS+v}(\theta)}-\frac{1}{\psi_{nS+v}(\theta)}\right)\frac{Y_{nS+v}}{\sigma_v^2\psi_{nS+v}(\theta)}\frac{\partial\psi_{nS+v}(\theta)}{\partial\theta}\right.\right.\\
&+\frac{1}{\sigma_v^2}\left(1-\frac{Y_{nS+v}}{\psi_{nS+v}(\theta)}\right)\left(\frac{1}{\widetilde{\psi}_{nS+v}(\theta)}-\frac{1}{\psi_{nS+v}(\theta)}\right)\frac{\partial\psi_{nS+v}(\theta)}{\partial\theta}\\
&\left.\left.+\left(1-\frac{Y_{nS+v}}{\psi_{nS+v}(\theta)}\right)\frac{1}{\sigma_v^2\widetilde{\psi}_{nS+v}(\theta)}\left(\frac{\partial\psi_{nS+v}(\theta)}{\partial\theta}-\frac{\partial\widetilde{\psi}_{nS+v}(\theta)}{\partial\theta}\right)\right]\right\|\\
&\leq\frac{M}{S\sqrt{N}}\sum_{n=0}^{N-1}\sum_{v=1}^{S}\rho^n\left(1+\xi_{nS+v}\right)\left\|1+\frac{1}{\psi_{nS+v}(\theta_0)}\frac{\partial\psi_{nS+v}(\theta_0)}{\partial\theta}\right\|.
\end{aligned}\tag{A.20}$$

Therefore, $(A.20)$ and the Markov inequality implies that for all $\varepsilon>0$,

$$\begin{aligned}
&P\left(\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\sum_{v=1}^{S}\rho^n\left(1+\xi_{nS+v}\right)\left\|1+\frac{1}{\psi_{nS+v}(\theta_0)}\frac{\partial\psi_{nS+v}(\theta_0)}{\partial\theta}\right\|>\varepsilon\right)\\
\leq\ &\frac{2}{\varepsilon}\left(1+E_{\theta_0}\left\|1+\frac{1}{\psi_{nS+v}(\theta_0)}\frac{\partial\psi_{nS+v}(\theta_0)}{\partial\theta}\right\|\right)\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\rho^n\to 0 \text{ as } N\to 0,
\end{aligned}$$

from which the result i) follows. The same argument shows result ii). □

**Lemma 6** *Under **A1**-**A6**,*

$$\frac{\partial^2 L_T\left(\theta^*,\boldsymbol{\sigma}^2\right)}{\partial\theta\partial\theta'} \underset{N\to\infty}{\overset{p}{\to}} \frac{1}{S} J\left(\theta_0,\boldsymbol{\sigma}^2\right)$$

*for any $\theta^*$ between $\widehat{\theta}_G$ and $\theta_0$.*

**Proof** Let $\mathscr{V}_k(\theta_0)$ $(k \in \mathbb{N}^*)$ be the open ball with center $\theta_0$ and radius $1/k$. Assume that $n$ is large enough so that $\theta^*$ belongs to $\mathscr{V}_k(\theta_0)$. By periodic stationarity and periodic ergodicity of $\left\{\sup_{\theta\in\mathscr{V}_k(\theta_0)}\left|\frac{\partial^2 l_{nS+v}\left(\theta,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_{nS+v}\left(\theta_0,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j}\right)\right|\right\}$ we have

$$\begin{aligned}
\left|\frac{\partial^2 L_T\left(\theta^*,\boldsymbol{\sigma}^2\right)}{\partial\theta_i\partial\theta_j} - J\left(\theta_0,\boldsymbol{\sigma}^2\right)_{i,j}\right| &= \left|\frac{\partial^2 L_T\left(\theta^*,\boldsymbol{\sigma}^2\right)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 L_T\left(\theta_0,\boldsymbol{\sigma}^2\right)}{\partial\theta_i\partial\theta_j}\right)\right| \\
&= \frac{1}{NS}\left|\sum_{n=0}^{N-1}\sum_{v=1}^{S}\frac{\partial^2 l_{nS+v}\left(\theta^*,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_{nS+v}\left(\theta_0,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j}\right)\right| \\
&\leq \frac{1}{NS}\sum_{n=0}^{N-1}\sum_{v=1}^{S}\sup_{\theta\in\mathscr{V}_k(\theta_0)}\left|\frac{\partial^2 l_{nS+v}\left(\theta,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_{nS+v}\left(\theta_0,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j}\right)\right| \\
&\underset{N\to\infty}{\overset{a.s.}{\to}} \frac{1}{S}\sum_{v=1}^{S} E\left(\sup_{\theta\in\mathscr{V}_k(\theta_0)}\left|\frac{\partial^2 l_v\left(\theta,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_v\left(\theta_0,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j}\right)\right|\right).
\end{aligned}$$

The Lebesgue dominated convergence theorem yields

$$\lim_{k\to\infty} E \sup_{\theta\in\mathscr{V}_k(\theta_0)}\left|\frac{\partial^2 l_v\left(\theta,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_v\left(\theta_0,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j}\right)\right| = E\lim_{k\to\infty}\sup_{\theta\in\mathscr{V}_k(\theta_0)}\left|\frac{\partial^2 l_v\left(\theta,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_v\left(\theta_0,\sigma_v^2\right)}{\partial\theta_i\partial\theta_j}\right)\right| = 0,$$

from which the proof of the lemma follows. □

**Lemma 7** *Under **A1**-**A6***

$$\sqrt{N}\frac{\partial L_T\left(\theta_0,\boldsymbol{\sigma}^2\right)}{\partial\theta} \underset{N\to\infty}{\overset{\mathscr{L}}{\to}} N\left(0,\frac{1}{S^2} I\left(\theta_0,\boldsymbol{\sigma}_0^2,\boldsymbol{\sigma}^2\right)\right).$$

**Proof** It is clear that $\sqrt{N}\frac{\partial L_T\left(\theta_0,\boldsymbol{\sigma}^2\right)}{\partial\theta} = \sum_{n=0}^{N-1}\sum_{v=1}^{S}\frac{1}{S\sqrt{N}}\frac{\partial l_{nS+v}\left(\theta_0,\sigma_v^2\right)}{\partial\theta}$ is a term of a square integrable periodically ergodic Martingale. Since by the periodic ergodic theorem (cf. Supplementary material)

$$\begin{aligned}
&\tfrac{1}{N}\sum_{n=0}^{N-1}\sum_{v=1}^{S}\tfrac{\partial l_{nS+v}\left(\theta_0,\sigma_v^2\right)}{\partial\theta}\tfrac{\partial l_{nS+v}\left(\theta_0,\sigma_v^2\right)}{\partial\theta'}\\
= &\tfrac{1}{N}\sum_{n=0}^{N-1}\sum_{v=1}^{S}\left(1-\xi_{nS+v}\right)^2\tfrac{1}{\sigma_v^4\psi_{nS+v}^2(\theta_0)}\tfrac{\partial\psi_{nS+v}(\theta)}{\partial\theta}\tfrac{\partial\psi_{nS+v}(\theta)}{\partial\theta'}\underset{N\to\infty}{\overset{a.s.}{\to}} I\left(\theta_0,\boldsymbol{\sigma}_0^2,\boldsymbol{\sigma}^2\right),
\end{aligned}$$

the result thus follows from the martingale central limit theorem (e.g. Billingsley, 1995). □

**Proof of Theorem 4.3** Set $U_{v,n}(\theta)=\frac{(Y_{nS+v}-\psi_{nS+v}(\theta))^2}{\psi_{nS+v}^2(\theta)}$, and denote by $o_{a.s.}(1)$ a term converging almost surely to 0 as $N\to\infty$. If we show

$$\widehat{\sigma}_v^2=\tfrac{1}{N}\sum_{n=0}^{N-1}U_{v,n}(\theta_0)+o_{a.s.}(1) \tag{A.21}$$

then the result $(4.9a)$ would follow from standard arguments.

Now a Taylor expansion of $U_{v,n}\left(\widehat{\theta}_G\right)$ around $\theta_0$ yields

$$\widehat{\sigma}_v^2=\tfrac{1}{N}\sum_{n=0}^{N-1}U_{v,n}\left(\widehat{\theta}_G\right)=\tfrac{1}{N}\sum_{n=0}^{N-1}U_{v,n}(\theta_0)+\left(\widehat{\theta}_G-\theta_0\right)\tfrac{1}{N}\sum_{n=0}^{N-1}\tfrac{\partial U_{v,n}(\theta^*)}{\partial\theta} \tag{A.22}$$

where $\theta^*$ is between $\widehat{\theta}_G$ and $\theta_0$. Note that

$$\tfrac{\partial U_{v,n}(\theta)}{\partial\theta}=2\left(\tfrac{Y_{nS+v}^2}{\psi_{nS+v}^2(\theta)}-\tfrac{Y_{nS+v}}{\psi_{nS+v}(\theta)}\right)\tfrac{1}{\psi_{nS+v}(\theta)}\tfrac{\partial\psi_{nS+v}(\theta)}{\partial\theta}.$$

Following the same lines of Francq and Zakoian (2019, p. 197) it can be easily seen that

$$E\left(\sup_{\theta\in\mathscr{V}(\theta_0)}\tfrac{Y_{nS+v}^2}{\psi_{nS+v}^2(\theta)}\right)<\infty,\ E\left(\sup_{\theta\in\mathscr{V}(\theta_0)}\tfrac{Y_{nS+v}}{\psi_{nS+v}(\theta)}\right)<\infty,\ E\left(\sup_{\theta\in\mathscr{V}(\theta_0)}\left\|\tfrac{1}{\psi_{nS+v}(\theta)}\tfrac{\partial\psi_{nS+v}(\theta)}{\partial\theta}\right\|\right)<\infty$$

for some neighborhood $\mathscr{V}(\theta_0)$ of $\theta_0$. Hence, by the ergodic theorem and the consistency of $\widehat{\theta}_G$ we get

$$\begin{aligned}
\limsup_{N\to\infty}\left\|N^{-1}\sum_{n=0}^{N-1}\tfrac{\partial U_{v,n}(\theta^*)}{\partial\theta}\right\| &\le \limsup_{N\to\infty}N^{-1}\sum_{n=0}^{N-1}\sup_{\theta\in\mathscr{V}(\theta_0)}\left\|\tfrac{\partial U_{v,n}(\theta)}{\partial\theta}\right\|\\
&= E\left(\sup_{\theta\in\mathscr{V}(\theta_0)}\left\|\tfrac{\partial U_{v,n}(\theta)}{\partial\theta}\right\|\right)<\infty.
\end{aligned}$$

Thus

$$\left(\widehat{\theta}_G - \theta_0\right) \frac{1}{N} \sum_{n=0}^{N-1} \frac{\partial U_{v,n}(\theta^*)}{\partial \theta} = o_{a.s.}(1),$$

and in view of $(A.22)$ we obtain $(A.21)$. Result $(4.9b)$ follows by a similar argument. □

**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.

Billingsley P. 1995. *Probability and measure.* 3rd edition, Wiley, New York.

Billingsley P. 1999. Convergence of probability measures. 2nd edition, Wiley, New York.

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, Toloi 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.

Bougerol P, Picard N. 1992a. Strict stationarity of generalised autoregressive processes. *Annals of Probability* 20: 1714–1730.

Bougerol P, Picard N. 1992b. Stationarity of GARCH processes and some nonnegative time series. *Journal of Econometrics* 52: 115–127.

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 edt.

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 Economet-*

*rics* 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.

# Supplementary material for: "Periodic Autoregressive Conditional Duration"

Abdelhakim Aknouche, Bader Almohaimeed and Stefanos Dimitrakopoulos

## 1 Definitions of periodic stationarity and periodic ergodicity

A positive real-valued stochastic process $\{Y_t,\ t \in \mathbb{Z}\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ is said to be *strictly periodically stationary* with period $S \in \mathbb{N}^*$ (henceforth $sps_S$) iff each one of its $S$ corresponding "sub-processes" $\{Y_{nS+v},\ n \in \mathbb{Z}\}$ $(1 \leq v \leq S)$ is strictly stationary in the standard sense. The simplest $sps_S$ process is an $ipd_S$ sequence. The periodic analog of the ergodic theorem for $sps_S$ processes (e.g. Boyles and Gardener, 1983) can be stated as follows. If $\{Y_t,\ t \in \mathbb{Z}\}$ is $sps_S$ with $E(Y_v) < \infty$ for all $1 \leq v \leq S$ then

$$\tfrac{1}{n}\sum_{t=1}^{n} Y_t \underset{n\to\infty}{\overset{a.s.}{\to}} \tfrac{1}{S}\sum_{v=1}^{S} Y_v^*, \tag{S.1}$$

for some random variables $(Y_v^*)_{1 \leq v \leq S}$ defined on $(\Omega, \mathcal{F}, P)$ and satisfying

$$Y_v^* = \lim_{n\to\infty} \tfrac{1}{n}\sum_{k=0}^{n-1} Y_{kS+v},\ a.s.$$

When for a given season $v_0 \in \{1, ..., S\}$ the sub-process $\{Y_{nS+v_0},\ n \in \mathbb{Z}\}$ is ergodic, the limiting random variable $Y_{v_0}^*$ is almost surely constant and then $Y_{v_0}^* = E(Y_{v_0})$, almost surely. The process $\{Y_t,\ t \in \mathbb{Z}\}$ is said to be *periodically ergodic* with period $S$ $(pe_S)$ iff all sub-processes $\{Y_{nS+v},\ n \in \mathbb{Z}\}$ $(v \in \{1, ..., S\})$ are ergodic in the usual sense. The simplest $pe_S$

process is an $ipd_S$ sequence. It follows that the limiting variable in $(S.1)$ simplifies to

$$\frac{1}{n}\sum_{t=1}^{n} Y_t \underset{n\to\infty}{\overset{a.s.}{\to}} \frac{1}{S}\sum_{v=1}^{S} E\left(Y_v\right),$$

the mean of the seasonal means. Like strict stationarity and ergodicity (see e.g. Billingsley, 1995, Theorem 36.4), strict periodic stationarity and periodic ergodicity are preserved under certain periodic transformations. Indeed, if $\{Y_t,\ t\in\mathbb{Z}\}$ is $sps_S$ and periodically ergodic and if $\{Z_t, t\in\mathbb{Z}\}$ is given by $Z_t = f_t\left(..., Y_{t-1}, Y_t, Y_{t+1}, ...\right)$, where $f_t$ is a function from $\mathbb{R}^{\mathbb{Z}}$ into $\mathbb{R}$, which is measurable, $S$-periodic over $t$ ($f_t = f_{t+nS}$ for all $n$ and $t$) then so is $\{Z_t, t\in\mathbb{Z}\}$.

# 2 Proof of Theorem 3.3, i)

The proof is similar to that of Lemma 2.3 of Berkes et *al* (2003). Let us first show that if $\gamma^S\left(A^0\right) < 0$ then there exists $\delta > 0$ and $n_0$ such that

$$E\left(\left\|A^0_{n_0S}A^0_{n_0S-1}...A^0_1\right\|^{\delta}\right) < 1. \qquad (S.2)$$

Since $\gamma^S\left(A^0\right) = \inf_{n\in\mathbb{N}^*}\left\{\frac{1}{n}E\left(\log\left\|A^0_{nS}A^0_{nS-1}...A^0_1\right\|\right)\right\}$ is strictly negative, there exists a positive integer $n_0$ such that

$$E\left(\log\left\|A^0_{n_0S}A^0_{n_0S-1}...A^0_1\right\|\right) < 0.$$

Using a multiplicative norm and by the $ipd_S$ property of the sequence $\{A^0_t, t\in\mathbb{Z}\}$ we have

$$\begin{aligned} E\left(\left\|A^0_{n_0S}A^0_{n_0S-1}...A^0_1\right\|\right) &= \left\|E\left(A^0_{n_0S}A^0_{n_0S-1}...A^0_1\right)\right\| \\ &\le \left\|E\left(A^0_SA^0_{S-1}...A^0_1\right)\right\|^{n_0} < \infty. \end{aligned}$$

Let $f(x) = E\left(\left\|A^0_{n_0S}A^0_{n_0S-1}...A^0_1\right\|^{x}\right)$. Since under (3.2) in the main paper

$$f'(0) = E\left(\log\left\|A^0_{n_0S}A^0_{n_0S-1}...A^0_1\right\|\right) \le \gamma^S < 0,$$

the function $f(x)$ decreases in a neighborhood of 0 and as $f(0) = 1$, it follows that there exists $0 < \delta < 1$ such that $(S.2)$ holds. Now from (3.3) in the main paper we have for some $v \in \{1, ..., S\}$

$$\left\|\underline{Y}_v\right\| \le \sum_{k=1}^{\infty}\left\|\prod_{j=0}^{k-1}A^0_{v-j}\right\|\left\|B^0_{v-k}\right\| + \left\|B^0_v\right\|.$$

Since $0 < \kappa < 1$, it follows that

$$\|\underline{Y}_v\|^{\kappa} \leq \sum_{k=1}^{\infty} \left\| \prod_{j=0}^{k-1} A_{v-j}^{0} \right\|^{\kappa} \left\| B_{v-k}^{0} \right\|^{\kappa} + \left\| B_{v}^{0} \right\|^{\kappa},$$

and then by the independence of $\{\xi_t, t \in \mathbb{Z}\}$

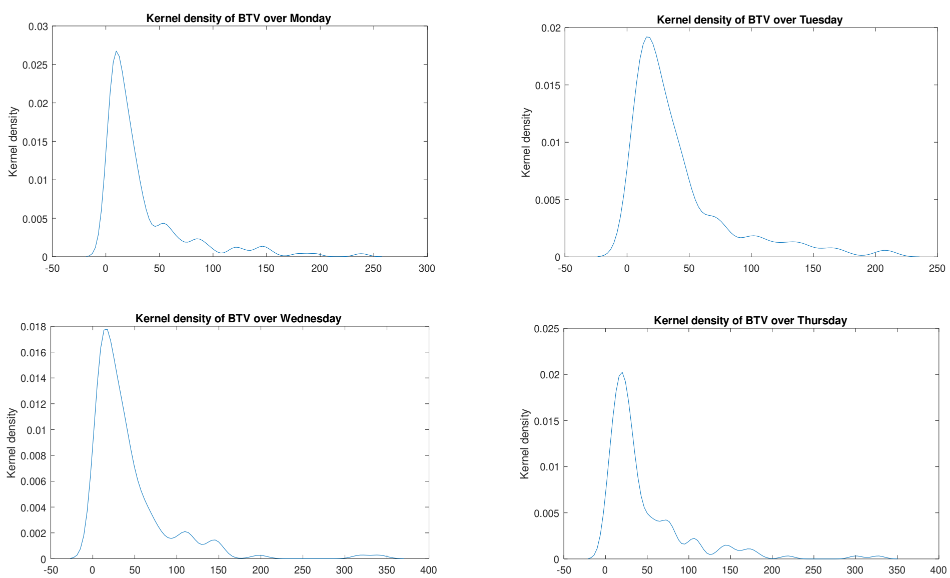
$$\begin{aligned} E\|\underline{Y}_v\|^{\kappa} &\leq \sum_{k=1}^{\infty} E\left( \left\| \prod_{j=0}^{k-1} A_{v-j}^{0} \right\|^{\kappa} \right) E\left( \left\| B_{v-k}^{0} \right\|^{\kappa} \right) + E\left( \left\| B_{v}^{0} \right\|^{\kappa} \right) \\ &\leq B^{0}(\kappa) \sum_{k=1}^{\infty} E\left( \left\| \prod_{j=0}^{k-1} A_{v-j}^{0} \right\|^{\kappa} \right) + E\left( \left\| B_{v}^{0} \right\|^{\kappa} \right), \end{aligned}$$

where $B^{0}(\kappa) = \max_{0 \leq v \leq S-1} E\left( \left\| B_{v-k}^{0} \right\|^{\kappa} \right)$. In view of $(S.2)$ there exists $a_v > 0$ and $0 < b_v < 1$ such that

$$E\left( \left\| \prod_{j=0}^{k-1} A_{v-j}^{0} \right\|^{\kappa} \right) \leq a_v b_v^{k} \leq c,$$

where $c = \max_{0 \leq v \leq S-1} \left\{ a_v b_v^{k} \right\}$. This proves that $E\|\underline{Y}_v\|^{\kappa} < \infty$ establishing the results.

# 3 Kernel densities of the Bitcoin trade volume across days



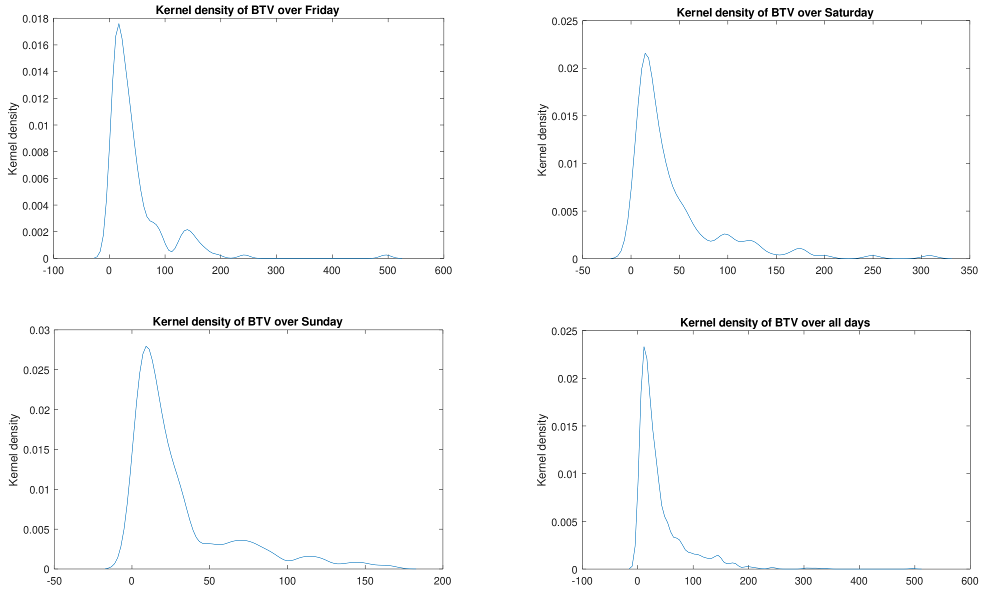
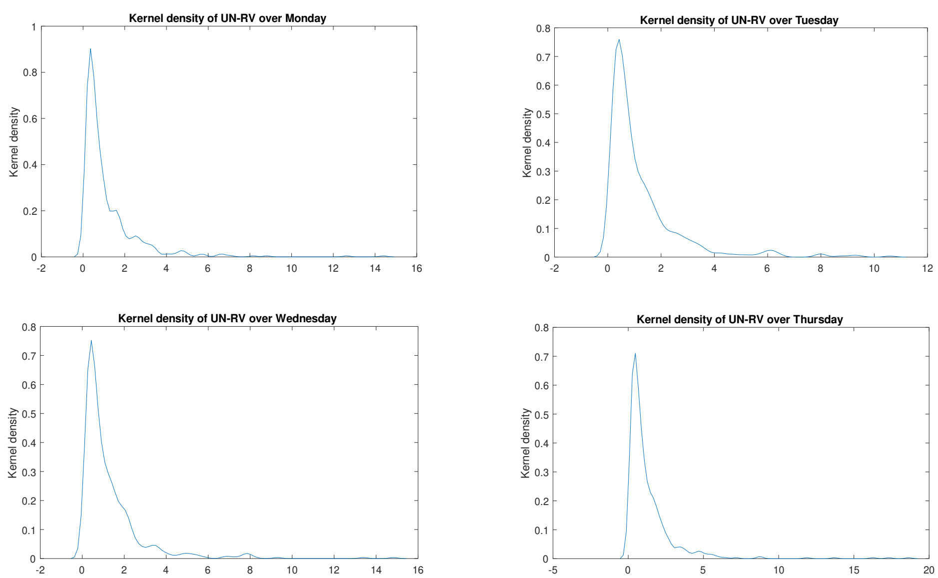



Figure S.1. Kernel densities of Bitcoin Trade Volume across days.

# 4 Kernel densities of the UN realized volatility across days



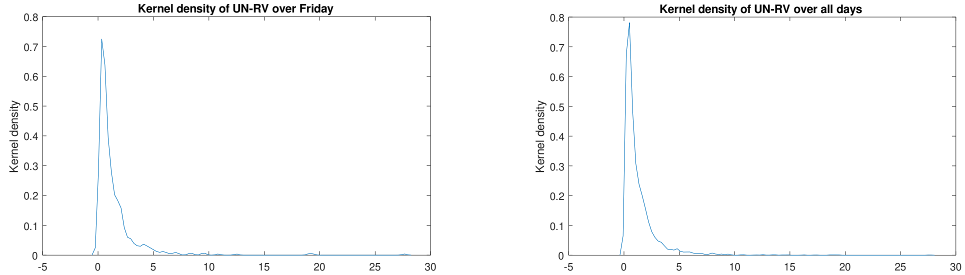



Figure S.2. Kernel densities of UN Realized Volatility across days.

# 5 Application to the S&P 500 volume

The third dataset is the daily S&P 500 volume (S&PV) over the sample period from January 03, 2011 to June, 19, 2020 with a total of $T = 2382$ observations. The time series plot of the index series is displayed in Figure S.3.

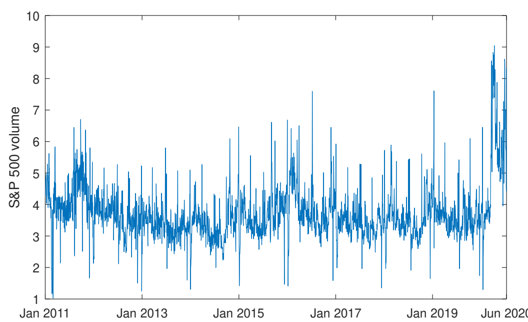



Figure S.3. S&P 500 volume series (S&PV).

We followed the same scheme as in the applications of the main paper. The conclusions were similar. The PACD dominates the ACD both in terms of the in-sample and out-of-sample forecasting criteria. So we only reports the results on Tables S.1-S.4.

| Day | Full series | Mon | Tue | Wed | Thu | Fri |
|---|---|---|---|---|---|---|
| Sample size | 2382 | 446 | 488 | 487 | 481 | 480 |
| Mean | 3.7125 | 2.2519 | 1.5130 | 1.5333 | 1.8267 | 1.4538 |
| StD | 0.8718 | 0.8128 | 0.7704 | 0.8133 | 0.8349 | 1.0651 |
| Kurtosis | 9.4152 | 12.2353 | 9.7505 | 9.5183 | 10.0866 | 7.1056 |
| Skewness | 1.7019 | 2.2519 | 1.5130 | 1.5333 | 1.8267 | 1.4538 |

Table S.1. Day-of-the-week pattern in daily S&PV series.

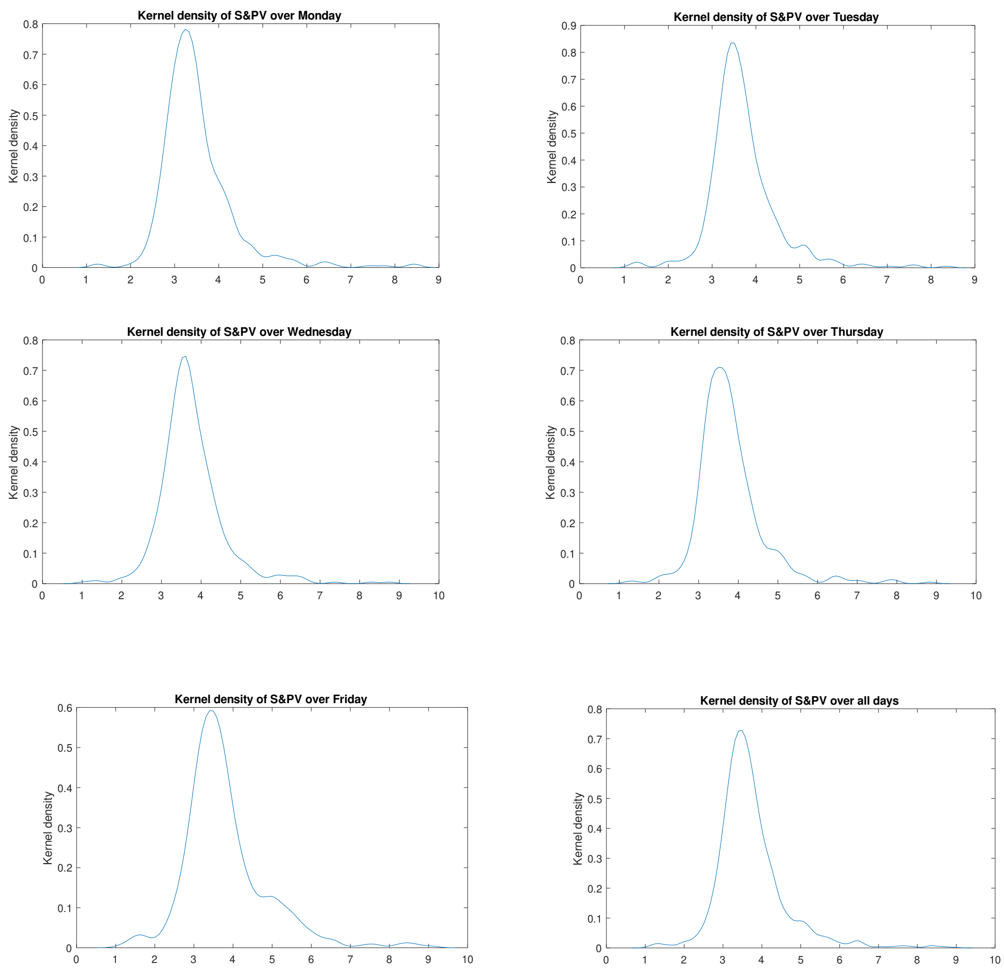



Figure S.4. Kernel densities of S&P 500 volume across days.

| $\widehat{\omega}$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\sigma}^2$ | $\widehat{\alpha}+\widehat{\beta}$ | IMSFE | IMAFE |
|---|---|---|---|---|---|---|
| 0.4822 (0.0521) | 0.4574 (0.0192) | 0.4123 (0.0264) | 0.02515 ($2.8e-06$) | 0.8698 | 0.3522 | 0.4004 |

Table S.2. ACD $(1,1)$ EQML estimates for the S&PV series.

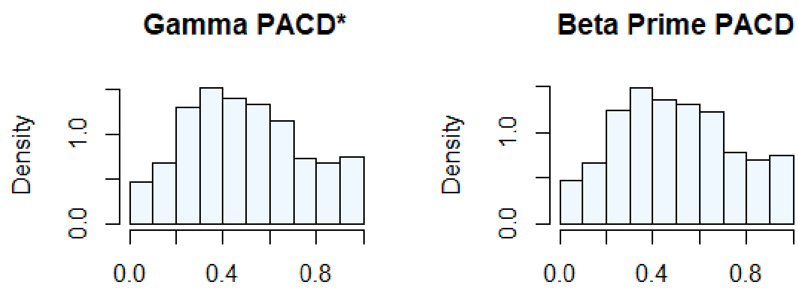
| Day | $v$ | $\widehat{\sigma}_v^2$ | $\widehat{\omega}_v$ | $\widehat{\alpha}_v$ | $\widehat{\beta}_v$ | $\widehat{\alpha}_v+\widehat{\beta}_v$ | IMSFE | IMAFE |
|---|---|---|---|---|---|---|---|---|
| Mon | 1 | 0.02327 ($8.8e-06$) | 0.1633 (0.0542) | 0.0211 (0.0288) | 0.8538 (0.0388) | 0.8748 | | |
| Tue | 2 | 0.0132 ($2.3e-06$) | 0.2731 (0.0991) | 0.6348 (0.0234) | 0.3446 (0.0366) | 0.9794 | | |
| Wed | 3 | 0.0156 ($4.0e-06$) | 0.0684 (0.1067) | 0.5904 (0.0363) | 0.4081 (0.0529) | 0.9985 | | |
| Thu | 4 | 0.0178 ($2.5e-05$) | 0.7527 (0.1046) | 0.4756 (0.0544) | 0.3290 (0.0609) | 0.8046 | | |
| Fri | 5 | 0.0463 ($3.6e-05$) | 0.5609 (0.1770) | 0.5713 (0.0782) | 0.2861 (0.0961) | 0.8574 | | |
| All | $\widehat{\phi}$ | | | | | 0.5902 | 0.3141 | 0.3744 |
| $\mathrm{W}^\theta$ | | 2571.9260 | | | | | | |
| $\mathrm{W}^\sigma$ | | 181.1893 | | | | | | |

Table S.3. PACD(1,1) 2S-GQML estimates for the S&PV series.

| $v \setminus s$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | 0 | 10.0809 | 5.1459 | 0.9666 | 13.1797 |
| 2 | **173.9991** | 0 | 0.8794 | 0.7592 | 29.0089 |
| 3 | **70.8388** | **57.2316** | 0 | 0.1705 | 24.0420 |
| 4 | **508.4607** | **86.3930** | **25.1698** | 0 | 13.5993 |
| 5 | **202.3575** | **34.8974** | **6.34912** | **2.0026** | 0 |

Table S.4. Pairwise Wald statistics for S&PV.

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




*S.pdf*

Figure S.5. PITs of the residuals $\widehat{\xi}_t$ (S&PV series) for two PACD innovation distributions.

| | $T_f$ | 1000 | 1200 | 1400 | 1600 | 1800 | 2000 |
|---|---|---|---|---|---|---|---|
| ACD | MSFE | 0.3750 | 0.4013 | 0.3925 | 0.4246 | 0.4246 | 0.6156 |
| | MAFE | 0.3959 | 0.4054 | 0.4027 | 0.4165 | 0.4165 | 0.4984 |
| | MQLI | 2.3145 | 2.3224 | 2.3099 | 2.3180 | 2.3180 | 2.3993 |
| $\text{PACD}_\text{E}$ | MSFE | 0.3697 | 0.3698 | 0.3582 | 0.3878 | 0.4420 | 0.5583 |
| | MAFE | 0.3833 | 0.3833 | 0.3809 | 0.3922 | 0.4148 | 0.4748 |
| | MQLI | 2.3214 | 2.3214 | 2.3089 | 2.3169 | 2.3468 | 2.3979 |
| $\text{PACD}_{\text{G}^*}$ | MSFE | 0.3534 | 0.3705 | 0.3580 | 0.3872 | 0.3872 | 0.5559 |
| | MAFE | 0.3769 | 0.3834 | 0.3804 | 0.3918 | 0.3918 | 0.4745 |
| | MQLI | 2.3138 | 2.3214 | 2.3089 | 2.3169 | 2.3169 | 2.3979 |

Table S.6. Out-of-sample forecasting performance of PACD and ACD on S&PV series.

# References


Billingsley, P. (1995). *Probability and measure*. 3rd edition, Wiley, New York.

Boyles, R.A. and Gardner, W.A. (1983). Cycloergodic properties of discrete-parameter nonstationary stochastic processes. *IEEE, Trans. Infor. Theory* **29**, 105-114.