Real-Time GARCH

Ekaterina Smetaninas

Group 7 : 范小龙　刘泽平　刘杰　史文轩　苏韩

WISE & SOE

December 17, 2021

Outline

Motivation

- There are several main approaches to modeling the volatility of discrete financial time series: GARCH models, stochastic volatility (SV) models, and hybrid models.
- GARCH models incorporate only past internal information (i.e., information generated only within the model itself) and are therefore deterministic.
- SV models describe the volatility process as stochastic by allowing for external information. in the form of unobserved random shocks that are independent from the shocks governing the returns process.

$$GARCH(1,1)\begin{cases}\varepsilon_t=\sigma_t z_t, \quad z_t\sim i.i.d.(0,1)\\ \sigma_t^2=\omega+\alpha_1\varepsilon_{t-1}^2+\beta_1\sigma_{t-1}^2\end{cases}$$

$$SV:\begin{cases}\varepsilon_t=\sigma_t z_t, \quad z_t\sim \mathrm{iid}(0,1)\\ \log\sigma_t=\omega+\phi\log\sigma_{t-1}+v_t\\ v_t\sim WN(0,1) \quad \& \quad v_t\perp z_t\\ \text{or } \log\sigma_t^2=\omega+\phi\log\sigma_{t-1}^2+v_t\end{cases}$$

What's the paper about?

- As a result, SV models can be more flexible in fitting the data; however, this comes at the cost of higher complexity involved in their estimation and inference.
- Contrasting with SV models, GARCH models are observation-driven. How to estimate? Quasi-maximum likelihood (QMLE), which accounts for their wider use among practitioners.
- In this paper, we show that it is possible to efficiently utilize all available internal information in GARCH models, in particular incorporating the current return.

Consider the following model

$$r_t = \epsilon_t \lambda_t$$

- where λ_t is $\mathcal{F}_t$ measurable
- r_t is the return series
- ϵ_t are i.i.d. random variables such that $E(\epsilon_t) = 0, E(\epsilon_t^2) = 1$
- $\mathcal{F}_t$ is the information set available at time t.

Remark:

- The new model therefore(Real-time GARCH, RT-GARCH for short) can be thought of as a link between GARCH and SV models
- An important advantage of this framework is that we allow the shape of the conditional distribution of returns to be time-varying.

Interpretation and Relation to GARCH Models

- **Current Information** is represented by the current squared return.
- Problem: If one is to forecast the future conditional variance at time $t+1$, the future return, r_{t+1}, will be required but is unobserved.
- Solution: Considering some function of the current return when forecasting. eg. current return scaled by its volatility.

GARCH-types Model: Only the second conditional moment of the error term will be required for forecasting, Considering joint process $(r_t; \lambda_t^2)$:

$$r_t = \lambda_t \epsilon_t \tag{1}$$

$$\lambda_t^2 = \underbrace{\alpha + \beta \lambda_{t-1}^2 + \gamma r_{t-1}^2}_{=b_{t-1}} + \varphi \underbrace{\frac{r_t^2}{\lambda_t^2}}_{=\epsilon_t^2}, \quad (\alpha, \beta, \gamma, \varphi) \geq 0 \tag{2}$$

- r_t: Return series
- ϵ_t: i.i.d. *rvs* with density function $f_\epsilon(\cdot)$ such that $E(\epsilon_t) = 0, E(\epsilon_t^2) = 1$
- $\alpha_0, \beta_0, \gamma_0, \varphi_0$: True parameters
- Note:
  - $var[r_t|\mathcal{F}_{t-1}] \neq \lambda_t^2$, λ_t is not independent of ϵ_t
  - In order to satisfied $\lambda_t^2 > 0$, here choose ϵ_t^2. Why? It also have some good properties.

Time-varying weights to returns on different days. It can be shown that Equation (2) can approximately be written in the following way:

$$\lambda_t^2 \approx \frac{\varphi r_t^2}{b_{t-1}} + \sum_{j=1}^{\infty}\left(\frac{\beta^j \varphi}{b_{t-1-j}} + \gamma\beta^{j-1}\right) r_{t-j}^2 \tag{3}$$

Where $b_{t-1} = \alpha + \beta\lambda_{t-1}^2 + \gamma r_{t-1}^2$

- Compared with the standard GARCH models: the weights are time-varying and depend on past volatility, which can be approximately taken to be b_{t-1}
- Intuition: The weight is bigger for a smaller past return and is smaller if the past return is large.

The weight of r_{t-j}^2 consist two part:

- $\gamma\beta^{j-1}$: Usual "GARCH weight",
- $(\beta^j\varphi)/b_{t-1-j}$: Additional time-varying weight, which assigns an extra weight if a particular realization of r_{t-j} is in the tails of the distribution.

Experiment: Enlarge the information content of the volatility process

$\sigma_t = 20\%, t < T; \sigma_t = 40\%, t \geq T$, $\epsilon_t \sim N(0,1)$

GARCH(1,1): $E(r_{T+k}^2) = E(\sigma_{T+k}^2) = \alpha + \gamma E(r_{T+k-1}^2) + \beta E(\lambda_{T+k-1}^2) = \frac{\alpha}{1-\beta} + \gamma\frac{1-\beta^k}{1-\beta}(40\%)^2 + \gamma\frac{\beta^k}{1-\beta}(20\%)^2 \quad k \geq 0$

RT-GARCH(1,1): $E(r_{T+k}^2) = \frac{\alpha+\varphi(3-2\beta)}{1-\beta} + \gamma\frac{1-\beta^k}{1-\beta}(40\%)^2 + \gamma\frac{\beta^k}{1-\beta}(20\%)^2$, where $E(r_t^2) = E(\lambda_t^2) + \varphi[E(\epsilon_t^4) - 1]$

Question: How many days following the jump will it take for the volatility process to adjust to its new level?

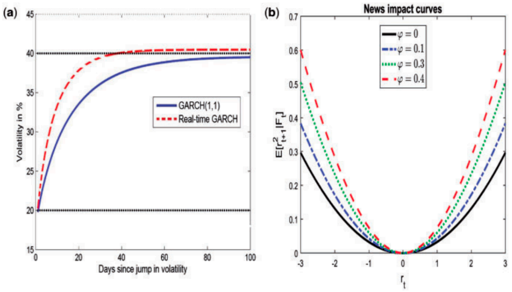


Figure 1. (a) Time scale of the volatility adjustment. For both graphs the parameter vector $[\alpha, \beta, \gamma, \varphi]$ is set to $[0, 0.92, 0.073, 0.035]$ for the RT-GARCH model, while $[\alpha, \beta, \gamma] = [0, 0.95, 0.045]$ for the standard GARCH(1,1) model. (b) News impact curves for different values of φ. For both graphs parameter values are in line with the ones from the estimated daily GE stock returns.

RT-GARCH model is at least two times faster when compared with the standard GARCH(1,1) model.

Another measure: **News Impact Curve**. The news impact curve of RT-GARCH:

$$E[r_{t+1}^2|\mathcal{F}_t] = \alpha + \varphi\kappa + \beta(\frac{\bar{b} + \sqrt{\bar{b}^2 + 4\varphi r_t^2}}{2}) + \gamma r_t^2 \tag{4}$$

with $\kappa = E(\epsilon_t^4)$, $\bar{b} = (\alpha + \beta\varphi + \kappa\gamma\varphi)/(1 - (\beta + \gamma))$

- $\varphi = 0$ is GARCH(1,1) model.
- For larger values of φ the volatility becomes even larger,

Unconditional Expectation

from (1) and (2) the unconditional expectations of r_t^2 and λ_t^2 are given by:

$$E[r_t^2] = \alpha + \beta E[\lambda_{t-1}^2] + \gamma E[r_{t-1}^2] + \varphi E[\epsilon_t^4] \quad (5)$$

$$E[\lambda_t^2] = E[\alpha + \beta\lambda_{t-1}^2 + \gamma r_{t-1}^2 + \varphi\epsilon_t^2] = \alpha + \beta E[\lambda_{t-1}^2] + \gamma E[r_{t-1}^2] + \varphi \quad (6)$$

$$\Rightarrow E[r_t^2] = E[\lambda_t^2] + \varphi(E[\epsilon_t^4] - 1) \quad (7)$$

Conditional Expectation

Theorem 1

- ϵ_t: symmetric around zero random variables such that $E(\epsilon_t) = 0, var(\epsilon_t) = 1$
- (r_t, λ_t^2) evolve according to Equations (1) - (2)
- $\mathcal{F}_{t-1} := \sigma(r_s, s \leq t-1)$ the r-algebra induced by the history of returns up to time $t-1$
- $\theta = (\alpha, \beta, \gamma, \varphi)'$: parameter vector, $\theta_0 = (\alpha_0, \beta_0, \gamma_0, \varphi_0)'$: True parameter vector
- $f_r(r|\mathcal{F}_{t-1})$: Conditional probability density function of the return series

$$f_r(r|\mathcal{F}_{t-1}) = \frac{r}{d(r, b_{t-1}, \theta)\sqrt{b_{t-1}^2 + 4r^2\varphi}} f_\epsilon(d(r, b_{t-1}, \theta)) \tag{8}$$

I jump to $l_t(\theta)$

- $f_\epsilon(\cdot)$: Probability density function of ϵ_t, $b_{t-1} = \alpha + \beta\lambda_{t-1}^2 + \gamma r_{t-1}^2$

Conditional Expectation

Theorem 1

- $d(r, b_{t-1}, \theta) = \begin{cases} \sqrt{\frac{\sqrt{b_{t-1}^2 + 4r^2\varphi} - b_{t-1}}{2\varphi}}, & \text{for } \varphi \neq 0 \\ r/\sqrt{b_{t-1}}, & \text{for } \varphi = 0 \end{cases}$, $\epsilon_t = d(r_t, b_{t-1}, \theta_0)$. Then we have

$$\lim_{r \to 0} \frac{r}{d(r, b_{t-1}, \theta)} = \sqrt{b_{t-1}} \text{ and } \lim_{r \to 0} f_r(r|\mathcal{F}_{t-1}) = \frac{1}{\sqrt{b_{t-1}}} f_\epsilon(0) \tag{9}$$

- $F(r|\mathcal{F}_{t-1}) = F_\epsilon(d(r, b_{t-1}, \theta))$: Conditional CDF of returns; $F_\epsilon(\cdot)$: CDF of ϵ_t

The conditional jth moment of returns, is given by

$$E[r_t^j|\mathcal{F}_{t-1}] = b_{t-1}^{j/2}[E(d(r, b_{t-1}, \theta)^j) + \frac{j\varphi}{2b_{t-1}} E(d(r, b_{t-1}, \theta)^{j+2})] \tag{10}$$

- **Remark 1**: If returns to be a martingale difference sequence(MDS), it is required that the third moment of ϵ_t is also zero.
- **Remark 2**: Since r_t is an odd function of ϵ_t, λ_t is an even function of ϵ_t, conditional and unconditional distribution of r_t is symmetric.
- **Remark 3**: The conditional density of the RT-GARCH model in Equation (8) nests the conditional density of the standard GARCH(1,1) model as its limiting case at $r = 0$. Intuition?
- **Remark 4**:

  The standardized conditional kurtosis of standard GARCH(1,1) model is

$$E[r_t^4|\mathcal{F}_{t-1}]/(E[r_t^2|\mathcal{F}_{t-1}])^2 = E[\epsilon_t^4]$$

  For the RT-GARCH model, we have

$$E\left[r_t^4|\mathcal{F}_{t-1}\right] / \left(E\left[r_t^2|\mathcal{F}_{t-1}\right]\right)^2 = \frac{b_{t-1}^2 E\left[\epsilon_t^4\right] + 2\varphi b_{t-1} E\left[\epsilon_t^6\right] + \varphi^2 E\left[\epsilon_t^8\right]}{\left(b_{t-1} + \varphi E\left[\epsilon_t^4\right]\right)^2}$$

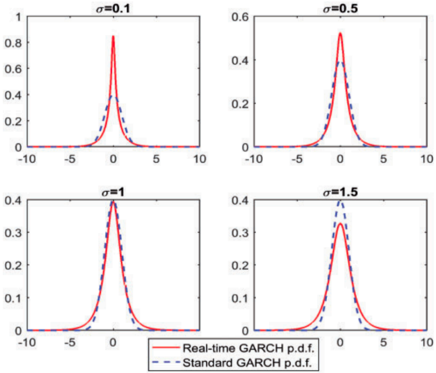


Figure 2. Conditional probability density function for different values of unconditional volatility. The parameter vector $\theta = [\alpha, \beta, \gamma, \varphi]'$ is set to $[0.003, 0.9, 0.04, 0.02]'$, which are typical parameter values.

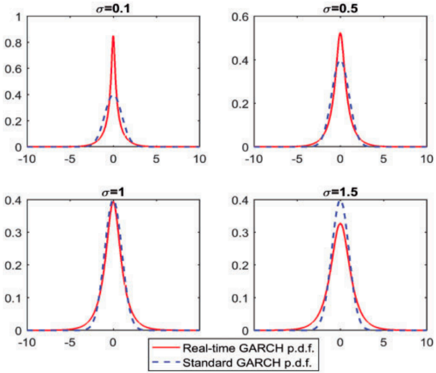


Figure 2. Conditional probability density function for different values of unconditional volatility. The parameter vector $\theta = [\alpha, \beta, \gamma, \varphi]'$ is set to $[0.003, 0.9, 0.04, 0.02]'$, which are typical parameter values.

- The return process described by the RT-GARCH with normal innovations is now able to account for **heavier tails** compared with the standard normal distribution.

Strictly Stationary of (r_t, λ_t^2)

Strictly Stationary of (r_t, λ_t^2) is important for developing estimation theory for the QMLE.

Theorem 2: Strictly Stationary of (r_t, λ_t^2)

- ϵ_t: i.i.d. symmetric around zero rvs such that $E(\epsilon_t) = 0, var(\epsilon_t) = 1$
- (r_t, λ_t^2) evolve according to Equations (1) - (2)
- $\alpha, \beta, \gamma > 0, \varphi \neq 0$

If the following conditions are satisfied:

$$-\infty \leqq E \log \left|\beta + \gamma\epsilon_0^2\right| < 0 \quad E\left(\log\left|\alpha + \varphi\epsilon_0^2\right|\right)^+ < \infty \tag{11}$$

then the process (r_t, λ_t^2) is strictly stationary.

Weak Stationarity of r_t^2, λ_t^2

The result of weak stationarity conditions of r_t^2, λ_t^2 can be used to derive the forecasting formulae for the conditional variance of returns.

Theorem 3: Weak Stationarity of λ_t^2

- ϵ_t: i.i.d. symmetric around zero *rvs* such that $\mathbb{E}(\epsilon_t) = 0, var(\epsilon_t) = 1$
- (r_t, λ_t^2) evolve according to Equations (1) - (2)

Then under the following conditions:

$$\begin{cases} \beta + \gamma < 1 \\ \alpha + \varphi + \gamma\varphi\left(\mathbb{E}\left[\epsilon_t^4\right] - 1\right) > 0 \end{cases} \text{(case 1) or } \begin{cases} \beta + \gamma > 1 \\ \alpha + \varphi + \gamma\varphi\left(\mathbb{E}\left[\epsilon_t^4\right] - 1\right) < 0 \end{cases} \text{(case 2)} \tag{12}$$

the process λ_t^2 is weakly stationary and its first unconditional moment is given by

$$E[\lambda_1^2] = \frac{\alpha + \varphi + \gamma\varphi(E[\epsilon_t^4] - 1)}{1 - (\beta + \gamma)} > 0 \tag{13}$$

Theorem 4: Weak Stationarity of r_t^2

- ϵ_t: i.i.d. symmetric around zero rvs such that $E(\epsilon_t) = 0, var(\epsilon_t) = 1$
- (r_t, λ_t^2) evolve according to Equations (1) - (2)

Then under the following conditions:

$$\begin{cases} \beta + \gamma < 1 \\ \alpha + \varphi E\left(\epsilon_t^4\right) + \varphi\beta\left(1 - E\left(\epsilon_t^4\right)\right) > 0 \end{cases} \quad \text{(case 3)} \tag{14}$$

$$\begin{cases} \beta + \gamma > 1 \\ \alpha + \varphi E\left(\epsilon_t^4\right) + \varphi\beta\left(1 - E\left(\epsilon_t^4\right)\right) < 0 \end{cases} \quad \text{(case 4)} \tag{15}$$

r_t is weakly stationary and its second unconditional moment is given by

$$E\left[r_1^2\right] = \frac{\alpha + \varphi E\left(\epsilon_t^4\right) + \varphi\beta\left(1 - E\left(\epsilon_t^4\right)\right)}{1 - (\beta + \gamma)} > 0 \tag{16}$$

It also holds that: $cov(r_t, r_s) = 0, t \neq s$

Stationary Condition of Fourth Moment of r_t

Theorem 5: Stationary Condition of Fourth Moment of r_t

- ϵ_t: i.i.d. symmetric around zero *rvs* such that $E(\epsilon_t) = 0, var(\epsilon_t) = 1$
- (r_t, λ_t^2) evolve according to Equations (1) - (2)

then r_t is fourth moment stationary if $\gamma^2 < \frac{1}{E[\epsilon_t^4]}$
with the unconditional fourth moment given by

$$E\left[r_1^4\right] = \frac{\xi_1 + E\left[\lambda_1^2\right]\xi_2 + 2\beta\gamma\mu_4\left[E\left[\lambda_1^2\right]\right]^2}{1 - \gamma^2\mu_4} \tag{17}$$

where $\mu_j := \varphi E\left[\epsilon_t^j\right]$ and constants ξ_1 and ξ_2 are given by (18) and $E\left[\lambda_1^2\right]$ is given by Equation (13).

$$\begin{cases} \xi_1 = \alpha^2\mu_4 + \mu_8 + 2\alpha\mu_6 + 4\varphi\gamma\left(\alpha\mu_4 + \beta\mu_4 + \mu_6\right) > 0 \\ \xi_2 = \mu_4\left(2\alpha\beta + \beta^2 + 2\alpha\gamma + 2\mu_6(\gamma + \beta)\right) > 0 \end{cases} \tag{18}$$

Leverage and Volatility Feedback Effects

- The RT-GARCH model described by Equations (2) has no leverage effect, meaning that when errors are symmetric about zero, $E(r_t) = 0, cov(r_t^2, r_j) = 0, \forall j$
- Different from GARCH-type models, the most recent information is represented by **current shocks** ϵ_t
- We therefore refer to "leverage effect" by differentiating the effect of positive and negative values of ϵ_t on λ_t^2
- Then the model become

$$r_t = \lambda_t \epsilon_t \tag{19}$$

$$\lambda_t^2 = \alpha + \beta \lambda_{t-1}^2 + \gamma r_{t-1}^2 + \varphi_1 \epsilon_t^2 \mathbf{1}_{(\epsilon_t > 0)} + \varphi_2 \epsilon_t^2 \mathbf{1}_{(\epsilon_t \leq 0)} \tag{20}$$

- It is also interesting to differentiate between the effect of positive and negative values of **past returns** on the conditional volatility.
- Given the differently defined leverage effect in our model, we refer to the different effects of the past positive and negative returns on conditional variance as "feedback effect".
- Then the RT-GARCH model with leverage and feedback effects is given by

$$r_t = \lambda_t \epsilon_t \tag{21}$$

$$\lambda_t^2 = \alpha + \beta\lambda_{t-1}^2 + \gamma_1 r_{t-1}^2 \mathbf{1}_{(r_t>0)} + \gamma_2 r_{t-1}^2 \mathbf{1}_{(r_t\leq 0)} + \varphi_1 \epsilon_t^2 \mathbf{1}_{(\epsilon_t>0)} + \varphi_2 \epsilon_t^2 \mathbf{1}_{(\epsilon_t\leq 0)} \tag{22}$$

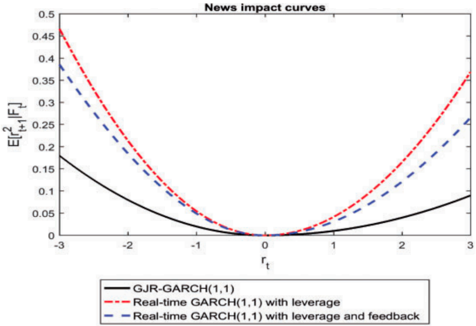


Figure 4. The figure displays the news impact curves for three models, estimated on the daily IBM data.

- For both specifications of the RT-GARCH model, volatility tends to respond **more to negative news** than in GJR-GARCH model.
- RT-GARCH model with leverage and feedback responds **slower to negative news** than the RT-GARCH just with the leverage effect.

Outline of the Estimation Theory

- $\theta = (\alpha, \beta, \gamma, \varphi)'$: Parameter vector, $\theta_0 = (\alpha_0, \beta_0, \gamma_0, \varphi_0)'$: True parameter vector
- we adopt a Gaussian specification, such that the log likelihood function can be written as:

$$L_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} l_t(\theta) \tag{23}$$

$$l_t(\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}d_t^2\left(r, b_{t-1}, \theta\right) + \log\left(\frac{\sqrt{b_{t-1}(\theta) + \varphi d_t^2\left(r, b_{t-1}, \theta\right)}}{b_{t-1}(\theta) + 2\varphi d_t^2\left(r, b_{t-1}, \theta\right)}\right) \tag{24}$$

I jump to $f_r(r|F_{t-1})$

- The author said the asymptotic distribution of $\hat{\theta}$ should be

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, V_\theta)$$

Testing

- One can consider testing the following null hypothesis:

$$\mathbb{H}_0 : \varphi = 0$$

- This test can be interpreted as the test for **constant standardized** conditional kurtosis of the returns against an alternative of a **time-varying** conditional kurtosis.
- Using likelihood quantities to calculate the likelihood-ratio(LRLeverage and Volatility Feedback Effects) test

$$LR = -2\ln(L_T(\theta^*)/L_T(\theta)) \xrightarrow{d} \chi_1^2, \text{ where } \theta = (\alpha, \beta, \gamma, \varphi)', \theta^* := \{\theta/\varphi\} \tag{25}$$

Volatility Forecasts with RT-GARCH

Theorem 6

Let the process:

- (r_t, λ_t^2) evolve according to Equations (20)
- ϵ_t is symmetric around zero i.i.d. random variables such that $E(\epsilon_t) = 0$, $\operatorname{var}(\epsilon_t) = 1$.

Then the k-step ahead, $k \geq 1$, volatility forecast is given by the following formula:

$$\begin{aligned} E\left[r_{t+k}^2 \mid \mathcal{F}_t\right] =& E\left[\hat{\lambda}_{t+k}^2 \mid \mathcal{F}_t\right] + (\hat{\varphi}_1 + \hat{\varphi}_2)\left(E\left[\epsilon_{t+k}^4 \mid \mathcal{F}_t\right] - 1\right) \\ =& E\left[\lambda_1^2\right] + \left(\hat{\beta} + \hat{\gamma}_1 + \hat{\gamma}_2\right)^k \left(E\left[\hat{\lambda}_t^2 \mid \mathcal{F}_t\right] - E\left[\lambda_1^2\right]\right) \\ &+ (\hat{\varphi}_1 + \hat{\varphi}_2)\left(E\left[\epsilon_{t+k}^4 \mid \mathcal{F}_t\right] - 1\right) \end{aligned}$$

where $E\left[\lambda_1^2\right]$ is given by

$$E\left[\lambda_1^2\right] = \frac{\alpha + (\varphi_1 + \varphi_2)\left[\eta\left(\gamma_1 + \gamma_2\right) + 1\right]}{1 - (\beta + \gamma_1 + \gamma_2)}$$

with $\eta \equiv E\left[\epsilon_t^4\right] - 1$ and $E\left[\hat{\lambda}_t^2 \mid \mathcal{F}_t\right]$ is an estimate of λ_t^2 at time t.

- To estimate and evaluate competing models, we use three datasets of open-to-close returns, namely IBM, General Electric (GE), and the S&P 500 index.
  - For IBM and GE, the data spans from January 2, 1998, till December 1, 2016
  - For the S&P500 the time span is from January 28, 2003, till December 1, 2016.
- We reserve two-third of the whole sample for the estimation and the rest of the sample for the forecast evaluation, where the latter includes the crisis period as well.
- Due to the likelihood of structural breaks during the financial crisis period we also present results for two subsamples: pre- and post-crisis periods.
- How to compare model ? Next slide. Remark: Simple and Exponential NoVaS methodologies of Politis (2007)

The conditional variance specification of different models

Table 1. The conditional variance specification of different models

RT-GARCH	$\lambda_t^2 = \alpha + \beta\lambda_{t-1}^2 + \gamma r_{t-1}^2 + \varphi\epsilon_t^2, \quad E[r_t^2] = E[\lambda_t^2] + 2\varphi$
RT-GARCH with leverage	$\lambda_t^2 = \alpha + \beta\lambda_{t-1}^2 + \gamma r_{t-1}^2 + \varphi_1\epsilon_t^2 1_{(\epsilon_t \geq 0)} + \varphi_2\epsilon_t^2 1_{(\epsilon_t < 0)}, \quad E[r_t^2] = E[\lambda_t^2] + 2(\varphi_1 + \varphi_2)$
RT-GARCH with leverage and feedback	$\lambda_t^2 = \alpha + \beta\lambda_{t-1}^2 + \gamma_1 r_{t-1}^2 1_{(r_t \geq 0)} + \gamma_2 r_{t-1}^2 1_{(r_t < 0)} + \varphi_1\epsilon_t^2 1_{(\epsilon_t \geq 0)} + \varphi_2\epsilon_t^2 1_{(\epsilon_t < 0)}$ $E[r_t^2] = E[\lambda_t^2] + 2(\varphi_1 + \varphi_2)$
GARCH(1,1) with standard normal errors	$\sigma_t^2 = \alpha + \beta\sigma_{t-1}^2 + \gamma r_{t-1}^2$
GARCH(1,2) with standard normal errors	$\sigma_t^2 = \alpha + \beta\sigma_{t-1}^2 + \gamma_1 r_{t-1}^2 + \gamma_2 r_{t-2}^2$
GARCH(1,1) with Student's *t*-distributed errors	$\sigma_t^2 = \alpha + \beta\sigma_{t-1}^2 + \gamma r_{t-1}^2$
GARCH(1,2) with Student's *t*-distributed errors	$\sigma_t^2 = \alpha + \beta\sigma_{t-1}^2 + \gamma_1 r_{t-1}^2 + \gamma_2 r_{t-2}^2$
APARCH(2,2) with Student's *t*-distributed errors	$\sigma_t^2 = \alpha_0 + \alpha_1[\lvert r_{t-1}\rvert - \gamma_1 r_{t-1}]^2 + \alpha_2[\lvert r_{t-2}\rvert - \gamma_2 r_{t-2}]^2 + \beta_1\sigma_{t-1}^2 + \beta_2\sigma_{t-2}^2$
Simple NoVaS	$\sigma_t^2 = \alpha s_{t-1}^2 + \alpha_0 X_t^2 + \sum_{i=1}^{p}\alpha_i X_{t-p}^2$ $s_{t-1}^2 = \frac{1}{t-1}\sum_{k=1}^{t-1} X_k^2, \alpha = 0, \alpha_i = \frac{1}{p+1}, 0 \leq i \leq p$
Exponential NoVaS	$\sigma_t^2 = \alpha s_{t-1}^2 + \alpha_0 X_t^2 + \sum_{i=1}^{p}\alpha_i X_{t-p}^2$ $s_{t-1}^2 = \frac{1}{t-1}\sum_{k=1}^{t-1} X_k^2, \alpha = 0, \alpha_i = c\prime^{e^{-ci}}, 0 \leq i \leq p, c' = \frac{1}{\sum_{i=0}^{p} e^{-ci}}$

Notes: In Simple NoVaS p is chosen to match the kurtosis (=3) of the normalized return series. In Exponential NoVaS initial value of p is chosen to be $\frac{n}{5}$; c is chosen to match the kurtosis (=3) of the normalized return series and p is adjusted by the maximization routine. APARCH(2,2) corresponds to the standard GJR(2,2) model whenever $0 \leq \gamma_i \leq 1, i = 1, 2$.

Parameter estimates of RT-GARCH models

Parameter estimates of RT-GARCH

Dataset	α	β	γ	φ
IBM	0.0006	0.8755	0.0780	0.0758
	$(18 * 10^{-4})$	$(9 * 10^{-4})$	$(14 * 10^{-4})$	$(21 * 10^{-4})$
GE	0.0001	0.9211	0.0627	0.0378
	$(14 * 10^{-4})$	$(38 * 10^{-3})$	$(2 * 10^{-5})$	$(17 * 10^{-4})$
S&P 500	0.0001	0.9124	0.0726	0.0138
	$(12 * 10^{-4})$	$(14 * 10^{-3})$	$(45 * 10^{-3})$	$(11 * 10^{-4})$

Parameter estimates of RT-GARCH with leverage

Dataset	α	β	γ	φ_1	φ_2
IBM	0.0003	0.8883	0.0703	0.0475	0.0886
	$(15 * 10^{-4})$	$(6 * 10^{-4})$	$(11 * 10^{-4})$	$(19 * 10^{-4})$	$(27 * 10^{-4})$
GE	0.0001	0.9273	0.0550	0.0237	0.0529
	$(2.7 * 10^{-4})$	$(38 * 10^{-4})$	$(4.2 * 10^{-4})$	$(2 * 10^{-4})$	$(48 * 10^{-3})$
S&P 500	0.0016	0.8995	0.0718	0.0003	0.0481
	$(25 * 10^{-4})$	$(15 * 10^{-3})$	$(6.7 * 10^{-4})$	$(27 * 10^{-4})$	$(8.1 * 10^{-4})$

Forecasts evaluation based on MSE loss (full sample)

Hansen[3]: P_{MC}

1-step ahead volatility forecasts

Model	IBM MSE	IBM p_{MCS}	GE MSE	GE p_{MCS}	S&P 500 MSE	S&P 500 p_{MCS}
RT-GARCH	5.9626	0.0940*	6.7355	0.2530*	2.5227	0.0400
RT-GARCH-L	5.8989	0.0990*	6.6591	0.9170*	2.0289	0.6220*
RT-GARCH-LF	6.0861	0.0820*	6.6274	1*	1.9370	1*
A-PARCH(2,2)-Student's *t* distribution	5.8152	0.6330*	6.6632	0.9170*	2.6657	0.0020
GARCH(1,1)-*N*(0, 1)	5.9074	0.0990*	6.8069	0.0070	2.3780	0.0450
GARCH(1,2)-*N*(0, 1)	5.9074	0.0990*	6.8069	0.0070	2.3780	0.0450
GARCH(1,1)- Student's *t* distribution	5.7074	0.7470*	6.8199	0.0030	2.3725	0.0450
GARCH(1,2)-Student's *t* distribution	5.6923	1*	6.9611	0.0030	2.5405	0.0310
Simple NoVaS	7.9097	0	8.6489	0	2.6415	0.0060
Exponential NoVaS	7.9305	0	8.7922	0	2.7714	0.0010

5-step ahead volatility forecasts

Model	IBM MSE	IBM p_{MCS}	GE MSE	GE p_{MCS}	S&P 500 MSE	S&P 500 p_{MCS}
RT-GARCH	5.4655	0.0660*	6.1762	0.1130*	1.7490	0.1820*
RT-GARCH-L	5.7604	0.0010	6.1410	0.1130*	1.6724	1*
RT-GARCH-LF	7.4039	0.0005	6.8305	0.0005	2.7361	0.0600*
A-PARCH(2,2)-Student's *t* distribution	5.5588	0.0200	6.2182	0.1130*	2.0496	0.1230*
GARCH(1,1)-*N*(0, 1)	5.5436	0.0200	6.2221	0.0440	1.7265	0.1820*
GARCH(1,2)-*N*(0, 1)	5.5436	0.0020	6.2221	0.0440	1.7265	0.1820*
GARCH(1,1)-Student's *t* distribution	5.2380	0.1460*	6.2346	0.0220	2.1047	0.1090*
GARCH(1,2)-Student's *t* distribution	5.1158	1*	5.9126	1*	2.1059	0.1090*
Simple NoVaS	7.9295	0	8.8844	0	2.8555	0.0220
Exponential NoVaS	7.9810	0	8.9874	0	2.9460	0.0050

Notes: p_{MCS} are the p-values from Model Confidence Set test of Hansen, Lunde, and Nason (2011). The

Forecasts evaluation based on QLIKE loss (full sample)

$$\text{QLIKE} = \log h + \frac{\hat{\sigma}^2}{h}$$

1-step ahead volatility forecasts

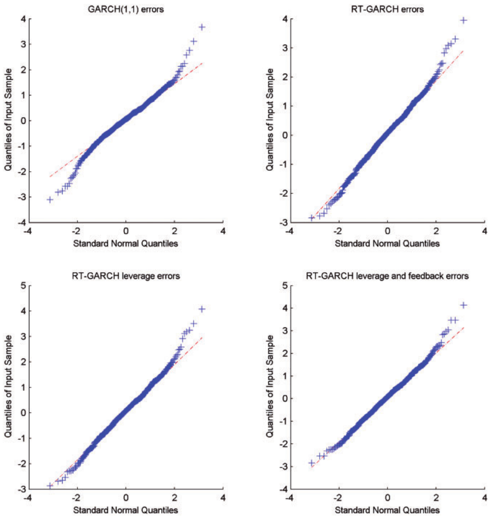
	IBM		GE		S&P 500	
Model	QLIKE	p_{MCS}	QLIKE	p_{MCS}	QLIKE	p_{MCS}
RT-GARCH	1.4628	0.5500★	1.4505	0.0610★	0.9426	0.3740★
RT-GARCH-L	1.4531	1★	1.4315	1★	0.9471	0.1550★
RT-GARCH-LF	1.4828	0.3800★	1.4384	0.6540★	1.0077	0.0010
A-PARCH(2,2)-Student's t distribution	1.5487	0.0120	1.4733	0.0320	0.9322	1★
GARCH(1,1)-N(0, 1)	1.5106	0.0320	1.4767	0.0270	0.9488	0.0450
GARCH(1,2)-N(0, 1)	1.5106	0.0320	1.4767	0.0270	0.9488	0.0450
GARCH(1,1)-Student's t distribution	1.5252	0.0400	1.4786	0.0180	0.9494	0.0440
GARCH(1,2)-Student's t distribution	1.5277	0.0400	1.4786	0.0180	0.9376	0.4540★
Simple NoVaS	3.9372	0	3.3669	0	1.4625	0.0005
Exponential NoVaS	3.9361	0	3.3633	0	1.5563	0.0005

5-step ahead volatility forecasts

	IBM		GE		S&P 500	
Model	QLIKE	p_{MCS}	QLIKE	p_{MCS}	QLIKE	p_{MCS}
RT-GARCH	1.3935	1★	1.3760	1★	0.8957	0.2950★
RT-GARCH-L	1.4834	0.0140	1.3932	0.0560★	1.0169	0.0090
RT-GARCH-LF	1.6347	0.0040	1.5198	0.0010	1.2279	0.0040
A-PARCH(2,2)-Student's t distribution	1.4492	0.0170	1.4034	0.0560★	0.8955	1★
GARCH(1,1)-N(0, 1)	1.4157	0.0280	1.4163	0.0110	0.9159	0.0310
GARCH(1,2)-N(0, 1)	1.4157	0.0280	1.4163	0.0110	0.9159	0.0310
GARCH(1,1)-Student's t distribution	1.4131	0.0300	1.3948	0.0560	0.9470	0.0120
GARCH(1,2)-Student's t distribution	1.4144	0.0300	1.3987	0.0560	0.9425	0.0120
Simple NoVaS	4.1445	0	3.6022	0	1.6679	0.0020
Exponential NoVaS	4.0942	0	3.5691	0	1.8015	0.0020

Notes: p_{MCS} are the p-values from Model Confidence Set test of Hansen, Lunde, and Nason (2011). The p-values that are marked with an ★ are those in the model confidence set $\widehat{\mathcal{M}}^{\star}_{95\%}$.

QQ plots of the implied error distribution for 3 Datasets



Summary for the results

- Most of the time the MCS for all datasets contains models with student-t innovations (which allows for heavier tails) and RT-GARCH models. However, RT-GARCH models perform no worse (or most of the time even better) with just the normal innovations.
- The possible reason for this is that RT-GARCH models account for a time-varying conditional kurtosis, therefore allowing the volatility to adjust to a new level faster than the other standard GARCH models.
- On the other hand, estimated on the full sample the RT-GARCH model with leverage and feedback effects (RT-GARCH-LF) seems to perform worse than the simple RT-GARCH or RT-GARCH-L, as it can potentially overfit the data due to the model’s higher complexity (i.e., higher number of parameters).

References I

Ekaterina Smetanina; PERSONAL WEBSITE: https://www.smetaninakatja.xyz/
Real-Time GARCH, 2017.
Journal of Financial Econometrics

Ekaterina Smetanina; PERSONAL WEBSITE: https://www.smetaninakatja.xyz/
ASYMPTOTIC THEORY FOR QMLE FOR THE REAL-TIME GARCH(1,1) MODEL, 2019.
Journal of Time Series Analysis

EHansen, Peter R and Lunde, Asger and Nason, James M
The model confidence set, 2011.
Econometrica

Thanks!