Noising the GARCH volatility: A random coefficient GARCH model

Aknouche, Abdelhakim and Almohaimeed, Bader and Dimitrakopoulos, Stefanos

Department of Mathematics, Faculty of Science, Qassim University, Saudi Arabia, Department of Mathematics, Faculty of Science, Qassim University, Saudi Arabia, Department of Statistics, Athens University of Economics and Business, Greece

15 March 2024

Noising the GARCH volatility: A random coefficient GARCH model

Abdelhakim Aknouche*, Bader Almohaimeed*, and Stefanos Dimitrakopoulos[1]**

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

**Department of Statistics, Athens University of Economics and Business, Greece

Abstract

This paper proposes a noisy GARCH model with two volatility sequences (an unobserved and an observed one) and a stochastic time-varying conditional kurtosis. The unobserved volatility equation, equipped with random coefficients, is a linear function of the past squared observations and of the past observed volatility. The observed volatility is the conditional mean of the unobserved volatility, thus following the standard GARCH specification, where its coefficients are equal to the means of the random coefficients. The means and the variances of the random coefficients as well as the unobserved volatilities are estimated using a three-stage procedure. First, we estimate the means of the random coefficients, using the Gaussian quasi-maximum likelihood estimator (QMLE), then, the variances of the random coefficients, using a weighted least squares estimator (WLSE), and finally the latent volatilities through a filtering process, under the assumption that the random parameters follow an Inverse Gaussian distribution, with the innovation being normally distributed. Hence, the conditional distribution of the model is the Normal Inverse Gaussian (NIG), which entails a closed form expression for the posterior mean of the unobserved volatility. Consistency and asymptotic normality of the QMLE and WLSE are established under quite tractable assumptions. The proposed methodology is illustrated with various simulated and real examples.

Keywords. Noised volatility GARCH, Randon coefficient GARCH, Markov switching GARCH, QMLE, Weighted least squares, filtering volatility, time-varying conditional kurtosis.

1 Introduction

Conditional variance/volatility models can be divided into two main categories, depending on whether the volatility is a function or not of the present shocks/noises. The first category, consists of observation-driven models (Cox, 1981), such as the generalized autoregressive conditional heteroscedastic (GARCH) model of Engle (1982) and Bollerslev (1986) and various extensions of it (Francq and Zakoian, 2019).

[1]Correspondence to: Stefanos Dimitrakopoulos, dimitrakopoulos.stefanos@outlook.com. We would like to pay tribute to Prof. Mike Tsionas for his contribution to this paper, who would have been a co-author of it, but unfortunately he passed away.

The second category, consists of parameter-driven models (Cox, 1981), such as the stochastic volatility (SV) models, introduced by Taylor (1982-1986). Markov Switching GARCH (MS-GARCH) models (Hamilton and Susmell, 1994; Gray, 1996; Klaassen, 2002; Haas et al, 2004) are often classified as parameter-driven models (Francq and Zakoian, 2008-2019). However, a distinct difference between these models and the SV models is that the MS-GARCH volatility is typically allowed to depend on past observations, whereas in SV models, the latent volatility process has an autoregressive structure that depends on its past latent values.

Observation-driven GARCH models are relatively simple to analyze and forecast, in the context of (Gaussian) quasi-maximum likelihood estimators (QMLEs). In particular, the volatility is deterministically obtained once the GARCH parameters have been estimated. However, the fact that GARCH volatility is a deterministic function of past observations might be restrictive. Such a restriction may create a sort of distortion towards large variabilities, which tends to ignore small volatilities; see Figure 1. Thus, the actual volatility path of a series may not be captured well by the GARCH model. In addition, the multiplicative form of the standard GARCH model generally implies a constant conditional kurtosis which is a non-negligible limitation (e.g. White et *al*, 2010). On the contrary, SV-type models have the advantage of integrating present shocks in the latent volatility equation but their estimation is non-trivial. In addition, in most SV models, the volatility does not depend on past observations, which can also be restrictive.

On the other hand, MS-GARCH models can overcome the limitations of the two aforementioned classes of volatility models. Indeed: i) the past of the observed process is integrated into the volatility specification and ii) the volatility depends on the present shocks, which are materialized by the regime sequence. Nonetheless, MS-GARCH models still have some drawbacks. First, their estimation and prediction is generally not straightforward, especially for MS-GARCH models that are characterized by path dependence (Gray, 1996; Haas et al, 2004; Francq and Zakoian, 2008-2019; Aknouche and Rabehi, 2010; Aknouche and Francq, 2022; Wee et al, 2022). Second, the regime sequence on which the parameters depend is generally discrete-valued and even finite, which can also be limitative. Finally, all volatility parameters depend on the same regime sequence, so a more flexible scenario, where each parameter has its own regime variable is ruled out.

Our contribution circumvents the limitations of the MS-GARCH models. In particular, we propose a multi-regime-variable random-coefficient GARCH (RC-GARCH) model that has a time-varying conditional kurtosis and two volatility sequences. The first volatility is the observed/predictive conditional variance sequence, which is nothing but the volatility equation of the standard GARCH model. So, the observed volatility is a deterministic function of past observations and can be estimated from the data using the standard Gaussian QMLE. The second one is the unobserved (latent/hidden) volatility which depends both on present shocks and past observations, as is the case with the MS-GARCH models. In contrast with the MS-GARCH models that are based on a single regime-specific, the parameters in our unobserved volatility equation are properly random (Nicholls and Quinn, 1982; Regis et al, 2022) so that each coefficient has its own (continuous) regime switching mechanism. The observed volatility can be seen as the conditional mean of the latent volatility, given the past observations. Hence, the means of the random coefficients constitute the coefficients in the observed volatility equation.

Most importantly, in our proposed RC-GARCH representation, the latent volatility is more heavy-tailed than the observed one, and, accordingly, the noise of the resulting RC-GARCH model is more light-tailed than that of the standard GARCH model. The latent volatility, therefore, can be seen as an elevated noisy version of the standard GARCH volatility.

For the estimation of the model parameters, we develop a three-stage frequentist method the asymptotic properties of which are established. The first stage estimates the means of the random coefficients using the standard Gaussian QMLE. In the second stage, the variances of the random coefficients are estimated using a weighted least squares estimate (WLSE), which is consistent and asymptotically normal (CAN) without any moment restrictions on the observed process (see also, Aknouche, 2015; Aknouche and Francq, 2023). Assuming the random coefficients are Inverse Gaussian distributed and the innovation is normally distributed, the unobserved volatility is estimated at the final stage through the posterior mean of the IG distribution, whose expression has a closed form due to the fact that the conditional distribution of the model is Normal Inverse Gaussian (NIG). Such a distribution is very flexible and can account for many stylized facts such as, asymmetry and heavy tailedness (e.g. Barndorff-Nielsen, 1997; Karlis, 2002; Rachev, 2003, 2008; Stentoft, 2008; Ayala, and Blazsek, 2019; Mozumder et al, 2024)

The structure of the paper is as follows. Section 2 defines the model and concisely study its stability properties. Section 3 presents the proposed estimation approach. Sections 4 and 5 illustrate our methodology with simulated and two real datasets, respectively. Section 6 concludes. The main proofs are displayed in the Appendix of this paper.

2 The proposed econometric specification

2.1 Noising the GARCH volatility: Some preliminaries

Let the standard Engle-Bollerslev GARCH model (Engle, 1982; Bollerslev, 1986)

$$Y_t = \delta_t \eta_t \quad \text{and} \quad \delta_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i} Y_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j} \delta_{t-j}^2 \tag{2.1}$$

where $\{\eta_t, t \in \mathbb{Z}\}$ is an iid sequence of real-valued variables with mean 0 and variance 1, with the volatility coefficients satisfying $\omega_0 > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$. Assume that (η_t) can be factorized as follows

$$\eta_t = \varepsilon_t \xi_t \tag{2.2}$$

where (ε_t) and (ξ_t) are independent, iid, such that (ε_t) is real-valued with mean 0 and variance 1, and (ξ_t) is positive-valued with $E\left(\xi_t^2\right) = 1$. Then, the standard GARCH model (2.1) could be written in the following representation

$$\begin{cases} Y_t = \delta_t \eta_t = \delta_t \overbrace{\xi_t \varepsilon_t}^{\eta_t} = \underbrace{\delta_t \xi_t}_{\sigma_t} \varepsilon_t = \sigma_t \varepsilon_t \\ \sigma_t = \delta_t \xi_t \\ \delta_t^2 = \omega_0 + \sum\limits_{i=1}^{q} \alpha_{0i} Y_{t-i}^2 + \sum\limits_{j=1}^{p} \beta_{0j} \delta_{t-j}^2 \end{cases} \tag{2.3}$$

In (2.3), the volatility (δ_t^2) is observed, given the true parameter $\theta_0 = (\omega_0, \alpha_{01}, ...\alpha_{0q}, \beta_{01}, ...\beta_{0p})'$ whereas the volatility (σ_t^2) is unobserved (latent/hidden), even with perfect knowledge of θ_0. In addition, δ_t^2 depends only on past observations $\mathscr{F}_{t-1}^Y := \sigma\{Y_{t-u}, u \geq 1\}$ and not on the present (latent) shock ξ_t as σ_t^2 does. Since $E\left(\sigma_t^2 \middle| \mathscr{F}_{t-1}^Y\right) = \delta_t^2$, the sequence $\left(\delta_t^2\right)$ can also be called predictive volatility.

Also, the new innovation (ε_t) of model (2.3) is less heavy-tailed than the innovation (η_t) of model (2.1), whereas the latent volatility (σ_t^2) is more heavy-tailed than $\left(\delta_t^2\right)$. Consequently, the standard GARCH volatility $\left(\delta_t^2\right)$ is less erratic than the latent $\left(\sigma_t^2\right)$ and seems to describe the true variability less well than $\left(\sigma_t^2\right)$. This can be seen from Figure 1, where we have generated a time series (Panel a) from a specific RC-GARCH$(1,1)$ model (see (2.5) below) along with the path of observed (Panel b), latent (Panel c) and filtered (Panel d) volatilities. In panel (b), we have annotated the observed volatility plot with artificial red curves. These curves are essentially created by large volatilities (distorted in the direction of the green lines), masking medium and small volatilities. Such a feature does not appear in the plots for the true and also the estimated/filtered latent volatilities, where medium and small volatilities are more visible. Finally, note that in (2.3), the unobserved volatility $\left(\sigma_t^2\right)$ has a multiplicative error model (MEM) representation (Engle and Russull, 1998; Engle, 2002; Aknouche and Francq, 2021; Aknouche et al, 2022b).

As in the MS-GARCH models, our latent volatility σ_t^2 also depends on past observations $\mathscr{F}_{t-1}^Y := \sigma\{Y_{t-u}, u \geq 1\}$. In fact, the noised volatility of the GARCH model (2.3) can be seen as an MS-GARCH model, yet, with a rather continuous regime sequence (ξ_t), since it can be rewritten as in the following specification

$$\begin{cases} Y_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \omega_t + \sum\limits_{i=1}^{q} \alpha_{it} Y_{t-i}^2 + \sum\limits_{j=1}^{p} \beta_{jt} \delta_{t-j}^2 \\ \delta_t^2 = \omega + \sum\limits_{i=1}^{q} \alpha_i Y_{t-i}^2 + \sum\limits_{j=1}^{p} \beta_j \delta_{t-j}^2 \end{cases} \tag{2.4}$$

in which the random coefficients $\omega_t = \xi_t\omega$, $\alpha_{it} = \xi_t\alpha_i$, and $\beta_{jt} = \xi_t\beta_j$ are "stochastically" proportional (i.e. fully positively correlated) to and are governed by the same regime variable ξ_t, the range of which can be uncountable. Equation (2.4) is, therefore, an iid regime-switching model with a single switching sequence (ξ_t). We call the procedure of passing from (2.1) to (2.3)/(2.4) as "noising the GARCH volatility" and name (2.4) the random coefficient (RC-GARCH) model.

In lieu of fully correlated random coefficients, which seems restrictive, the random coefficients of the RC-GARCH model we propose are mutually independent. The resulting specification is, thus, a multi-switching sequence (vector regime switching), where each coefficient has its own distribution. In conventional MS-GARCH models, all regimes are governed by the same Markov mechanism. The range of regimes in our model, though, can be finite, countable, or uncountable. In addition, our model has a stochastic time-varying kurtosis that can be estimated from data, unlike the standard GARCH model, in which the conditional kurtosis is constant. Finally, the parameters of our model are essentially the means and variances of the random coefficients, not necessarily having fully-specified distributions. Once these parameters are estimated, the distributions of the random coefficients can be recovered through some parametric assumptions.

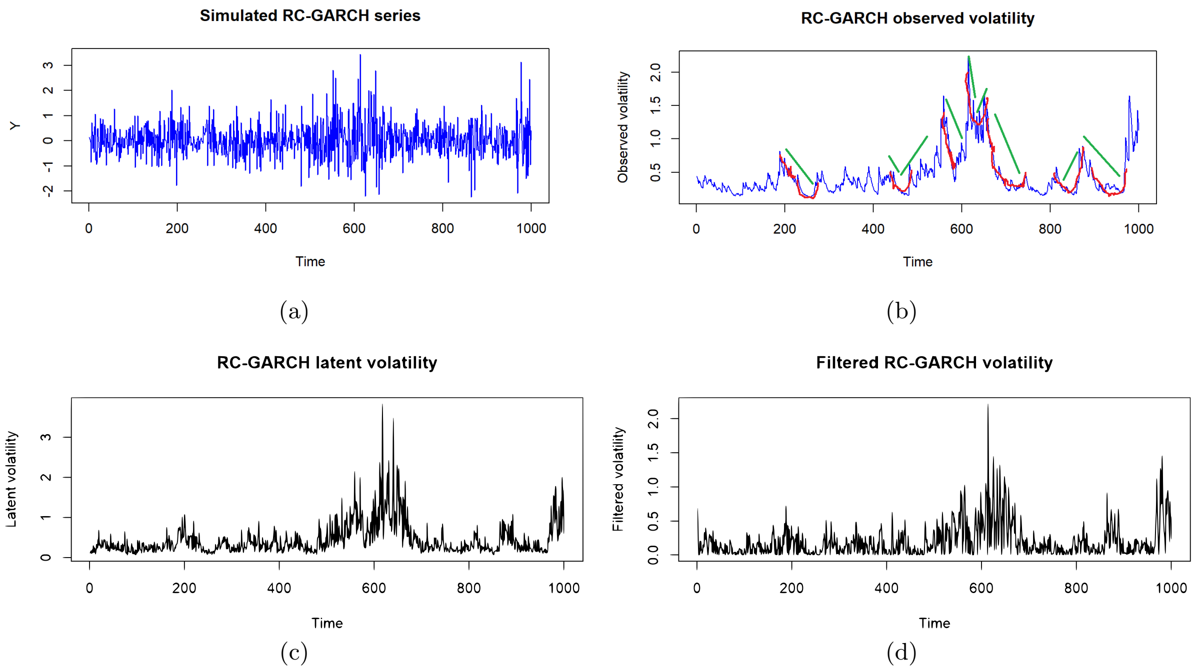


Figure 1: Simulated RC-GARCH series with $\omega_0 = 0.01, \alpha_0 = 0.1$, and $\beta_0 = 0.85$.

2.2 The RC-GARCH model

Let $\{\varepsilon_t, t \in \mathbb{Z}\}$ be an iid sequence of real-valued variables with mean 0, variance 1, and and $E\left(\varepsilon_t^4\right) := \kappa > 0$. Let also the nonnegative iid sequences $\{\omega_t, t \in \mathbb{Z}\}$, $\{\alpha_{it}, t \in \mathbb{Z}\}$ $(i = 1, ..., q)$, and $\{\beta_{it}, t \in \mathbb{Z}\}$ $(j = 1, ..., p)$ with means $\omega_0 > 0$, $\alpha_{0i} \geq 0$ and $\beta_{0j} \geq 0$, and variances $\sigma^2_{0\omega}$, $\sigma^2_{0\alpha_i}$, and $\sigma^2_{0\beta_j}$, respectively. Assume that $\{\varepsilon_t, t \in \mathbb{Z}\}$, $\{\omega_t, t \in \mathbb{Z}\}$, $\{\alpha_{it}, t \in \mathbb{Z}\}$, and $\{\beta_{it}, t \in \mathbb{Z}\}$ are mutually independent.

The observable process $\{Y_t, t \in \mathbb{Z}\}$ is said to be a random coefficient GARCH (RC-GARCH) if it is given by

$$Y_t = \sigma_t \varepsilon_t \tag{2.5a}$$

$$\sigma_t^2 = \omega_t + \sum_{i=1}^{q} \alpha_{it} Y_{t-i}^2 + \sum_{j=1}^{p} \beta_{jt} \delta_{t-j}^2 \tag{2.5b}$$

where

$$\delta_t^2 := Var\left(Y_t | \mathscr{F}_{t-1}^Y\right) = E\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) \tag{2.5c}$$

is the observed conditional variance which, by taking the conditional expectation with respect to $\mathscr{F}_{t-1}^Y$, satisfies the following standard (Bollerslev's) GARCH dynamics

$$\delta_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i} Y_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j} \delta_{t-j}^2. \tag{2.5d}$$

Let $\mathscr{F}_t^{\varepsilon,\phi} = \sigma\left\{\left(\varepsilon_{t-u}, \phi_{t-u+1}'\right)', u \geq 0\right\}$ be the complete σ-algebra generated by the past and the presence of the random inputs of (2.5) up to time t, where $\phi_t = (\omega_t, \alpha_{1t}, ..., \alpha_{qt}, \beta_{1t}, ..., \beta_{pt})'$. Then

$$\sigma_t^2 := Var\left(Y_t | \mathscr{F}_{t-1}^{\varepsilon,\phi}\right) \tag{2.6}$$

is referred to as the complete (or latent) volatility of the model (2.5). Comparing the complete and observed volatilities in (2.5b) and (2.5d), respectively, we have

$$\sigma_t^2 - \delta_t^2 = \omega_t - \omega_0 + \sum_{i=1}^{q} \left(\alpha_{it} - \alpha_{0i}\right) Y_{t-i}^2 + \sum_{j=1}^{p} \left(\beta_{jt} - \beta_{0j}\right) \delta_{t-j}^2.$$

Therefore, from the mutual independence of the random coefficients, the conditional variance of the latent volatility σ_t^2 has the following linear-in-parameter GARCH-type representation

$$\begin{aligned} Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) &= E\left(\left(\sigma_t^2 - \delta_t^2\right)^2 | \mathscr{F}_{t-1}^Y\right) \\ &= \sigma_{0\omega}^2 + \sum_{i=1}^{q} \sigma_{0\alpha_i}^2 Y_{t-i}^4 + \sum_{j=1}^{p} \sigma_{0\beta_j}^2 \delta_{t-j}^4 \end{aligned} \tag{2.7}$$

in terms of Y_{t-i}^4 and δ_{t-j}^4. Thus, the conditional variance of the squared RC-GARCH process has the form

$$Var\left(Y_t^2 | \mathscr{F}_{t-1}^Y\right) = \kappa\left(Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) + \delta_t^4\right) - \delta_t^4. \tag{2.8}$$

In particular, the conditional kurtosis of the RC-GARCH model given by

$$\kappa_t := \frac{E\left(Y_t^4 | \mathscr{F}_{t-1}^Y\right)}{\left(Var\left(Y_t | \mathscr{F}_{t-1}^Y\right)\right)^2} = \kappa \left(\frac{Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right)}{\delta_t^4} + 1 \right)$$
$$= \kappa (1 + \sigma_{0\omega}^2 \frac{1}{\delta_t^4} + \sum_{i=1}^{q} \sigma_{0\alpha_i}^2 \frac{Y_{t-i}^4}{\delta_t^4} + \sum_{j=1}^{p} \sigma_{0\beta_j}^2 \frac{\delta_{t-j}^4}{\delta_t^4}) \quad (2.9)$$

is stochastically time-varying and has a linear representation (in terms of $\frac{Y_{t-i}^4}{\delta_t^4}$ and $\frac{\delta_{t-j}^4}{\delta_t^4}$), unlike the standard Bollerslev's GARCH model in which κ_t is restrictively constant.

Note that the RC-GARCH process given by (2.5) can be seen as an extended regime-switching GARCH model in which the coefficients, components of ϕ_t, are not necessarily governed by the same law, as is the case for standard Markov-Switching GARCH (MS-GARCH) models (Francq and Zakoian, 2005-2008; Aknouche and Francq, 2022). In fact, each random coefficient can have its own distribution, the range of which can be finite, countably infinite, or even uncountable. In addition, a Markov structure could be assumed for these random coefficients, which makes the RC-GARCH model (2.5) potentially flexible. Note finally that the proposed RC-GARCH model (2.5) is non path-dependent Markov switching and is similar to the representation of Gray (1996) in the sense that (2.5b) is used instead of the following path-dependent recursion

$$\sigma_t^2 = \omega_t + \sum_{i=1}^{q} \alpha_{it} Y_{t-i}^2 + \sum_{j=1}^{p} \beta_{jt} \sigma_{t-j}^2$$

where the latent lagged value σ_{t-j}^2 is replaced by its conditional mean δ_{t-j}^2.

We now study the existence of a causal stationary and ergodic solution to equation (2.5) following the conventional stochastic recurrence equation (SRE) approach (Francq and Zakoian, 2019). Combining (2.5a), (2.5b), and (2.5d) we obtain the following stochastic recurrence equation

$$Z_t = A_t Z_{t-1} + B_t, \quad (2.10)$$

driven by the iid sequence $\{(A_t, B_t), t \in \mathbb{Z}\}$, where $Z_t = (Y_t^2, ..., Y_{t-q+1}^2, \delta_t^2, ..., \delta_{t-p+1}^2)'$,

$B_t = (\omega_t\varepsilon_t^2, 0_{(q-1)\times 1}, \omega_0, 0_{(p-1)\times 1})'$, and

$$A_t = \begin{pmatrix} \alpha_{1t}\varepsilon_t^2 & \cdots & \alpha_{q-1,t}\varepsilon_t^2 & \alpha_{qt}\varepsilon_t^2 & \beta_{1t}\varepsilon_t^2 & \cdots & \beta_{p-1,t}\varepsilon_t^2 & \beta_{pt}\varepsilon_t^2 \\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_{01} & \cdots & \alpha_{0,q-1} & \alpha_{0q} & \beta_{01} & \cdots & \beta_{0,p-1} & \beta_{0p} \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

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

$$\gamma(\mathbf{A}) = \inf\left\{\tfrac{1}{t}E\log\|A_t...A_2A_1\|,\ t \geq 1\right\}$$

be the largest Lyapunov exponent associated with the iid-driven SRE (2.10) (Bougerol and Picard, 1992). Let also

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_{01} & \cdots & \beta_{0,p-1} & \beta_{0p} \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}.$$

The following result gives conditions for equation (2.10) to have a unique strictly stationary and ergodic solution.

Proposition 2.1 *i) Assume* $E\left(\log\left(\varepsilon_t^2\right)\right) < \infty$, $E(\log(\omega_t)) < \infty$, $E(\log(\alpha_{it})) < \infty$ *and* $E(\log(\beta_{jt})) < \infty$ *(*$i = 1,...,q,\ j = 1,...,p$*). A necessary and sufficient condition for model* (2.10) *to have a unique nonanticipative strictly stationary and ergodic solution is that*

$$\gamma(\mathbf{A}) < 0. \tag{2.11}$$

Such a solution is given for all $t \in \mathbb{Z}$ *by*

$$Z_t = \sum_{j=0}^{\infty}\prod_{i=0}^{j-1} A_{t-i}B_{t-j}, \tag{2.12}$$

where the series in the right hand side of (2.12) *converges absolutely almost surely.*

ii) If (2.10) *admits a strictly stationary solution then*

$$\rho(\boldsymbol{\beta}) < 1. \tag{2.13}$$

In the special case where $p = q = 1$, another simple and equivalent stationarity condition for (2.10) is as follows

$$E\left(\log\left|\alpha_{1t}\varepsilon_{t-1}^2 + \beta_{1t}\right|\right) < 0,$$

while (2.13) reduces to $0 \leq \beta_{01} < 1$.

Conditions for the existence of second and fourth moments of the model (2.5) are given as follows.

Proposition 2.2 *Assume* $E\left(\varepsilon_t^2\right) < \infty$, $E\left(\omega_t\right) < \infty$, $E\left(\alpha_{it}\right) < \infty$ *and* $E\left(\beta_{jt}\right) < \infty$ $(i = 1, ..., q,$ $j = 1, ..., p)$. *A sufficient condition for the process, given by* (2.1), *to be strictly stationary and ergodic with* $E\left(Y_t^2\right) < \infty$ *is that*

$$\rho\left(E\left(A_t\right)\right) < 1 \tag{2.14}$$

where

$$E\left(A_t\right) = \begin{pmatrix} \alpha_{01} & \cdots & \alpha_{0,q-1} & \alpha_{0q} & \beta_{01} & \cdots & \beta_{0,p-1} & \beta_{0p} \\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_{01} & \cdots & \alpha_{0,q-1} & \alpha_{0q} & \beta_{01} & \cdots & \beta_{0,p-1} & \beta_{0p} \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

Using a similar device by Chen and An (1998) and Francq and Zakoian (2019), condition (2.14) reduces to the following

$$\sum_{i=1}^{q}\alpha_{0i} + \sum_{j=1}^{p}\beta_{0j} < 1.$$

The unconditional mean of the process is given under (2.14) by

$$E\left(Y_t^2\right) = \frac{\omega_0}{1 - \sum_{i=1}^{q}\alpha_{0i} - \sum_{j=1}^{p}\beta_{0j}}.$$

Proposition 2.3 *Assume* $E\left(\varepsilon_t^4\right) < \infty$, $E\left(\omega_t^2\right) < \infty$, $E\left(\alpha_{it}^2\right) < \infty$, *and* $E\left(\beta_{jt}^2\right) < \infty$ $(i = 1, ..., q,$ $j = 1, ..., p)$. *A sufficient condition for the process, given by* (2.1), *to be strictly stationary and ergodic with* $E\left(Y_t^4\right) < \infty$ *is that*

$$\rho\left(E\left(A_t \otimes A_t\right)\right) < 1. \tag{2.15}$$

When $p = q = 1$, the eigenvalues of $E\left(A_t \otimes A_t\right)$ are

$$\left\{\kappa\alpha_{01}^2 + 2\alpha_{01}\beta_{01} + \beta_{01}^2 + \kappa\sigma_{0\alpha_1}^2 + \sigma_{0\beta_1}^2, 0\right\},$$

so condition (2.15) is

$$\kappa\alpha_{01}^2 + 2\alpha_{01}\beta_{01} + \beta_{01}^2 + \kappa\sigma_{0\alpha_1}^2 + \sigma_{0\beta_1}^2 < 1.$$

In particular, when all slope parameters are not random, i.e. $\sigma_{0\alpha_1}^2 = \sigma_{0\beta_1}^2 = 0$, we obtain the fourth moment condition for the standard GARCH$(1,1)$ model (cf. Francq and Zakoian, 2019).

3 Parameter estimation

The parameters of the RC-GARCH model are now estimated given a realization $Y_1, ..., Y_n$ generated from (2.5). These parameters are of three types, namely: i) the random coefficient means $\theta_0 = (\omega_0, \alpha_{01}, ...\alpha_{0q}, \beta_{01}, ...\beta_{0p})'$, ii) the random coefficient variances $\Lambda_0 = \left(\sigma_{0\omega}^2, \sigma_{0\alpha_1}^2, ..., \sigma_{0\alpha_q}^2, \sigma_{0\beta_1}^2, ..., \sigma_{0\beta_p}^2\right)'$, and iii) the unobserved conditional variances $\sigma_1^2, ..., \sigma_n^2$ which are augmented parameters. To estimate the model parameters we use a three-stage procedure, where each stage deals with each block of parameters in the mentioned order. In particular, the Gaussian QMLE is first used to estimate θ_0. In principle, no assumption on the distribution of the innovation ε_t is needed. Second, a weighted least squares estimate is used for Λ_0 and requires the specification of the fourth moment $\kappa = E\left(\varepsilon_t^4\right)$. For the latent volatilities $\sigma_1^2, ..., \sigma_n^2$, we use the posterior means $E\left(\sigma_t^2 | Y_1, ..., Y_t\right)$ $(1 \leq t \leq n)$. To get closed form results, the random coefficients are assumed to be Inverse Gaussian (IG) distributed, while the innovation is assumed to be normally distributed $\mathcal{N}(0,1)$. As such, the conditional distribution $Y_t | \mathscr{F}_{t-1}^Y$ of the model is Normal Inverese Gaussian (NIG) distributed (Barndorff-Nielsen, 1997), where the conditional posterior mean $\sigma_t^2 | Y_1, ..., Y_t$ can be easily obtained in closed form (Karlis, 2022). The NIG distribution (see Appendix) has many advantages over the normal distribution, such as allowing for asymmetry and heavy tailedness and is very flexible in modelling financial time series (Bardorff-Nielsen, 1997; Karlis, 2002; Rachev, 2003; Blazsek et *al*, 2018).

3.1 Estimating the random coefficient means

First, the parameter vector $\theta_0 = (\omega_0, \alpha_{01}, ...\alpha_{0q}, \beta_{01}, ...\beta_{0p})'$ is estimated from the data using the Gaussian QMLE. Then, the observable volatilities $\delta_1^2, ..., \delta_n^2$ are estimated from $(2.1d)$. For all generic $\theta = (\omega, \alpha_1, ...\alpha_q, \beta_1, ...\beta_p)' \in \Theta \subset \mathbb{R}^{p+q+1}$ let

$$\delta_t^2(\theta) = \omega + \sum_{i=1}^{q} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \delta_{t-j}^2(\theta), \qquad t \in \mathbb{Z} \tag{3.1}$$

be the generic observed volatility, which exists and is stationary and ergodic whenever (2.11) and the following condition

$$\sum_{j=1}^{p} \beta_j < 1, \ \forall \theta \in \Theta, \tag{3.2}$$

are satisfied. Given arbitrary initial values $Y_0, ...Y_{1-q}, \widetilde{\delta}_0^2, ..., \widetilde{\delta}_{1-p}^2$, let $\widetilde{\delta}_t^2(\theta)$ be an observable approximation to (3.1) given by

$$\widetilde{\delta}_t^2(\theta) = \omega + \sum_{i=1}^{q} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \widetilde{\delta}_{t-j}^2(\theta), \qquad t \geq 1. \tag{3.3}$$

The Gaussian QMLE of θ_0 is a solution to the following problem

$$\widehat{\theta}_n = \arg\min_{\theta} \widetilde{L}_n(\theta) \tag{3.4}$$

where

$$\widetilde{L}_n(\theta) = \tfrac{1}{n} \sum_{t=1}^{n} \widetilde{\ell}_t(\theta) \quad \text{and} \quad \widetilde{\ell}_t(\theta) = \log \widetilde{\delta}_t^2(\theta) + \tfrac{Y_t^2}{\widetilde{\delta}_t^2(\theta)}. \tag{3.5}$$

Based on the standard asymptotic GARCH theory (Francq and Zakoian, 2004-2019) we will show that $\widehat{\theta}_n$ is consistent and asymptotically Normal under the following standard assumptions.

A1 Θ is a compact.

A2 Conditions (2.11) and (3.2) are satisfied.

A3 The distribution of ε_t^2 is non-degenerate and $E\left(\varepsilon_t^2\right) = 1$.

A4 The polynomials $A_{\theta_0}(z) = \sum_{i=1}^{q} \alpha_{0i} z^i$ and $B_{\theta_0}(z) = 1 - \sum_{i=1}^{p} \beta_{0i} z^i$ have no common roots, $A_{\theta_0}(z) \neq 1$, and $\alpha_{0q} + \beta_{0p} \neq 0$.

A5 θ_0 is in the interior of Θ.

A6 $E\left(\varepsilon_t^4\right) = \kappa < \infty$.

Set

$$I := E\left(\tfrac{(\kappa-1)\delta_t^4(\theta_0) + \kappa Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right)}{\delta_t^8(\theta_0)} \tfrac{\partial \delta_t^2(\theta_0)}{\partial \theta} \tfrac{\partial \delta_t^2(\theta_0)}{\partial \theta'} \right) \text{ and } J := E\left(\tfrac{1}{\delta_t^4(\theta_0)} \tfrac{\partial \delta_t^2(\theta_0)}{\partial \theta} \tfrac{\partial \delta_t^2(\theta_0)}{\partial \theta'} \right). \tag{3.6}$$

Theorem 3.1 *Under* ***A1-A4***,

$$\widehat{\theta}_n \underset{n \to \infty}{\overset{a.s.}{\to}} \theta_0. \tag{3.7}$$

If, in addition, ***A5-A6*** *are satisfied then*

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \underset{n \to \infty}{\overset{D}{\to}} \mathcal{N}\left(0, J^{-1} I J^{-1}\right), \tag{3.8}$$

where J is invertible.

When all random parameters are degenerate, it follows that $\sigma_t^2 = \delta_t^2$ and

$$Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) = Var\left(\delta_t^2 | \mathscr{F}_{t-1}^Y\right) = 0$$

since δ_t^2 is $\mathscr{F}_{t-1}^Y$-measurable. Thus, $E\left(\frac{\partial l_t(\theta_0)}{\partial \theta}\frac{\partial l_t(\theta_0)}{\partial \theta'}\right)$ reduces to $(\kappa - 1)\, E\left(\frac{1}{\delta_t^4(\theta_0)}\frac{\partial \delta_t^2(\theta_0)}{\partial \theta}\frac{\partial \delta_t^2(\theta_0)}{\partial \theta'}\right)$, which is the covariance matrice of the Gaussian QMLE of the standard GARCH model (Francq and Zakoian, 2004-2019). Consistent estimates of I and J are given, respectively, by

$$\widehat{I}_n = \frac{1}{n}\sum_{t=1}^{n} \frac{\left(Y_t^2 - \widehat{\delta}_t^2\right)^2}{\widehat{\delta}_t^8}\frac{\partial \widehat{\delta}_t^2}{\partial \theta}\frac{\partial \widehat{\delta}_t^2}{\partial \theta'},\quad \widehat{J}_n = \frac{1}{n}\sum_{t=1}^{n}\frac{1}{\widehat{\delta}_t^4}\frac{\partial \widehat{\delta}_t^2}{\partial \theta}\frac{\partial \widehat{\delta}_t^2}{\partial \theta'}, \tag{3.9a}$$

where

$$\widehat{\delta}_t^2 = \widetilde{\delta}_t^2\left(\widehat{\theta}_n\right),\ 1 \leq t \leq n. \tag{3.9b}$$

3.2 Estimating the random coefficient variances

At this stage, the distribution of ε_t and hence of $Y_t|\sigma_t^2$ has to be specified. It is assumed that ε_t is normally distributed with mean zero and unit variance ($\varepsilon_t \sim \mathcal{N}(0,1)$) and hence $\kappa = E\left(\varepsilon_t^4\right) = 3$. Then, $\Lambda_0 = (\sigma_{0\omega}^2, \sigma_{0\alpha_1}^2, ..., \sigma_{0\alpha_q}^2, \sigma_{0\beta_1}^2, ..., \sigma_{0\beta_p}^2)'$ will be estimated from a regression built from equation (2.8). Let $e_t = \left(Y_t^2 - \delta_t^2\right)^2 - Var\left(Y_t^2 | \mathscr{F}_{t-1}^Y\right)$ so that (Nichols and Quinn, 1982)

$$\left(Y_t^2 - \delta_t^2\right)^2 = Var\left(Y_t^2 | \mathscr{F}_{t-1}^Y\right) + e_t. \tag{3.10a}$$

Then, from (2.8) and (2.7), we have $Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) = M_t'\Lambda_0$ and $Var\left(Y_t^2 | \mathscr{F}_{t-1}^Y\right) = (\kappa - 1)\,\delta_t^4 + M_t'\Lambda_0$, so $(3.10a)$ becomes

$$\frac{\left(Y_t^2 - \delta_t^2\right)^2 - (\kappa - 1)\delta_t^4}{\kappa \delta_t^4} = \frac{1}{\delta_t^4}M_t'\Lambda_0 + \frac{e_t}{\delta_t^4}, \tag{3.10b}$$

where $E\left(\frac{e_t}{\delta_t^4} | \mathscr{F}_{t-1}^Y\right) = \frac{1}{\delta_t^4}E\left(e_t | \mathscr{F}_{t-1}^Y\right) = 0$ and

$$M_t = (1, Y_{t-1}^4, ..., Y_{t-q}^4, \delta_{t-1}^4, ..., \delta_{t-p}^4)'. \tag{3.11}$$

From the regression $(3.10b)$, a WLS estimate of Λ_0 is given by

$$\widehat{\Lambda}_n = \left(\sum_{t=1}^{n}\frac{1}{\widehat{\delta}_t^8}\widehat{M}_t\widehat{M}_t'\right)^{-1}\sum_{t=1}^{n}\widehat{M}_t\frac{\left(Y_t^2 - \widehat{\delta}_t^2\right)^2 - (\kappa - 1)\widehat{\delta}_t^4}{\kappa \widehat{\delta}_t^8} \tag{3.12}$$

where $\widehat{\delta}_t^2 = \widetilde{\delta}_t^2(\widehat{\theta}_n)$ is evaluated from (3.9b) and

$$\widehat{M}_t = (1, Y_{t-1}^4, ..., Y_{t-q}^4, \widehat{\delta}_{t-1}^4, ..., \widehat{\delta}_{t-p}^4)'.$$

To study the consistency and asymptotic normality of $\widehat{\Lambda}_n$, define

$$A = E\left(\frac{1}{\delta_t^8(\theta_0)} M_t M_t'\right) \tag{3.13a}$$

$$B = \frac{1}{\kappa^2} E\left(\frac{e_t^2}{\delta_t^{16}} M_t M_t'\right) = \frac{1}{\kappa^2} E\left(\frac{Var\left(\left(Y_t^2 - \delta_t^2\right)^2 | \mathscr{F}_{t-1}^Y\right)}{\delta_t^{16}} M_t M_t'\right). \tag{3.13b}$$

Clearly, these matrices are finite and A is invertible. Consider the following moment assumption.

A7: $E\left(\varepsilon_t^8\right) < \infty$.

Theorem 3.2 *Under **A1-A4** and **A6**,*

$$\widehat{\Lambda}_n \underset{n\to\infty}{\overset{a.s.}{\to}} \Lambda_0. \tag{3.14}$$

*If, in addition, **A7** holds then*

$$\sqrt{n}\left(\widehat{\Lambda}_n - \Lambda_0\right) \underset{n\to\infty}{\overset{D}{\to}} \mathcal{N}\left(0, A^{-1}BA^{-1}\right). \tag{3.15}$$

Assuming that $\varepsilon_t \sim \mathcal{N}(0,1)$, all moments of ε_t are finite, so the eigth moment assumption **A7** does not really appear stringent. From (3.10a) and (2.8), a consistent estimates of A and B in (3.13) are, respectively,

$$\widehat{A}_n = \frac{1}{n}\sum_{t=1}^{n} \frac{1}{\widehat{\delta}_t^8} \widehat{M}_t \widehat{M}_t' \text{ and } \widehat{B}_n = \frac{1}{n}\sum_{t=1}^{n} \frac{\left(\left(Y_t^2 - \widehat{\delta}_t^2\right)^2 - (\kappa-1)\widehat{\delta}_t^4 - \kappa M_t'\widehat{\Lambda}_n\right)^2}{\kappa^2 \widehat{\delta}_t^{16}} \widehat{M}_t \widehat{M}_t'. \tag{3.16}$$

3.3 Estimating/filtering the unobserved volatilities

Finally, the unobserved volatilities $\sigma_1^2, ..., \sigma_n^2$ are estimated using the smoothed volatility

$$\widehat{\sigma}_t^2 = E\left(\sigma_t^2 | Y_1, ..., Y_n\right), \ t = 1, ..., n, \tag{3.17}$$

that we obtain from the smoothed/filtered distribution $f\left(\sigma_t^2 | Y_1, ..., Y_t\right)$. Consider the RC-GARCH model (2.5). We first need to specify the distribution of the innovation ε_t and the random coefficients $\theta_t := (\omega_t, \alpha_{1t}, ..., \alpha_{qt}, \beta_{1t}, ..., \beta_{pt})'$. We, thus, assume that

$$\varepsilon_t \sim \mathcal{N}(0,1) \text{ so that } Y_t | \sigma_t^2 \sim \mathcal{N}\left(0, \sigma_t^2\right). \tag{3.18}$$

Then, the random coefficients are assumed to be IG distributed (see Appendix A, $(A.1)$), that is

$$\omega_t \sim \mathcal{IG}(\omega_0, \lambda_\omega) \text{ with mean } \omega_0 \text{ and shape } \lambda_\omega \text{ so } Var(\omega_t) = \tfrac{\omega_0^3}{\lambda_\omega} \tag{3.19a}$$

$$\alpha_{it} \sim \mathcal{IG}(\alpha_{0i}, \lambda_{\alpha_i}) \text{ with mean } \alpha_{0i} \text{ and shape } \lambda_{\alpha_i} \text{ so } Var(\alpha_{it}) = \tfrac{\alpha_{0i}^3}{\lambda_{\alpha_i}} \tag{3.19b}$$

$$\beta_{jt} \sim \mathcal{IG}\left(\beta_{0j}, \lambda_{\beta_j}\right) \text{ with mean } \beta_{0j} \text{ and shape } \lambda_{\beta_j} \text{ so } Var(\beta_{jt}) = \tfrac{\beta_{0j}^3}{\lambda_{\beta_j}}. \tag{3.19c}$$

From the summability property of the IG distribution (see Appendix A) and the mutual independence of $\{\omega_t, t \in \mathbb{Z}\}$, $\{\alpha_{it}, t \in \mathbb{Z}\}$ $(i = 1, ..., q)$, and $\{\beta_{it}, t \in \mathbb{Z}\}$ $(j = 1, ..., p)$, which entails the conditional independence of ω_t, $\alpha_{it} Y_{t-i}^2$, and $\beta_{jt}\delta_{t-j}^2$ $(i = 1, ..., q,\ j = 1, ..., p)$ given $\mathscr{F}_{t-1}^Y$, the conditional distribution of $\sigma_t^2 | \mathscr{F}_{t-1}^Y$ is thus

$$\sigma_t^2 | \mathscr{F}_{t-1}^Y \sim \mathcal{IG}\left(\delta_t^2, \Delta_t^2\right). \tag{3.20a}$$

In view of (3.19), δ_t^2 is given by $(2.5d)$ and

$$\Delta_t^2 = \lambda_\omega + \sum_{i=1}^{q} \lambda_{\alpha_i} Y_{t-i}^2 + \sum_{j=1}^{p} \lambda_{\beta_j} \delta_{t-j}^2 \tag{3.20b}$$

where $\lambda_\omega = \frac{\omega_0^3}{\sigma_{0\omega}^2}$, $\lambda_{\alpha_i} = \frac{\alpha_{0i}^3}{\sigma_{0\alpha_i}^2}$, and $\lambda_{\beta_j} = \frac{\beta_{0j}^3}{\sigma_{0\beta_j}^2}$.

Consequently, the conditional distribution of the model given by

$$f\left(Y_t | \mathscr{F}_{t-1}^Y\right) = \int_{(0,\infty)} f\left(Y_t, \sigma_t^2 | \mathscr{F}_{t-1}^Y\right) d\sigma_t^2 = \int_{(0,\infty)} f\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) f\left(y_t | \sigma_t^2\right) d\sigma_t^2, \tag{3.21}$$

is a continuous mixture of normal distributions with Inverse Gaussian mixings. This distribution is called Normal Inverse Gaussian (NIG, see Appendix A). It has a closed form density (see Appendix, $(A.2)$) and is also given in the following hierarchical mixture (see Appendix, $(A.3)$)

$$\begin{cases} \sigma_t^2 | \mathscr{F}_{t-1}^Y \sim \mathcal{IG}\left(\delta_t^2, \Delta_t^2\right) \\ Y_t | \sigma_t^2 \sim \mathcal{N}\left(0, \sigma_t^2\right) \end{cases} \Longrightarrow Y_t | \mathscr{F}_{t-1}^Y \sim \mathcal{NIG}\left(\Delta_t^2, 0, \delta_t^2, 0\right). \tag{3.22}$$

Due to the the non-anticpativeness of the model, the posterior smoothed volatility $\sigma_t^2 | \mathscr{F}_n^Y$ is the same as the posterior filtered volatility $\sigma_t^2 | \mathscr{F}_t^Y$ whose density is given by

$$f\left(\sigma_t^2 | \mathscr{F}_t^Y\right) = f\left(\sigma_t^2 | Y_t, \mathscr{F}_{t-1}^Y\right) = \tfrac{f\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) f\left(Y_t | \sigma_t^2\right)}{f\left(Y_t | \mathscr{F}_{t-1}^Y\right)}.$$

Thus, the unobserved volatility σ_t^2 can be estimated by the smoothed volatility $E\left(\sigma_t^2 | \mathscr{F}_t^Y\right)$ given by

$$E\left(\sigma_t^2 \middle| \mathscr{F}_t^Y\right) = \frac{1}{f\left(Y_t \middle| \mathscr{F}_{t-1}^Y\right)} \int_{(0,\infty)} \sigma_t^2 f\left(\sigma_t^2 \middle| \mathscr{F}_{t-1}^Y\right) f\left(Y_t \middle| \sigma_t^2\right) d\sigma_t^2. \tag{3.23}$$

Using the result of Karlis (2002, formula (4)) for the NIG distribution, a closed form for the IG posterior mean in (3.23) is given by $(A.4)$ (see Appendix A). Thus, our estimate $\widehat{\sigma}_t^2$ of σ_t^2 is obtained while replacing the true parameters in the expression (3.23) and $(A.5)$ by their estimates obtained in the first and second stages, giving

$$\widehat{\sigma}_t^2 = \widehat{E}\left(\sigma_t^2 \middle| \mathscr{F}_t^Y\right) = \frac{\sqrt{\widehat{\delta}_t^2 + Y_t^2}}{\widehat{\Delta}_t} \frac{K_0\left(\widehat{\Delta}_t \sqrt{\widehat{\delta}_t^2 + Y_t^2}\right)}{K_1\left(\widehat{\Delta}_t \sqrt{\widehat{\delta}_t^2 + Y_t^2}\right)} \tag{3.24}$$

where $\widehat{\delta}_t^2$ is given by $(3.9b)$, $K_r(y)$ denotes the modified Bessel function of the third kind of order r evaluated at y, and

$$\widehat{\Delta}_t^2 = \widehat{\lambda}_\omega + \sum_{i=1}^{q} \widehat{\lambda}_{\alpha_i} Y_{t-i}^2 + \sum_{j=1}^{p} \widehat{\lambda}_{\beta_j} \widehat{\delta}_{t-j}^2. \tag{3.25}$$

The estimates $\widehat{\lambda}_\omega = \frac{\widehat{\omega}^3}{\widehat{\sigma}_\omega^2}$, $\widehat{\lambda}_{\alpha_i} = \frac{\widehat{\alpha}_i^3}{\widehat{\sigma}_{\alpha_i}^2}$, and $\widehat{\lambda}_{\beta_j} = \frac{\widehat{\beta}_j^3}{\widehat{\sigma}_{\beta_j}^2}$ $(i = 1, ..., q,\ j = 1, ..., p)$ are obtained from (3.4), (3.19), $(3.9b)$, and (3.12). Note finally that Δ_t^2 can be interpreted as a "conditional" heavy-tail parameter (see Appendix A; Barndorff-Nielsen and Prause, 2001).

3.4 Summary

The following algorithm summarizes the three-stage method to estimate the RC-GARCH parameters (2.5).

Algorithm 3.1 (Three-stage method)

Given an observed series $Y_1, ..., Y_n$:

Stage I

1- Estimate $\theta_0 = (\omega_0, \alpha_{01}, ..., \beta_{0p})'$ using the Gaussian QMLE $\widehat{\theta}_n$ given by (3.4).

2- Estimate the observable volatilities $\widehat{\delta}_1^2, ..., \widehat{\delta}_n^2$ from (3.3), where $\widehat{\delta}_t^2 = \widetilde{\delta}_t^2(\widehat{\theta}_n)$ $(1 \leq t \leq n)$.

3- Estimate the asymptotic variance of $\widehat{\theta}_n$ and hence its asymptotic standard error (ASE) from (3.9).

Stage II

4- Estimate the variances of the random coefficients $\Lambda_0 = \left(\sigma_{0\omega}^2, \sigma_{0\alpha_1}^2, ..., \sigma_{0\alpha_q}^2, \sigma_{0\beta_1}^2, ..., \sigma_{0\beta_p}^2\right)'$ using the WLSE $\widehat{\Lambda}_n$ from (3.12).

5- Estimate the asymptotic variance and then the ASE of $\widehat{\Lambda}_n$ from (3.16).

Stage III

6- Estimate the latent volatilities $\sigma_1^2, ..., \sigma_n^2$ from (3.24)-(3.25), using the posterior mean of the Inverse Gaussian distribution. □

4 Simulated data

The finite-sample performances of the QML and WLS estimators given by Algorithm 3.1 are assessed for the RC-GARCH$(1,1)$ model via a Monte Carlo simulation study. To this end, three cases of the RC-GARCH model are considered. In the first case, ε_t is Gaussian, whereas the random coefficients $\phi_t = (\omega_t, \alpha_t, \beta_t)$ are inverse Gaussian distributed ; see Table 1. In the second case, ε_t is Gaussian, while the random coefficients are Poisson distributed; see Table 2. Finally, in the third case, ε_t is Gaussian, where the random coefficients are exponentially distributed; see Table 3.

We run the QMLE and WLSE on 1000 sample-paths generated from the RC-GARCH$(1,1)$ model with sample size $n \in \{1000, 3000, 5000\}$, and $\theta_0 = (\omega_0, \alpha_0, \beta_0)' = (0.01, 0.15, 0.80)'$. This choice is close to the estimated values obtained in the real applications. The variance parameters $\Lambda_0 = \left(\sigma^2_{0\omega}, \sigma^2_{0\alpha}, \sigma^2_{0\beta}\right)$ are deduced accordingly from the distribution of ϕ_t in each case (see Tables 1-3). For the QMLE, we use the nonlinear optimization function "nlimb", while for the WLSE, the constrained nonnegative least squares function "nnls". In fact, without any nonnegativity constraint, the WLS estimates can give negative values.

		QMLE			WLSE		
n	(θ_0, Λ_0)	ω_0	α_0	β_0	$\sigma^2_{0\omega}$	$\sigma^2_{0\alpha}$	$\sigma^2_{0\beta}$
		0.0100	0.1500	0.8000	0.0100	0.3375	0.2560
1000	Mean	0.0113	0.1527	0.7898	0.0097	0.2929	0.2772
	StD	0.0058	0.0558	0.0627	0.0331	0.0822	0.0709
	ASE	0.0048	0.0475	0.0554	0.0144	0.0649	0.0635
3000	Mean	0.0103	0.1507	0.7980	0.0082	0.3598	0.2699
	StD	0.0031	0.0330	0.0366	0.0212	0.0776	0.0565
	ASE	0.0027	0.0294	0.0328	0.0115	0.0529	0.0501
5000	Mean	0.0102	0.1509	0.7983	0.0086	0.3325	0.2522
	StD	0.0021	0.0233	0.0254	0.0170	0.0358	0.0473
	ASE	0.0021	0.0229	0.0256	0.0104	0.0226	0.0388

Table 1. QMLE and WLSE results for 1000 RC-GARCH(1,1) series with sample size n, $\omega_t \sim \mathcal{IG}(\omega_0, 0.0001)$, $\alpha_t \sim \mathcal{IG}(\alpha_0, 0.01)$, and $\beta_t \sim \mathcal{IG}(\beta_0, 2)$.

n	(θ_0, Λ_0)	QMLE			WLSE		
		ω_0	α_0	β_0	$\sigma^2_{0\omega}$	$\sigma^2_{0\alpha}$	$\sigma^2_{0\beta}$
		0.0100	0.1500	0.8000	0.0100	0.1500	0.8000
1000	Mean	0.0111	0.1513	0.7901	0.0109	0.1482	0.7770
	StD	0.0052	0.0488	0.0592	0.0509	0.0549	0.0421
	ASE	0.0048	0.0459	0.0553	0.0430	0.0464	0.0388
3000	Mean	0.0105	0.1523	0.7950	0.0100	0.1480	0.8099
	StD	0.0026	0.0277	0.0311	0.0316	0.0327	0.0364
	ASE	0.0025	0.0276	0.0309	0.0295	0.0303	0.0252
5000	Mean	0.0103	0.1504	0.7970	0.0099	0.14971	0.8065
	StD	0.0020	0.0203	0.0235	0.0122	0.0284	0.0291
	ASE	0.0019	0.0215	0.0240	0.0086	0.0206	0.0257

Table 2. QMLE and WLSE results for 1000 RC-GARCH(1,1) series with sample size n, $\omega_t \sim \mathcal{P}(\omega_0)$, $\alpha_t \sim \mathcal{P}(\alpha_0)$, and $\beta_t \sim \mathcal{P}(\beta_0)$.

n	(θ_0, Λ_0)	QMLE			WLSE		
		ω_0	α_0	β_0	$\sigma^2_{0\omega}$	$\sigma^2_{0\alpha}$	$\sigma^2_{0\beta}$
		0.0100	0.1500	0.8000	0.0001	0.0225	0.6400
1000	Mean	0.0117	0.1535	0.7851	0.0029	0.0366	0.6350
	StD	0.0051	0.0379	0.0534	0.0056	0.0364	0.0438
	ASE	0.0044	0.0372	0.0485	0.0051	0.0338	0.0445
3000	Mean	0.0100	0.1481	0.8011	0.0013	0.0229	0.6421
	StD	0.0024	0.0216	0.0283	0.0018	0.0311	0.0390
	ASE	0.0022	0.0207	0.0258	0.0031	0.0269	0.0369
5000	Mean	0.0102	0.1515	0.7991	0.0013	0.0275	0.6408
	StD	0.0018	0.0163	0.0198	0.0027	0.0275	0.0274
	ASE	0.0017	0.0167	0.0189	0.0019	0.0279	0.0226

Table 3. QMLE and WLSE results for 1000 RC-GARCH(1,1) series with sample size n, $\omega_t \sim \Gamma\left(1, \frac{1}{\omega_0}\right)$, $\alpha_t \sim \Gamma\left(1, \frac{1}{\alpha_0}\right)$, and $\beta_t \sim \Gamma(1, \frac{1}{\beta_0})$.

For each instance, the mean, StDs (standard-deviations), and ASEs (asymptotic standard errors) of estimates over the 1000 sample-paths are shown in Tables 1-3. A few conclusions can be drawn. Firstly, the true values of the parameters are well estimated, given their smaller ASEs, which are quite close to their StDs, especially for the QMLE part. Secondly, the results overall confirm the asymptotic theory of Section 3 (Theorems 3.1-3.2). Indeed, the larger the sample size, the more accurate the estimate is in terms of bias and standard errors. Thirdly, the QMLE gives slightly more accurate results, especially in terms of bias, StDs, and ASEs.

5 Empirical data

5.1 Intel stock returns

In our first empirical application, we fit the RC-GARCH(1,1) model to the daily returns of the Intel stock ranging from 12/15/72 to 12/31/08. In total, we have $n = 9097$ observations. The series, taken from Tsay (2010), exhibits conventional stylized facts of stock return series, such as dependence without correlation, high persistence, and volatility clustering (see Figure 2).

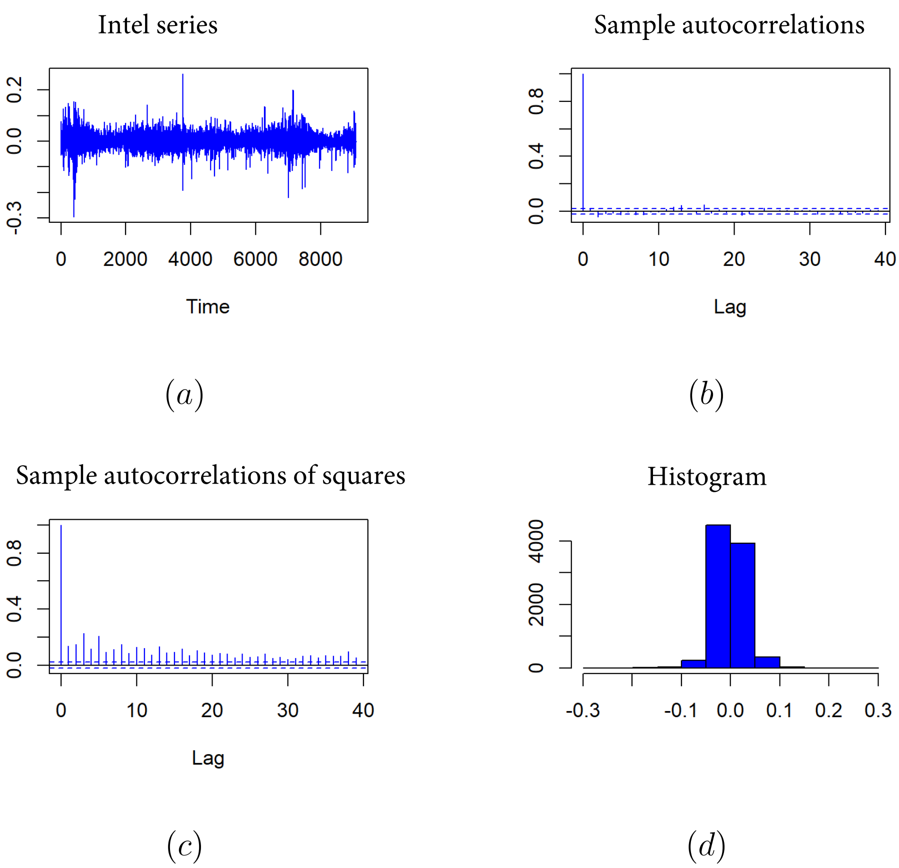


Figure 2: Intel stock return series: (a) The series, (b) sample autocorrelation, (c) sample autocorrelation of squares, (d) histogram.

Applying the first two stages of Algorithm 3.1, we obtain the estimated RC-GARCH(1,1) model given by Table 4, that displays the estimated means and variances of the random coefficients as well as their asymptotic standard errors (ASE) in parentheses.

	ω_t	α_{1t}	β_{1t}		
QMLE	$\widehat{\omega}_n$	$\widehat{\alpha}_{1n}$	$\widehat{\beta}_{1n}$	$\widehat{\alpha}_{1n}+\widehat{\beta}_{1n}$	
	7.4e-06 (1.9e-06)	0.0520 (0.0069)	0.9397 (0.0071)	0.9918	
WLSE	$\widehat{\sigma}^2_{\omega n}$	$\widehat{\sigma}^2_{\alpha n}$	$\widehat{\sigma}^2_{\beta n}$		FMC
	5.7e-08 (1.1e-07)	0.0255 (0.0177)	0.6447 (0.4031)		1.710

Table 4. QML and WLS estimates for the RC-GARCH$(1,1)$;Intel series.

The parameter estimate $\widehat{\alpha}_{1n}+\widehat{\beta}_{1n} \simeq 0.9918$ indicates a strong persistence, while the estimated RC-GARCH model remains strictly stationary with a finite second moment. In addition, the estimated indicator of the fourth moment condition,

$$FMC := 3\widehat{\alpha}_{1n}^2 + 2\widehat{\alpha}_{1n}\widehat{\beta}_{1n} + \widehat{\beta}_{1n}^2 + \kappa\widehat{\sigma}^2_{0\alpha_1} + \sigma^2_{0\beta_1} \simeq 1.710,$$

is larger than one, so the estimated RC-GARCH model has an infinite fourth moment and hence an infinite unconditional kurtosis. Nevertheless, the conditional (excess) kurtosis $\widehat{\kappa}_t - 3$ (see (2.9)) is finite and its estimated values are plotted in Figure 3 (Panel (d)). In the same figure, we plotted the observed (predictive) conditional volatility $\widehat{\delta}_t^2$ (Figure 3 (a)), which is nothing but the volatility of the standard GARCH model, and the filtered volatility $\widehat{\sigma}_t^2$ (Figure 3 (b)) obtained from Stage 3 of the Algorithm 3.1. It can be seen that the filtered volatility $\widehat{\sigma}_t^2$ is more erratic than the predictive volatility $\widehat{\delta}_t^2$ and that the latter constitute an envelope of the former. In addition, as mentioned in the introduction, the filtering volatility $\widehat{\sigma}_t^2$ does not seem to contain curves, in the sense of large volatilities, as does the predictive (or standard GARCH) volatility $\widehat{\delta}_t^2$.

Thus, with the two estimated volatilities, we have a better and more insightful picture of the evolution of the variability of the series. Note also that the normalized residuals $\widehat{\varepsilon} := \frac{Y_t}{\widehat{\sigma}_t}$ (see Figure 3 (c)) look like an independent noise (see also the Figures in the supplementary material) as the sample autocorrelations of residuals and their squares do not show any significant spikes. Finally, the estimated conditional excess kurtosis $\widehat{\kappa}_t - 3$ (Figure 3 (d)) seems to be in accordance with the estimated predictive volatility of the model.

5.2 Cisco stock returns

The second application concerns the daily returns of Cisco stock (Tsay, 2010) for the period from 01/02/2001 to 12/31/2008 involving $n = 2011$ observations (see Figure 4). The parameter estimates

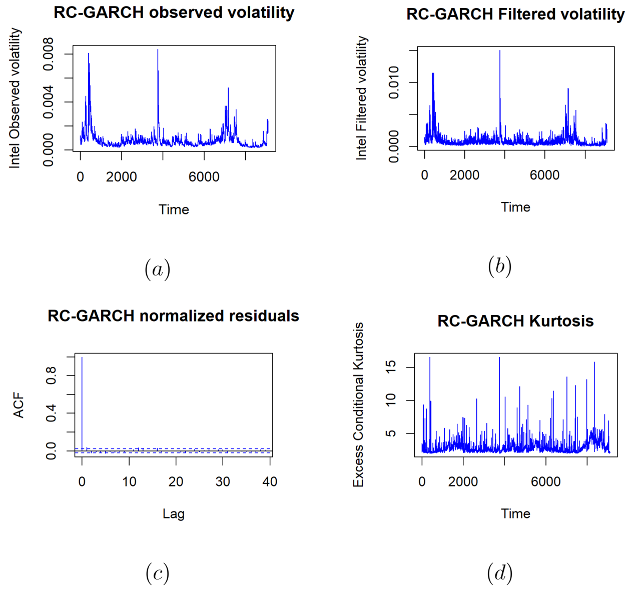


Figure 3: Estimated RC-GARCH for the Intel series. (a) Predictive volatility, (b) smoothed volatility, (c) ACF of residuals, (d) conditional excess kurtosis.

are reported in Table 5. Similar conclusions as in the previous application can be drawn: the estimated model is highly persistent, has a finite second moment and an infinite fourth moment.

	ω_t	α_{1t}	β_{1t}		
QMLE	$\widehat{\omega}_n$	$\widehat{\alpha}_{1n}$	$\widehat{\beta}_{1n}$	$\widehat{\alpha}_{1n} + \widehat{\beta}_{1n}$	
	3.2e-06 (1.8e-06)	0.0341 (0.0077)	0.9609 (0.0082)	0.9950	
WLSE	$\widehat{\sigma}^2_{\omega n}$	$\widehat{\sigma}^2_{\alpha n}$	$\widehat{\sigma}^2_{\beta n}$		FMC
	5.6e-08 (1.1e-07)	0.1229 (0.0101)	1.3650 (0.9704)		2.7260

Table 5. QML and WLS estimates for the RC-GARCH$(1, 1)$; Cisco series.

Finally, the predictive and smoothed volatilities are plotted in Figure 5 (panel (a) and panel (b), respectively). The smoothed volatility is more erratic and captures small and large volatilities better than does the predictive (observed) volatility. The conditional excess kurtosis (panel (d)) exceptionally shows very large picks, which are probably due to the fourth-order instability of the model.

6 Conclusion

This paper proposed a random coefficient GARCH model with time-varying conditional kurtosis and two conditional volatility sequences, one observed, which is driven by past observations and one latent, which is driven by past and present random inputs. Our formulation, which is path-independent, mimics the Markov switching specification of Gray (1996) and is different from earlier random coefficient GARCH models introduced by Kazakevicius et *al* (2004), Klivecka (2004) and Thavaneswaran

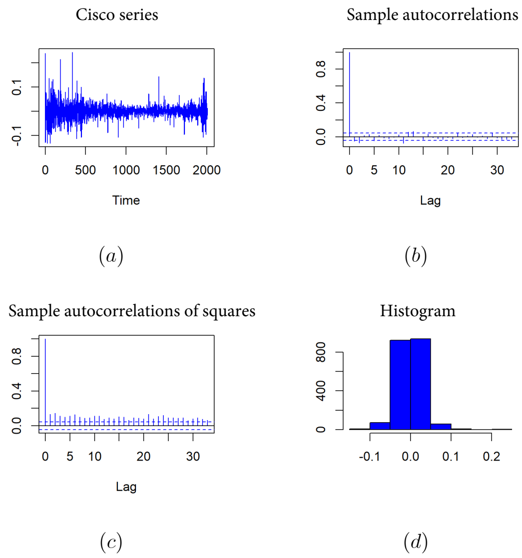


Figure 4: Cisco stock return series: (a) The series; (b) sample autocorrelation, (c) sample autocorrelation of squares, (d) histogram.

et *al* (2005). The observable/predictive volatility of our model is the same as that of the standard GARCH model and represents the conditional mean of the latent volatility, which is the main interest of this article. The proposed model equipped with two volatility equations can shed more light on the evolution of the variability of the underlying series.

Regarding estimation, the QMLE for the means of the random coefficients is consistent and asymptotically normal (CAN) with a different covariance matrix than the QMLE of the standard GARCH. In addition, the WLSE for the variances of the random coefficients is also CAN and is given in closed form regardless of the distribution of the model. Finally, the latent volatility filtering/smoothing is based on the assumption that the conditional model is NIG distributed with IG distributed random coefficients and normal innovations. The NIG hypothesis allows for closed-form posteriors, is very flexible, and can account for heavy tailedness and asymmetry.

Further extensions of this paper are possible. First of all, the asymmetry parameter was set to zero although it could be considered as an unknown parameter to be estimated. Also, alternative estimation methods could be used, such as the Bayesian approach or the EM algorithm (Karlis, 2002; Aknouche et al, 2022a). Finally, other random-coefficient GARCH models could be considered, such as the random coefficient EGARCH and the random coefficient score-driven model. These aspects of analysis could be analyzed in a future research agenda.

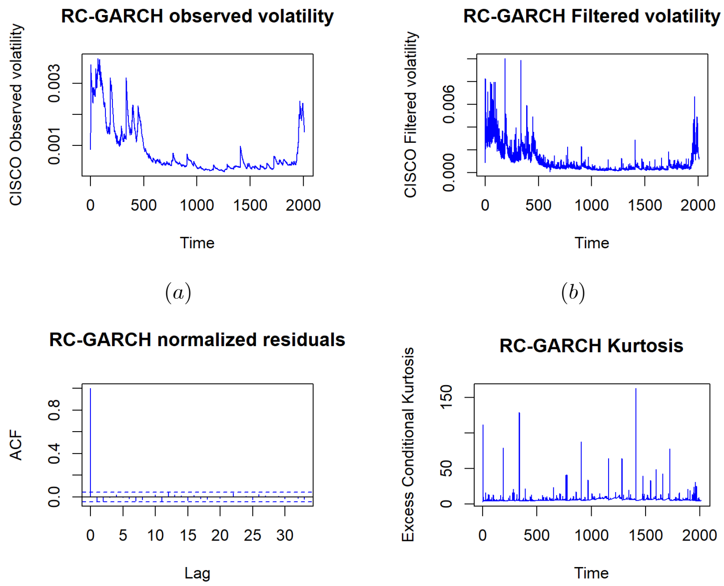


Figure 5: Estimated RC-GARCH for Cisco series. (a) Predictive volatility, (b) smoothed volatility, (c) ACF of residuals, (d) conditional excess kurtosis.

Appendix

Appendix A: Inverse Gaussian and Normal Inverse Gaussian distributions

A continuous random variable Z is said to have an Inverse Gaussian (IG) distribution with mean $\rho > 0$ and shape $\lambda > 0$ ($Z \sim \mathcal{IG}(\rho, \lambda)$) if its probability density function is given by

$$f(z; \rho, \lambda) = \frac{\sqrt{\lambda}}{\sqrt{2\pi z^3}} \exp\left(-\frac{\lambda (z-\rho)^2}{2\rho z}\right), \; z > 0. \tag{A.1}$$

An equivalent form is given in terms of the mean ρ and the dispersion (1/shape) $\phi = \frac{1}{\lambda}$. Another reparametrization is given in terms of $\rho = \frac{\xi}{\psi}$ and $\lambda = \xi^2$ so $\xi = \sqrt{\lambda}$ and $\psi = \frac{\sqrt{\lambda}}{\rho}$ (e.g. Barndorff-Nielsen, 1978). The mean and variance of the IG distribution are $E(X) = \rho$ and $Var(X) = \frac{\rho^3}{\lambda}$. The IG distribution is linear in the sense that if $Z_1 \sim \mathcal{IG}(\rho_1, \lambda_1)$ and $Z_2 \sim \mathcal{IG}(\rho_2, \lambda_2)$ are independent then $aZ_1 + bZ_2 \sim \mathcal{IG}(a\rho_1 + b\rho_2, a\lambda_1 + b\lambda_2)$ $(a, b > 0)$.

A continuous mixture of normal distributions with Inverse Gaussian mixings leads to the Normal Inverse Gaussian (NIG) distribution that has a closed form. A continuous random variable Y is said to have a NIG distribution with parameters α, ρ, μ, β ($Y \sim \mathcal{NIG}(\alpha, \beta, \rho, \mu)$, $\alpha, \rho > 0$, $|\beta| \leq \alpha$, $\mu \in \mathbb{R}$) if its probability density function is given by (Barndorff-Nielsen, 1978)

$$f(y; \alpha, \beta, \rho, \mu) = \frac{\alpha \rho K_1\left(\alpha \sqrt{\rho^2 + (y-\mu)^2}\right)}{\pi \sqrt{\rho^2 + (y-\mu)^2}} \exp(\rho \sqrt{\alpha^2 - \beta^2} + \beta (y - \mu)) \tag{A.2}$$

where K_1 is the modified Bessel function of the third kind of order one. In terms of the hierarchical mixture form, the $\mathcal{NIG}$ distribution is defined as follows (Barndorff-Nielsen, 1997; Karlis, 2002; Murphy, 2007)

$$\begin{cases} Z|\rho, \alpha, \beta \sim \mathcal{IG}\left(\rho, \sqrt{\alpha^2 - \beta^2}\right) \\ Y|Z, \mu, \beta \sim \mathcal{N}(\mu + \beta Z, Z) \end{cases} \Longrightarrow Y \sim NIG(\alpha, \beta, \rho, \mu).$$

In particular, when $\beta = 0$ the hierarchical form of the $\mathcal{NIG}(\alpha, 0, \rho, \mu)$ distribution becomes

$$\begin{cases} Z|\alpha, \rho \sim \mathcal{IG}(\rho, \alpha) \\ Y|Z, \mu \sim \mathcal{N}(\mu, Z) \end{cases} \Longrightarrow Y \sim \mathcal{NIG}(\alpha, 0, \rho, \mu). \tag{A.3}$$

The mean and variance of the NIG variable are $E(Y) = \mu + \frac{\rho\beta}{\sqrt{\alpha^2 - \beta^2}}$ and $Var(Y) = \frac{\rho\alpha^2}{\sqrt{(\alpha^2 - \beta^2)^3}}$. The NIG distribution is closed under affine transformations: If $Y \sim \mathcal{NIG}(\alpha, \beta, \rho, \mu)$ then (Paolella, 2007) $aY + b \sim \mathcal{NIG}\left(\frac{\alpha}{|a|}, \frac{\beta}{a}, |a|\rho, a\mu + b\right)$. The main advantages of the NIG distribution over the normal distribution is that it allows for asymmetry (with parameter β) and heavy tailedness (with parameter α); see Barndorff-Nielsen, (1997). Note that μ is a location parameter while ρ is a scale parameter. Another advantage of the NIG distribution is that the posterior mean of the IG distribution $E(Z|Y)$

can be obtained through $(A.3)$ in a closed form (cf. Karlis, 2002, formula (4)) as follows

$$E\left(Z|Y\right)=\frac{\rho\sqrt{1+\left(\frac{Y-\mu}{\rho}\right)^2}}{\alpha}\frac{K_0\left(\alpha\rho\sqrt{1+\left(\frac{Y-\mu}{\rho}\right)^2}\right)}{K_1\left(\alpha\rho\sqrt{1+\left(\frac{Y-\mu}{\rho}\right)^2}\right)} \tag{A.4}$$

where $K_r\left(y\right)$ denotes the modified Bessel function of the third kind of order r evaluated at y.

In R, we use the function dnig() of the package *fBasics* for the density of the NIG distribution. Moreover, to evaluate $(A.4)$, we use the function besselK of the package *base*.

Appendix B Proofs

Proof of Propositions 2.1-2.3 The proofs of Propositions 2.1-2.3 are standard and follow the same lines of the stability proofs for GARCH models (see e.g. Francq and Zakoian, 2019). Hence, they are omitted but they are available upon request.

Proof of Theorem 3.1 The proof is similar to that of QMLE's consistency and asymptotic normality for the GARCH model (Francq and Zakoian, 2004-2019). So, only the relevant steps of the proof are provided. Define $L_n\left(\theta\right)$ and ℓ_t as $\widetilde{L}_n\left(\theta\right)$ and $\widetilde{\ell}$ in (3.5) while substituting $\widetilde{\delta}_t^2\left(\theta\right)$ in (3.3) by $\delta_t^2\left(\theta\right)$ given by (3.1). Concerning the consistency result (3.7), the following intermediary lemmas are proved under **A1-A4** in the same way as in Francq and Zakoian (2004).

a) $\lim_{n\to\infty}\sup_{\theta\in\Theta}\left|\widetilde{L}_n\left(\theta\right)-L_n\left(\theta\right)\right|=0\quad a.s.$

b) $E\left(\ell_t\left(\theta_0\right)\right)<\infty$*,* $E\left(\ell_t\left(\theta\right)\right)$ *is minimized at* $\theta=\theta_0$*, and* $E\left(\ell_t\left(\theta_0\right)\right)=E\left(\ell_t\left(\theta\right)\right)\Rightarrow\theta=\theta_0$.

c) *For any* $\theta\neq\theta_0$*, there is a neighborhood* $\mathcal{V}\left(\theta\right)$ *so that*

$$\limsup_{n\to\infty}\inf_{\theta^*\in\mathcal{V}(\theta)}\widetilde{L}_n\left(\theta^*\right)>\liminf_{n\to\infty}\widetilde{L}_n\left(\theta_0\right)\quad a.s.$$

The proof of the asymptotic normality result (3.8) can be split into the following lemmas.

d) $\sqrt{n}\sup_{\theta\in\Theta}\left\|\frac{\partial\widetilde{L}_n(\theta)}{\partial\theta}-\frac{\partial L_n(\theta)}{\partial\theta}\right\|\underset{n\to\infty}{\overset{a.s.}{\to}}0.$

e) $\sqrt{n}\frac{\partial L_n(\theta_0)}{\partial\theta}\underset{n\to\infty}{\overset{D}{\to}}\mathcal{N}\left(0,I\right).$

f) $\frac{\partial^2L_n(\theta^*)}{\partial\theta\partial\theta'}\underset{n\to\infty}{\overset{a.s.}{\to}}J$*, where* θ^* *is between* $\widehat{\theta}_n$ *and* θ_0.

Result d) is proved in the same way as in Francq and Zakoian (2004). So only e) and f) are established.

Regarding e), the sequence $\left\{\sqrt{n}\frac{\partial L_n(\theta_0)}{\partial\theta},t\in\mathbb{Z}\right\}$ is a square-integrable martingale with respect to $\{\mathscr{F}_t,t\in\mathbb{Z}\}$ with

$$n^{1/2}\frac{\partial L_n(\theta_0)}{\partial\theta}=n^{-1/2}\sum_{t=1}^{n}\left(1-\frac{Y_t^2}{\delta_t^2(\theta_0)}\right)\frac{1}{\delta_t^2(\theta)}\frac{\partial\delta_t^2(\theta_0)}{\partial\theta}.$$

Since the ergodic theorem under (2.11) entails

$$\sum_{t=1}^{n} n^{-1}\left(1-\frac{Y_t^2}{\delta_t^2(\theta_0)}\right)^2 \frac{1}{\delta_t^4(\theta)}\frac{\partial\delta_t^2(\theta_0)}{\partial\theta}\frac{\partial\delta_t^2(\theta_0)}{\partial\theta'}$$
$$\underset{n\to\infty}{\overset{a.s.}{\to}} E\left(\frac{1}{\delta_t^4(\theta_0)}E\left(\left(1-\frac{Y_t^2}{\delta_t^2(\theta_0)}\right)^2 | \mathscr{F}_{t-1}^Y\right)\frac{\partial\delta_t^2(\theta_0)}{\partial\theta}\frac{\partial\delta_t^2(\theta_0)}{\partial\theta'}\right) \quad \text{(A.5)}$$

where

$$\begin{aligned} E\left(\left(1-\frac{Y_t^2}{\delta_t^2(\theta_0)}\right)^2 | \mathscr{F}_{t-1}^Y\right) &= \frac{1}{\delta_t^4(\theta_0)}E\left(\left(\delta_t^2-\sigma_t^2\varepsilon_t^2\right)^2 | \mathscr{F}_{t-1}^Y\right) \\ &= \frac{Var\left(Y_t^2 | \mathscr{F}_{t-1}^Y\right)}{\delta_t^4(\theta_0)} \\ &= \frac{1}{\delta_t^4(\theta_0)}\left((\kappa-1)\,\delta_t^4+\kappa Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right)\right) \quad \text{(A.6)} \end{aligned}$$

the result e) thus follows from $(A.5)$, $(A.6)$, and the central limit theorem for square-integrable martingales (e.g. Billingsley, 2008; Francq and Zakoian, 2019).

To prove f), the Taylor expansion of the criterion (3.5) at θ_0, the almost convergence of $\widehat{\theta}_n$ to θ_0, and the ergodic theorem yield

$$\begin{aligned} n^{-1}\sum_{t=1}^{n}\frac{\partial^2\ell_t\left(\theta_{ij}^*\right)}{\partial\theta_i\partial\theta_j} &= n^{-1}\sum_{t=1}^{n}\frac{\partial^2\ell_t(\theta_0)}{\partial\theta_i\partial\theta_j}+o_{a.s.}(1) \underset{n\to\infty}{\overset{a.s.}{\to}} E\left(\frac{\partial^2\ell_t(\theta_0)}{\partial\theta_i\partial\theta_j}\right) \\ &= E\left(\left(1-\frac{\sigma_t^2\varepsilon_t^2}{\delta_t^2(\theta_0)}\right)\frac{1}{\delta_t^2(\theta_0)}\frac{\partial^2\delta_t^2(\theta_0)}{\partial\theta\partial\theta'}\right) \\ &+ E\left(\left(\frac{2Y_t^2}{\delta_t^2(\theta_0)}-1\right)\frac{1}{\delta_t^2(\theta_0)}\frac{\partial\delta_t^2(\theta_0)}{\partial\theta}\frac{\partial\delta_t^2(\theta_0)}{\partial\theta'}\right) \\ &= J, \end{aligned}$$

which completes the proof. □

Proof of Theorem 3.2 i) We first prove (3.14). Under **A1-A4**, the strong consistency of $\widehat{\theta}_n$ entails $\delta_t^2-\widehat{\delta}_t^2 \underset{t\to\infty}{\overset{a.s.}{\to}} 0$ and hence $\left\|\widehat{M}_t-M_t\right\| \underset{t\to\infty}{\overset{a.s.}{\to}} 0$, where $\|.\|$ denotes the Euclidian norm in $\mathbb{R}^{p+q+1}$. Therefore, a standard argument shows that (3.12) becomes

$$\widehat{\Lambda}_n=\left(\sum_{t=1}^{n}\frac{1}{\delta_t^8}M_tM_t'\right)^{-1}\sum_{t=1}^{n}M_t\frac{\left(Y_t^2-\delta_t^2\right)^2-(\kappa-1)\delta_t^4}{\kappa\delta_t^8}+o_{a.s.}(1),$$

which, in turns, using (3.10), gives

$$\widehat{\Lambda}_n-\Lambda_0=\left(\frac{1}{n}\sum_{t=1}^{n}\frac{1}{\delta_t^8}M_tM_t'\right)^{-1}\frac{1}{n}\sum_{t=1}^{n}M_t\frac{e_t}{\kappa\delta_t^8}+o_{a.s.}(1). \quad \text{(A.7)}$$

Now under (2.11), the ergodic theorem yields

$$\frac{1}{n}\sum_{t=1}^{n}\frac{1}{\delta_t^8}M_tM_t' \underset{n\to\infty}{\overset{a.s.}{\to}} A \tag{A.8}$$

and, further under **A6**,

$$\frac{1}{n}\sum_{t=1}^{n}M_t\frac{e_t}{\kappa\delta_t^8} \underset{n\to\infty}{\overset{a.s.}{\to}} E\left(M_t\frac{e_t}{\kappa\delta_t^8}\right)=E\left(M_t\frac{1}{\kappa\delta_t^8}E\left(e_t|\,\mathscr{F}_{t-1}^Y\right)\right)=0. \tag{A.9}$$

Thus, (3.14) follows from $(A.7)$-$(A.9)$.

ii) To show (3.15), we first rewrite $(A.7)$ as follows

$$\sqrt{n}\left(\widehat{\Lambda}_n-\Lambda_0\right)=\left(\frac{1}{n}\sum_{t=1}^{n}\frac{1}{\delta_t^8}M_tM_t'\right)^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}M_t\frac{e_t}{\kappa\delta_t^8}+o_{a.s.}(1)\,. \tag{A.10}$$

The ergodic theorem shows under **A7** that

$$\begin{aligned}\sum_{t=1}^{n}\left(\frac{1}{\sqrt{n}}M_t\frac{e_t}{\kappa\delta_t^8}\right)\left(\frac{1}{\sqrt{n}}M_t\frac{e_t}{\kappa\delta_t^8}\right)' &= \frac{1}{n\kappa^2}\sum_{t=1}^{n}\frac{e_t^2}{\delta_t^{16}}M_tM_t'\\ &\underset{n\to\infty}{\overset{a.s.}{\to}}\frac{1}{\kappa^2}E\left(\frac{Var\left(\left(Y_t^2-\delta_t^2\right)^2|\,\mathscr{F}_{t-1}^Y\right)}{\delta_t^{16}}M_tM_t'\right).\end{aligned} \tag{A.11}$$

From $(A.11)$ and **A5-A7**, the central limit theorem for square-integrable martingales implies

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}M_t\frac{e_t}{\kappa\delta_t^8}\underset{n\to\infty}{\overset{D}{\to}}\mathcal{N}\left(0,B\right). \tag{A.12}$$

The result (3.15) thus follows from $(A.10)$, $(A.8)$, and $(A.12)$. □

References

Aknouche, A. (2015). Quadratic random coefficient autoregression with linear-in-parameters volatility. *Statistical Inference for Stochastic Processes* **18**, 99-125.

Aknouche, A., Almohaimeed, B. and Dimitrakopoulos, S. (2022a) Forecasting transaction counts with integer-valued GARCH models. *Studies in Nonlinear Dynamics & Econometrics* **26**, 529-539.

Aknouche, A., Almohaimeed, B. and Dimitrakopoulos, S. (2022b). Periodic autoregressive conditional duration. *Journal of Time Series Analysis* **43**, 5–29.

Aknouche, A. and Francq, C. (2021). Count and duration time series with equal conditional stochastic and mean orders. *Econometric Theory* **37**, 248–280.

Aknouche, A. and Francq, C. (2022). Stationarity and ergodicity of Markov-switching positive conditional mean models. *Journal of Time Series Analysis* **43**, 436-459.

Aknouche, A. and Francq, C. (2023). Two-stage weighted least squares estimator of the conditional mean of observation-driven time series models. *Journal of Econometrics* **237**, 105-174.

Aknouche, A. and Rabehi, N. (2010). On an independent and identically distributed mixture bilinear time series model. *Journal of Time Series Analysis* **31**, 113-131.

Ayala, A. and Blazsek, S. (2019). Score-driven currency exchange rate seasonality as applied to the Guatemalan Quetzal/US Dollar. *SERIEs* **10**, 65–92.

Barndorff-Nielsen, O.E. (1978). Hyperbolic Distributions and distributions on Hyperbolae. *Scandinavian Journal of Statistics* **5**, 151-157.

Barndorff-Nielsen, O.E. (1997). Normal Inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* **24**, 1–13.

Barndorff-Nielsen, O.E., Mikosch T. and Resnick, S.I. (2013). *Lévy Processes: Theory and Applications*. Birkhäuser.

Barndorff-Nielsen, O.E. and Prause, K., (2001). Apparent scaling. *Finance and Stochastics* **5**, 103–113.

Billingsley, P. (2008). *Probability and measure*. John Wiley and Sons, second edition.

Blazsek, S., Ho, H.-C. and Liu. S.-P. (2018). Score-Driven Markov-Switching EGARCH Models: An Application to Systematic Risk Analysis. *Applied Economics* **50**, 6047–6060.

Bollerslev, T. (1982). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 7-327.

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

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.

Cox, D.R. (1981). Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics* **8**, 93-115.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**, 987-1007.

Engle, R.F. (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. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* **10**, 605-637.

Francq, C. and Zakoian, J.-M. (2005). The L2-structures of standard and switching-regime GARCH models. *Stochastic Processes and their Applications* **115**, 1557–1582.

Francq, C. and Zakoian, J.-M. (2008). Deriving the autocovariances of powers of Markov-switching GARCH models, with applications to statistical inference. *Computational Statistics and Data Analysis* **52**, 3027–3046.

Francq, C. and Zakoian, J. -M. (2019). *GARCH Models, Structure, Statistical inference and nancial applications.* John Wiley & Sons, Ltd., Publication.

Gray, S. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics* **42**, 27–62.

Haas, M., Mittnik, S. and Paolella, M. (2004). A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics* **2**, 493–530.

Hamilton, J.D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics* **64**, 307–333.

Karlis, D. (2002). An EM type algorithm for maximum likelihood estimation of the normal–inverse Gaussian distribution. *Statistics & Probability Letters* **57**, 43–52.

Kazakevicius, V., Leipus, R. and Viano, M.-C. (2004). Stability of random coefficient ARCH models and aggregation schemes, *Journal of Econometrics* **120**, 139–158.

Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics* **27**, 363–94.

Klivecka, A. (2004). Random coefficient GARCH(1,1) model with i.i.d. coefficients. *Lithuanian Mathematical Journal* **44**, 374–384.

Murphy, K.P. (2007). Conjugate Bayesian analysis of the Gaussian distribution. *Preprint.*

Mozumder, S., Talukdar, B. Kabir M.H. and Li, B. (2024). Non-linear volatility with normal inverse Gaussian innovations: ad-hoc analytic option pricing. *Review of Quantitative Finance and Accounting* **62**, 97–133,

Nicholls, F. and Quinn, B.G. (1982). *Random coefficient autoregressive models*: *An Introduction.* Springer-Verlag, New York.

Paolella, M.S. (2007). *Intermediate Probability: A computational Approach.* John Wiley & Sons.

Rachev, S.T. (2003). *Handbook of heavy tailed distributions in Finance*, Volume 1: Handbooks in Finance, North Holland.

Regis, M., Serra P. and van den Heuvel E.R. (2022). Random autoregressive models: A structured overview. *Econometric Reviews* **41**, 207–230.

Stentoft, L. (2008). American Option Pricing Using GARCH Models and the Normal Inverse Gaussian Distribution. *Journal of Financial Econometrics* **6**, 540–582.

Taylor, S.J. (1982). Financial returns modelled by the product of two stochastic processes- a study of daily sugar prices, 1961-1979. *In "Time Series Analysis: Theory and Practice"*, Anderson O.D. (ed.). North-Holland, Amsterdam, 203-226.

Taylor, S.J. (1986). *Modelling Financial Time Series.* Wiley, Chichester.

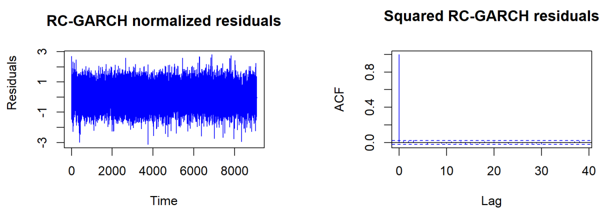
Thavaneswaran, A., Appadoo, S.S. and Samanta, M. (2005). Random coefficient GARCH models. *Mathematical and Computer Modelling* **41**, 723–733.

Tsay, R.S. (2010). *Analysis of financial time series: Financial Econometrics*, 3rd edition, Wiley.

Wee, D.C.H., Chen, F., Dunsmuir, W.T.M. (2022). Likelihood inference for Markov switching GARCH(1,1) models using sequential Monte Carlo. *Econometrics and Statistics*, **21**, 50 – 68.

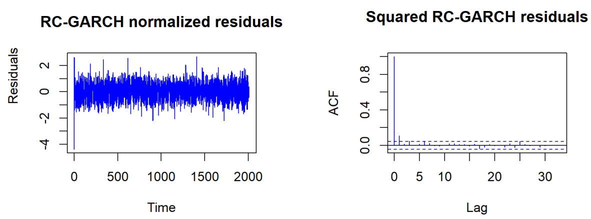
White, H., Kim, T.H. and Manganelli, S. (2010). Modeling Autoregressive Conditional Skewness and Kurtosis with Multi-Quantile CAViaR. *In Volatility and Time Series Econometrics: Essays in Honor of Robert Engle. Oxford University Press.* https://doi.org/10.1093/acprof:oso/9780199549498.003.0012.

Supplementary material



(a) (b)

Figure 6: (a) Residuals, (b) ACF of squared residuals; Intel series.



(a) (b)

Figure 7: (a) Residuals, (b) ACF of squared residuals; Cisco series.