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/predictive 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 predictive volatility. The predictive 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 volatility 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 and of the random coefficients. 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. It is shown that the filtered volatility can improve both in-sample and out-of-sample forecasts of the predictive volatility, even when the future observations are unknown and are replaced by their predictions.

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

[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.

MSCS 62P20, 62M10, 91B84

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 (cf. Francq and Zakoian, 2019). 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*a*) 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. Moreover, GARCH volatility is synonymous with the concept of conditional variance given past information, which makes it easily interpretable. However, 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; Smetanina, 2017). Moreover, GARCH volatility, by construction, does not incorporate contemporaneous/current information in it, rendering it unable to adapt quickly to new events (e.g. Breitung and Hafner, 2016). This can also create a kind of distortion towards large variabilities, as it tends to ignore small volatilities; see Figure 1. Thus, the standard GARCH model of Engle-Bollerslev could potentially capture the actual volatility path of a series even better, if the current available information could be integrated into the volatility equation. In particular, current information could be used in ex-post volatility improving out-of-sample volatility forecasting (e.g. Zhang and Zhao, 2023).

On the contrary, SV-type models are able to integrate present shocks in the volatility equation, making it latent/unobserved. The main motivation of using latent/stochastic volatility comes from the idea that the arrival of information is random and unobserved (Ghysels et al., 1996). However, estimation of the SV models is non-trivial. In addition, the volatility does not depend on past observations, which can also be restrictive. In some variants of the SV model, the volatility is no longer the conditional variance, which can lead to misinterpretation issues. For example, in the stochastic volatility model with leverage in which the regression error and the volatility error are correlated (e.g. Jacquier et al, 2004), the scale volatility process cannot be the conditional variance (Yu, 2005). In this case, a comparison between the predictive performance of this type of volatility, which is just a scale factor in a multiplicative error structure, and that of a conditional-variance-volatility, would be unfair, if not biased.

A special SV model in which the two error terms are fully dependent was investigated by Breitung and Hafner (2016) with the aim of combining the advantages of both the SV and GARCH models. Their model was named Now-Casting volatility (henceforth NC-GARCH). Smetanina (2017) proposed a real-time GARCH (RT-GARCH) model in which the squared model innovation is incorporated into the GARCH volatility equation. In both the NC-GARCH and RT-GARCH models, the full dependence of the two error terms allows for easy parameter estimation using the standard QMLE, which is a significant statistical advantage over the standard SV model. Moreover, both models incorporate current information into the (so-called) volatility, which is also another important addition to the standard GARCH model. Finally, the conditional kurtosis in these models is time-varying, which is yet another advancement. However, the full dependence between the two innovation terms in the NC-GARCH and RT-GARCH models induces some interpretability issues. Firstly, the contemporaneous volatility is no longer the conditional variance but just a scale factor in the multiplicative form assumed by the two models. Secondly, with respect to the actual conditional variance, both the NC-GARCH and RT-GARCH models can be reformulated within the class of (exponential) GARCH models with a non-multiplicative structure (see Section 2 and Supplementary Material). Both models are therefore observation-driven and this is the reason for which the QMLE is obtained easily, as is the case with the Engle-Bollerslev GARCH model. On the other hand, comparison of the contemporaneous volatility in these models with that of other classical GARCH type models is not fair, since in the former the volatility is not the conditional variance. Finally, the volatility of volatility induced by real-time GARCH models (cf. Ding, 2023; Wu et al, 2023) is not properly the conditional variance of conditional variance.

MS-GARCH models can overcome the limitations of the GARCH and SV 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. In addition, the volatility in the MS-GARCH models is always the conditional variance given past information and present regime. Nonetheless, MS-GARCH models (or some classes of them) still have some limitations. First, the estimation and prediction of MS-GARCH models that are characterized by path dependence (Francq and Zakoian, 2008-2019; Aknouche and Francq, 2022; Wee et al, 2022) is generally not straightforward. Yet, there exist non-path dependent MS-GARCH models (Gray, 1996; Klaassen, 2002; Haas et al, 2004a-2004b), whose estimation is relatively easy to perform compared to path-dependent MS-GARCH models. Second, the regime sequence on which the parameters depend is generally discrete-valued and even finite. 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, although the MS-GARCH of Haas et al (2004a) could be extended to have this property.

Another competing class of regime switching GARCH models but with deterministic switching is that of time-varying GARCH models (tvGARCH) in which the parameters vary deterministically over time. There are in fact many versions of the tvGARCH model such as that of Dahlhauss and Subba Rao (2006) and Roan and Ramanathan (2013), and that of Amado and Terasvirta (2013, 2017); see also Campos-Martins and Sucarrat (2024). The resulting tvGARCH models are nonstationary but locally stationary in the sense of Dahlhauss (1997). The estimation of the tvGARCH models is relatively easy to perform using M-estimation based methods (Dahlhauss and Subba Rao, 2006; Amado and Teräsvirta, 2013) and/or non-parametric local polynomial estimation (Rohan and Ramanathan, 2013), as these models belong to the class of observation-driven models. However, as is the case with any observation-driven volatility model, the present information is not integrated into the volatility equation.

Our paper aspires to contribute to the class of MS-GARCH models. In particular, we propose a multi-regime-variable random-coefficient GARCH (RC-GARCH) model that has two volatility sequences and a time-varying conditional kurtosis. The first volatility is the observable/predictive conditional variance sequence, which is nothing but the volatility equation of the standard GARCH model. So, the predictive 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 specific regime, the parameters in the unobserved volatility equation we propose evolve solely (Nicholls and Quinn, 1982; Regis et al, 2022), so that each coefficient has its own (continuous-valued) regime switching mechanism. The predictive 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 predictive volatility equation.

Most importantly, in the proposed RC-GARCH representation, the latent volatility is more heavy-tailed than the predictive 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 and can be better estimated with the filtered volatility, which is its conditional mean given past and present information. In contrast with the predictive volatility, the filtered volatility allows us to integrate the current observation and could improve the forecasting ability of the predictive volatility as shown in Section 4 as well as in our simulations given in an online material. Finally, the volatility of volatility generated by the RC-GARCH model is indeed the conditional variance of conditional variance.

For the estimation of the model parameters, we develop a three-stage method, where asymptotic properties of the first two stages 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 in a closed form 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 and Francq, 2023). Assuming the random coefficients to be Inverse Gaussian (IG) distributed and the innovation to be normally distributed, the unobserved volatility is estimated in the final stage through the filtered volatility, which is the posterior mean of the IG distribution. This filtered volatility has a closed form expression due to the fact that the conditional distribution of the model is Normal Inverse Gaussian (NIG). Such a distribution, which is a (continuous-valued) Gaussian mixture with IG mixings, 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 studies its stability properties. Section 3 presents the proposed estimation approach. In particular, a test for randomness of the coefficients is proposed. Section 4 illustrates the methodology with a real dataset, respectively, where it is shown that the filtered volatility can improve the forecasting ability of the predictive volatility. Section 5 concludes.

The main proofs of this paper are displayed in an Online Supplementary Material. It also contains various simulated examples, a procedure for predicting the latent volatility using the filtered volatility, either when future observations are available or predicted, a second empirical application, and an account of a general class of volatility models that encompasses all competing volatility models in relation to the RC-GARCH model, namely the MS-GARCH, the time-varying GARCH (tvGARCH), the real-time GARCH, and the Now-Casting GARCH models.

2 The proposed econometric specification

2.1 Noising the GARCH volatility: Some preliminaries

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

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

where $\{\eta_t, t \in \mathbb{Z}\}$ is an iid (independent and identically distributed (iid)) sequence of real-valued variables with mean 0 and variance 1, and the volatility coefficients satisfy $\omega_0 > 0$, $\alpha_{0i} \geq 0$ for all $i \in \{1, 2, ..., q\}$ and $\beta_{0j} \geq 0$ for all $j \in \{1, 2, ..., j\}$. Assume that $\{\eta_t\}$ 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 observable, given the true parameter $\theta_0 = (\omega_0, \alpha_{01}, ...\alpha_{0q}, \beta_{01}, ...\beta_{0p})'$ whereas the volatility (σ_t^2) is unobservable (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 $E\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) = \delta_t^2$, the sequence $\left(\delta_t^2\right)$ can also be called predictive volatility and is an estimate of σ_t^2. Another estimate of σ_t^2 is the filtered volatility $\varrho_t^2 := E\left(\sigma_t^2 | \mathscr{F}_t^Y\right)$ which depends on the past and present observations.

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 expression (2.5) below) along with the path of predictive (Panel b), latent (Panel c) and filtered (Panel d) volatilities. In panel (b), we have annotated the predictive volatility plot with artificial red curves. These curves are essentially created by large volatilities (distorted in the direction of the green arrows), masking medium and small volatilities. Such a feature does not appear in the plots for the unobserved and also the filtered 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, 2022).

As in the MS-GARCH models, the 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 σ_t^2 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

$$\begin{cases} Y_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \omega_{0t} + \sum\limits_{i=1}^{q} \alpha_{0it} Y_{t-i}^2 + \sum\limits_{j=1}^{p} \beta_{0jt} \delta_{t-j}^2 \\ \delta_t^2 = \omega_0 + \sum\limits_{i=1}^{q} \alpha_{0i} Y_{t-i}^2 + \sum\limits_{j=1}^{p} \beta_{0j} \delta_{t-j}^2 \end{cases} \tag{2.4}$$

in which the random coefficients $\omega_{0t} = \xi_t^2 \omega_0$, $\alpha_{0it} = \xi_t^2 \alpha_{0i}$, and $\beta_{0jt} = \xi_t^2 \beta_{0j}$ 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 with a single regime sequence.

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 (regime switching vector), where each coefficient has its own distribution. In conventional MS-GARCH models, all regimes are governed by the same Markov mechanism. In addition, the proposed model has a stochastic time-varying kurtosis that can be estimated from the data, unlike the standard GARCH model, in which the conditional kurtosis is constant. Finally, the parameters of this 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.

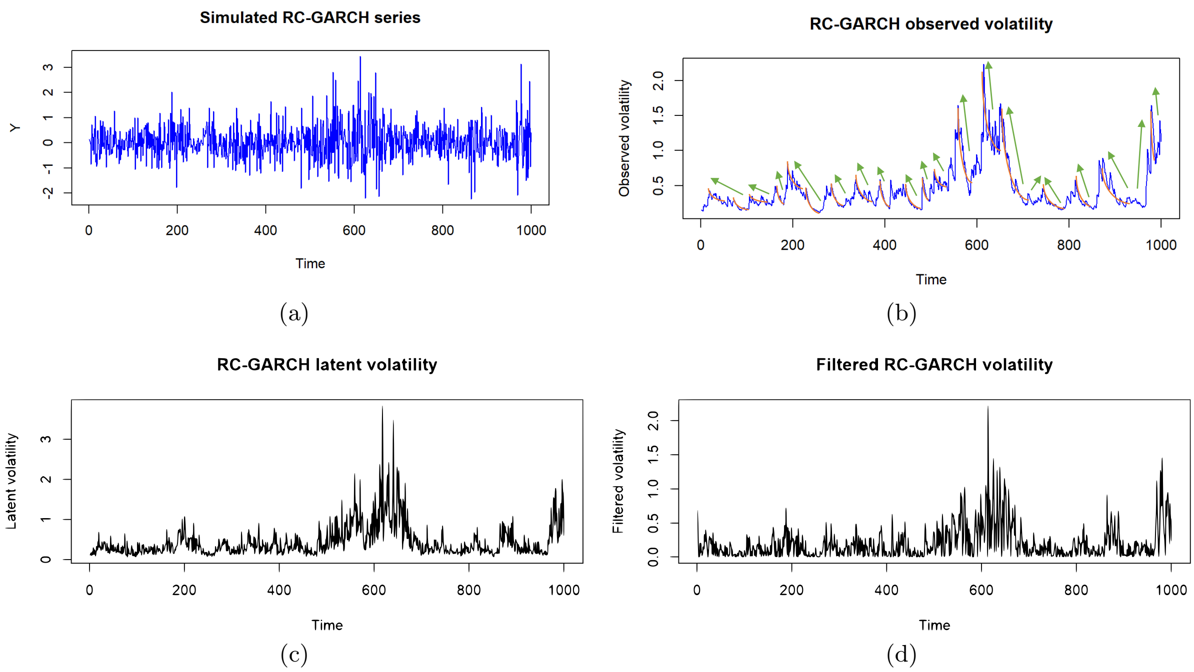


Figure 1: Simulated RC-GARCH series with $\omega_0 = 0.01, \alpha_0 = 0.1$, and $\beta_0 = 0.85$.

2.2 The RC-GARCH model

Let $\{\varepsilon_t, t \in \mathbb{Z}\}$ be an iid sequence of real-valued (random) variables with mean 0, variance 1, and $E\left(\varepsilon_t^4\right) := \kappa < \infty$ [1]. In addition, let $\{\omega_t, t \in \mathbb{Z}\}$, $\{\alpha_{it}, t \in \mathbb{Z}\}$ $(i = 1, ..., q)$, and $\{\beta_{it}, t \in \mathbb{Z}\}$ $(j = 1, ..., p)$ be iid sequences of non-negative random variables with means $\omega_0 > 0$, $\alpha_{0i} \geq 0$ and $\beta_{0j} \geq 0$, and

[1] If $E\left(\varepsilon_t^4\right) < \infty$, then $E\left(\varepsilon_t^2\right) = 1$ implies that $E\left(\varepsilon_t^4\right) > 0$. Assume that $E\left(\varepsilon_t^2\right) = 1$. Assume, in contradiction, that $E\left(\varepsilon_t^4\right) = 0$. Then, $\varepsilon_t = 0$ a.s. which implies that $E\left(\varepsilon_t^2\right) = 0$. Hence, $E\left(\varepsilon_t^4\right) > 0$.

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

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

As emphasized by the Co-Editor, an important question about the model (2.5) is the uniqueness of σ_t^2. Indeed, as in (2.2), assuming that $\varepsilon_t = \xi_t \varepsilon_t^*$, where ξ_t and ε_t^* are independent, model (2.5) can be rewritten as

$$\begin{aligned} Y_t &= \sigma_t^* \varepsilon_t^* \\ \sigma_t^{*2} &= \omega_t^* + \sum_{i=1}^{q} \alpha_{it}^* Y_{t-i}^2 + \sum_{j=1}^{p} \beta_{jt}^* \delta_{t-j}^2 \end{aligned}$$

where $\sigma_t^* = \sigma_t \xi_t$, $\omega_t^* = \xi_t \omega_t$, $\alpha_{it}^* = \xi_t \alpha_{it}$, and $\beta_{jt}^* = \xi_t \beta_{jt}$. However, in the latter representation the coefficients ω_t^*, α_{it}^*, and β_{jt}^* are no longer mutually independent since they all depend on ξ_t. So the uniqueness of model (2.5) is ensured by the mutual independence of its coefficients. Let $\mathscr{F}_{t-1}^{Y,\phi} = \sigma\left\{\left(Y_{t-u}, \phi'_{t-u+1}\right)', u \geq 1\right\}$ be the *complete* σ-algebra generated by the past observations up to time $t-1$ and the past and present 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

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

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

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

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

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

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

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

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

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

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 Engle-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 with the standard Markov-Switching GARCH (MS-GARCH) models (Haas et al, 2004a; Francq and Zakoian, 2005-2008; Aknouche and Francq, 2022). Moreover, the RC-GARCH model, being a general mixture GARCH model with not necessarily finite-valued mixings, is therefore related to the finite mixture GARCH model of Haas et al (2004b) which, in turn, is a particular case of the MS-GARCH of Haas et al (2004a).

In addition, a Markov structure could be assumed for these random coefficients, but this makes the RC-GARCH model more complex and many simple and closed-form formula for the RC-GARCH model (2.5) are lost. Note finally that the proposed RC-GARCH model (2.5) is not a 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 lagged latent volatility σ_{t-j}^2 is replaced by its conditional mean $\delta_{t-j}^2 = E\left(\sigma_{t-j}^2 \middle| \mathscr{F}_{t-j-1}^Y\right)$.

2.3 Now-casting and real-time GARCH models

As highlighted in the introduction, our main motivation in this paper is to introduce a GARCH model whose volatility also depends on current information. There are, actually, two classes of volatility models achieving this goal, namely the Now-Casting GARCH (NC-GARCH) model of Breitung and Hafner (2016) and the real-time GARCH (RT-GARCH) model of Smetanina (2017). We, thus, give in this subsection a brief description of these two models and the main difference between them and the RC-GARCH model we propose.

In its simplest form, the RT-GARCH model of Smetanina (2017) is given by (we consider the RT-GARCH(1,1) for simplicity)

$$\begin{aligned} Y_t &= \lambda_t \eta_t \\ \lambda_t^2 &= \omega + \alpha Y_{t-1}^2 + \beta \lambda_{t-1}^2 + \varphi \eta_t^2 \end{aligned}$$

where $(\omega, \alpha, \beta, \varphi)' \in (0, \infty) \times [0, \infty)^3$ and the sequence (η_t) is iid with mean zero, unit variance, and a symmetric density. As emphasized by Smetanina (2017), the volatility sequence (λ_t) is not the conditional variance given the past (and present) information. So $\left(\lambda_t^2\right)$ is just a scale factor in the multiplicative form $Y_t = \lambda_t \eta_t$. Taking $\delta_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \lambda_{t-1}^2$ so that $\lambda_t^2 = \delta_t^2 + \varphi \eta_t^2$, and putting $h_t := Var\left(Y_t \middle| \mathscr{F}_{t-1}\right) = \delta_t^2 + \varphi E\left(\eta_t^4\right)$, the RT-GARCH model can be written as a standard but non-MEM exponential GARCH (EGARCH) model as follows

$$\begin{aligned}
Y_t^2 &= h_t\eta_t^2 + \varphi\eta_t^2\left(\eta_t^2 - E\left(\eta_t^4\right)\right) \\
h_t &= \omega_1 + \alpha Y_{t-1}^2 + \beta h_{t-1} + \phi\eta_{t-1}^2 \\
\lambda_t^2 &= h_t + \varphi\left(\eta_t^2 - E\left(\eta_t^4\right)\right)
\end{aligned}$$

where $\omega_1 = \omega + \varphi E\left(\eta_t^4\right)(1-\beta)$ and $\phi = \beta\varphi$. When $\varphi = 0$, the Engel-Bollerslev GARCH model is retrieved. In the latter representation, $Y_t = \sqrt{h_t\eta_t^2 + \varphi\eta_t^2\left(\eta_t^2 - E\left(\eta_t^4\right)\right)} := g\left(h_t, \eta_t; \varphi\right)$ is not a multiplicative function of $\sqrt{h_t}$ and η_t, or more precisely is a multiplicative term plus the additive term $\varphi\eta_t^2\left(\eta_t^2 - E\left(\eta_t^4\right)\right)$ that makes the conditional Kurtosis not constant contrary to what happens with the standard multiplicative GARCH model.

Thus, the RT-GARCH model provides two types of volatilities: First, the conditional variance given past information (up to time $t-1$) $h_t = Var\left(Y_t|Y_{t-1}, Y_{t-2}, ...\right)$ in the spirit of the Engle-Bollerslev GARCH model. Second, the real-time/contemporaneous volatility $\lambda_t^2 = Var\left(Y_t|Y_{t-1}, Y_{t-2}, ...\right) + \varphi\left(\eta_t^2 - E\left(\eta_t^4\right)\right)$ which is the conditional variance plus the additional term $\varphi\left(\eta_t^2 - E\left(\eta_t^4\right)\right)$. Unless $\varphi = 0$, the real-time volatility λ_t^2 is not the conditional variance given past or past-present information and therefore has only a formal effect in the multiplicative relationship $Y_t = \lambda_t\eta_t$. The name "real-time GARCH" probably comes from the fact that the volatility equation for the scaling factor λ_t^2 contains the current innovation term $\varphi\eta_t^2$. However, if the considered volatility is the conditional variance $h_t = Var\left(Y_t|Y_{t-1}, Y_{t-2}, ...\right)$ given past information, then the term "real-time" loses all its meaning and the model does not fit into the framework of parameter driven models such as the (standard) SV, the MS-GARCH and the RC-GARCH models in which the "complete" volatility $\sigma_t^2 = Var\left(Y_t|Y_{t-1}, Y_{t-2}, ...; e_t\right)$ is the conditional variance given past observations and present random parameters.

On the other hand, since the real-time volatility λ_t^2 is not the conditional variance, comparing it to the conditional variance of a standard GARCH model does not seem to be fair. In fact, the conditional variance of the RT-GARCH model given past-present information $Var\left(Y_t|Y_{t-1}, Y_{t-2}, ...; \eta_t\right)$ is zero and therefore makes no sense. Moreover, the so-called volatility of volatility (e.g. Ding, 2020-2023) which is $Var\left(\lambda_t^2|Y_{t-1}, Y_{t-2}, ...\right)$ is not the conditional variance of the conditional variance, as is the case with the RC-GARCH model. Thus, the relevance of the RT-GARCH model in volatility forecasting using

h_t (as shown by Smetanina, 2017) seems to come mainly from the non-MEM structure and/or the EGARCH term $\phi\eta_{t-1}^2$ in the conditional variance equation. Extensions of the RT-GARCH model are given by Ding (2020-2023) and Wu et al (2023).

Similar conclusions can be drawn for the NC-GARCH model of Breitung and Hafner (2016) which is given by the following equation

$$\begin{aligned} Y_t &= \sigma_t \eta_t \\ \log\left(\sigma_t^2\right) &= \omega + \beta \log\left(\sigma_{t-1}^2\right) + \varphi e_t \end{aligned}$$

where $e_t := \left(\log\left(\eta_t^2\right) - E\left(\log\left(\eta_t^2\right)\right)\right)$, $(\omega, \beta, \varphi)' \in \mathbb{R}^3$, and the sequence (η_t) is iid with mean zero, unit variance, and a symmetric density. Since in this model, the innovation term e_t is fully dependent with η_t, the so-called volatility σ_t^2 cannot be the conditional variance of Y_t even given the past of e_t up to time t. Taking $\log(h_t) := \omega + \beta \log\left(\sigma_{t-1}^2\right)$, so that $\log\left(\sigma_t^2\right) = \log(h_t) + \varphi e_t$, Breitung and Hafner (2016)'s model becomes

$$\begin{aligned} Y_t^2 &= h_t \eta_t^2 \exp \varphi \left(\log\left(\eta_t^2\right) - E\left(\log\left(\eta_t^2\right)\right)\right) \\ \log(h_t) &= \omega + \beta \log(h_{t-1}) + \beta\varphi e_{t-1} \end{aligned}$$

which is a specific EGARCH model but with a non-multiplicative structure. Note that when $\varphi = 0$ the standard GARCH(0,1) model is retrieved. See the Supplementary Material for more details about these models and their relationships with other volatility models.

2.4 Stability properties

We now study the existence of a causal/nonanticipative 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

$$Z_t = A_t Z_{t-1} + B_t, \tag{2.10}$$

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

$$\gamma(\mathbf{A}) = \inf\left\{\tfrac{1}{t} E \log \|A_t ... A_2 A_1\|, \ t \geq 1\right\}$$

be the largest Lyapunov exponent associated with the iid-driven SRE (2.10) (Bougerol and Picard, 1992). Consider 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* $E\left(\log\left(\varepsilon_t^2\right)\right) < \infty$, $E\left(\log\left(\omega_t\right)\right) < \infty$, $E\left(\log\left(\alpha_{it}\right)\right) < \infty$ *and* $E\left(\log\left(\beta_{jt}\right)\right) < \infty$ *(*$i = 1, ..., q, \ j = 1, ..., p$*). A necessary and sufficient condition for model* (2.10) *to have a unique nonanticipative/causal strictly stationary and ergodic solution is that*

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

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

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

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

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

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 (Francq and Zakoian, 2019). Proofs of Propositions 2.1-2.3 are given in the online material.

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 mean $\varrho_t^2 := E\left(\sigma_t^2|Y_1, ..., Y_t\right)$ $(1 \leq t \leq n)$, also called the filtered volatility. 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 Inverse 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 (e.g. Karlis, 2002). The NIG distribution (see the online material) 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 predictive 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 predictive volatility, which exists and is strictly 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

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

Based on the standard asymptotic GARCH theory (Francq and Zakoian, 2004-2019) it will be shown 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

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

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*

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

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 matrix 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

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

For proof of Theorem 3.1, see the online material.

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 the volatility of volatility equations (2.7)-(2.8). Consider $e_t = \left(Y_t^2 - \delta_t^2\right)^2 - Var\left(Y_t^2|\mathscr{F}_{t-1}^Y\right)$ 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

$$Var\left(Y_t^2|\mathscr{F}_{t-1}^Y\right) = \kappa M_t'\Lambda_0 + (\kappa-1)\,\delta_t^4,$$

so $(3.10a)$ becomes

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

where $E\left(\frac{e_t}{\kappa\delta_t^4}|\mathscr{F}_{t-1}^Y\right) = \frac{1}{\kappa\delta_t^4}E\left(e_t|\mathscr{F}_{t-1}^Y\right) = 0$ 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) under 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 eight moment assumption **A7** which remains quite strong, can be valid in applications. From (3.10a) and (2.8), consistent estimates of A and B in (3.13) are, respectively,

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

For proof of Theorem 3.2, see the online material.

3.3 Estimating/filtering the unobserved volatilities and random coefficients

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

$$\varrho_t^2 = E\left(\sigma_t^2 | Y_1, ..., Y_t\right), \; t = 1, ..., n, \tag{3.17}$$

that we obtain from the filtered distribution $f\left(\sigma_t^2|Y_1,...,Y_t\right)$. Note that the unobserved volatility σ_t^2 can also be estimated by the predictive volatility $\delta_t^2 = E\left(\sigma_t^2|Y_1,...,Y_{t-1}\right)$ but the latter does not use the current observation Y_t as is the case with ϱ_t^2 given by (3.17).

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 the online material), that is

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

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

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

The precise choice of the IG distribution is due to the fact that closed-form formulas can be obtained for the posteriors and because the resulting distribution of $Y_t|\mathscr{F}_{t-1}^Y$ is very flexible. From the summability property of the IG distribution (see the online material) and the mutual independence of $\{\omega_t, t\in\mathbb{Z}\}$, $\{\alpha_{it}, t\in\mathbb{Z}\}$ $(i=1,...,q)$, and $\{\beta_{jt}, t\in\mathbb{Z}\}$ $(j=1,...,p)$, 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 the online material). It has a closed form density (see the online material) and is also given in the following hierarchical mixture (see the online material)

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

The posterior/filtered volatility $\sigma_t^2|\mathscr{F}_t^Y$ has the density

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

which is nothing but the generalized inverse Gaussian (GIG) distribution (cf. the online material). In fact, the GIG distribution is conjugate for the normal distribution, so the distribution of $\sigma_t^2|\mathscr{F}_t^Y$ is given by (cf. the online material)

$$\begin{cases} \sigma_t^2|\mathscr{F}_{t-1}^Y \sim \mathcal{IG}\left(\delta_t^2, \Delta_t^2\right) \\ Y_t|\mathscr{F}_{t-1}^Y, \phi_t \sim \mathcal{N}\left(0, \sigma_t^2\right) \end{cases} \Longrightarrow \sigma_t^2|\mathscr{F}_t^Y \sim \mathcal{GIG}\left(-1, \tfrac{\Delta_t}{\delta_t^2}\sqrt{\Delta_t^2 + Y_t^2}, \tfrac{\Delta_t}{\delta_t^2}\right)$$

where $\mathcal{GIG}(\tau, \varphi, \gamma)$ denotes the Generalized Inverse Gaussian distribution with parameters τ, φ, γ (cf. the online material). Thus, the filtered volatility given by

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

can be obtained in a closed form. Using the result of Barndorff-Nielssen (1978) for the GIG distribution (see also Karlis, 2002, formula (4)), a closed form formula for the IG posterior mean in (3.23) is given in the online material. Hence, the estimate $\widehat{\varrho}_t^2$ of ϱ_t^2 is obtained while replacing the true parameters in the expression (3.23) by their estimates obtained in the first and second stages, giving

$$\widehat{\varrho}_t^2 = \widehat{E}\left(\sigma_t^2|\mathscr{F}_t^Y\right) = \tfrac{\widehat{\delta}_t^2\sqrt{\widehat{\Delta}_t^2 + Y_t^2}}{\widehat{\Delta}_t} \frac{K_0\left(\frac{\widehat{\Delta}_t}{\widehat{\delta}_t^2}\sqrt{\widehat{\Delta}_t^2 + Y_t^2}\right)}{K_{-1}\left(\frac{\widehat{\Delta}_t}{\widehat{\delta}_t^2}\sqrt{\widehat{\Delta}_t^2 + Y_t^2}\right)} \tag{3.24}$$

where $\widehat{E}$ denotes the expectation in which the true parameters are replaced by their estimates, $\widehat{\delta}_t^2$ is given by (3.9b), $K_r\left(y\right)$ denotes the modified Bessel function of the third kind of order r evaluated at

y, and from (3.20b)

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

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 the online material; Barndorff-Nielsen and Prause, 2001).

It is also possible to make inference about the regimes taken by each random coefficient using its posterior mean given the past and present observations. In view of the independence of the random coefficients, their inverse Gaussian priors given by (3.19), and the normality assumption (3.18) about the model innovation, the posterior distribution of the coefficients is generalized inverse Gaussian. For instance, for the random coefficient ω_t we have

$$\begin{cases} \omega_t | \mathscr{F}_{t-1}^Y \sim \mathcal{IG}(\omega_0, \lambda_\omega) \\ Y_t | \mathscr{F}_{t-1}^Y, \phi_t \sim \mathcal{N}(0, \sigma_t^2) \end{cases} \Longrightarrow \omega_t | \mathscr{F}_t^Y \equiv \omega_t | Y_t \sim \mathcal{GIG}\left(-1, \frac{\sqrt{\lambda_\omega}}{\omega_0}\sqrt{\lambda_\omega + Y_t^2}, \frac{\sqrt{\lambda_\omega}}{\omega_0}\right).$$

Likewise,

$$\begin{aligned} \alpha_{it}|Y_t &\sim \mathcal{GIG}\left(-1, \frac{\sqrt{\lambda_{\alpha_i}}}{\alpha_0}\sqrt{\lambda_{\alpha_i} + Y_t^2}, \frac{\sqrt{\lambda_{\alpha_i}}}{\alpha_0}\right),\ 1 \le i \le q \\ \beta_{jt}|Y_t &\sim \mathcal{GIG}\left(-1, \frac{\sqrt{\lambda_{\beta_j}}}{\beta_0}\sqrt{\lambda_{\beta_j} + Y_t^2}, \frac{\sqrt{\lambda_{\beta_j}}}{\beta_0}\right),\ 1 \le j \le p. \end{aligned}$$

Therefore, similarly to $\sigma_t^2 | \mathscr{F}_t^Y$, the posterior means of the random coefficients are given from the mean of the GIG distribution by (cf. the online material)

$$\begin{aligned} E\left(\omega_t | \mathscr{F}_t^Y\right) &= E\left(\omega_t | Y_t\right) = \frac{\omega_0\sqrt{\lambda_\omega + Y_t^2}}{\sqrt{\lambda_\omega}} \frac{K_0\left(\frac{\sqrt{\lambda_\omega}}{\omega_0}\sqrt{\lambda_\omega + Y_t^2}\right)}{K_{-1}\left(\frac{\sqrt{\lambda_\omega}}{\omega_0}\sqrt{\lambda_\omega + Y_t^2}\right)} \\ E\left(\alpha_{it} | \mathscr{F}_t^Y\right) &= E\left(\alpha_{it} | Y_t\right) = \frac{\alpha_0\sqrt{\lambda_{\alpha_i} + Y_t^2}}{\sqrt{\lambda_{\alpha_i}}} \frac{K_0\left(\frac{\sqrt{\lambda_{\alpha_i}}}{\alpha_0}\sqrt{\lambda_{\alpha_i} + Y_t^2}\right)}{K_{-1}\left(\frac{\sqrt{\lambda_{\alpha_i}}}{\alpha_0}\sqrt{\lambda_{\alpha_i} + Y_t^2}\right)},\ 1 \le i \le q \\ E\left(\beta_{jt} | \mathscr{F}_t^Y\right) &= E\left(\beta_{jt} | Y_t\right) = \frac{\beta_0\sqrt{\lambda_{\beta_j} + Y_t^2}}{\sqrt{\lambda_{\beta_j}}} \frac{K_0\left(\frac{\sqrt{\lambda_{\beta_j}}}{\beta_0}\sqrt{\lambda_{\beta_j} + Y_t^2}\right)}{K_{-1}\left(\frac{\sqrt{\lambda_{\beta_j}}}{\beta_0}\sqrt{\lambda_{\beta_j} + Y_t^2}\right)},\ 1 \le j \le p. \end{aligned} \tag{3.26}$$

The estimated posterior means are obtained while replacing the true parameters by their estimates.

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 predictive 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 filtered volatilities $\widehat{\varrho}_1^2, ..., \widehat{\varrho}_n^2$ from (3.24)-(3.25), using the posterior mean of the Inverse Gaussian distribution.

7- Estimate the posterior mean of the random coefficients from (3.26).

3.4 Testing the randomness of coefficients

As for any random coefficient model, an important step in building a RC-GARCH model is to test for the randomness of its coefficients. This may validate the random-coefficient structure of the model. We, thus, use Wald tests for the null hypothesis of the nullity of the variance parameter Λ_0 (or some of its components) against its opposite as alternative, based on the asymptotic distribution of the WLSE $\widehat{\Lambda}_n$. Such a null hypothesis writes for each component of Λ_0 as follows

$$H_0^i : \Lambda_{0i} = 0, \;\; i = 1, ..., p + q + 1. \tag{3.27}$$

A feasible Wald statistic for testing (3.27), based on the asymptotic distribution of the WLSE as given by (3.15), is defined by

$$W_{i,n} = n \frac{\widehat{\Lambda}_{n,i}^2}{\widehat{g}_i}, \;\; i = 1, ..., p + q + 1, \tag{3.28}$$

where $\widehat{g}_i$ is the ith diagonal element of the matrix $\widehat{\Sigma} = \widehat{A}^{-1}\widehat{B}\widehat{A}^{-1}$ with $\widehat{A}$ and $\widehat{B}$ being given by (3.16). Under H_0^i in (3.27) and the assumptions of Theorem 3.2 it can be seen that

$$W_{i,n} \overset{D}{\to} \chi^2_{(1)} \quad \text{as} \quad n \to \infty.$$

Given a size $\alpha \in (0,1)$, let $\chi_{i,\alpha}$ be the critical value such that $P(W_{i,n} > \chi_{i,\alpha}) \to \alpha$ as $n \to \infty$. The null H_0^i is thus rejected if $W_{i,n} > \chi_{i,\alpha}$.

If the null (3.27) is not rejected, this does not imply that all volatility coefficients are not random, and it may be possible that at least one of the coefficients is random. Letting $M = (1, ..., 1)'$ be a $1 \times (p+q+1)$-vector of ones, a global hypothesis over all coefficients can be considered as the null

$$H_0 : M'\Lambda_0 = 0, \tag{3.29}$$

for which we use again the Wald test. A Wald statistic for (3.29) is given by

$$W_n = \left(M\widehat{\Lambda}_n\right)' \left(M\tfrac{1}{n}\widehat{\Sigma}M'\right)^{-1} M\widehat{\Lambda}_n. \tag{3.30}$$

Under H_0 and the above assumptions of Theorem 3.2 it can be seen that

$$W_n \overset{D}{\to} \chi^2_{(1)} \quad \text{as} \quad n \to \infty.$$

4 Empirical data: An application to Cisco stock returns

The empirical application concerns the daily returns of Cisco stock for the period 01/02/2001 to 12/31/2008 that consists of $n = 2011$ observations (see Figure 2). 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).

Applying the first two stages of Algorithm 3.1, gives the estimated RC-GARCH(1,1) model for which the estimated means and variances of the random coefficients as well as their asymptotic stan-

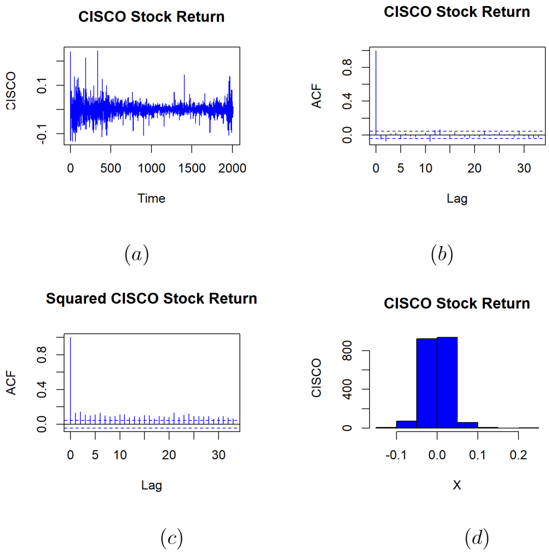


Figure 2. CISCO stock return series: (a) The CISCO return series of RCISCO; (b) sample autocorrelation, (c) sample autocorrelation of squares, (d) histogram.

dard errors (ASE) in parentheses are displayed in Table 1.

	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 (9.7e-08)	0.1229 (0.0480)	1.3650 (0.9231)		2.7260

Table 1. QML and WLS estimates for the RC-GARCH(1, 1).

The parameter estimate $\widehat{\alpha}_{1n}+\widehat{\beta}_{1n} \simeq 0.9950$ 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,

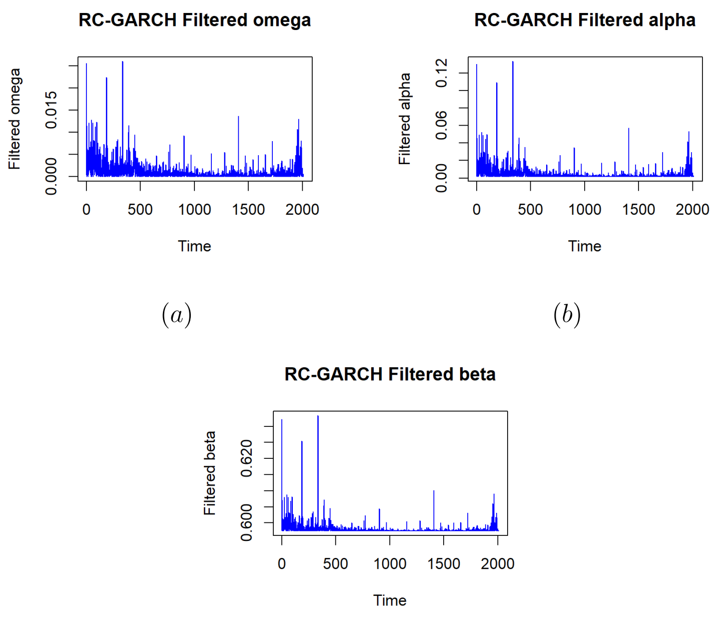
$$FMC := 3\widehat{\alpha}^2_{1n} + 2\widehat{\alpha}_{1n}\widehat{\beta}_{1n} + \widehat{\beta}^2_{1n} + \kappa\widehat{\sigma}^2_{0\alpha_1} + \widehat{\sigma}^2_{0\beta_1} \simeq 2.7260,$$

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 4 (Panel (d)). Note that all parameters are significant and, in particular, the variance parameters as confirmed by Table 2, which reports the Wald statistics defined in (3.28). The hypothesis that the parameter ω_t is random is only accepted at the level 0.50, which is highly stringent. However, the randomness of the parameter α_t is accepted at any level even at 0.01. The hypothesis that β_t is not random is accepted at the level 0.1. This show that the randomness of the RC-GARCH model is not rejected at least at the level 0.10, regarding the global Wald statistics. The same holds for the Intel application (see Supplementary Material).

$W_{\omega,n}$	$W_{\alpha,n}$	$W_{\beta,n}$
0.3349	6.5446	2.1867

Table 2. Individual Wald Statistics for testing the randomness of coefficients.

Next, the posterior means $E(\omega_t|Y_t)$, $E(\alpha_t|Y_t)$ and $E(\beta_t|Y_t)$ $(1 \leq t \leq n)$ of the random coefficients are plotted in Figure 3. It can be seen that, as expected, each posterior mean exhibits a behavior that is consistent with the continuous regime assumption. To get an idea about the behavior of the random coefficients, the graphs of the simulated random coefficients $(\omega_t, \alpha_t, \beta_t)$ obtained from the estimated RC-GARCH(1,1) model as well as the random persistence $\alpha_t + \beta_t$ are given in the Supplementary Material, where it is shown that the generated α_t and β_t coefficients as well as their persistence can flexibly exceed unity by a large margin.



(c)

Figure 3. Posterior means of the random coefficients $(\omega_t, \alpha_t, \beta_t)$: (a) $E(\omega_t|Y_t)$, (b) $E(\alpha_t|Y_t)$, (c) $E(\beta_t|Y_t)$, $1 \leq t \leq n$.

Figure 4 shows the predictive volatility $\widehat{\delta}_t^2$ (Figure 3 (a)), which is nothing but the volatility of the standard GARCH model, and the filtered volatility $\widehat{\varrho}_t^2$ (Figure 4 (b)) obtained from Stage 3 of the Algorithm 3.1. The filtered volatility $\widehat{\varrho}_t^2$ is more erratic and captures small and large volatilities better than the predictive volatility does. In addition, as mentioned in the introduction, the filtering volatility $\widehat{\varrho}_t^2$ does not seem to contain curves in the sense of large volatilities as the predictive (or standard GARCH) volatility $\widehat{\delta}_t^2$ does. The conditional excess kurtosis (panel (d)) exceptionally shows very large picks, which are probably due to the fourth-order instability of the model. Panel (c) shows the time plot of the volatility of volatility generated by the RC-GARCH model, whose behavior is consistent with that of the predictive volatility.

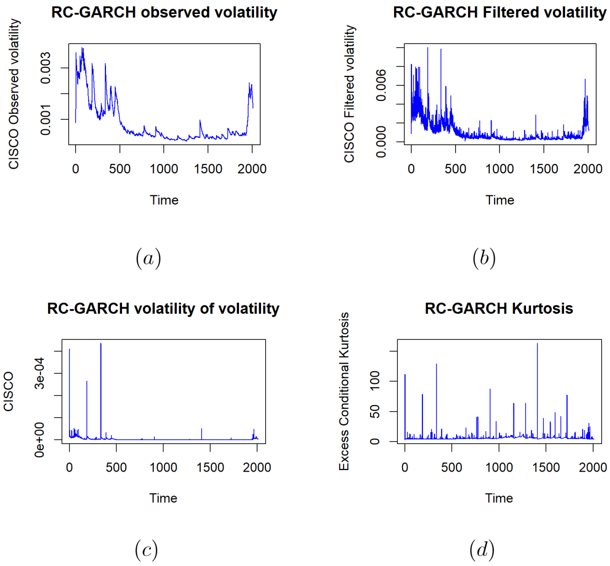


Figure 4. Estimated RC-GARCH for CISCO series. (a) Predictive volatility, (b) filtered volatility, (c) volatility of volatility, (d) conditional excess kurtosis.

We finally assess the out-of-sample forecasting ability of four competing volatilities. The first one is the Normal GARCH predictive volatility $\widehat{\delta}_t^2$, the second one is the NIG filtered volatility $\widehat{\varrho}_t^2$ in which the future returns are available, the third one is the NIG filtered volatility $\widehat{\varrho}_t^{*2}$ in which the future returns Y_{t+j} ($j \geq 1$) are estimated by 0 (see also online material), while the fourth one is the MS-GARCH volatility (MS). We have not included the out-of-sample volatility forecasts generated by the tvGARCH model since they are not available from the package tvGARCH, while error messages are thrown, when estimating the tvGARCH model based on truncated series. We, thus, estimate the four volatility models on the basis of the first n_c observations of the series, where $1 < n_c < n$. Then, we compute the one-step ahead volatility forecast on the period $(n_c + 1, ..., n)$ for each model and obtain the three criteria: i) the MSFE$= \frac{1}{n-n_c} \sum\limits_{t=n_c+1}^{n} (Y_t^2 - \widehat{h}_t)^2$, ii) the MAFE $= \frac{1}{n-n_c} \sum\limits_{t=n_c+1}^{n} \left| Y_t^2 - \widehat{h}_t \right|$, and iii) the MQLI$= \frac{1}{n-n_c} \sum\limits_{t=n_c+1}^{n} (\log \widehat{h}_t + \frac{Y_t^2}{\widehat{h}_t})$. The one-step ahead GARCH predictive volatility forecast is simply $\widehat{\delta}_t^2$ $(n_c + 1 \leq t \leq n)$, while that of the filtered volatility is obtained from expression (A.5) (given in the online material) with $h = 1$. Finally, the one-step ahead MS-GARCH volatility forecast are obtained using the function "*predict()*" of the package MSGARCH of Ardia et al (2019). Table 3 shows the computed values of the above criteria for the four models and for various truncated series with sample size $n_c \in \{1450, 1500, 1600, 1700, 1800, 1900\}$ (cf. Table 3).

	n_c	1450	1500	1600	1700	1800	1900
	MSFE	2.12e-06	2.29e-06	2.78e-06	3.58e-06	4.78e-06	8.57e-06
$\widehat{\delta}_t^2$	MAFE	0.00061	0.00063	0.00072	0.00085	0.00103	0.00157
	MQLI	−6.6401	−6.5706	−6.6338	−6.1700	−5.9546	−5.3450
	MSFE	9.59e-07	1.05e-06	1.29e-06	1.65e-06	2.25e-06	4.09e-06
$\widehat{\varrho}_t^2$	MAFE	0.00036	0.00038	0.00043	0.00051	0.00064	0.00099
	MQLI	−7.2763	−7.1797	−6.9818	−6.8116	−6.6594	−6.0851
	MSFE	**1.29e-08**	**1.45e-08**	**1.84e-08**	**2.45e-08**	**3.60e-08**	**6.85e-08**
$\widehat{\varrho}_t^{*2}$	MAFE	**0.00008**	**0.00008**	**0.00009**	**0.00012**	**0.00015**	**0.00024**
	MQLI	**−9.6907**	**−9.6465**	**−9.5180**	**−9.3342**	**−8.9941**	**−8.4441**
	MSFE	2.74e-06	2.16e-06	3.79e-06	3.45e-06	4.60e-06	8.39e-06
MS	MAFE	0.00075	0.00061	0.00086	0.00081	0.00098	0.00149
	MQLI	−6.8789	−6.6866	−6.7176	−6.2358	−6.0537	−5.3804

Table 3. Out-of-sample volatility forecasting performance of the filtered volatility, the predictive volatility, and the MS-GARCH volatility (MS). $\widehat{\varrho}_t^2$: filtered volatility using available returns. $\widehat{\varrho}_t^{*2}$: filtered volatility using predictive returns.

From Table 3 some conclusions can be drawn. i) First, it can be seen that regardless of the chosen time-cut n_c, the filtered volatilities $\widehat{\varrho}_t^2$ and $\widehat{\varrho}_t^{*2}$ give better out-of-sample forecasts with respect to the above-mentioned criteria. In particular $\widehat{\varrho}_t^{*2}$ outperforms $\widehat{\varrho}_t^2$ in terms of all criteria and all n_c. Moreover, $\widehat{\varrho}_t^2$ dominates both $\widehat{\delta}_t^2$ and the MS-GARCH volatility in terms of all criteria and all n_c. ii) Second, except the cases $n_c \in \{1450, 1600\}$, the MS-GARCH volatility outperforms the standard GARCH volatility for all other n_c and all criteria.

For a more meaningful comparison between the four volatility forecasts in Table 3, we resort to the Model Confidence Set method of Hansen et *al* (2011) using the R package MCS of Bernardi and Catania (2014). Regarding the mean square loss function (MSFE), we found that the forecasts obtained by the filtered volatility $\widehat{\varrho}_t^{*2}$ constitutes the Superior Set Models for all n_c, with the GARCH and MS-GARCH volatilites being excluded from this set.

All in all, confirming the results of the simulation Section (see online material), our CISCO application showed that the filtered volatility is able to improve the in-sample GARCH volatility forecasts and also the out-of-sample forecasts, even when the future data are not available and are replaced by their predictions.

It should be noted that, the comparison of predictive and filtered volatilities is not used here in the sense of a competition between them, but only to see how filtered volatility, based on predictive volatility, could complement the latter, which is important, especially in volatility forecasting. In fact, filtered volatility can be used in nowcasting, as it could improve volatility forecasting when current or immediate future observations are available/predicted.

5 Conclusion

This work proposed a random coefficient GARCH (RC-GARCH) model with a time-varying conditional kurtosis and a latent conditional volatility sequence driven by past observations and present iid random inputs. The proposed formulation, which is path-independent, mimics the Markov switching specification of Gray (1996) in a continuous-valued regime framework, and is different from earlier random coefficient GARCH models introduced by Kazakevicius et al (2004), Klivecka (2004), and Thavaneswaran et al (2005). The latent volatility, which is the main focus of this paper, can be estimated in two ways. First, using the predictive/observable volatility, which is the conditional mean of the latent volatility given past observations. It is, therefore, entirely determined by past observations and is exactly the same as the volatility of the standard Engle-Bollerslev GARCH model. Second, the latent volatility can be estimated using the filtered volatility, which is the conditional mean of the latent volatility given past and present observations. This filtered volatility has the advantage of also incorporating the current observation and thus can better describe latent volatility. Through simulation and empirical studies, we found that the filtered volatility can increase the in-sample forecasting ability and model fit via the conditional NIG distribution and also the out-of-sample ability if the data were available or even predicted. The same conclusion holds for the application of the Intel returns, given in the Supplementary Material. Thus, the RC-GARCH model equipped with the two volatility estimates for the latent volatility can shed more light on the evolution of the variability of the underlying series. In particular, the volatility and volatility of volatility generated by the model are understood in the natural sense of conditional variance and therefore do not have an abusive meaning as is the case with some SV and real-time volatility models in which volatility is no longer the conditional variance. Furthermore, as suggested by the co-Editor, the RC-GARCH model also provides a standard GARCH volatility, unlike most parameter-driven models.

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 a closed form regardless of the distribution of the model. Finally, assuming that the random coefficients are IG distributed and the innovation is Normal, the conditional model is NIG distributed. The NIG hypothesis allows for closed-form posteriors, is very flexible, and can account for heavy tailedness and asymmetry. Moreover, the latent volatility filtering process is obtained simply in a closed form unlike most parameter-driven models (SV, MS-GARCH) that require highly computational methods to filter the latent volatility.

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). Other random-coefficient GARCH models could be considered, such as the random coefficient EGARCH, the random coefficient asymmetric power GARCH, the random coefficient tvGARCH (in the sense of Amado and Teräsvirta, 2013), and the random coefficient score-driven model. Finally, multivariate extensions of the RC-GARCH model seem appealing. These aspects of analysis could be analyzed in a future research agenda.

Acknowledgements

The authors are deeply grateful to the Co-Editor, Professor Christian Francq and two referees for their insightful comments and suggestions that led to a significant improvement of the earlier version of the manuscript. In particular, the suggestions of: conditional variance interpretation of volatility, testing the randomness of coefficients, filtering the random coefficients, and in-sample and out-of-sample forecasting comparison with MS-GARCH and tvGARCH models, are greatly appreciated.

Supplementary material for: 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

1 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) $\phi = \frac{1}{\lambda}$. The original parametrization (e.g. Barndorff-Nielsen, 1978-1997; Karlis, 2002) has been expressed in terms of the parameters $\varphi = \sqrt{\lambda}$ and $\gamma = \frac{\sqrt{\lambda}}{\rho}$ so that $\rho = \frac{\varphi}{\gamma}$ and $\lambda = \varphi^2$, giving

$$f(z;\varphi,\gamma) = \frac{\varphi}{\sqrt{2\pi z^3}} \exp(\varphi\gamma) \exp\left(-\frac{1}{2}\left(\frac{\varphi^2}{z} + \gamma^2 z\right)\right),\ z > 0.$$

The mean and variance of the IG distribution are $E(Z) = \rho$ and $Var(Z) = \frac{\rho^3}{\lambda}$, and in terms of the IG$(\varphi,\gamma)$ parametrization by $E(Z) = \frac{\varphi}{\gamma}$ and $Var(Z) = \frac{\varphi}{\gamma^3}$. 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

[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.

Inverse Gaussian (NIG) distribution that has a closed form. A continuous random variable Y is said to have a NIG distribution with parameters $\alpha, \varphi, \mu, \beta$ ($Y \sim \mathcal{NIG}(\alpha, \beta, \varphi, \mu)$, $\alpha, \varphi > 0$, $|\beta| \leq \alpha$, $\mu \in \mathbb{R}$) if its probability density function is given by (Barndorff-Nielsen, 1978)

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

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

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

In particular, when $\beta = 0$ and $\mu = 0$, and using the mean-shape parametrization of the $\mathcal{IG}(\rho, \lambda)$ with $\varphi = \sqrt{\lambda}$ and $\gamma = \frac{\sqrt{\lambda}}{\rho}$, the above hierarchical form of the $\mathcal{NIG}\left(\frac{\sqrt{\lambda}}{\rho}, 0, \sqrt{\lambda}, 0\right)$ distribution becomes

$$\begin{cases} Z|\rho, \lambda \sim \mathcal{IG}(\rho, \lambda) \\ Y|Z \sim \mathcal{N}(0, Z) \end{cases} \Longrightarrow Y \sim \mathcal{NIG}\left(\frac{\sqrt{\lambda}}{\rho}, 0, \sqrt{\lambda}, 0\right). \tag{A.3}$$

The mean and variance of the NIG variable are $E(Y) = \mu + \frac{\varphi\beta}{\gamma}$ and $Var(Y) = \frac{\varphi(\gamma^2 + \beta^2)}{\gamma^3}$. 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 advantage 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 and IG distributions is that the posterior mean of the IG distribution $E(Z|Y)$ can be obtained in a closed form using the Generalized Inverse Gaussian (GIG) distribution. A ransom variable Z is said to have a $\mathcal{GIG}(\tau, \varphi, \gamma)$ distribution with parameters τ, φ and γ if

$$f(z; \tau, \varphi, \gamma) = \left(\frac{\gamma}{\varphi}\right)^{\tau} \frac{z^{\tau-1}}{2K_\tau(\varphi\gamma)} \exp\left(-\frac{1}{2}\left(\frac{\varphi^2}{z} + \gamma^2 z\right)\right), \; z > 0.$$

In terms of the mean-shape representation of the $\mathcal{IG}(\rho, \lambda)$ distribution, the $\mathcal{GIG}(\tau, \rho, \lambda)$ distribution takes the form

$$f(z;\tau,\rho,\lambda) = \rho^{-\tau} \frac{z^{\tau-1}}{2K_\tau\left(\frac{\lambda}{\rho}\right)} \exp\left(-\frac{1}{2}\left(\frac{\lambda}{z} + \frac{\lambda}{\rho^2} z\right)\right), \; z > 0.$$

The r-th moment of the $\mathcal{GIG}(\tau,\varphi,\gamma)$ distribution is given by

$$E(Z^r) = \left(\frac{\varphi}{\gamma}\right)^r \frac{K_{\tau+r}(\varphi\gamma)}{K_\tau(\varphi\gamma)}.$$

For $\tau = -\frac{1}{2}$, the $\mathcal{GIG}\left(-\frac{1}{2},\varphi,\gamma\right)$ coincides with the $\mathcal{IG}(\varphi,\gamma)$ distribution. The main advantage of the GIG distribution is that it is conjugate to the normal distribution: If $Y|Z,\mu,\beta \sim N(\mu,\mu+\beta Z)$ and $Z|\tau,\varphi,\gamma \sim GIG(\tau,\gamma,\varphi)$ then,

$$Z|Y,\mu,\beta,\tau,\varphi,\gamma \sim GIG\left(\tau - \frac{1}{2}, \sqrt{\varphi^2 + (Y-\mu)^2}, \sqrt{\gamma^2+\beta^2}\right)$$

In particular, when $\beta = \mu = 0$, it follows that

$$\begin{array}{c} Y|Z \sim \mathcal{N}(0,Z) \\ Z|\tau,\gamma,\varphi \sim \mathcal{GIG}(\tau,\gamma,\varphi) \end{array} \Rightarrow Z|Y,\tau,\varphi,\gamma \sim \mathcal{GIG}\left(\tau-\frac{1}{2}, \sqrt{\varphi^2+Y^2}, \gamma\right)$$

and the posterior mean is given by

$$E(Z|Y,\tau,\varphi,\gamma) = \frac{\sqrt{\varphi^2+Y^2}}{\gamma} \frac{K_{\tau+1/2}\left(\gamma\sqrt{\varphi^2+Y^2}\right)}{K_{\tau-1/2}\left(\gamma\sqrt{\varphi^2+Y^2}\right)},$$

where $K_r(y)$ denotes the modified Bessel function of the third kind of order r evaluated at y. Taking $\tau = -\frac{1}{2}$, the posterior mean of the particular $\mathcal{IG}(\varphi,\gamma)$ distribution writes as (cf. Barndorff-Nielsen, 1997; Karlis, 2002, formula (4))

$$E(Z|Y,\varphi,\gamma) := E\left(Z|Y,\tau=-\frac{1}{2},\gamma,\varphi\right) = \frac{\sqrt{\varphi^2+Y^2}}{\gamma} \frac{K_0\left(\gamma\sqrt{\varphi^2+Y^2}\right)}{K_{-1}\left(\gamma\sqrt{\varphi^2+Y^2}\right)}.$$

In terms of the mean-shape parametrization $\varphi = \sqrt{\lambda}$ and $\gamma = \frac{\sqrt{\lambda}}{\rho}$, the posterior mean of the IG(ρ,λ) distribution takes the form

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

In R, we use the function dnig() of the package fBasics for the density of the NIG distribution, and the invgauss() function of the package actuar for the IG(ρ, λ) density under the mean-shape representation. Moreover, to evaluate $(A.4)$, we use the function besselK of the package *base.*

2 Forecasting volatilities

Since the volatility σ_t^2 is not observable even with a perfect knowledge of the parameters θ_0 and Λ_0, it is only estimated using available data. We have seen that there are two estimates of σ_t^2: the predictive volatility $\delta_t^2 = E\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right)$ given data up to time $t-1$, which is exactly the standard GARCH volatility, and the filtered volatility $\varrho_t^2 = E\left(\sigma_t^2 | \mathscr{F}_t^Y\right)$, which uses data up to time t. As is the case with the MS-GARCH model of Haas et *al* (2004*a*), the RC-GARCH model allows for a closed-form volatility. Since the model offers two types of (predictive, filtered) volatilities, two volatility forecasts can be provided. The predictive volatility $\left(\delta_t^2\right)$ is exactly that of a standard GARCH model, so the one-step and multi-step ahead forecasts δ_{t+h}^2 are obtained with the same formula as with the standard GARCH model (e.g. Francq and Zakoian, 2019). Then, if $Y_{t+1}, ..., Y_{t+h}$ are already available, the filtered volatility forecast ϱ_{t+h}^2 can be obtained while adapting (3.24) expression of the main paper as follows

$$\widehat{\varrho}_{t+h}^2 := \widehat{E}\left(\sigma_{t+h}^2 | \mathscr{F}_{t+h}^Y\right) = \frac{\widehat{\delta}_{t+h}^2 \sqrt{\widehat{\Delta}_{t+h}^2 + Y_{t+h}^2}}{\widehat{\Delta}_{t+h}} \frac{K_0\left(\frac{\widehat{\Delta}_{t+h}}{\widehat{\delta}_{t+h}^2}\sqrt{\widehat{\Delta}_{t+h}^2 + Y_{t+h}^2}\right)}{K_{-1}\left(\frac{\widehat{\Delta}_{t+h}}{\widehat{\delta}_{t+h}^2}\sqrt{\widehat{\Delta}_{t+h}^2 + Y_{t+h}^2}\right)}, \; h \geq 1, \tag{A.5}$$

where $\widehat{\Delta}_{t+h}^2$ is updated from (3.25) expression of the main paper. Of course, $Y_{t+1}, ..., Y_{t+h}$ are generally not available, but predictions of them could be obtained to anticipate filtered volatility forecasts. We denote by $\widehat{\varrho}_{t+h}^{*2}$ the filtered volatility obtained exactly from (A.5) when $Y_{t+1}, ..., Y_{t+h}$ are replaced by their predictions $\widehat{Y}_{t+1}, ..., \widehat{Y}_{t+h}$, where $\widehat{Y}_{t+j} = E\left(Y_{t+j} | \mathscr{F}_t^Y\right)$ $(1 \leq j \leq h)$. Thus, $\widehat{\varrho}_{t+h}^2$ and $\widehat{\varrho}_{t+h}^{*2}$ can be seen as a complement to the volatility forecast $\widehat{\delta}_{t+h}^2$, if $Y_{t+1}, ..., Y_{t+h}$ can be obtained/predicted.

The following algorithm summarizes the main steps for obtaining RC-GARCH volatility forecasts.

Algorithm 3.2 Volatility forecasts

Given a RC-GARCH series $Y_1, ..., Y_n$, and initial values $Y_0, ..., Y_{1-q}$, $\delta_0^2, ..., \delta_{1-p}^2$.

Step1: Compute the predictive volatilities $\widehat{\delta}_1^2, ..., \widehat{\delta}_n^2$ given the parameter estimate $\widehat{\theta}_n$ from

$$\widehat{\delta}_t^2 = \widehat{E}\left(\sigma_t^2 | Y_1, ..., Y_{t-1}\right) = \widehat{\omega} + \sum_{i=1}^{q} \widehat{\alpha}_i Y_{t-i}^2 + \sum_{j=1}^{p} \widehat{\beta}_j \widehat{\delta}_{t-j}^2, \;\; 1 \leq t \leq n.$$

Step 2: Compute the filtered volatilities $\widehat{\varrho}_t^2 = \widehat{E}\left(\sigma_t^2 | Y_1, ..., Y_t\right)$, $1 \leq t \leq n$ given the parameter estimates $\widehat{\theta}_n$ and $\widehat{\Lambda}_n$ using expression (3.24) in the main paper.

Step 3: Compute $\widehat{\delta}_{t+h}^2 = \widehat{E}\left(\sigma_{t+h}^2 | Y_1, ..., Y_t\right)$ $(h \geq 1)$ using the standard GARCH volatility forecast formula.

Step 4: Given new/predicted observations $Y_{n+1}, ..., Y_{n+h}$:

Compute $\widehat{\varrho}_{t+h}^2 = \widehat{E}\left(\sigma_{t+h}^2 | Y_1, ..., Y_{t+h}\right)$ from expression $(A.5)$.

Use $\widehat{\delta}_t^2$ and $\widehat{\varrho}_t^2$ $(1 + n \leq t \leq n + h)$ as (predictive and filtered) volatility forecasts.

3 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 $L_n(\theta)$ and ℓ_t as $\widetilde{L}_n(\theta)$ and $\widetilde{\ell}$ in (3.5) while substituting $\widetilde{\delta}_t^2(\theta)$ in (3.3) by $\delta_t^2(\theta)$ 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) $\lim_{n\to\infty} \sup_{\theta\in\Theta} \left|\widetilde{L}_n(\theta) - L_n(\theta)\right| = 0 \quad a.s.$

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

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

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

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

f) $\frac{\partial^2 L_n(\theta^*)}{\partial\theta\partial\theta'} \underset{n\to\infty}{\overset{a.s.}{\to}} J$*, 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 here.

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

$$\begin{aligned}
&\sum_{t=1}^{n} n^{-1}\left(1-\frac{Y_t^2}{\delta_t^2(\theta_0)}\right)^2\frac{1}{\delta_t^4(\theta)}\frac{\partial\delta_t^2(\theta_0)}{\partial\theta}\frac{\partial\delta_t^2(\theta_0)}{\partial\theta'}\\
&\underset{n\to\infty}{\overset{a.s.}{\to}} E\left(\frac{1}{\delta_t^4(\theta_0)}E\left(\left(1-\frac{Y_t^2}{\delta_t^2(\theta_0)}\right)^2|\mathscr{F}_{t-1}^Y\right)\frac{\partial\delta_t^2(\theta_0)}{\partial\theta}\frac{\partial\delta_t^2(\theta_0)}{\partial\theta'}\right)
\end{aligned}\tag{A.6}$$

where

$$\begin{aligned}
E\left(\left(1-\frac{Y_t^2}{\delta_t^2(\theta_0)}\right)^2|\mathscr{F}_{t-1}^Y\right) &= \frac{1}{\delta_t^4(\theta_0)}E\left(\left(\delta_t^2-\sigma_t^2\varepsilon_t^2\right)^2|\mathscr{F}_{t-1}^Y\right)\\
&= \frac{Var\left(Y_t^2|\mathscr{F}_{t-1}^Y\right)}{\delta_t^4(\theta_0)}\\
&= \frac{1}{\delta_t^4(\theta_0)}\left((\kappa-1)\,\delta_t^4+\kappa Var\left(\sigma_t^2|\mathscr{F}_{t-1}^Y\right)\right)
\end{aligned}\tag{A.7}$$

the result e) thus follows from $(A.6)$, $(A.7)$, 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 R^{p+q+1}.

Therefore, a standard argument shows that (3.12) becomes

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

which, in turns using (3.10), gives

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

Now under (2.11), the ergodic theorem yields

$$\frac{1}{n}\sum_{t=1}^{n} \frac{1}{\delta_t^8} M_t M_t' \underset{n\to\infty}{\overset{a.s.}{\to}} A \tag{A.9}$$

and, further under **A6**,

$$\frac{1}{n}\sum_{t=1}^{n} M_t \frac{e_t}{\kappa\delta_t^8} \underset{n\to\infty}{\overset{a.s.}{\to}} E\left(M_t \frac{e_t}{\kappa\delta_t^8}\right) = E\left(M_t \frac{1}{\kappa\delta_t^8} E\left(e_t \middle| \mathscr{F}_{t-1}^Y\right)\right) = 0. \tag{A.10}$$

Thus, (3.14) follows from $(A.8)$-$(A.10)$.

ii) To show (3.15), we first rewrite $(A.8)$ as follows

$$\sqrt{n}\left(\widehat{\Lambda}_n - \Lambda_0\right) = \left(\frac{1}{n}\sum_{t=1}^{n} \frac{1}{\delta_t^8} M_t M_t'\right)^{-1} \frac{1}{\sqrt{n}}\sum_{t=1}^{n} M_t \frac{e_t}{\kappa\delta_t^8} + o_{a.s.}(1). \tag{A.11}$$

The ergodic theorem shows under **A7** that

$$\begin{aligned}\sum_{t=1}^{n} \left(\frac{1}{\sqrt{n}} M_t \frac{e_t}{\kappa\delta_t^8}\right)\left(\frac{1}{\sqrt{n}} M_t \frac{e_t}{\kappa\delta_t^8}\right)' &= \frac{1}{n\kappa^2}\sum_{t=1}^{n} \frac{e_t^2}{\delta_t^{16}} M_t M_t' \\ &\underset{n\to\infty}{\overset{a.s.}{\to}} \frac{1}{\kappa^2} E\left(\frac{Var\left(\left(Y_t^2-\delta_t^2\right)^2 \middle| \mathscr{F}_{t-1}^Y\right)}{\delta_t^{16}} M_t M_t'\right).\end{aligned} \tag{A.12}$$

From $(A.12)$ and **A5**-**A7**, the central limit theorem for square-integrable martingales implies

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} M_t \frac{e_t}{\kappa\delta_t^8} \underset{n\to\infty}{\overset{D}{\to}} \mathcal{N}(0, B). \tag{A.13}$$

The result (3.15) thus follows from $(A.11)$, $(A.9)$, and $(A.13)$. □

4 Related volatility models

Let the general multiplicative volatility model

$$
\begin{aligned}
Y_t &= \sigma_t \eta_t \\
\psi_t^2 &= \omega + \sum_{i=1}^{q} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{p} \beta \psi_{t-j}^2 + \varphi e_t
\end{aligned}
$$

where (η_t) and (e_t) are iid, and η_t has mean zero and unit variance. For simplicity of exposition we consider the case $p = q = 1$ giving:

$$Y_t = \sigma_t \eta_t \tag{A.14a}$$

$$\psi_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \psi_{t-1}^2 + \varphi e_t. \tag{A.14b}$$

The model (A.14) encompasses many popular volatility models.

i) Engle-Bollerslev GARCH model

When $\psi_t = \sigma_t$, $(\omega, \alpha, \beta)' \in (0, \infty) \times [0, \infty)^2$, and $\varphi = 0$, model (A.14) reduces to the standard Engle-Bollerslev GARCH(1.1) model (Engle, 1982; Bollerslev, 1986)

$$Y_t = \sigma_t \eta_t \tag{A.15a}$$

$$\sigma_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \sigma_{t-1}^2. \tag{A.15b}$$

In this case, the volatility

$$\sigma_t^2 = Var\left(Y_t | Y_t, Y_{t-1}, ...\right) \tag{A.16}$$

is the conditional variance of Y_t given the past $(Y_{t-u}, u \geq 1)$ up to time $t-1$. The model (A.15) belongs to the class of observation-driven volatility models (Cox, 1981; Francq and Zakoian, 2019).

ii) Standard stochastic volatility model

When $\psi_t^2 = \log \sigma_t^2$, $\alpha = 0$, $(\omega, \beta, \varphi) \in \mathbb{R}^2 \times [0, \infty)$, and (e_t) is real-valued with mean 0 and unit variance, the model reduces to the standard stochastic volatility (SV) model (Taylor, 1982-1986), which simply writes as

$$Y_t = \sigma_t \eta_t \tag{A.17a}$$

$$\log\left(\sigma_t^2\right) = \omega + \beta \log\left(\sigma_{t-1}^2\right) + \varphi e_t \tag{A.17b}$$

In this case, σ_t^2 is no longer the conditional variance of Y_t given past process (Y_t) up to time $t-1$, i.e. in the sense of (A.16). However, if (η_t) and (e_t) are independent/uncorrelated then

$$E\left(Y_t|\sigma_t\right) = E\left(Y_t|e_t, e_{t-1}, ...\right) = \sigma_t E\left(\eta_t\right) = 0, \tag{A.18}$$

so (Y_t) is a $\mathscr{F}_t^e$-martingale difference ($\mathscr{F}_t^e = \sigma\left\{e_{t-u}, u \geq 0\right\}$ being the $\sigma-$algebra generated by $\left\{e_{t-u}, u \geq 0\right\}$), and

$$E\left(Y_t^2|\sigma_t\right) = \sigma_t^2 E\left(\eta_t^2\right) = \sigma_t^2 = Var\left(Y_t|e_t, e_{t-1}, ...\right), \tag{A.19}$$

and therefore σ_t^2 is the conditional variance of Y_t given the past and present values of (e_t) up to time t. Even if e_t is not observable, the variance $\sigma_t^2 = Var\left(Y_t|e_t, e_{t-1}, ...\right)$ can be estimated. So in this case, the name volatility, in the sense of the conditional variance for σ_t^2, makes sense.

Note that when η_t and e_t are correlated/dependent, as it happens for the SV model with leverage (cf. Jacquier et al, 2004), (Y_t) is no longer a martingale difference in the sense of (A.18) and σ_t^2 is not the conditional variance, neither in the sense of (A.16) nor in the sense (A.19). Finally, the SV model (A.17) belongs to the class of *parameter-driven* models (Cox, 1981).

iii) Now-Casting GARCH model

When $\psi_t^2 = \log \sigma_t^2$, $\alpha = 0$, $(\omega, \beta, \varphi) \in \mathbb{R}^3$, and $e_t = \log\left(\eta_t^2\right) - E\left(\log\left(\eta_t^2\right)\right)$ is real-valued with mean zero and a symmetric distribution, the model (A.14) reduces to the now-casting GARCH (NC-GARCH) model of Breitung and Hafner (2016), which is given by

$$Y_t = \sigma_t \eta_t \tag{A.20a}$$

$$\log \sigma_t^2 = \omega + \beta \log \sigma_{t-1}^2 + \varphi\left(\log\left(\eta_t^2\right) - E\left(\log\left(\eta_t^2\right)\right)\right). \tag{A.20b}$$

Since in this model, the volatility innovation term $e_t = \log\left(\eta_t^2\right) - E\left(\log\left(\eta_t^2\right)\right)$ is fully dependent with the model innovation term η_t, the so-called volatility σ_t^2 cannot be the conditional variance of Y_t

neither in the sense of (A.16) nor in the sense of (A.19), i.e. even given the past and present values of (e_t) up to time t. Taking

$$\log h_t := \omega + \beta \log \sigma_{t-1}^2,$$

it follows from the second equation of the model (A.20) that

$$\log \sigma_t^2 = \log h_t + \varphi e_t,$$

Hence, Breitung and Hafner's model (A.20) becomes

$$\begin{aligned} \log Y_t^2 &= \log \sigma_t^2 + \log \eta_t^2 \\ \log \sigma_t^2 &= \log h_t + \varphi e_t \end{aligned}$$

where, since $E(e_t) = 0$, it holds

$$E\left(\log \sigma_t^2 | \sigma_{t-1}^2\right) = \log h_t.$$

Hence, $\log h_t = \log \sigma_t^2 - \varphi e_t$ so the NC-GARCH model (A.20) writes as

$$Y_t^2 = h_t \eta_t^2 \exp\left(\varphi e_t\right) \quad \text{(A.21a)}$$

$$\log h_t = \omega + \beta \log h_{t-1} + \beta \varphi e_{t-1} \quad \text{(A.21b)}$$

which is a non-MEM exponential GARCH(1.1). Thus, the model (A.20) written as (A.21) can be seen as an observation-driven volatility model.

iv) Real Time GARCH model

When $\psi_t = \lambda_t$, $(\omega, \alpha, \beta, \varphi) \in (0, \infty) \times [0, \infty)^3$, $e_t = \eta_t^2$, and η_t has a symmetric density, the model (A.14) reduces to the RT-GARCH of Smetanina (2017), which is given by

$$Y_t = \lambda_t \eta_t \quad \text{(A.22a)}$$

$$\lambda_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \lambda_{t-1}^2 + \varphi \eta_t^2. \quad \text{(A.22b)}$$

Since the volatility innovation term $\varphi \eta_t^2$ is fully dependent with the model innovation term η_t, the so-called volatility λ_t^2 cannot be the conditional variance of Y_t neither in the sense of (A.16) nor in the sense of (A.19). This resembles the case of the now-casting GARCH of Breitung and Hafner (2016). Taking $\delta_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \lambda_{t-1}^2$ so that $\lambda_t^2 = \delta_t^2 + \varphi \eta_t^2$ and $h_t = \delta_t^2 + \varphi E\left(\eta_t^4\right)$, the RT-GARCH model can be written as a non-multiplicative EGARCH model as follows

$$Y_t^2 = h_t \eta_t^2 + \varphi \eta_t^2 \left(\eta_t^2 - E\left(\eta_t^4\right)\right) \tag{A.23a}$$

$$h_t = Var\left(Y_t \middle| \mathscr{F}_{t-1}\right) \tag{A.23b}$$

$$h_t = \omega_1 + \alpha Y_{t-1}^2 + \beta h_{t-1} + \phi \eta_{t-1}^2 \tag{A.23c}$$

$$\lambda_t^2 = h_t + \varphi \left(\eta_t^2 - E\left(\eta_t^4\right)\right) \tag{A.23d}$$

where $\omega_1 = \omega + \varphi E\left(\eta_t^4\right)(1-\beta)$ and $\phi = \beta\varphi$. Note that (A.22) written in the form of (A.23) turns out to be an observation-driven model.

v) Time-varying GARCH model with decomposed volatility

When $\psi_t^2 = \sigma_t^2 = g_{n,t}\delta_t^2$ and $\delta_t^2 = \omega + \alpha \frac{Y_{t-1}^2}{g_{n,t-1}} + \beta \delta_{t-1}^2$, where $g_{n,t}$ is a given deterministic function, then model (A.14) becomes a time-varying GARCH (TV(k)-GARCH$(1,1)$) model as proposed by Amado and Teräsvirta (2013, 2017) and is given by

$$Y_t = \sigma_t \eta_t \qquad 1 \le t \le n \tag{A.24a}$$

$$\sigma_t^2 = g_{n,t}\delta_t^2 \tag{A.24b}$$

$$\delta_t^2 = \omega + \alpha \frac{Y_{t-1}^2}{g_{n,t-1}} + \beta \delta_{t-1}^2. \tag{A.24c}$$

The volatility $\sigma_t^2 = Var\left(Y_t | Y_{t-1}, Y_{t-2}, \ldots\right)$ is decomposed into the multiplication of two components: The first one, the long-term/nonstationary component $g_{n,t}$ is a deterministic function given, for example, in terms of logistic transition functions as follows

$$g_{n,t} = a_0 + a_1 \left(1 + \exp\left(-\gamma \prod_{l=1}^{k} \left(\tfrac{t}{n} - c_l\right)\right)\right)^{-1},$$

where the corresponding parameters a_0, a_1, γ, $(c_l)_{l=1,k}$ are known, respectively, as the intercept, size, speed, and locations, and $\frac{t}{n} \in [0,1]$ is called the transition variable. The second one, the stationary short-term/rescaled component δ_t^2 is a standard GARCH volatility in terms of $\frac{Y_{t-1}^2}{g_{n,t-1}}$ and δ_{t-1}^2. Other forms of the transition variable and the transition logistic functions can be exhibited, see Amado and Teräsvirta (2013, 2017) and Campos-Martins and Sucarrat (2024). Note that model (A.24) is an observation-driven volatility model.

vi) General time-varying GARCH model

Taking $\psi_t = \sigma_t$ with

$$\sigma_t^2 = \omega\left(\tfrac{t}{n}\right) + \alpha\left(\tfrac{t}{n}\right) Y_{t-1}^2 + \beta\left(\tfrac{t}{n}\right)\sigma_{t-1}^2, \quad 1 \leq t \leq n,$$

where $\omega(.)$, $\alpha(.)$, and $\beta(.)$ are non-negative deterministic functions, model (A.14) becomes a tvGARCH$(1,1)$ model and is given explicitly by

$$Y_t = \sigma_t \eta_t \qquad\qquad 1 \leq t \leq n \tag{A.25a}$$

$$\sigma_t^2 = \omega\left(\tfrac{t}{n}\right) + \alpha\left(\tfrac{t}{n}\right) Y_{t-1}^2 + \beta\left(\tfrac{t}{n}\right)\sigma_{t-1}^2. \tag{A.25b}$$

Model (A.25) was proposed by Dahlhauss and Subba Rao (2006) in the case $\beta(.) = 0$ and extended to the GARCH case by Rohan and Ramanathan (2013). The tvGARCH model (A.25) belongs to the class of observation-driven models. The solution of (A.25) is nonstationary but locally stationary in the sense of Dahlhauss (1997).

vii) Present-regime Markov Switching GARCH model

In (A.14), set $\psi_t = \sigma_t(S_t)$ with S_t being a finite homogeneous stationary and ergodic Markov chain with a transition probability $P_{ij} = P(S_t = j | S_t = i)$ and a stationary distribution $P(S_t = j) = \pi_j$, $i, j \in \{1, ..., K\}$. If $(A.14b)$ is replaced by the following recursion

$$\sigma_t^2(s) = \omega_s + \alpha_s Y_{t-1}^2 + \beta_s \sigma_{t-1}^2(s), \quad s \in \{1, ..., K\}$$

then model (A.14) becomes a Markov Switching GARCH (MS-GARCH$_K(1,1)$) model in the sense of Haas et al (2004a). This model is also called present-regime MS-GARCH (cf. Aknouche and Francq, 2022) and is given by

$$Y_t = \sigma_t(S_t)\eta_t \tag{A.26a}$$

$$\sigma_t^2(s) = \omega_s + \alpha_s Y_{t-1}^2 + \beta_s \sigma_{t-1}^2(s). \tag{A.26b}$$

Model (A.26) is also called a finite Markov mixture GARCH and can be seen as a parameter-driven model. In the particular case where (S_t) is iid, the model (A.26) is called iid mixture GARCH (cf. Haas et al, 2004b). Note finally that whenever (S_t) and (η_t) are independent,

$$\begin{aligned}\sigma_t^2 \quad : & = \sigma_t^2 (S_t) \\ & = Var\left(Y_t | \mathscr{F}_{t-1}^{Y,S}\right)\end{aligned}$$

can be seen as the conditional variance of Y_t given $\mathscr{F}_{t-1}^{Y,S} = \sigma\{(Y_{t-u}, S_{t-u+1}), u \geq 1\}$, the σ-algebra generated by the past of (Y_t) up to time $t-1$, and the past and present of (S_t) up to time t.

5 Details of obtaining (2.7)-(2.9) and (3.10b)

First, we have

$$\begin{aligned}Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) &= E\left(\left(\sigma_t^2 - \delta_t^2\right)^2 | \mathscr{F}_{t-1}^Y\right) \\ &= E([\omega_t - \omega_0 + \sum_{i=1}^q (\alpha_{it} - \alpha_{0i}) Y_{t-i}^2 + \sum_{j=1}^p (\beta_{jt} - \beta_{0j}) \delta_{t-j}^4]^2 | \mathscr{F}_{t-1}^Y) \\ &= E\left((\omega_t - \omega_0)^2\right) + \sum_{i=1}^q E(\alpha_{it} - \alpha_{0i})^2 Y_{t-i}^4 + \sum_{j=1}^p E(\beta_{jt} - \beta_{0j})^2 \delta_{t-j}^4 \\ &= \sigma_\omega^2 + \sum_{i=1}^q \sigma_{\alpha_i}^2 Y_{t-i}^4 + \sum_{j=1}^p \sigma_{\beta_j}^2 \delta_{t-j}^4. \end{aligned} \tag{A.27}$$

Hence,

$$\begin{aligned}Var\left(Y_t^2 | \mathscr{F}_{t-1}^Y\right) &= E\left(\left(Y_t^2 - E\left(Y_t^2 | \mathscr{F}_{t-1}^Y\right)\right)^2 | \mathscr{F}_{t-1}^Y\right) \\ &= E\left(\left(\sigma_t^2 \varepsilon_t^2 - \delta_t^2\right)^2 | \mathscr{F}_{t-1}^Y\right) \\ &= E\left(\varepsilon_t^4\right) E\left(\sigma_t^4 | \mathscr{F}_{t-1}^Y\right) - 2E\varepsilon_t^2 \delta_t^2 E\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) + \delta_t^4 \\ &= \kappa\left(Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) + \left(E\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right)\right)^2\right) - 2\delta_t^4 + \delta_t^4 \\ &= \kappa Var\left(\sigma_t^2 | \mathscr{F}_{t-1}^Y\right) + (\kappa - 1)\delta_t^4. \end{aligned} \tag{A.28}$$

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

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

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

$$\left(Y_t^2 - \delta_t^2\right)^2 = Var\left(Y_t^2|\mathscr{F}_{t-1}^Y\right) + e_t. \tag{A.29}$$

Then, from (A.27) and (A.28) we have $Var\left(\sigma_t^2|\mathscr{F}_{t-1}^Y\right) = M_t'\Lambda_0$ and

$$Var\left(Y_t^2|\mathscr{F}_{t-1}^Y\right) = \kappa M_t'\Lambda_0 + (\kappa-1)\,\delta_t^4,$$

so (A.29) becomes

$$\begin{aligned}\left(Y_t^2 - \delta_t^2\right)^2 &= Var\left(Y_t^2|\mathscr{F}_{t-1}^Y\right) + e_t \\ &= \kappa M_t'\Lambda_0 + (\kappa-1)\,\delta_t^4 + e_t\end{aligned}$$

where

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

Hence

$$\left(Y_t^2 - \delta_t^2\right)^2 - (\kappa-1)\,\delta_t^4 = \kappa M_t'\Lambda_0 + e_t$$

so that

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

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

6 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
	Mean	0.0113	0.1527	0.7898	0.0097	0.2929	0.2772
1000	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
	Mean	0.0103	0.1507	0.7980	0.0082	0.3598	0.2699
3000	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
	Mean	0.0102	0.1509	0.7983	0.0086	0.3325	0.2522
5000	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 $\beta_t \sim \mathcal{IG}(\beta_0, 2)$.

		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	0.1500	0.8000
	Mean	0.0111	0.1513	0.7901	0.0109	0.1482	0.7770
1000	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
	Mean	0.0105	0.1523	0.7950	0.0100	0.1480	0.8099
3000	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
	Mean	0.0103	0.1504	0.7970	0.0099	0.14971	0.8065
5000	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 $\beta_t \sim \mathcal{P}(\beta_0)$.

		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	0.0001	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 of the main paper (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.

Since the predictive volatility of the RC-GARCH model is exactly that of the standard GARCH model which is well known, our main interest will be in the filtered volatility which also incorporates the current observation and is obtained in a closed form. So a crucial question is to see if this filtered volatility can improve the forecasting ability of the predictive GARCH volatility. We, thus, first assess the in-sample volatility forecasting ability of the predictive and filtered volatilities given respectively by (2.5d) and (3.23). We use the true parameters since the estimated parameters could alter the vision about the ability of each kind of volatility. We generate 1000 NIG RC-GARCH series with sample size 3000 for each of which we obtain the predictive volatility $\left(\delta_t^2\right)$ and the filtered volatility $\left(\varrho_t^2\right)$. The true parameters are chosen so as to be close to those estimated in the real application (cf. Section 4 of the main paper). Since the actual volatility $\left(\sigma_t^2\right)$ is unavailable in practice, we use the squared series Y_t^2 as a proxy (Table 4, panel (a)). We also use the true volatility σ_t^2 in comparison (Table 4, panel (b)) We compute for each volatility and each replication the following criteria: i) the mean square forecast error, with proxy MSFE$=\frac{1}{n}\sum_{t=1}^{n}(Y_t^2-h_t)^2$ and with true volatility $\text{MSFE}^{\text{T}}=\frac{1}{n}\sum_{t=1}^{n}(\sigma_t^2-h_t)^2$, and ii) the mean absolute forecast error with proxy $\text{MAFE}=\frac{1}{n}\sum_{t=1}^{n}\left|Y_t^2-h_t\right|$ and with ture volatility $\text{MAFE}^{\text{T}}=\frac{1}{n}\sum_{t=1}^{n}\left|\sigma_t^2-h_t\right|$, where the generic $h_t\in\left\{\delta_t^2,\varrho_t^2\right\}$. Then, the sample mean of each one of MSFEs, MSFE$^{\text{T}}$s, MAFEs, and MAFE$^{\text{T}}$s over the 1000 replications are obtained and are denoted, respectively, by MMSFE, MMSFE$^{\text{T}}$, MMAFE, and MMSFE$^{\text{T}}$ (cf. Table 4 panels (a) and (b)). The best criteria are reporteed in bold. It can be observed from Table 4 that the filtered volatility ϱ_t^2 gives the best MMSFE, MMSFE$^{\text{T}}$, MMAFE, and MMSFE$^{\text{T}}$ compared to the predictive volatility δ_t^2. Moreover, as expected, the criteria computed on the basis of the true volatility σ_t^2 are smaller than those obtained using the proxy Y_t^2.

	(a)	(a)	(b)	(b)
	MMSFE	MMAFE	MMSFE$^{\text{T}}$	MMAFE$^{\text{T}}$
Predictive volatility δ_t^2	0.000137	0.001809	0.000101	0.001776
Filtered volatility ϱ_t^2	0.000095	0.001193	0.000061	0.001163

Table 4. In-sample forecasting comparison between the predictive and filtered volatilities for 1000 NIG RC-GARCH series with $n=3000$, $\omega_0=0.00001$, $\lambda_\omega=$ 1e-8, $\alpha_0=0.05$, $\lambda_\alpha=0.00005$, $\beta_0=0.94$, $\lambda_\beta=0.65$.
(a) Using the square Proxy X_t^2. (b) Using the true volatility σ_t^2.

We now assess the out-of-sample forecasting ability of the predictive and filtered volatilities. We generate 1000 replications of the NIG RC-GARCH model with sample size $n=3000$ and the same parameters as in Table 4. For each replication we compute the one-step ahead predictive volatility δ_t^2 for $t\in\{n_c+1,...,n\}$ where n_c $(1<n_c<n)$ is the sample size of the truncated series and belongs to $\{2000,2200,2400,2600,2800\}$. We compute the filtered volatility ϱ_t^2 for $t\in\{n_c+1,...,n\}$ and also obtain the modified filtered volatility ϱ_t^{*2} in which the future observations (Y_{n_c+j}) are estimated by their conditional means $\widehat{Y}_{n_c+j}=E\left(Y_{n_c+j}\middle|\mathscr{F}_{n_c}^{Y}\right)=0$ $(1\le j\le n-n_c)$. Then, we obtain the

three above criteria for each replication: i) the MSFE$=\frac{1}{n-n_c}\sum_{t=n_c+1}^{n}(Y_t^2-h_t)^2$, ii) the MAFE $=\frac{1}{n-n_c}\sum_{t=n_c+1}^{n}\left|Y_t^2-h_t\right|$, and iii) the MQLI$=\frac{1}{n-n_c}\sum_{t=n_c+1}^{n}(\log h_t+\frac{Y_t^2}{h_t})$, as well as their sample means over the replications, giving MMSFE, MMAFE and MMQLI (cf. Table 5).

	n_c	2000	2200	2400	2600	2800
δ_t^2	MMSFE	0.000199	0.000361	0.000227	0.000653	0.000601
	MMAFE	0.002997	0.003560	0.004221	0.005937	0.010391
	MMSFE$^{\mathrm{T}}$	0.000152	0.000199	0.000206	0.000410	0.000604
	MMAFE$^{\mathrm{T}}$	0.002949	0.003473	0.004120	0.005825	0.010319
ϱ_t^2	MMSFE	0.000112	0.000218	0.000103	0.000388	**0.000244**
	MMAFE	0.001938	**0.002298**	**0.002706**	**0.003801**	**0.006587**
	MMSFE$^{\mathrm{T}}$	**0.000078**	**0.000097**	**0.000087**	**0.000217**	**0.000249**
	MMAFE$^{\mathrm{T}}$	**0.002048**	**0.002359**	**0.002740**	**0.003791**	**0.006519**
ϱ_t^{*2}	MMSFE	**0.000056**	**0.000091**	**0.000084**	**0.000155**	**0.000244**
	MMAFE	**0.001792**	0.002547	0.002964	0.004098	0.007011

Table 5. Out-of-sample forecasting comparison between the predictive and filtered volatilies for 1000 NIG RC-GARCH series with $n=3000$, $\omega_0=0.00001$, $\lambda_\omega=$ 1e-8, $\alpha_0=0.05$, $\lambda_\alpha=0.00005$, $\beta_0=0.94$, and $\lambda_\beta=0.65$.

It can be seen from Table 5 that the filtered volatilities ϱ_t^2 and ϱ_t^{*2} outperform the predictive volatility δ_t^2 for all criteria and all time cut n_c even if the true observations are replaced by their predictions. Regarding MMAFE, ϱ_t^2 gives better volatility forecasts while with respect to MMSFE, ϱ_t^{*2} surprisingly outperforms ϱ_t^2. Overall, the filtered volatilities improve the forecasting ability of the predictive volatility (which is that of the standard Engle-Bollerslev GARCH model), even when the true observations are replaced by their predictions.

Finally, to see the effect of the absence of coefficient randomness on the QMLE and WLS estimates, we generate 1000 Normal GARCH replications with sample size $n\in\{1000,3000,5000\}$ and $(\omega_t,\alpha_t,\beta_t)=(\omega_0,\alpha_0,\beta_0)$. For each replication we obtain the QML and WLS estimates and then the mean, StD and ASE of estimates over the 1000 replications.

		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	0.0000	0.0000	0.0000
1000	Mean	0.0112	0.1485	0.7927	0.0004	0.0035	0.0043
	StD	0.0038	0.0289	0.0401	0.0008	0.0012	0.0019
	ASE	0.0042	0.0294	0.0415	0.0007	0.0010	0.0018
3000	Mean	0.0105	0.1502	0.7968	0.0002	0.0018	0.0028
	StD	0.0018	0.0185	0.0226	0.0004	0.0010	0.0013
	ASE	0.0021	0.0167	0.0220	0.0003	0.0008	0.0071
5000	Mean	0.0102	0.1495	0.8010	0.0000	0.0004	0.0018
	StD	0.0013	0.0140	0.0156	0.0002	0.0003	0.0009
	ASE	0.0015	0.0127	0.0166	0.0001	0.0006	0.0045

Table 6. QMLE and WLSE results for 1000 RC-GARCH(1,1) series with sample size n, and degenerated coefficients $(\omega_t, \alpha_t, \beta_t) = (\omega_0, \alpha_0, \beta_0)$.

The same conclusions can be drawn from Table 6 as was the case with the previous simulations. The parameters are well estimated, and in particular the variance parameters are close to zero, and their ASE and StD are also close to zero as the sample size increases. Moreover, the mean parameters are very well estimated since the QMLE becomes in this case the MLE. Note finally that unreported simulations showed that in most replications (more that 95% of cases) the global test of randomness rejected the randomness of parameters at any reasonable level, even with 0.01 significance level.

7 Application to the Intel stock returns

We fit the RC-GARCH$(1, 1)$ model to the daily returns of the Intel stock (RINTEL) spanning 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 S.1).

The parameter estimates are reported in Table S.1. Conclusions similar to those for the CISCO

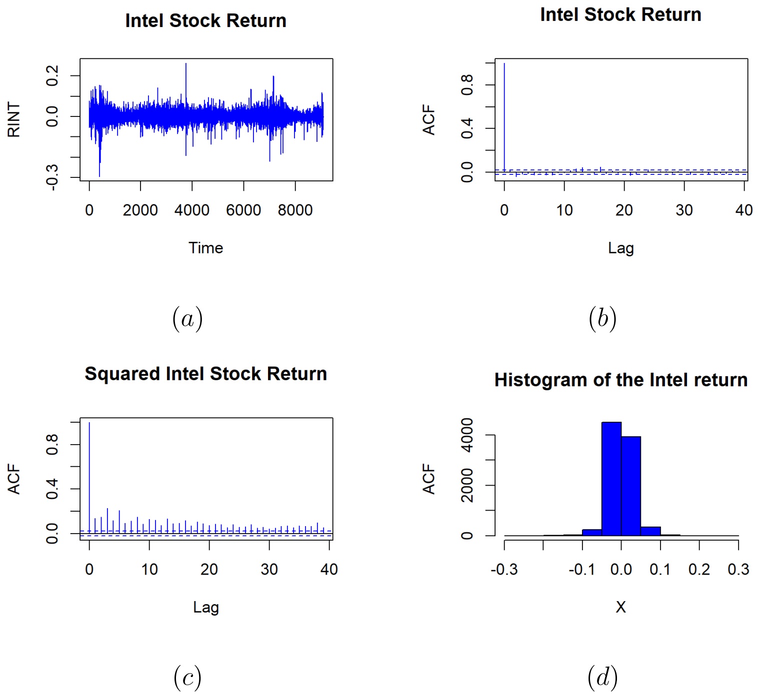


Figure S.1 Intel stock return series: (a) The RINTEL series; (b) sample autocorrelation (c) sample autocorrelation of squares, (d) histogram.

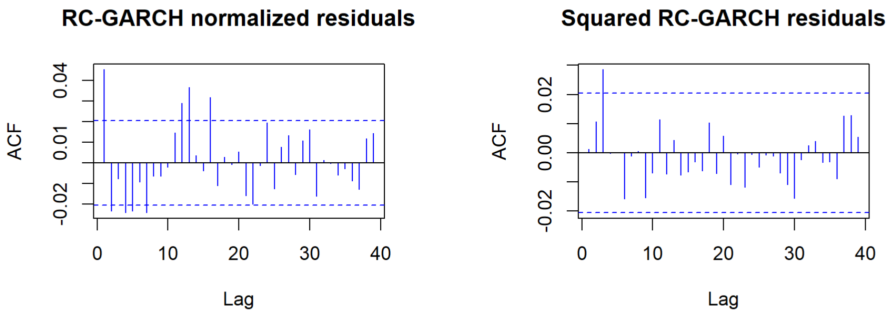
application can be drawn: the estimated model is highly persistent, has a finite second moment and an infinite fourth moment.

	ω_t	α_{1t}	β_{1t}		
QMLE	$\widehat{\omega}_n$	$\widehat{\alpha}_{1n}$	$\widehat{\beta}_{1n}$	$\widehat{\alpha}_{1n} + \widehat{\beta}_{1n}$	
	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 S.1. QML and WLS estimates for the RC-GARCH(1.1); Intel series.

The sample autocorrelations of the normalized residuals and squared normalized residuals $\widehat{\varepsilon} := \frac{Y_t}{\widehat{\varrho}_t}$ (see Figure S.2 (a)-(b)) look like an independent noise as the sample autocorrelations of residuals and their squares do not show significant spikes.

As in the CISCO application, the variance parameters are all significant as confirmed by Table S.2. The hypothesis of randomness of each coefficient cannot be rejected at any reasonable level and the value $W_n = 2.8291$ of the global Wald statistic suggests that the randomness of the model cannot



(c) (d)

Figure S.2. (a) Residuals, (b) ACF of squared residuals; RINTEL series.

be rejected at reasonable level.

$W_{\omega,n}$	$W_{\alpha,n}$	$W_{\beta,n}$
0.2622	2.0016	2.6501

Table S.2. Individual Wald Statistics for testing

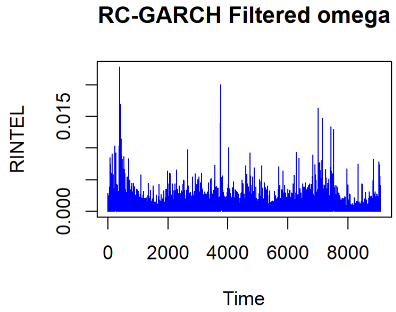
the randomness of coefficients; RINTEL series.

The plots of the posterior means of the random coefficients as displayed in Figure S.3 show that each posterior mean exhibits a proper behavior that is more pronounced than for the RC-GARCH coefficients in the CISCO empirical application.

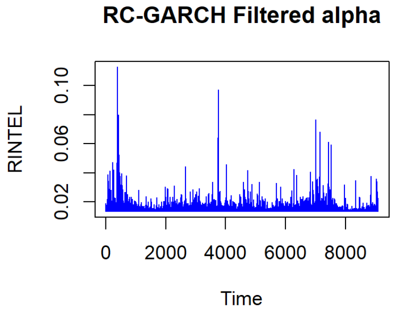
The graphs of the simulated random coefficients $(\omega_t, \alpha_t, \beta_t)$ from the estimated RC-GARCH(1.1) model using the RINTEL series as well as the random persistence $\alpha_t + \beta_t$ are plotted in Figure S.4. It can be seen that the generated α_t and β_t coefficients as well as their persistence can be largely greater than the unity.

The predictive and smoothed volatilities are plotted in Figure S.5 (panel (a) and panel (b), respectively). The conditional excess kurtosis $\widehat{\kappa}_t - 3$ (Figure S.5 panel (d)) seems to be in accordance with the estimated predictive volatility of the model. Also, the volatility of volatility plotted in panel (c) of Figure S.5, has a consistent behavior with the predictive volatility.

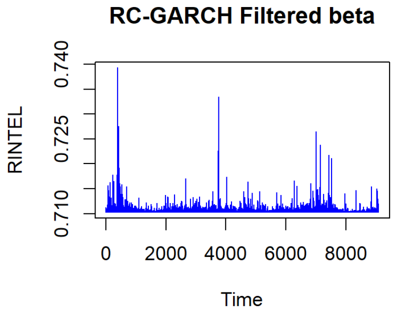
Figure S.7 shows the probability integral transform (PIT) of the RINTEL series $(Y_t)_{1\leq t\leq n}$ with respect to three conditional distributions: (a) the standard Normal GARCH$(1,1)$ with Normal conditional distribution $\mathcal{N}(0, \widehat{\delta}_t^2)$, (b) The RC-GARCH(1.1) with Normal Inverse Gaussian conditional distribution $Y_t \sim \mathcal{NIG}(\widehat{\Delta}_t \widehat{\delta}_t^{-2}, 0, \widehat{\Delta}_t, 0)$, and (c) the MS-GARCH$_2$(1.1) model with Normal mixings (cf. Haas et al, 2004a; Ardia et al, 2019). The estimation results for the tvGARCH model are not



(a)



(b)



(c)

Figure S.3. Posterior means of the random coefficients $(\omega_t, \alpha_t, \beta_t)$:
(a) $E(\omega_t|Y_t)$, (b) $E(\alpha_t|Y_t)$, (c) $E(\beta_t|Y_t)$, $1 \leq t \leq n$; Intel series.

available for the RINTEL series since the tvGARCH package give error message. The parameter estimates of MS-GARCH$_2$(1.1) model as well as their ASEs in parenthesis are reported in Table S.3 while the MS-GARCH$_2$(1.1) volatility plot is displayed in Figure S.6.

$\widehat{P}_{11}$	$\widehat{P}_{21}$	$\widehat{\omega}_1$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\omega}_2$	$\widehat{\alpha}_2$	$\widehat{\beta}_2$
0.9190 (0.0738)	0.6699 (0.0541)	0.0000 (0.0000)	0.0192 (0.0063)	0.9722 (0.0016)	0.0003 (0.0001)	0.1703 (0.2250)	0.8248 (0.0066)

Table S.3. Estimated MS-GARCH$_2$(1.1) model; the Intel returns.

Likelihood: 20476.094, AIC $= -40936.1889$, BIC $= -40879.2633$.

It can be seen from Figure S.7 that the returns fit better with the NIG RC-GARCH model and the MS-GARCH model, followed by the Normal GARCH model. However all PITs for all models show a kind of mixture distribution due to significant picks at the probability 0.5.

Table S.4 shows the MSFE, MAFE and MQLI computed for the three models as did for the CISCO application. It can be observed that the filtered volatility $\widehat{\varrho}_t^2$ provides the best MSFE, MAFE, and MQLI compared to the GARCH(1.1) predictive volatility $\widehat{\delta}_t^2$ and also the MS-GARCH$_2$ $(1,1)$ volatility.

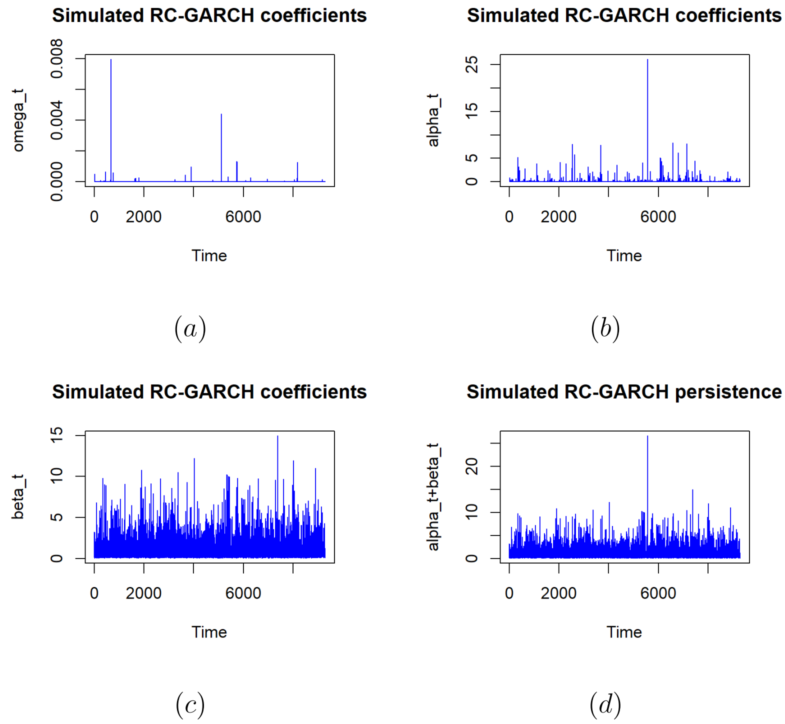


Figure S.4 Simulated random coefficients $(\omega_t, \alpha_t, \beta_t)$ and the persistence $\alpha_t + \beta_t$ from the estimated RC-GARCH $(1,1)$ model using the RINTEL series.

In addition, the MS-GARCH volatility outperforms the GARCH(1.1) regarding all criteria. However, all criteria for the three volatility models are quite close to each other.

	MSFE	MSAE	MQLI
Predictive volatility $\widehat{\delta}_t^2$	7.2e-06	0.00089	-6.4391
Filtered volatility $\widehat{\varrho}_t^2$	**4.3e-06**	**0.00060**	$\mathbf{-6.9334}$
MS-GARCH$_2$ $(1,1)$ volatility	5.2e-06	0.00087	-6.2829

Table S.4. In-sample forecast comparison of the GARCH(1.1), the RC-GARCH(1.1) the MS-GARCH(1.1) volatilities.

Regarding the out-of-sample forecasting ability of the predictive volatility, Table S.5 shows the computed values of the above criteria for the four models and for various truncated series with sample size $n_c \in \{6000, 7000, 8000, 8400, 8800\}$. Conclusions similar to the CISCO application can be made here as well. The filtered volatilities $(\widehat{\varrho}_t^2, \widehat{\varrho}_t^{*2})$ based on true or predictive observations provide the best MSFE, MAFE, and MQLI for all n_c. Moreover, the forecasts of $\widehat{\varrho}_t^{*2}$ are better that those of $\widehat{\varrho}_t^2$, and

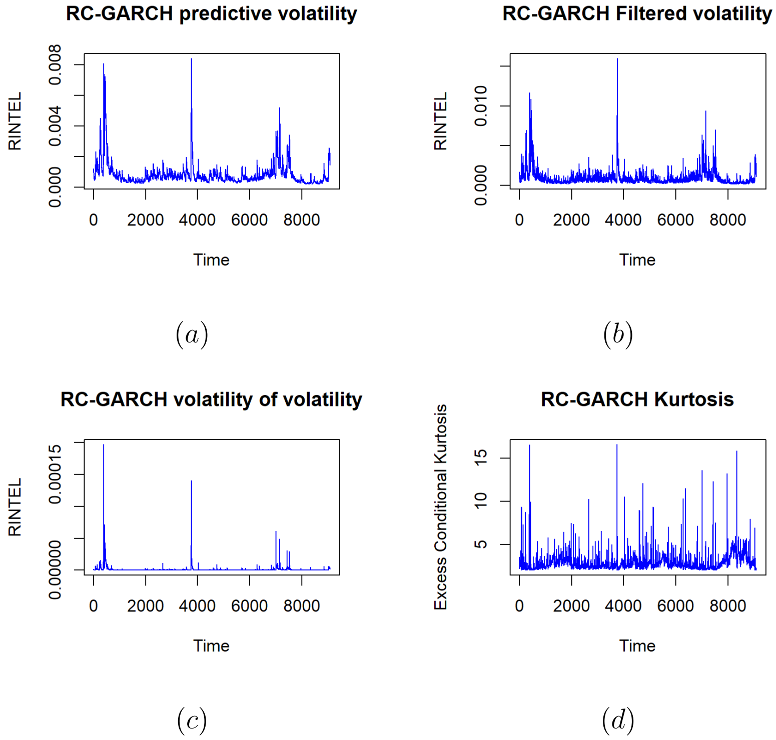


Figure S.5. Estimated RC-GARCH for RCISCO series. (a) Predictive volatility, (b) filtered volatility, (c) volatility of volatility, (d) conditional excess kurtosis.

the MS-GARCH volatility outperforms the GARCH volatility for all criteria and all n_c.

	n_c	6000	7000	8000	8400	8800
	MSFE	4.20e-06	5.13e-06	1.26e-06	1.64e-06	3.5e-06
$\widehat{\delta}_t^2$	MAFE	0.00085	0.00088	0.00052	0.00062	0.00102
	MQLI	−6.3331	−6.4250	−6.9263	−6.7295	−6.0596
	MSFE	2.95e-06	3.62e-06	8.09e-07	1.06e-06	2.25e-06
$\widehat{\varrho}_t^2$	MAFE	0.00067	0.00069	0.00039	0.00048	0.00078
	MQLI	−6.6557	−6.7462	−7.2302	−7.0392	−6.4440
	MSFE	**1.75e-08**	**2.22e-08**	**2.05e-08**	**1.67e-08**	**2.12e-08**
$\widehat{\varrho}_t^{*2}$	MAFE	**0.00013**	**0.00015**	**0.00014**	**0.00012**	**0.00012**
	MQLI	**−8.9539**	**−8.8375**	**−8.9041**	**−9.0932**	**−9.1636**
	MSFE	3.53e-06	4.38e-06	1.08e-06	1.27e-06	3.01e-06
MS	MAFE	0.00075	0.00079	0.00049	0.00053	0.00091
	MQLI	−6.5482	−6.6157	−7.0459	−6.9483	−6.2912

Table S.5. Out-of-sample volatility forecasting performance of the filtered volatility, the predictive volatility, and the MS-GARCH volatility (MS). $\widehat{\varrho}_t^2$: filtered volatility using available returns. $\widehat{\varrho}_t^{*2}$: filtered volatility using predictive returns.

Finally, using the model confidence set test of Hansen et al (2011) through the R package "*MSC*" of Bernardi and Catania (2014), the filtered volatilities $(\widehat{\varrho}_t^2, \widehat{\varrho}_t^{*2})$ constitute the superior model for all

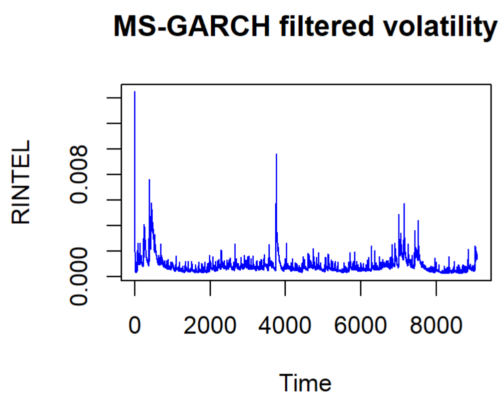


Figure S.6. MS-GARCH$_2$ (1, 1) volatility; Intel returns.

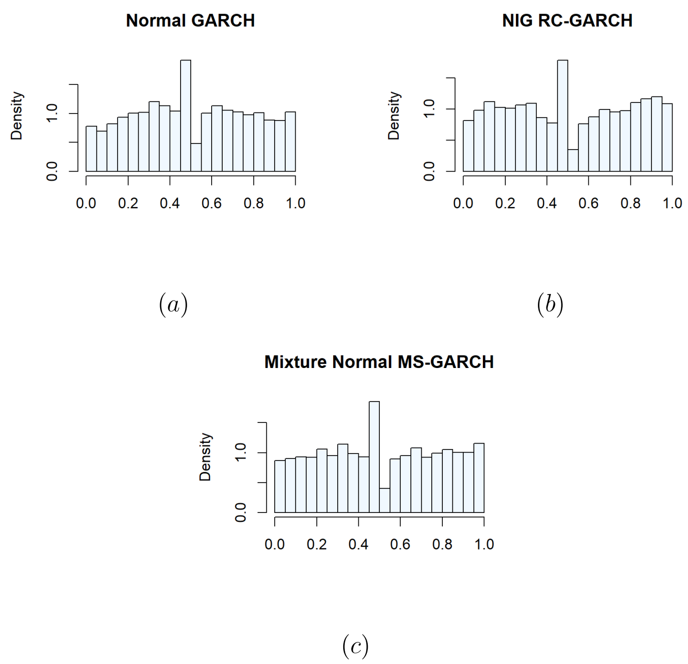


Figure S.7 PITs of the RINTEL with respect to: (a) The Normal GARCH(1.1), (b) the NIG RC-GARCH(1.1), (c) the mixture Normal MS-GARCH$_2$(1.1).

criteria and all n_c.

8 Additional empirical analysis for the CISCO application

8.1 In-sample performance

We now assess the in-sample performance of the RC-GARCH(1,1) model in terms of both model fit and volatility forecasting. We, thus, compare the RC-GARCH(1,1) model with two popular volatility models, namely the two-regime Markov-switching GARCH(1,1) (MS-GARCH$_2$(1,1)) model of Haas et al (2004a) and the time-varying GARCH model of Amado and Teräsvirta (2013, 2014, 2017) with a single logistic transition function (TV(1)-GARCH(1,1)). To this end, we use the MSGARCH R package of Ardia et al (2019) for the MS-GARCH model and the *tvgarch* package of Campos-Martins and Sucarrat (2024) for the tvGARCH model. The two models are described in this Supplementary Material. The choice of the MS-GARCH model is natural, as both the RC-GARCH and MS-GARCH models are mixture GARACH models. As pointed out in the introduction of the main paper, the tvGARCH model could also be viewed as a (nonstationary) regime switching GARCH model but with a deterministic regime sequence. For the MS-GARCH(1,1) model, an instance with two regimes (MS-GARCH$_2$(1,1)) is estimated. For more than two regimes, the estimated transition probabilities of the MS-GARCH$_2$(1,1) are not significant. The same holds for the tvGARCH(1,1) model with a multiple logistic transition function.

Figure S.8 shows the probability integral transform (PIT) of the return series $(Y_t)_{1\leq t\leq n}$ with respect to four conditional distributions: (a) the standard Normal GARCH$(1,1)$ with Normal conditional distribution $\mathcal{N}(0,\widehat{\delta}_t^2)$, (b) the RC-GARCH(1,1) with Normal Inverse Gaussian conditional distribution $Y_t \sim \mathcal{NIG}(\widehat{\Delta}_t\widehat{\delta}_t^{-2},0,\widehat{\Delta}_t,0)$, (c) the MS-GARCH$_2$(1,1) model with Normal mixings (cf. Haas et al, 2004a; Ardia et al, 2019), and finally, (d) the tvGARCH(1,1) model with a Normal conditional distribution $\mathcal{N}(0,\widehat{\delta}_t^2\widehat{g}_{n,t})$, where $g_{n,t}=\widehat{\delta}_0+\widehat{\delta}_1\frac{1}{1+\exp\left(\widehat{\gamma}\left(\frac{t}{n}-\widehat{c}\right)\right)}$ is the estimated deterministic time-varying component (cf. Amado and Teräsvirta, 2013-2017; Campos-Martins and Sucarrat, 2024). The parameter estimates of both TV(1)-GARCH(1,1) and MS-GARCH(1,1) models as well as their volatility plots are reported in the Supplementary Material. It can be seen from Figure S.8 that the returns fit well with the NIG RC-GARCH model and the MS-GARCH model (as the corresponding PITs are close to a straight line) followed by the tvGARCH model, and finally by the standard GARCH model. Thus, the PIT suggests that the NIG distribution could be a good model for the CISCO returns.

The PIT results (Figure S.8) are also verified by the in-sample volatility forecasting ability of the

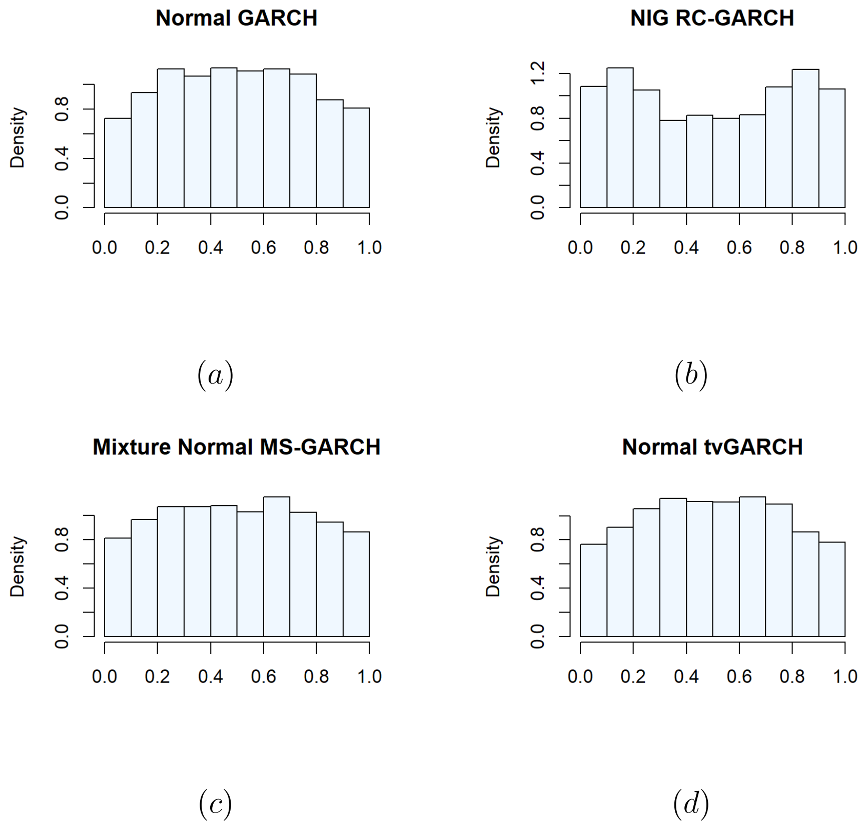


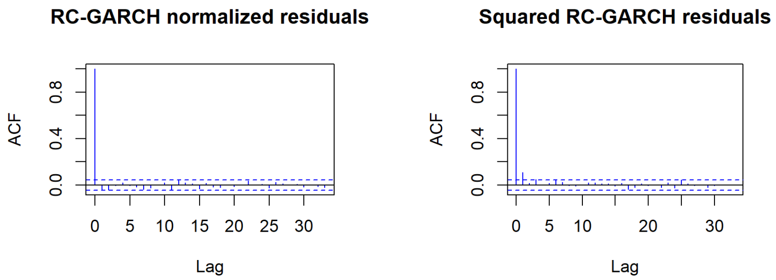
Figure S.8: PITs of the CISCO data with respect to: (a) The Normal GARCH(1,1), (b) the NIG RC-GARCH(1,1), (c) the Normal TV(1)-GARCH(1,1), (d) the mixture Normal $MS-GARCH_2(1,1)$.

four models. Since the actual volatility is unobservable, we use the squared return Y_t^2 as a proxy (as did in the simulation above), which is unbiased and commonly used in the literature (see e.g. Charles and Olivier, 2017), although it is noisy (Lopez, 2001). We compute for each model the three criteria also used in our simulation exercises (see above): i) the mean square forecast error MSFE$=\frac{1}{n}\sum_{t=1}^{n}(Y_t^2-\widehat{h}_t)^2$, ii) the mean absolute forecast error, given by MAFE $=\frac{1}{n}\sum_{t=1}^{n}\left|Y_t^2-\widehat{h}_t\right|$, and iii) the mean QLIKE (cf. Patton, 2011; Aknouche and Francq, 2023) MQLI$=\frac{1}{n}\sum_{t=1}^{n}(\log\widehat{h}_t+\frac{Y_t^2}{\widehat{h}_t})$, where $\widehat{h}_t$ is the estimated volatility generated by each model (cf. Table S.6). Following Patton (2011), the MSFE criterion is robust for forecast comparison if the true volatility is replaced by some consistent proxy such as the squared return. From Table S.6, it can be seen that the filtered volatility $\widehat{\varrho}_t^2$ provides the best MSFE, MAFE, and MQLI compared to the predictive volatility (which is that of a GARCH model), the tvGARCH volatility, and MS-GARCH volatility. The tvGARCH volatility is better than the GARCH and MS-GARCH regarding the MQLI, but the GARCH volatility outperforms the tvGARCH and MS-GARCH regarding the criteria MSFE and MAFE. The MS-GARCH provides the worst (in-sample) volatility forecast, although all criteria for the four models are quite close to each other.

	MSFE	MSAE	MQLI
Predictive volatility $\widehat{\delta}_t^2$	7.22e-06	0.00089	-6.4391
Filtered volatility $\widehat{\varrho}_t^2$	**4.28e-06**	**0.00068**	$\mathbf{-6.6335}$
TV(1)-GARCH (1, 1) volatility	7.28e-06	0.00091	-6.4422
MS-GARCH$_2$ (1, 1) volatility	7.94e-06	0.00102	-6.4026

Table S.6. In-sample forecast comparison of the RC-GARCH(1,1), the MS-GARCH(1,1), and the tvGARCH(1,1).

8.2 Some additional graphs



(a) (b)

Figure S.9. (a) ACF of residuals. (b) ACF of squared residuals; CISCO

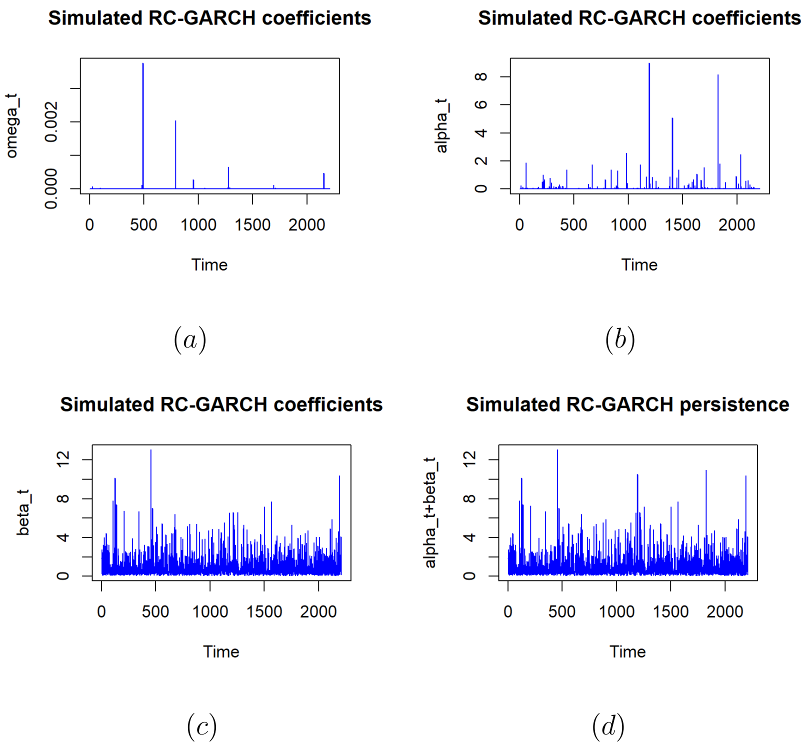


Figure S.10. Simulated random coefficients $(\omega_t, \alpha_t, \beta_t)$ and the persistence $\alpha_t + \beta_t$ from the estimated RC-GARCH(1,1) model based on the CISCO series.

8.3 Estimation results for the TV(1)-GARCH(1.1) model

Using the package "*tvGARCH*" of Campos-Martins and Sucarrat (2024), we estimate a TV(1)-GARCH(1.1) model based on the CiSCO series. The method used is the estimation by parts of Amado and Terasvirta (2013). The parameter estimates are reported in Table S.7. For higher-orders TV(k)-GARCH(1.1) models with $k \geq 2$, the estimated models are not available from the package "tvGARCH", which delivers error messages.

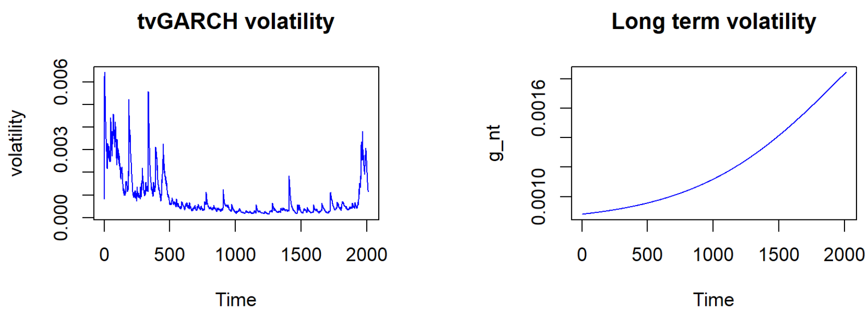
$\widehat{a}_0$	$\widehat{a}_1$	$\widehat{\gamma}$	$\widehat{c}$	$\widehat{\omega}$	$\widehat{\alpha}$	$\widehat{\beta}$
0.0008	0.0020 (0.0007)	3.5391 (1.0112)	0.9955 (0.0649)	0.0108 (0.0102)	0.0761 (0.0586)	0.9126 (0.0642)

Table S.7. Estimated TV(1)-GARCH(1.1) model; CISCO returns.

Likelihood: 4629.672.

The generated volatility by the estimated TV(1)-GARCH(1.1) model is displayed in Figure S.11 (a), while in panel (b) of the same figure we display the graph of the estimated long-term deterministic component

$$\widehat{g}_{n,t} = \widehat{a}_0 + \widehat{a}_1 \left(1 + \exp\left(-\widehat{\gamma}\left(\tfrac{t}{n} - \widehat{c}\right)\right)\right)^{-1}.$$



(a) (b)

Figure S.11. (a) Volatility induced by the TV(1)-GARCH(1.1) model.
(b) Estimated long-term function $\widehat{g}_{n,t}$; CISCO series.

8.4 Estimation results for the MS-GARCH$_2$(1.1) model

On the basis of the CISCO series, the parameter estimates of two-regime MS-GARCH$_2$(1.1) models and their ASEs in parenthesis are reported in Table S.8, while the MS-GARCH$_2$(1.1) volatility plot is displayed in Figure S.12.

$\widehat{P}_{11}$	$\widehat{P}_{21}$	$\widehat{\omega}_1$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\omega}_2$	$\widehat{\alpha}_2$	$\widehat{\beta}_2$
0.9936 (0.0000)	0.0694 (0.0000)	0.0000 (0.0000)	0.0219 (0.0000)	0.9750 (0.0000)	0.0007 (0.0000)	0.0787 (0.0000)	0.8553 (0.0000)

Table S.8. Estimated MS-GARCH$_2$(1.1) model; CISCO returns.

Likelihood: 4724.2557, AIC $= -9432.5114$, BIC $= -9387.6603$.

The volatility generated by the MS-GARCH(1.1) model for the RCISCO series is plotted in Figure S.12.

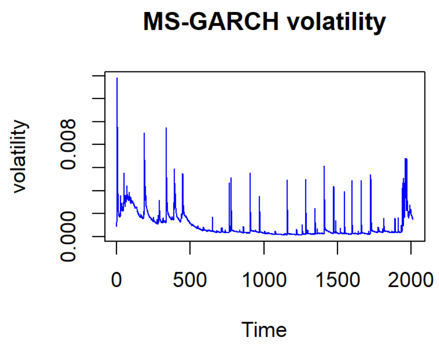


Figure S.12. Volatility induced by the MS-GARCH$_2$(1.1) model.

References

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

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

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

Aknouche A, 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.

Amado C, Teräsvirta T. 2013. Modelling volatility by variance decomposition. *Journal of Econometrics* 175: 142–153.

Amado C, Teräsvirta T. 2014. Modelling changes in the unconditional variance of long stock return series. *Journal of Empirical Finance* 25:15–35.

Amado C, Teräsvirta T. 2017. Specification and testing of multiplicative time-varying GARCH models with applications. *Econometric Reviews* 36: 421–446.

Ardia D, Bluteau K, Boudt K, Catania L, Trottier D. 2019. Markov-Switching GARCH models in R: The MSGARCH package. *Journal of Statistical Software* 91: 1–38.

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

Bernardi M, Catania L. 2014. The model confidence set package for R. *ArXiv:1410. 8504v1.*

Barndorff-Nielsen OE. 1978. Hyperbolic Distributions and distributions on Hyperbolae. *Scandinavian Journal of Statistics* 5: 151–157.

Barndorff-Nielsen OE. 1997. Normal Inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* 24: 1–13.

Barndorff-Nielsen OE, 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 HC, Liu SP. 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, Picard N. 1992. Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* 52: 115–127.

Breitung J, Hafner CM. 2016. A simple model for now-casting volatility series. *International Journal of Forecasting* 32: 1247–1255.

Campos-Martins S, Sucarrat G. 2024. Modeling nonstationary financial volatility with the R package tvgarch. *Journal of Statistical Software* 108: 1–38.

Charles A, Olivier D. 2017. Forecasting crude-oil market volatility: Further evidence with jumps. *Energy Economics* 67: 508–519

Chen M, An HZ. 1998. A note on the stationarity and the existence of moments of the *GARCH* model. *Statistica Sinica* 8: 505–510.

Corsi F, Mittnik S. 2008. The volatility of realized volatility. *Econometric Reviews* 27: 46–78.

Cox DR. 1981. Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics* 8: 93–115.

Dahlhaus R. 1997. Fitting time series models to nonstationary processes. *Annals of Statistics* 25:1–37.

Dahlhaus R, Subba Rao S. 2006. Statistical inference for time-varying ARCH processes. *Annals of Statistics* 34: 1075–1114.

Ding YD. 2020. Weak diffusion limit of real-time GARCH models: the role of current return information. *Cambridge Working Paper in Economics CWPE20112*, University of Cambridge.

Ding YD. 2023. A simple joint model for returns, volatility and volatility of volatility. *Journal of Econometrics* 232: 521–543.

Engle RF. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50: 987–1007.

Engle RF. 2002. New frontiers for ARCH models. *Journal of Applied Econometrics* 17: 425–446.

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

Francq C, Zakoian JM. 2004. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* 10: 605–637.

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

Francq C, Zakoian JM. 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, Zakoian JM. 2019. *GARCH Models, structure, statistical inference and financial applications.* John Wiley & Sons, Ltd., Publication.

Fryzlewicz P, Sapatinas T, Subba Rao S. 2008. Normalized least-squares estimation in time-varying ARCH models. *Annals of Statistics* 36: 742–786.

Ghysels E, Harvey AC, Renault E. 1996. *Stochastic volatility. Handbook of Statistics* 14: 119-191.

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, Paolella M. 2004a. A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics* 2: 493–530.

Haas M, Mittnik S, Paolella M. 2004b. Mixed Normal conditional heteroskedasticity. *Journal of Financial Econometrics* 2: 211–250.

Hamilton JD, Susmel R. 1994. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics* 64: 307–333.

Hansen PR, Lunde A, Nason JM. 2011. The model confidence set. *Econometrica* 79: 453–497.

Jacquier E, Polson NG, Rossi PE. 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics* 122: 185–212.

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, Viano MC. 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.

Lopez JA. 2001. Evaluating the predictive accuracy of volatility models. *Journal of Forecasting* 20: 87–109.

Mozumder S, Talukdar B, Kabir MH, 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.

Murphy KP. 2007. Conjugate Bayesian analysis of the Gaussian distribution. *Preprint*.

Nicholls F, Quinn BG. 1982. *Random coefficient autoregressive models*: *An Introduction*. Springer-Verlag, New York.

Nilsson O. 2016. On stochastic volatility models as an alternative to GARCH type models. https://www.diva-portal.org/smash/get/diva2:941043/FULLTEXT01.pdf.

Paolella MS. 2007. *Intermediate Probability: A computational Approach*. John Wiley & Sons.

Patton AJ. 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160: 246–256.

Rachev ST. 2003. *Handbook of heavy tailed distributions in Finance*, Volume 1: Handbooks in Finance, North Holland.

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

Rohan N, Ramanathan T. 2013. Nonparametric estimation of a time-varying GARCH model. *Journal of Nonparametric Statistics* 25: 33–52.

Rohan N, Ramanathan T. 2013. Nonparametric estimation of a time-varying GARCH model. *Journal of Nonparametric Statistics* 25: 33–52.

Smetanina E. 2017. Real-Time GARCH. *Journal of Financial Econometrics* 15: 561–601.

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

Taylor SJ. 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 SJ. 1986. *Modelling Financial Time Series.* Wiley, Chichester.

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

Tsay RS. 2010. *Analysis of financial time series: Financial Econometrics*, 3rd edition, Wiley.

Wee DCH, Chen F, Dunsmuir WTM. 2022. Likelihood inference for Markov switching GARCH(1,1) models using sequential Monte Carlo. *Econometrics and Statistics* 21: 50–68.

White H, Kim TH, 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.

Wu X, Zhao A, Chen T. 2023. A Real-Time GARCH-MIDAS model. *Finance Research Letters* 56: 104103.

Yu J. 2005. On leverage in a stochastic volatility model. *Journal of Econometrics* 127: 165–178.

Zhang Z, Zhao R. 2023. Improving the asymmetric stochastic volatility model with ex-post volatility: the identification of the asymmetry. *Quantitative Finance* 23: 35–51.