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**, and Stefanos Dimitrakopoulos[1]***

*Department of Mathematics, College of Science, Qassim University (Saudi Arabia) & Faculty of Mathematics, University of Science and Technology Houari Boumediene (Algeria)

**Department of Mathematics, College of Science, Qassim University, Saudi Arabia.

***Division of Economics, Leeds University, UK

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]Correspondence to: Stefanos Dimitrakopoulos, s.dimitrakopoulos@leeds.ac.uk.

The authors are grateful to Professor Christian Francq for providing the UN realized volatility series.

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 (McCullag 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.

A well-known model for positive-valued time series data is the *autoregressive conditional duration* (ACD), introduced by Engel 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 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, 2019; 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; Hejer and Veltic 2007; Caporin et *al*, 2017; Gallo and Ortanto, 2018). In this paper, we extend the literature on 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* (henceforth ipd_S) 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, accounts 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 intraday 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 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 ω_t, $\alpha_{t1}, ..., \alpha_{tq}$, $\beta_{t1}, ..., \beta_{tp}$ ($p, q \in \mathbb{N} = \{0, 1, ...\}$) be positive real parameters S-periodic over time, i.e. $\omega_t = \omega_{t+kS}$, $\alpha_{ti} = \alpha_{t+kS,i}$ $(i = 1, ..., q)$ and $\beta_{tj} = \beta_{t+kS,j}$ $(j = 1, ..., p)$ 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 $\mathrm{Var}(\xi_t) = \sigma_t^2 > 0$. Assume that $\{\xi_t, t \in \mathbb{Z}\}$ is *ipd*$_S$ in the sense that $\xi_t \overset{D}{=} \xi_{t+S}$ 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 a 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

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

where the innovation term ξ_t is independent of ψ_{t-j} for all $j \geq 1$. To ensure the almost sure (a.s.) positivity of ψ_t, it is assumed that $\omega_t > 0$, $\alpha_{ti} \geq 0$, and $\beta_{tj} \geq 0$, 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 $1 \leq v \leq S$. Then, equation (2.1b) can be written as follows

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

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

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

and

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

The PACD model, thus, follows the quadratic variance-to-mean relationship (i.e. the conditional variance is proportional to the squared conditional mean), where $\sigma_t^2 > 0$ is the variance of ξ_t and is S-periodic by construction (from the ipd_S property of the innovation sequence $\{\xi_t, t \in \mathbb{Z}\}$). The specification (2.1) is a multiplicative error model (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 Engel and Russell (1998). No specification for the distribution of $\{\xi_t, t \in \mathbb{Z}\}$ is imposed apart the semiparametric quadratic variance-to-mean function (2.2b). However, a useful family of conditional distributions satisfying (2.2b) is the Gamma distribution with shape $\frac{1}{\sigma_t^2}$ and scale $\frac{1}{\sigma_t^2 \psi_t}$, that is

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

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

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

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 $\sigma_t^2 \equiv 1$ is constant, so $\xi_t \sim \Gamma(1,1)$, which corresponds to the *exponential* PACD. As in the time-invariant case, the periodic ACD model can be seen as a squared periodic GARCH (PGARCH) model as proposed by Ghysels and Bollerslev (1996). Indeed, consider the following real-valued PGARCH(p,q) process given by

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

and

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

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

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

where $h_t = \psi_t$ satisfies (2.5b) and $\eta_t = z_t \sqrt{\xi_t}$ is a term of an ipd_S sequence. Hence $\{X_t, t \in \mathbb{Z}\}$ is a PGARCH model in the sense of (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

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

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 $E\left(\varepsilon_t^2\right) = E\left(\psi_t^2\right) E\left(\xi_t - 1\right)^2 = E\left(\psi_t^2\right) \sigma_t^2$.

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_ψ is a cumulative probability distribution (with positive support) with mean ψ, and ψ_t is given by (2.1b).

3 Periodic ergodicity and finite moment conditions

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

we obtain the following stochastic recurrence equation (SRE)

$$\underline{Y}_t = A_t \underline{Y}_{t-1} + B_t \tag{3.1}$$

driven by the ipd_S sequence $\{(A_t, B_t), t \in \mathbb{Z}\}$, where $\underline{Y}_t = (Y_t, ..., Y_{t-q+1}, \psi_t, ..., \psi_{t-p+1})'$, $B_t = \left(\omega_t \xi_t, 0_{(q-1)\times 1}, \omega_t, 0_{(p-1)\times 1}\right)'$, and

$$A_t = \begin{pmatrix} \alpha_{t1}\xi_t & \cdots & \alpha_{t,q-1}\xi_t & \alpha_{tq}\xi_t & \beta_{t1}\xi_t & \cdots & \beta_{t,p-1}\xi_t & \beta_{tp}\xi_t \\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_{t1} & \cdots & \alpha_{t,q-1} & \alpha_{tq} & \beta_{t1} & \cdots & \beta_{t,p-1} & \beta_{tp} \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

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

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

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

$$\beta_t = \begin{pmatrix} \beta_{t1} & \cdots & \beta_{t,p-1} & \beta_{tp} \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix},$$

and denote by $\rho(A)$ the spectral radius of the squared matrix A, i.e. the maximum modulus of the eigenvalues of A. 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 $t \in \mathbb{Z}$ by

$$\underline{Y}_t = \sum_{j=0}^{\infty} \prod_{i=0}^{j-1} A_{t-i} B_{t-j}, \tag{3.3}$$

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*

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

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

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

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

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 $E(Y_t) < \infty$ is that*

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

Some remarks are in order:

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

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

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

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

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

$$E\left(\psi_v^{\kappa}\right)<\infty \quad \text{and} \quad E\left(Y_v^{\kappa}\right)<\infty. \tag{3.7}$$

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

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

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

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) $E\left(Y_v\right)<\infty$ $(1\leq v\leq S)$ *if and only if* (3.6) *holds.*

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

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

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

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

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

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

For the particular exponential PACD$(1,1)$ model, $Y_t|\mathscr{F}_{t-1}\sim\Gamma\left(1,\frac{1}{\psi_t}\right)$, just replace in

Proposition 3.1 the moments $E\left(\xi_v^2\right)$, $E\left(\xi_v^3\right)$ and $E\left(\xi_v^4\right)$ by 1, 6 and 24 respectively, and σ_v^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

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

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 $\sigma_0^2 = \left(\sigma_{01}^2, ..., \sigma_{0S}^2\right)'$ with $\sigma_{0v}^2 = Var\left(\xi_{nS+v}\right)$ $(1 \leq v \leq S)$ also belongs to a parametric space $\Delta \subset \mathbb{R}_+^S$. 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

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

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

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

4.1 Exponential and profile Gamma QMLEs

The true conditional distribution of (4.1) is unknown due to the unpecification of the law of ξ_v $(1 \leq v \leq S)$. Thus, a quasi-maximum likelihood estimate (QMLE) which does not require any precise knowledge of the conditional distribution is suitable for estimating the parameter θ_0 involved in the conditional mean. Among many possible QMLEs, the one computed on the basis of the exponential distribution (EQMLE in short) is especially useful for positive duration data because it reduces to the maximum likelihood estimate when ξ_v is exponentially distributed (Aknouche and Francq, 2020). A more general QMLE, which can be more efficient than the EQMLE in the periodic time-varying innovation context is the one computed on the basis of the Gamma distribution with arbitrary fixed variance parameters. Let $(\sigma_t^2)_t$ be fixed known positive numbers, S-periodic over t, i.e. $\sigma_{v+kS}^2 = \sigma_v^2$, for all $k \in \{0, ..., N-1\}$. The profile Gamma likelihood associated with $\sigma^2 = (\sigma_1^2, ..., \sigma_S^2)' > 0$ is given, ignoring constants, by

$$\widetilde{L}_T(\theta) = \tfrac{1}{T} \sum_{t=1}^{T} \widetilde{l}_t(\theta), \tag{4.3a}$$

$$\widetilde{l}_t(\theta) = \tfrac{1}{\sigma_t^2} \left(\tfrac{Y_t}{\widetilde{\psi}_t(\theta)} + \log \widetilde{\psi}_t(\theta) \right), \quad t \geq 1. \tag{4.3b}$$

The profile Gamma QMLE (GQMLE) $\widehat{\theta}_G$ of θ_0 is, then, the minimizer of $\widetilde{L}_T(\theta)$ over Θ,

$$\widehat{\theta}_G = \arg\min_{\theta \in \Theta} \widetilde{L}_T(\theta). \tag{4.4}$$

When $\sigma^2 = (1, ..., 1)'$, the GQMLE defined by (4.4) reduces to the EQMLE and is denoted by $\widehat{\theta}_E$ (Aknouche and Francq, 2020).

Let $\gamma^S(A^0)$ be the top Lyapunov exponent associated with $(A_t^0, t \in \mathbb{Z})$ where the matrix A_t^0 is just A_t defined in (3.1) with θ_0 in place of θ. To establish the strong consistency of $\widehat{\theta}_G$ we need to the following assumptions.

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

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

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

A4 ξ_v is non-degenerate for all $1 \leq v \leq S$.

As seen in Section 3, $\gamma^S(A^0) < 0$ in **A1** ensures periodic stationarity and periodic ergodicity of the PACD model (4.1). The condition $\rho\left(\prod_{v=0}^{S-1} \beta_{S-v}\right) < 1$ is imposed for the invertibility of equation (4.2b) for any $\theta \in \Theta$. The compactness assumption **A2** is standard while **A3** and **A4** are made to guarantee the identifiability of the model.

Theorem 4.1 *Let* $\left(\widehat{\theta}_G\right)$ *be a sequence of EQMLEs defined by* (4.3). *Under **A1-A4**,*

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

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

A5 θ_0 *belongs to the interior of* Θ.

A6 *The matrices*

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

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

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

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

where

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

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

Remark 4.1

i) When $\sigma^2=(1,...,1)':=\mathbf{1}$, the EQMLE has a covariance matrix in a "sandwich" form and is, in general, not asymptotically efficient unless $\sigma_0^2=\mathbf{1}$ 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 $\sigma^2=\mathbf{1}$ then $J\left(\theta_0,\mathbf{1}\right)=I\left(\theta_0,\mathbf{1}\right)$ 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 ξ_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 $\sigma^2=(1,...,1)'$ and apply the EQMLE. We would have $I\left(\theta_0,\mathbf{1}\right)=\sigma_0^2J\left(\theta_0,\mathbf{1}\right)$ and the covariance matrix would be equal to $\Sigma=\sigma_0^2J\left(\theta_0,\mathbf{1}\right)^{-1}$. 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 $\widetilde{L}_t\left(\theta\right)=\frac{1}{\sigma^2}\frac{1}{T}\sum\limits_{t=1}^{T}\left(\frac{Y_t}{\widetilde{\psi}_t(\theta)}+\log\left(\widetilde{\psi}_t\left(\theta\right)\right)\right)$ and the resulting GQMLE then reduces to maximizing $\frac{1}{T}\sum\limits_{t=1}^{T}\left(\frac{Y_t}{\widetilde{\psi}_t(\theta)}+\log\left(\widetilde{\psi}_t\left(\theta\right)\right)\right)$ 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) When the profile variance σ^2 coincides with the true variance σ_0^2 we would have $J(\theta_0, \sigma_0^2) = I(\theta_0, \sigma_0^2)$ and $\Sigma = J(\theta_0, \sigma_0^2)^{-1}$, where the GQMLE is the most efficient among all QMLEs belonging to the exponential family. As σ_0^2 is generally unknown, a crucial step is to get a consistent estimate $\widehat{\sigma}_0^2$ 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.

vi) Theorem 4.1 and 4.2 also hold 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' $E\left(Y_t^{1+\epsilon}\right) < \infty$ 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 σ_0^2 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 *iid* 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 θ_0, the regressand in (4.7a) is unobservable. If we replace θ_0 by a consistent estimate, say the GQMLE in (4.4), then we get the following approximate regression but with observable regressand

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

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

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

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

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

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

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

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

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

A consistent estimate of the limiting variance Λ_v in (4.9b) is given by

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

where $\widehat{\xi}_{v+nS} = \frac{Y_{v+nS}}{\widehat{\psi}_{v+nS}}$ is the residual of model (2.1). With (4.9b) and (4.10), the asymptotic matrices in (4.5) may also be estimated. A consistent estimate of Σ is

$$\widehat{\Sigma} = \widehat{J}^{-1}\widehat{I}^{-1}\widehat{J}^{-1} \tag{4.11}$$

where

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

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 σ^2. To improve the efficiency of the

GQMLE, we can replace in (4.4) the profile variances σ^2 by the OLS estimates $\widehat{\sigma}^2 = (\widehat{\sigma}_1^2, ..., \widehat{\sigma}_S^2)'$ 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 $\sigma^2 > \mathbf{0}$, for example $\sigma^2 = (1, ..., 1)'$.

ii) Get the profile GQMLE $\widehat{\theta}_G$ from (4.4).

iii) Estimate the variance innovation σ_0^2 using $\widehat{\sigma}^2$ in (4.8).

iv) Consider the 2S-GQMLE as a solution of the following problem

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

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

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

$$\sqrt{N}\left(\widehat{\theta}_G^* - \theta_0\right) \xrightarrow{D} \mathscr{N}\left(0, J\left(\theta_0, \sigma^2\right)^{-1}\right) \quad \text{as} \quad N \to \infty \text{ for all } \sigma^2 > \mathbf{0}.$$

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 GPQMLE $\widehat{\theta}_G$ and therefore than the EQMLE $\widehat{\theta}_E$.

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 two distributions for the innovation ξ_t in (2.1), namely i) the exponential distribution ($\xi_t \sim \mathscr{E}(1) \equiv \Gamma(1,1)$) so that $Y_t|\mathscr{F}_{t-1} \sim \Gamma(1, 1/\psi_t)$, and ii) the Gamma distribution ($\xi_t \sim \Gamma\left(\sigma_{0t}^{-2}, \sigma_{0t}^{-2}\right)$) so that $Y_t|\mathscr{F}_{t-1} \sim \Gamma\left(\sigma_{0t}^{-2}, \sigma_{0t}^{-2}/\psi_t\right)$, where σ_{0v}^{-2} $(1 \leq v \leq S)$.

For these two cases we take $S = 5$, which is representative of many real daily trading measurements, such as trading volumes and realized volatilities. The true conditional mean parameters, which are reported in Tables 5.1 and 5.2, are chosen so that the PACD model to be stable in the sense of Section 3, while implying fairly persistent series that are in accordance with the empirical evidence. For each case and for each series we compute the EQMLE and the two-stage GQMLE (2S-GQMLE), using 1000 Monte Carlo replications.

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

		EQMLE			2S-GQMLE			
v	θ_v^0	ω_v^0	α_v^0	β_v^0	σ_{0v}^2	ω_v^0	α_v^0	β_v^0
	True	0.5	0.6	0.35	1	0.5	0.6	0.35
1	Mean	0.5191	0.5981	0.3484	0.9817	0.5191	0.5980	0.3484
	Std	0.3163	0.0709	0.0698	0.0968	0.3168	0.0710	0.0699
	True	0.9	0.4	0.5	1	0.9	0.4	0.5
2	Mean	0.8941	0.4003	0.5039	0.9917	0.8940	0.4001	0.5040
	Std	0.3624	0.0723	0.0921	0.1001	0.3628	0.0723	0.0924
	True	1.5	0.5	0.5	1	1.5	0.5	0.5
3	Mean	1.4558	0.5006	0.5063	0.9837	1.4559	0.5005	0.5063
	Std	0.4795	0.0784	0.1010	0.0992	0.4806	0.0783	0.1011
	True	0.45	0.45	0.45	1	0.45	0.45	0.45
4	Mean	0.4521	0.4490	0.4453	0.9859	0.4525	0.4492	0.4452
	Std	0.4000	0.0635	0.0795	0.1041	0.4007	0.0636	0.0798
	True	0.7	0.55	0.40	1	0.7	0.55	0.40
5	Mean	0.6797	0.5512	0.4051	0.9792	0.6793	0.5510	0.4053
	Std	0.3828	0.0726	0.0804	0.0939	0.3832	0.0728	0.0806

Table 5.1. EQMLE and 2S-GQMLE results for 1000 PACD(1,1) series with $n = 2000$ generated from the exponential $\Gamma(1, 1/\psi_t)$ distribution.

Means and standard deviations of the estimates $\widehat{\theta}_E$ and $\widehat{\theta}_G^*$ over the 1000 replications are reported in Table 5.1 for the exponential PACD(1,1) model and in Table 5.2 for the homolog Gamma PACD(1,1) model. It can be observed from Tables 5.1-5.2 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.

		EQMLE			2S-GQMLE			
v	θ_v^0	ω_v^0	α_v^0	β_v^0	σ_{0v}^2	ω_v^0	α_v^0	β_v^0
	True	0.2	0.4	0.5	0.5	0.2	0.4	0.5
1	Mean	0.2099	0.4022	0.4991	0.4976	0.2020	0.4012	0.5006
	Std	0.1573	0.0408	0.0768	0.0428	0.1550	0.0399	0.0754
	True	0.9	0.3	0.6	0.3	0.9	0.3	0.6
2	Mean	0.8843	0.3045	0.6046	0.2981	0.8934	0.3036	0.6009
	Std	0.1462	0.0654	0.0891	0.0240	0.1452	0.0553	0.0798
	True	0.3	0.5	0.4	1.5	0.3	0.5	0.4
3	Mean	0.3341	0.5030	0.3815	1.4729	0.3417	0.5058	0.3777
	Std	0.2856	0.1130	0.1240	0.1582	0.2826	0.1100	0.1201
	True	0.4	0.45	0.45	1	0.4	0.45	0.45
4	Mean	0.4040	0.4526	0.4534	0.9809	0.3995	0.4525	0.4541
	Std	0.2645	0.0702	0.0971	0.0935	0.2599	0.0681	0.0933
	True	0.5	0.55	0.35	2	0.5	0.55	0.35
5	Mean	0.4997	0.5546	0.3494	1.9424	0.4923	0.5546	0.3533
	Std	0.2729	0.0921	0.1002	0.2247	0.2487	0.0834	0.0838

Table 5.2. EQMLE and 2S-GQMLE results for 1000 PACD(1,1) series with $T = 2000$ generated from the Gamma $\Gamma\left(1/\sigma_{0t}^2, 1/\sigma_{0t}^2\psi_t\right)$ distribution.

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 dataset was obtained from the webpage www.blockchain.com. This series spans from July,

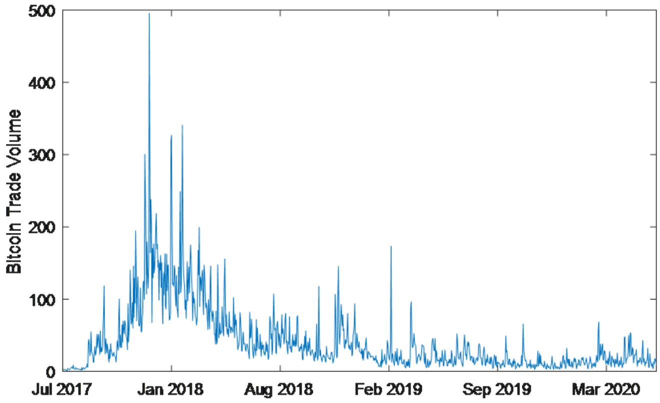


Figure 6.1: Daily Bitcoin trading volume (BTV).

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.

In the context of Bitcoin prices, Mbanga (2019) found evidence of the presence of the day-of-the-week pattern. 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 BVT 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 BVT 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). 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, 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 BVT 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)	0.659 56		
All	$\prod_{v=1}^{7}(\widehat{\alpha}_v + \widehat{\beta}_v)$					0.9025	651.9418	13.1164

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. The persistence parameters over the days show locally explosive behaviors except for Saturday and Sunday. However, the whole persistence parameter, $\prod_{v=1}^{7}(\widehat{\alpha}_v + \widehat{\beta}_v) = 0.9025$ (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.

To 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 $1 < T_f < T$. 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 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 6.4 shows these computed values of these criteria for the two models and for various truncated series with sample size T_f. It can be observed that irrespective of the chosen sample size, the PACD yields better out-of-sample forecasts with regard to the aforementioned criteria. Overall, the PACD$(1,1)$ outperforms the ACD$(1,1)$, both in terms of in-sample and out-of-sample forecasting power.

	T_f	500	600	700	800	900	1000
ACD	MSFE	207.8341	165.8957	89.6771	88.3917	94.4021	104.8815
	MAFE	8.6659	7.8318	6.9127	6.6050	6.9080	7.6159
	MQLI	3.8983	3.7735	3.6879	3.6052	3.6959	3.8487
PACD	MSFE	193.1508	151.0290	79.2446	80.8776	80.4284	93.9384
	MAFE	8.0187	7.1074	6.1546	5.9031	6.02684	6.6908
	MQLI	3.8759	3.7495	3.6605	3.5754	3.6635	3.8144

Table 6.4. Out-of-sample forecasting performance of the PACD and ACD; BVT 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. The plot of the index series is displayed in Figure 6.2.

Table 6.5 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

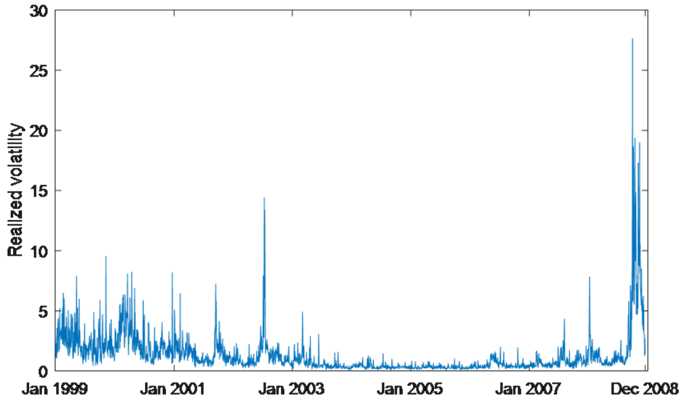


Figure 6.2: Daily UN realized volatility UN (UN-RV).

material). These facts suggests 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 6.5. 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 6.6 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 6.6. EQML estimates for the ACD(1,1); UN-RV series.

Table 6.7 displays the 2S-GQML estimates of the PACD(1,1) based on the UN-RV series. These estimates are quite different across the days and are all significant. Also, 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 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.

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	$\prod_{v=1}^{7}(\widehat{\alpha}_v + \widehat{\beta}_v)$					0.8897	1.0570	0.4757

Table 6.7. 2S-GQML estimates for the PACD(1,1); UN-RV series.

We finally compare the out-of-sample forecasting performance of the two models, using the same devices as before. From Table 6.8 it can be concluded that for all truncated series (with sample size T_f), the PACD gives more 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).

	T_f	1100	1200	1500	1600	1800	2000
ACD	MSFE	1.0803	1.1501	1.4973	1.6645	2.1461	2.9891
	MAFE	0.3662	0.3644	0.4219	0.4548	0.5509	0.6966
	MQLI	0.4485	0.4534	0.5903	0.6690	0.9126	1.1470
PACD	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 6.8. 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, giving more accurate results compared to the exponential QMLE. The proposed estimates are also consistent and asymptotically normal 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.

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.

Although our model is named periodic ACD in reference to the ACD proposed by Engel 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. These issues are left for future research.

Proofs

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

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

where $\mathbb{A}_{nS+v} = \prod_{i=0}^{S-1} A_{nS+v-i}$ and $\mathbb{B}_{nS+v} = \sum_{j=0}^{S-1}\prod_{i=0}^{j-1} A_{nS+v-i}B_{nS+v-j}$ are such that the sequence $\{(\mathbb{A}_{nS+v}, \mathbb{B}_{nS+v}), n \in \mathbb{Z}\}$ 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, 1992a)

$$\begin{aligned} \gamma_v^{(S)} &= \inf\left\{\tfrac{1}{n}E\log\left\|\mathbb{A}_{nS+v}\mathbb{A}_{(n-1)S+v}...\mathbb{A}_{S+v}\right\|,\ n \geq 1\right\} \\ &= \lim_{n\to\infty}\tfrac{1}{n}\log\left\|A_{nS+v}A_{nS+v-1}...A_{v+1}\right\| \quad a.s. \end{aligned} \tag{$A.2$}$$

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

$$\underline{Y}_{nS+v} = \sum_{j=0}^{\infty}\prod_{i=0}^{j-1}\mathbb{A}_{(n-i)S+v}\mathbb{B}_{(n-j)S+v}, \quad n \in \mathbb{Z}, v \in \{0,...,S-1\} \tag{$A.3$}$$

which is exactly (3.3), where the series in equality $(A.3)$ converges absolutely $a.s.$ This shows that $\{\underline{Y}_t, t \in \mathbb{Z}\}$ 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 $v \in \{0,...,S-1\}$

$$\gamma_v^{(S)} = \lim_{n\to\infty}\tfrac{1}{n}\log\left\|A_{nS+v}A_{nS+v-1}...A_{v+1}\right\| = \lim_{n\to\infty}\tfrac{1}{n}\log\left\|A_{nS}A_{nS-1}...A_1\right\| := \gamma^{(S)}.$$

ii) If model (3.1) admits a nonanticipative strictly periodically stationary solution $\{\underline{Y}_t, t \in \mathbb{Z}\}$ then from the non-negativity of the coefficients of A_t in (3.1) it follows that for all $k > 1$,

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

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

$$\lim_{j\to\infty}\prod_{i=0}^{j-1}A_{v-i}e_m=0, a.s.\text{ for all } 1\leq m\leq r, \tag{A.4}$$

where $r=p+q$ and $(e_m)_{1\leq m\leq r}$ is the canonical basis of $\mathbb{R}^r$. Since A_t has the same zero-structure as the matrix A_t in Bougerol and Picard (1992b), then $(A.4)$ follows from their results using the same argument.

ii) By the nonnegativity of $\{A_t, t\in\mathbb{Z}\}$ we have

$$\gamma^S(A)\geq\gamma^S(\beta):=\log\rho\left(\prod_{v=0}^{S-1}\beta_{S-v}\right). \tag{A.5}$$

If (3.1) admits a strictly periodically stationary solution then $\gamma^S(A)<0$. In view of $(A.5)$, it follows that

$$\gamma^S(\beta)<0 \tag{A.6}$$

which in turn implies (3.4). $\square$

Proof of Theorem 3.2 Theorem 3.2 is a particular case of Theorem 3.3, ii). $\square$

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 $\left\{\widetilde{Y}_t, t\in\mathbb{Z}\right\}$ by

$$\begin{cases}\widetilde{Y}_t=A_t\widetilde{Y}_{t-1}+B_t & t\geq 1\\ \widetilde{Y}_t=0 & t\leq 0,\end{cases} \tag{A.7}$$

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, 1968), 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 $(A.7)$ we get the following first-order

S-periodic difference equation

$$V_{nS+v} = E\left(A_v^{\otimes 2}\right) V_{nS+v-1} + \left[E(A_v \otimes B_v) + E(B_v \otimes A_v)\right] E\left(\widetilde{Y}_{nS+v}\right) + vec(E\left(B_v B_v'\right)) \quad (A.8)$$

where $E\left(A_t^{\otimes 2}\right)$, $E(A_t \otimes B_t)$ and $vec(E\left(B_t B_t'\right))$ are finite S-periodic matrices over t. Since, the last two terms of the right-hand side of $(A.8)$ are bounded, it follows that $\lim_{n\to\infty} V_{nS+v}$ exists for every $1 \leq v \leq S$ as long as (3.8) holds, from which follows the proof for $m = 2$. For general m, the proof is similar. $\square$

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

$$y_t = a_t y_{t-1} + b_t, \qquad t \in \mathbb{Z}, \qquad (A.9)$$

be an ordinary difference equation with *S-periodic positive coefficients* $a_t = a_{t+S} > 0$ *and* $b_t = b_{t+S} > 0$ for all $t \in \mathbb{Z}$.

Lemma 1 *The real-valued ordinary difference equation* $(A.9)$ *admits a unique solution* $\{y_t, t \in \mathbb{Z}\}$ *if and only if*

$$\prod_{v=1}^{S} a_v < 1.$$

Proof of Proposition 3.1 It is well-known that if $Y_t|\mathscr{F}_{t-1} \sim \Gamma\left(\frac{1}{\sigma_t^2}, \frac{1}{\sigma_t^2 \psi_t}\right)$ then the conditional moments up to the fourth order are given by

$$\begin{aligned}
E\left(Y_t|\mathscr{F}_{t-1}\right) &= \psi_t & (A.10a)\\
E\left(Y_t^2|\mathscr{F}_{t-1}\right) &= \left(1+\sigma_t^2\right)\psi_t^2 & (A.10b)\\
E\left(Y_t^3|\mathscr{F}_{t-1}\right) &= \left(1+\sigma_t^2\right)\left(1+2\sigma_t^2\right)\psi_t^3 & (A.10c)\\
E\left(Y_t^4|\mathscr{F}_{t-1}\right) &= \left(1+\sigma_t^2\right)\left(1+2\sigma_t^2\right)\left(1+3\sigma_t^2\right)\psi_t^4. & (A.10d)
\end{aligned}$$

In view of $(A.10)$ it turns out that the conditional moment of Y_t of order i is a polynomial in ψ_t with degree i $(i = 1, 2, 3, 4)$. Hence $E\left(Y_t^i\right) < \infty$ if and only if $E\left(\psi_t^i\right) < \infty$ $(i = 1, 2, 3, 4)$, conditions for which are given as follows.

i) Expanding $E\left(\psi_t\right)$ using $(2.1b)$ with $p = q = 1$ and $(A.10a)$, we find the following linear

S-periodic difference equation

$$E\left(\psi_t\right) = \left(\alpha_t + \beta_t\right) E\left(\psi_{t-1}\right) + \omega_t, \quad t \in \mathbb{Z}. \tag{A.11}$$

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 $E\left(Y_v^2\right)$ $(1 \leq v \leq S)$, expanding $E\left(\psi_t^2\right)$ using (2.1b) and (3.10a)-(3.10b), we find the following linear periodic difference equation

$$E\left(\psi_t^2\right) = \left(\alpha_t^2 E\left(\xi_{t-1}^2\right) + 2\alpha_t\beta_t + \beta_t^2\right) E\left(\psi_{t-1}^2\right) + K_t^{(1)}, \ t \in \mathbb{Z}, \tag{A.12}$$

where

$$K_t^{(1)} = \left(2\alpha_t\omega_t + 2\beta_t\omega_t\right) E\left(\psi_{t-1}\right) + \omega_t^2$$

is finite if and only if $E\left(\psi_{t-1}\right) < \infty$, 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 $E\left(\psi_t^3\right)$ using (2.1b) and $(A.10a)$-$(A.10c)$ we get the following linear periodic difference equation

$$E\left(\psi_t^3\right) = \left(\alpha_t^3 E\left(\xi_{t-1}^3\right) + 3\alpha_t^2\beta_t\left(\sigma_{t-1}^2 + 1\right) + \beta_t^3 + 3\alpha_t\beta_t^2\right) E\left(\psi_{t-1}^3\right) + K_t^{(2)}, \ t \in \mathbb{Z} \tag{A.13}$$

where

$$K_t^{(2)} = \left(3\alpha_t^2\omega_t E\left(\xi_{t-1}^2\right) + 6\alpha_t\beta_t\omega_t + 3\beta_t^2\omega_t\right) E\left(\psi_{t-1}^2\right) + \left(3\beta_t\omega_t^2 + 3\alpha_t\omega_t^2\right) E\left(\psi_{t-1}\right) + \omega_t^3$$

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

iv) Expanding $E\left(\psi_t^4\right)$ using (2.1b) and $(A.10a)$-$(A.10d)$ we get the following linear periodic difference equation

$$\begin{aligned} E\left(\psi_t^4\right) = & \left(\alpha_t^4 E\left(\xi_{t-1}^4\right) + 4\alpha_t^3\beta_t\left(1+\sigma_{t-1}^2\right)\left(1+2\sigma_{t-1}^2\right) + 6\alpha_t^2\beta_t^2\left(1+\sigma_{t-1}^2\right) + 4\alpha_t\beta_t^3 + \beta_t^4\right) E\left(\psi_{t-1}^4\right) \\ & + K_t^{(3)} \end{aligned} \tag{A.14}$$

where

$$K_t^{(3)} = 4\alpha_t^3\omega_t E\left(Y_{t-1}^3\right) + 12\alpha_t^2\beta_t\omega_t E\left(Y_{t-1}^2\psi_{t-1}\right) + 12\alpha_t\beta_t^2\omega_t E\left(Y_{t-1}\psi_{t-1}^2\right)$$
$$+4\beta_t^3\omega_t E\left(\psi_{t-1}^3\right) + 6\alpha_t^2\omega_t^2 E\left(Y_{t-1}^2\right) + 12\alpha_t\beta_t\omega_t^2 E\left(Y_{t-1}\psi_{t-1}\right) + 6\beta_t^2\omega_t^2 E\left(\psi_{t-1}^2\right)$$
$$+4\alpha_t\omega_t^3 E\left(Y_{t-1}\right) + 4\beta_t\omega_t^3 E\left(\psi_{t-1}\right) + \omega_t^4$$

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 $L_T(\theta)$ and $l_t(\theta)$ be defined in the same way as $\widetilde{L}_T(\theta)$ and $\widetilde{l}_t(\theta)$ in $(4.3a)$ and $(4.3b)$, respectively, with $\psi_t(\theta)$ in place of $\widetilde{\psi}_t(\theta)$.

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

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

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

$$\underline{\psi}_t = \beta_t\underline{\psi}_{t-1} + \alpha_t, \ t \in \mathbb{Z}, \tag{A.15}$$

where $\underline{\psi}_t = \left(\psi_t(\theta), \psi_{t-1}(\theta), ..., \psi_{t-p+1}(\theta)\right)'$ and $\alpha_t = \left(\omega_t + \sum\limits_{i=1}^{q}\alpha_{ti}Y_{t-i}, 0, ..., 0\right)'_{1\times p}$. By **A1** and the assumption **A2** of compactness of Θ it follows that

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

Iterating $(A.15)$ gives

$$\underline{\psi}_t = \sum_{k=0}^{t-1}\prod_{i=0}^{k-1}\beta_{t-i}\underline{\alpha}_{t-k} + \prod_{i=0}^{t}\beta_{t-i}\underline{\psi}_0, \ t \in \mathbb{Z}.$$

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

$$\sup_{\theta\in\Theta}\left\|\underline{\psi}_t - \widetilde{\underline{\psi}}_t\right\| = \sup_{\theta\in\Theta}\left\|\sum_{k=t-q}^{t-1}\prod_{i=0}^{k-1}\beta_{t-i}\left(\underline{\alpha}_{t-k} - \widetilde{\underline{\alpha}}_{t-k}\right) + \prod_{i=0}^{t-1}\beta_{t-i}\left(\underline{\psi}_0 - \widetilde{\underline{\psi}}_0\right)\right\| \le M\rho^t. \tag{A.17}$$

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

$$\begin{aligned}\sup_{\theta\in\Theta}\left|L_T(\theta)-\widetilde{L}_T(\theta)\right| &\leq \frac{1}{T}\sum_{t=1}^{T}\sup_{\theta\in\Theta}\frac{1}{\sigma_t^2}\left[\left|\frac{\widetilde{\psi}_t-\psi_t(\theta)}{\widetilde{\psi}_t\psi_t(\theta)}\right|Y_t+\left|\log\left(\frac{\psi_t(\theta)}{\widetilde{\psi}_t(\theta)}\right)\right|\right]\\ &\leq \max_{1\leq v\leq S}\sup_{\theta\in\Theta}\left(\frac{\omega_v^{-2}}{\sigma_v^2}\right)\frac{M}{T}\sum_{t=1}^{T}\rho^t Y_t+\max_{1\leq v\leq S}\sup_{\theta\in\Theta}\left(\frac{\omega_v^{-2}}{\sigma_v^2}\right)\frac{M}{T}\sum_{t=1}^{T}\rho^t.\end{aligned}$$

The existence of $E\left(Y_t^\delta\right)$ (cf, Theorem 3.3, i)) implies, by the Borel-Cantelli lemma, that $\rho^t Y_t \to 0$ $a.s.$ and the conclusion follows by Césaro's lemma. $\square$

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

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

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

If $\frac{\alpha_v(L)}{\beta_v(L)}\neq\frac{\alpha_v^0(L)}{\beta_v^0(L)}$ for some $1\leq v\leq S$ then there exists a deterministic periodic time-varying combination of Y_{t-j}, $j\geq 1$. This contradicts **A4** which assumes $(\xi_t, t\in\mathbb{Z})$ non-degenerate, since by (2.6) we have $Y_t=E(Y_t|\mathscr{F}_{t-1})+\psi_t(\xi_t-1)$. Therefore,

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

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 $1\leq v\leq S$. $\square$

Lemma 3 *Under **A1***

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

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

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

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

Hence

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

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

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

showing that $\sum_{v=1}^{S} E\left(l_v(\theta)\right)$ is minimized at θ_0.

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

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

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

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

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

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

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

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

and by $(A.18)$ the result follows. □

Proof of Theorem 4.1 The proof of the theorem is completed by standard compactness arguments using Lemmas 2-4. $\square$

Proof of Theorem 4.2 The proof of Theorem 4.2 is based on a Taylor expansion of $\frac{\partial \widetilde{L}_T(\theta)}{\partial \theta}$ at θ_0 which, by **A5** and the strong consistency of $\widehat{\theta}_G$, yields

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

where θ^* is between $\widehat{\theta}_G$ and θ_0. The derivatives $\frac{\partial L_T(\theta)}{\partial \theta}$ and $\frac{\partial^2 L_T(\theta)}{\partial \theta \partial \theta'}$ are given by

$$\begin{aligned}
\frac{\partial L_T(\theta)}{\partial \theta} &= \frac{1}{T}\sum_{t=1}^{T}\frac{\partial l_t(\theta)}{\partial \theta} = \frac{1}{T}\sum_{t=1}^{T}\left(1 - \frac{Y_t}{\psi_t(\theta)}\right)\frac{1}{\sigma_t^2 \psi_t(\theta)}\frac{\partial \psi_t(\theta)}{\partial \theta} \\
\frac{\partial^2 L_T(\theta)}{\partial \theta \partial \theta'} &= \frac{1}{T}\sum_{t=1}^{T}\left[\left(1 - \frac{Y_t}{\psi_t(\theta)}\right)\frac{1}{\sigma_t^2 \psi_t(\theta)}\frac{\partial^2 \psi_t(\theta)}{\partial \theta \partial \theta'} + \left(\frac{2Y_t}{\psi_t(\theta)} - 1\right)\frac{1}{\sigma_t^2 \psi_t^2(\theta)}\frac{\partial \psi_t(\theta)}{\partial \theta}\frac{\partial \psi_t(\theta)}{\partial \theta'}\right].
\end{aligned}$$

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

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

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

for some neighborhood $V(\theta_0)$ of θ_0.

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

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

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

$$\begin{aligned}
\left| \frac{1}{\widetilde{\psi}_t(\theta)} - \frac{1}{\psi_t(\theta)} \right| &\leq \frac{M\rho^t}{\psi_t(\theta)}\frac{\psi_t(\theta)}{\widetilde{\psi}_t(\theta)}(1+M)\rho^t, \\
\sup_{\theta \in \Theta}\left| \frac{\partial \psi_t(\theta)}{\partial \theta} - \frac{\partial \widetilde{\psi}_t(\theta)}{\partial \theta} \right| &\leq M\rho^t,\ \sup_{\theta \in \Theta}\left| \frac{\partial^2 \psi_t(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 \widetilde{\psi}_t(\theta)}{\partial \theta \partial \theta'} \right| \leq M\rho^t,
\end{aligned}$$

and

$$\sup_{\theta\in V(\theta_0)} \sqrt{N}\left\| \frac{\partial L_T(\theta)}{\partial\theta} - \frac{\partial \widetilde{L}_T(\theta)}{\partial\theta}\right\| = \sup_{\theta\in V(\theta_0)} \frac{1}{S\sqrt{N}} \left\| \sum_{t=1}^{T} \left[\left(\frac{1}{\widetilde{\psi}_t(\theta)} - \frac{1}{\psi_t(\theta)} \right) \frac{Y_t}{\sigma_t^2 \psi_t(\theta)} \frac{\partial \psi_t(\theta)}{\partial\theta} \right.\right.$$

$$\left.\left. + \frac{1}{\sigma_t^2}\left(1 - \frac{Y_t}{\psi_t(\theta)}\right)\left(\frac{1}{\psi_t(\theta)} - \frac{1}{\widetilde{\psi}_t(\theta)}\right)\frac{\partial \psi_t(\theta)}{\partial\theta} + \left(1 - \frac{Y_t}{\psi_t(\theta)}\right)\frac{1}{\sigma_t^2 \widetilde{\psi}_t(\theta)}\left(\frac{\partial \psi_t(\theta)}{\partial\theta} - \frac{\partial \widetilde{\psi}_t(\theta)}{\partial\theta}\right)\right]\right\|$$

$$\leq \frac{M}{S\sqrt{N}} \sum_{v=1}^{S}\sum_{n=0}^{N-1} \rho^n \left(1 + \xi_{v+nS}\right) \left\| 1 + \frac{1}{\psi_t(\theta_0)} \frac{\partial \psi_t(\theta_0)}{\partial\theta} \right\|. \qquad \text{(A.20)}$$

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

$$P\left(\frac{1}{\sqrt{N}} \sum_{v=1}^{S}\sum_{n=0}^{N-1} \rho^n \left(1 + \xi_{v+nS}\right) \left\| 1 + \frac{1}{\psi_t(\theta_0)} \frac{\partial \psi_t(\theta_0)}{\partial\theta} \right\| > \varepsilon \right)$$

$$\leq \frac{2}{\varepsilon}\left(1 + E_{\theta_0}\left\| 1 + \frac{1}{\psi_t(\theta_0)} \frac{\partial \psi_t(\theta_0)}{\partial\theta} \right\|\right) \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \rho^n \to 0 \text{ as } N \to 0,$$

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

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

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

for any θ^* between $\widehat{\theta}_G$ and θ_0.

Proof Let $V_k(\theta_0)$ ($k \in \mathbb{N}^*$) be the open ball with center θ_0 and radius $1/k$. Assume that n is large enough so that θ^* belongs to $V_k(\theta_0)$. By periodic stationarity and periodic ergodicity of $\left\{\sup_{\theta\in V_k(\theta_0)} \left| \frac{\partial^2 l_t(\theta)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_t(\theta_0)}{\partial\theta_i\partial\theta_j}\right)\right|\right\}$ we have

$$\begin{aligned}
\left| \frac{\partial^2 L_T(\theta^*)}{\partial\theta_i\partial\theta_j} - J\left(\theta_0, \sigma^2\right)_{i,j} \right| &= \left| \frac{\partial^2 L_T(\theta^*)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 L_T(\theta_0)}{\partial\theta_i\partial\theta_j}\right)\right| = \frac{1}{T}\left| \sum_{t=1}^{T} \frac{\partial^2 l_t(\theta^*)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_t(\theta_0)}{\partial\theta_i\partial\theta_j}\right)\right| \\
&\leq \frac{1}{NS}\sum_{t=1}^{T} \sup_{\theta\in V_k(\theta_0)} \left| \frac{\partial^2 l_t(\theta)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_t(\theta_0)}{\partial\theta_i\partial\theta_j}\right)\right| \\
&\underset{N\to\infty}{\overset{a.s.}{\to}} \frac{1}{S}\sum_{v=1}^{S} E\left(\sup_{\theta\in V_k(\theta_0)} \left| \frac{\partial^2 l_v(\theta)}{\partial\theta_i\partial\theta_j} - E\left(\frac{\partial^2 l_v(\theta_0)}{\partial\theta_i\partial\theta_j}\right)\right|\right).
\end{aligned}$$

The Lebesgue dominated convergence theorem yields

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

from which the proof of the lemma follows. □

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

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

Proof It is clear that $\sqrt{N}\frac{\partial L_T(\theta_0)}{\partial\theta} = \sum_{t=1}^{T} \frac{1}{S\sqrt{N}} \frac{\partial l_t(\theta_0)}{\partial\theta}$ is a term of a square integrable periodically ergodic Martingale. Since by the periodic ergodic theorem (cf. Supplementary material)

$$\frac{1}{N}\sum_{t=1}^{T} \frac{\partial l_t(\theta_0)}{\partial\theta}\frac{\partial l_t(\theta_0)}{\partial\theta'} = \frac{1}{N}\sum_{t=1}^{T} (1-\xi_t)^2 \frac{1}{\sigma_t^4\psi_t^2(\theta_0)}\frac{\partial\psi_t(\theta)}{\partial\theta}\frac{\partial\psi_t(\theta)}{\partial\theta'} \underset{N\to\infty}{\overset{a.s.}{\to}} I\left(\theta_0, \sigma^2\right),$$

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

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

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

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

Now a Taylor expansion of $U_{v,n}\left(\widehat{\theta}_G\right)$ around θ_0 yields

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

where θ^* is between $\widehat{\theta}_G$ and θ_0. Note that

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

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

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

for some neighborhood $\mathscr{V}(\theta_0)$ of θ_0. Hence, by the ergodic theorem and the consistency of $\widehat{\theta}_G$

we get

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

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. □

References

Aknouche, A. and Francq, C. (2020). Count and duration time series with equal conditional stochastic and mean orders. *Econometric Theory*, forthcoming.

Aknouche, A. and Francq, C. (2019). Two-stage weighted least squares estimator of the conditional mean of observation-driven time series models. MPRA paper 97382.

Aknouche, A., Demmouche, N., Dimitrakopoulos, S. and Touche, N. (2020). Bayesian analysis of periodic asymmetric power GARCH models. *Studies in Nonlinear Dynamics and Econometrics*, forthcoming.

Ambach, D. and Croonenbroeck, C. (2015). Obtaining superior wind power predictions from a periodic and heteroscedastic wind power prediction tool. *In Stochastic Models, Statistics and Their Applications, edt*, 225–232.

Ambach, D. and Schmid, W. (2015). Periodic and long range dependent models for high frequency wind speed data. *Energy* **82**, 277–293.

Andersen, T. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance* **4**, 115–158.

Bhogal, S.K. and Variyam, R.T. (2019). Conditional duration models for high-frequency data: A review on recent developments. *Journal of Economic Surveys* **33**, 252–273.

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

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

Bollerslev, T. and Ghysels, E. (1996). Periodic autoregressive conditional heteroskedasticity. *Journal of Business & Economic Statistics* **14**, 139–152.

Bollerslev, T., Cai, J. and Song, F.M. (2000). Intraday periodicity, long memory volatility, and macroeconomic announcement effects in the US Treasury bond market. *Journal of Empirical Finance* **7**, 37–55.

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

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

Boynton, W., Oppenheimer, H.R. and Reid, S.F. (2009). Japanese day-of-the-week return patterns: New results. *Global Finance Journal* **20**, 1–12.

Bracher, J. and Held, L. (2017), Periodically stationary multivariate autoregressive models. *arXiv preprint, arXiv:1707.04635.*

Caporin, M., Rossi, E., and 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. and An, H.Z. (1998). A note on the stationarity and the existence of moments of the GARCH model. *Statistica Sinica* **8**, 505-510.

Chou, R.Y. (2005). Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) Model. *Journal of Money, Credit, and Banking* **37**, 561–582.

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

Francq, C. and Zakoian, J.-M. (2019). *GARCH models: structure, statistical inference and financial applications.* John Wiley & Sons, 2nd edt.

Francq, C., Roy, R. and Saidi, A. (2011). Asymptotic Properties of Weighted Least Squares Estimation in Weak PARMA Models. *Journal of Time Series Analysis* **32**, 699–723

Franses, P.H. and Paap. R. (2000). Modeling day-of-the-week seasonality in the S&P 500

Index. *Applied Financial Economics* **10**, 483–488.

Hautsch, N. (2012). Econometrics of Financial High-Frequency Data. Berlin, Heidelberg: Springer.

Gallo, G.M. and 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. and Trognon, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica* **52**, 681-700.

Hujer, R. and Vuletic, S. (2007). Econometric analysis of financial trade processes by discrete mixture duration models. *Journal of Economic Dynamics and Control* **31**, 635–667.

Lanne, M. (2006). A mixture multiplicative error model for realized volatility. *Journal of Financial Econometrics* **4**, 594–616.

Li, Q. (2019). Location multiplicative error models with quasi maximum likelihood estimation. *Journal of Time Series Analysis*, DOI: 10.1111/jtsa.12513.

Lund, R. and Basawa, I.V. (2000). Recursive prediction and likelihood evaluation for periodic ARMA models. *Journal of Time Series Analysis* **21**, 75–93.

Mbanga, C.L. (2019). The day-of-the-week pattern of price clustering in Bitcoin. *Applied Economics Letters* **26**, 1-6.

McCullagh, P. and Nelder, J.A. (1989). Generalized Linear Models, 2nd edn. Chapman and Hall, London.

Mikosch, T. and Starica, C. (2004). Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects. *Review of Economics and Statistics* **86**, 378–390.

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, A.J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* **160**, 246–256.

Rossi, E. and Fantazzini, D. (2015). Long memory and periodicity in intraday volatility. *Journal of Financial Econometrics* **13**, 922–961.

Tsiakas, I. (2006). Periodic stochastic volatility and fat tails. *Journal of Financial Econometrics* **4**, 90–135.

Wooldridge, J.M. (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.

Ziel, F., Steinert, R. and Husmann, S. (2015). Efficient modeling and forecasting of electricity spot prices. *Energy Economics* **47**, 98–111.

Ziel, F., Croonenbroeck, C. and Ambach, D. (2016). Forecasting wind power- Modeling periodic and non-linear effects under conditional heteroscedasticity. *Applied Energy* **177**, 285–297.

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

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

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

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

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 $\gamma^S(A) < 0$ then there exists $\delta > 0$ and n_0 such that

$$E\left(\left\|A_{n_0S}A_{n_0S-1}...A_1\right\|^{\delta}\right) < 1. \qquad (S.2)$$

Since $\gamma^S(A) = \inf_{n\in\mathbb{N}^*}\left\{\frac{1}{n}E\left(\log\left\|A_{nS}A_{nS-1}...A_1\right\|\right)\right\}$ is strictly negative, there exists a positive integer n_0 such that

$$E\left(\log\left\|A_{n_0S}A_{n_0S-1}...A_1\right\|\right) < 0.$$

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

$$\begin{aligned} E\left(\left\|A_{n_0S}A_{n_0S-1}...A_1\right\|\right) &= \left\|E\left(A_{n_0S}A_{n_0S-1}...A_1\right)\right\| \\ &\leq \left\|E\left(A_SA_{S-1}...A_1\right)\right\|^{n_0} < \infty. \end{aligned}$$

Let $f(x) = E\left(\left\|A_{n_0S}A_{n_0S-1}...A_1\right\|^x\right)$. Since under (3.2) in the main paper $f'(0) = E\left(\log\left\|A_{n_0S}A_{n_0S-1}...A_1\right\|\right)$ $\gamma^S < 0$, 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 $v \in \{1, ..., S\}$

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

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

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

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

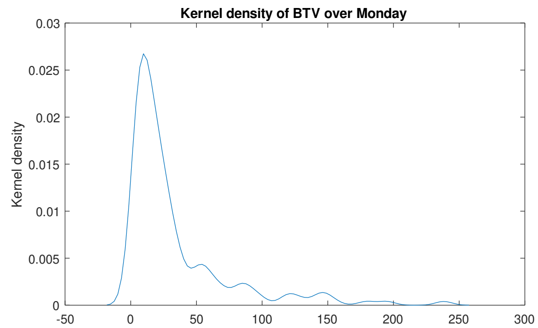
$$\begin{aligned} E\left\|\underline{Y}_v\right\|^{\kappa} &\leq \sum_{k=1}^{\infty} E\left(\left\| \prod_{j=0}^{k-1} A_{v-j} \right\|^{\kappa}\right) E\left(\left\|B_{v-k}\right\|^{\kappa}\right) + E\left(\left\|B_v\right\|^{\kappa}\right) \\ &\leq B(\kappa) \sum_{k=1}^{\infty} E\left(\left\| \prod_{j=0}^{k-1}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

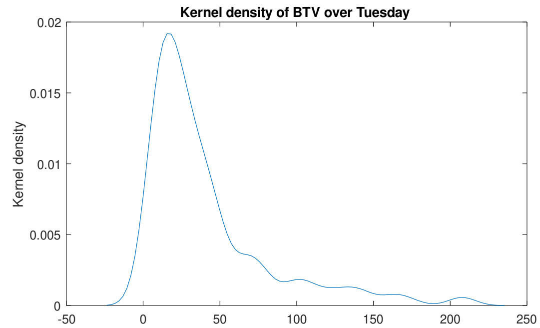
where $B(\kappa) = \max\limits_{0 \leq v \leq S-1} E\left(\left\|B_{v-k}\right\|^{\kappa}\right)$. In view of $(S.2)$ there exists $a_v > 0$ and $0 < b_v < 1$ such that

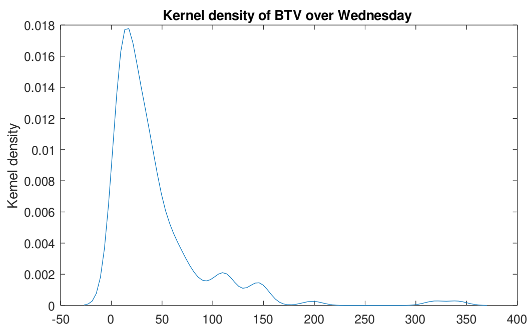
$$E\left(\left\| \prod_{j=0}^{k-1} A_{v-j} \right\|^{\kappa}\right) \leq a_v b_v^k \leq c,$$

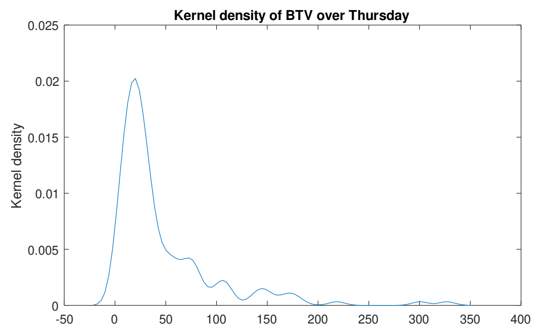
where $c = \max\limits_{0 \leq v \leq S-1} \left\{a_v b_v^k\right\}$. This proves that $E\left\|\underline{Y}_v\right\|^{\kappa} < \infty$ establishing the results.

3 Kernel densities of the Bitcoin trade volume across days









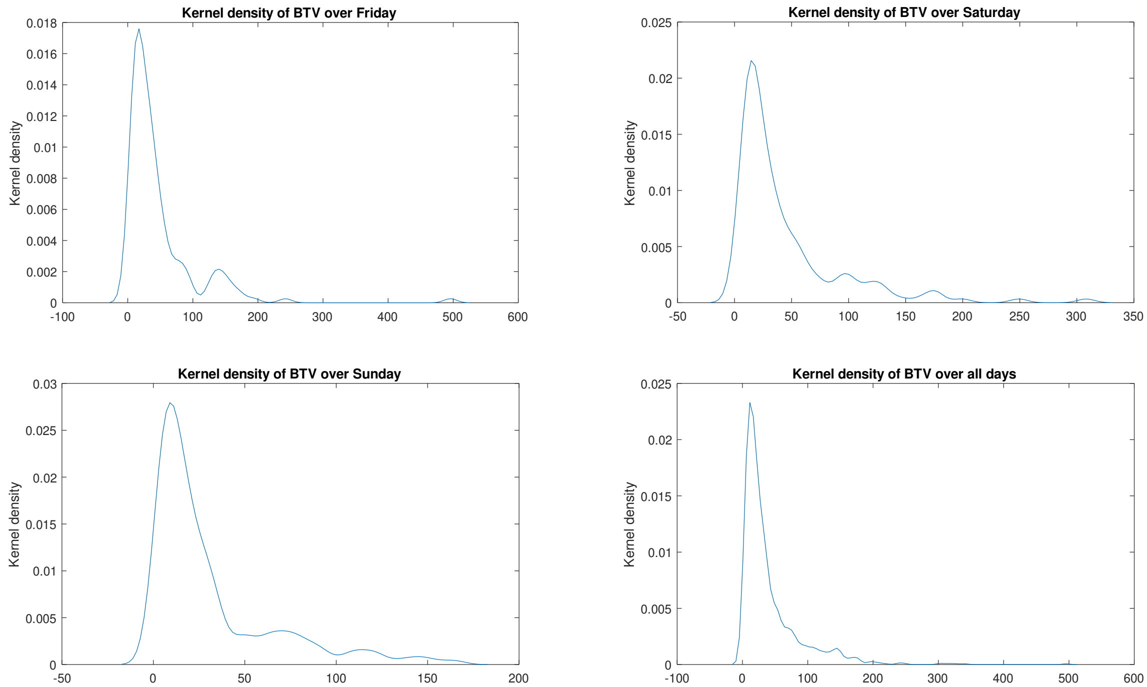


Figure S.1. Kernel densities of Bitcoin Trade Volume across days

4 Kernel densities of the UN realized volatility across days

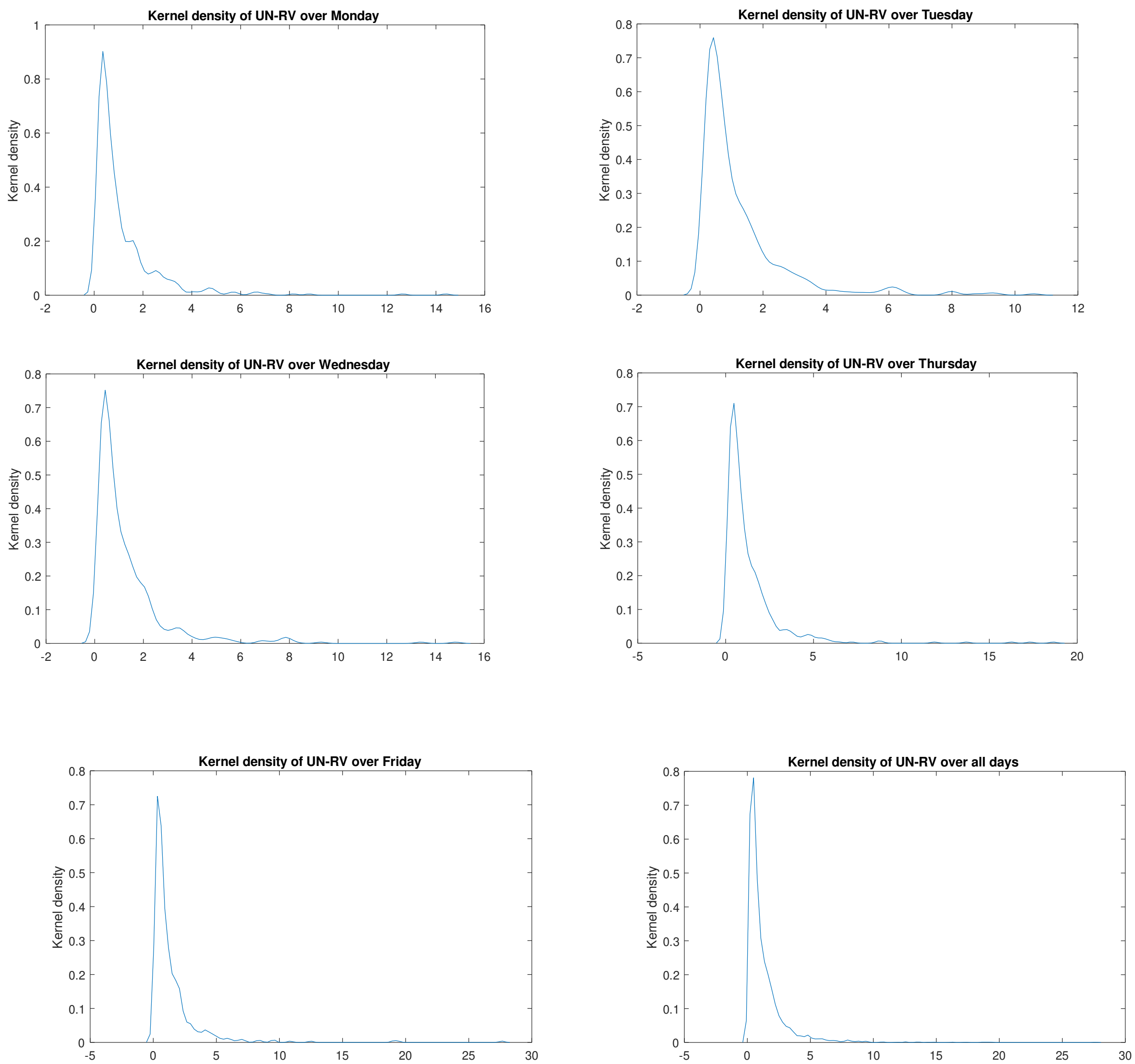


Figure S.2. Kernel densities of UN Realized Volatility across days.

5 Application to the S&P 500 volume

The third dataset is the daily S&P 500 volume over the sample period from January 04, 1999 to December, 31, 2008 with a total of $T = 2382$ observations. The time series plot of the index series is displayed in Figure S.1.

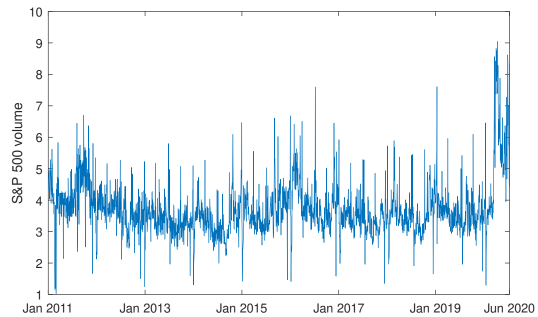


Figure S.3. S&P 500 volume series (S&PV).

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.

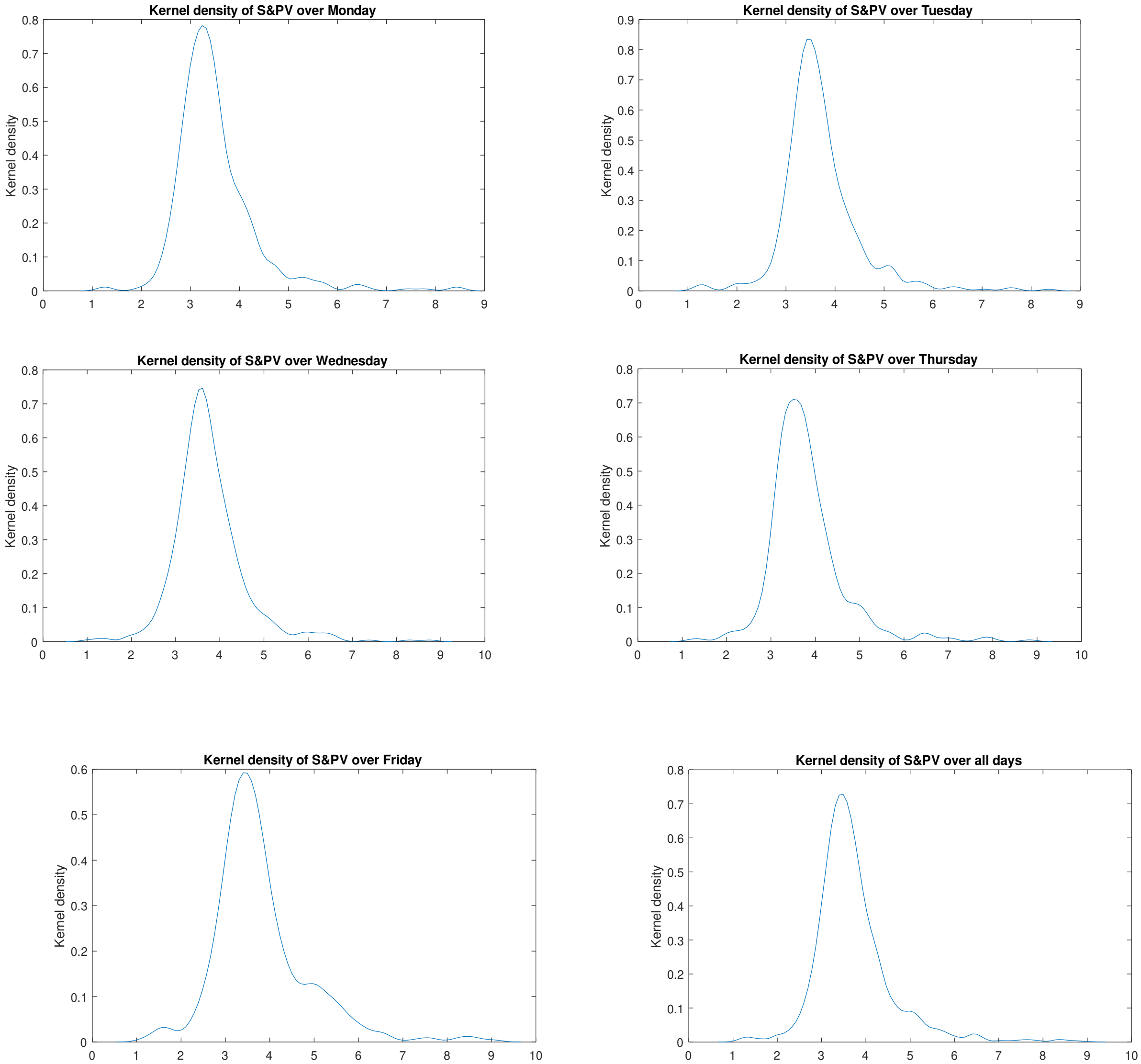


Figure S.4. Kernel densities of S&P 500 volume across days

$\widehat{\omega}$	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\sigma}^2$	$\widehat{\alpha}+\widehat{\beta}$	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	$\prod_{v=1}^{7}(\widehat{\alpha}_v + \widehat{\beta}_v)$					0.5902	0.3141	0.3744

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

	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
PACD	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.4. 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.