Erasmus University Rotterdam
Erasmus School of Economics (ESE)

Econometrics and Operational Research

# Bachelor Thesis [Quantitave Finance]

## Real Time approach in Volatility modelling

A replication and further extension on RT-GARCH and Beta-$t$-EGARCH

by

Thomas de Jong
Student Nr.: 481017

Date final version: July 4, 2021
Supervisor: Rutger-Jan Lange
Second assessor: Kleen, O.

**Abstract**

This research is a replication of the RT-GARCH model proposed by Smetanina (2017) and the notes on this model by Lange (2021). By means of an empirical analysis and investigation of both the robustness in terms of responsiveness to extreme returns as well as the predictive performance, it can be concluded the 'real time' feature that is present within the RT-GARCH adds significant value. This research also investigates the Beta-$t$-EGARCH and a two component extension and compares them with the RT-GARCH. The RT-GARCH is able to outperform Beta-$t$-EGARCH on predictive level and shows to be a better fit. An extension on the Beta-$t$-EGARCH by adding the squared current returns, shows to be a better emperical fit, with a significantly higher log-likelihood value. For one step ahead forecasts the model does not outperform both the RT-GARCH as the Beta-$t$-EGARCH models.

# Contents

# 1 Introduction

This research will be of interest for all parties trading on equity markets. The importance of volatility estimation cannot be overstated, since it is of vital importance for equity traders and investors. Volatility is a measure of dispersion around the mean return of an asset. Research found that higher volatility corresponds to a higher probability of a declining market and vice versa.

Over the years, a lot of estimation methods for volatility of asset returns have been formulated. The most well-known being the GARCH-type models as described by Engle and Bollerslev (1986). In this type of modelling the volatility is estimated by only making use of past information. That is, $\sigma_t|\mathcal{I}_{t-1}$, where $\sigma_t$ is the volatility dependent on information before time $t$. The information set up to $t-1$, is denoted by $\mathcal{I}_{t-1}$. This potential shortcoming of only using past information has been tackled by proposing so called Real-time GARCH models (RT-GARCH) in context of Smetanina (2017). Here, an extension on the GARCH model by adding the squared current returns as explanatory variable is being proposed. Making the original GARCH model nested within RT-GARCH. Smetanina (2017) shows that the RT-GARCH model outperforms simple GARCH-type models because of a number of reasons. Most importantly, the RT-GARCH model is able to account for rapid changes in volatility because the distribution of the returns has a time-varying kurtosis. Also, because of this feature, the RT-GARCH provides a better empirical fit, especially in the tails of the distribution. This research replicates these findings and concludes this to be true. The findings are found true since the RT-GARCH model has a higher log-likelihood value than the standard GARCH model for the S&P500 stock returns over a period of twenty years. The News Impact Curve shows the responsiveness of the RT-GARCH model to be higher than the original GARCH model by measuring the effects of returns on volatility.

Ding (2021) adds to this RT-model of Smetanina (2017), by taking the so-called volatility of volatility into account. This formulation, by the name of augmented RT-GARCH (or ART-GARCH), allows for conditional heteroskedasticity in the volatility process. It is shown that the weights on past squared returns are more flexible than the RT-GARCH, because not only the predictable part of the process, is time-varying, but also the unpredictable part. Where the unpredictable part, measured as a coefficient of $\varepsilon_{t-j}$ is constant for RT-GARCH. By means of the News Impact Curve (NIC) Ding (2021) shows that for smaller returns the volatility responds faster than the RT-GARCH, while for bigger returns, the volatility responds slower than RT-GARCH. This is a good result since it is desirable that volatility responds fast to 'normal' returns and downweigh large abnormal returns. So eventhough the model is still quadratic, it is not as sensitive for outliers and thus more robust than the RT-GARCH. Ding (2021) also show by means of an empirical analysis that this model is able to outperform the simple GARCH and RT-GARCH on a predictive level.

Another variation on the GARCH-type models is the Beta-$t$-EGARCH first formulated by Harvey and Chakravarty (2008). Here, a dynamic scale is driven by a robust variance and location measure, by looking at the past values of the score of the $t$-distribution. This enables the model to handle potential outliers better and thus robustifies the model. So, just like the ART-GARCH model, the Beta-$t$-EGARCH model is able to handle potential outliers better than a simple GARCH model. However, as said, the Beta-$t$-EGARCH model only looks at past scores of the $t$-distribution and does not take into account present values like the RT-GARCH or ART-GARCH can do. Harvey and Lange (2018) have made a further extension on this model by the name of two-component Beta-$t$-EGARCH. This model makes it possible to distinguish between long- and short-run effects of returns on volatility.

This research takes an interest in the RT-GARCH and the Beta-t-EGARCH and investigates which of the two is best able to predict volatility. Therefore, the research question that the research investigates is the following:

*Which of the RT-GARCH and Beta-t-EGARCH models is best able to predict volatility?*

To extend on these models, the research estimates and evaluates the two component Beta-$t$-EGARCH and asymmetric RT-GARCH. Evaluation will be done by means of the test statistic of Quaedvlieg (2021).

The sub-questions to help answer the main research question will be:

1. *How accurate are volatility forecasts?* The research evaluates this by different loss functions within the multi-horizon test statistic by Quaedvlieg (2021). Namely, squared error and absolute error.

2. *Which model is more responsive to high returns?* The research investigates this by drawing a NIC for the given models. In these curves the effect of shocks on the volatility can be measured. This gives insight on the responsiveness of the models.

3. *What parameter and implied epsilon estimates does a combination of the two models in the form of the RT-Beta-t-EGARCH give?* This extension attempts to combine the 'real time' feature of the RT-GARCH with the score and location measures of the Beta-$t$-EGARCH.

An empirical analysis will be conducted on both the RT-GARCH as the Beta-t-EGARCH models. Firstly, in section 2, an explanation of the data of the S&P500 is given. The derivation of the returns, the properties of the log returns as well as its unconditional distribution is provided. Secondly, in section 3, the models will be explained together with important assumptions and results made in the relevant literature. Section 3 will also explain what methods the research uses to evaluate the given models and provides an explanation of the derivation of the five minute intra-day Realised Variance (RV). To evaluate said models, two main benchmarks are used. The first being the simple GARCH model and the second being the VIX. The VIX is a commonly used volatility benchmark estimator and is calculated by averaging the weighted prices of puts and calls on the market. For evaluating the estimated models, the research follows the multi-horizon forecast comparison method discussed by Quaedvlieg (2021). The methods described there, are derived from the simple Diebold-Mariano test, but performed over multiple horizons to see whether a model is overall better performing instead of solely on specific forecast horizons. However, this method introduces a new problem. That is finding a just conditionally unbiased proxy. Patton (2011) claims that forecast comparison techniques rely on a volatility proxy that is an imperfect estimator of the true conditional variance. This can then lead to wrong conclusions on which model is preferred. Patton (2011) goes on that less noisy volatility proxies, such as the intra-daily range and RV, lead to less distortion. In section 4 the research provides all results and their interpretation so that a conclusion can be given in section 5. Lastly, this research proposes a new model which adds real time value to the Beta-$t$-EGARCH model as a further extension. This model has real time features as well as dynamic measures closely related to the beta-$t$-EGARCH model as described by Harvey and Chakravarty (2008). The proposed model is a combination of the RT-GARCH and Beta-$t$-EGARCH and is named RT-Beta-$t$-EGARCH.

## 2 Data

The data for this research is collected from Yahoo! Finance. The S&P500 index for the period January 2000 to December 2021 is used. These are in total 5114 daily observations. The dataset collected from Yahoo! Finance provided daily open and close rate of the S&P500. To compute the daily returns, this research uses the close-to-close returns following formula

$$r_t = \frac{C_t - C_{t-1}}{C_{t-1}}, \quad \forall t \in [1, .., T], \tag{1}$$

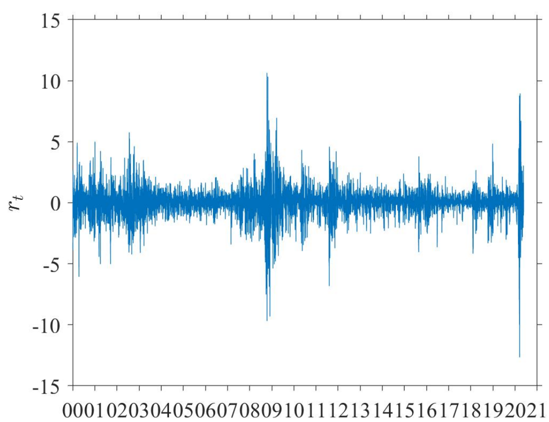
where $C_t$ is the closing rate of the S&P500 for day $t$. In Figure 1a are the log returns of the sample. The log returns being

$$\text{log returns} = \log(\frac{C_{t-1}}{C_t}), \quad \forall t \in [1, .., T]. \tag{2}$$

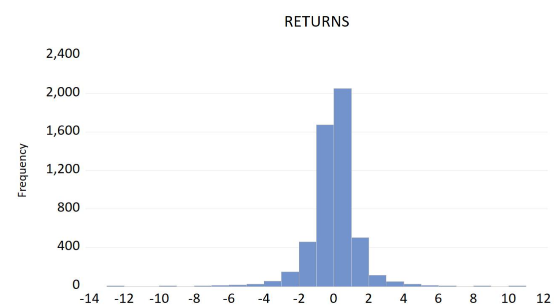
Looking at the graph, a set of periods in time are more volatile than others. The financial crisis of 2008 and the COVID-19 crisis of last year are both represented by high log returns.

For estimation of the models, the full sample is used. When forecasting, the research makes use of a 'pseudo' out-of-sample method, meaning all parameters are estimated over all data and forecasts are done using full sample parameter estimates. To calculate the Realized Variance (RV), the research uses a 5-min intraday high frequency data from FirstRate Data LLC. For smaller intraday frequency, Lunde and Hansen (2004) showed the presence of microstructure noise.

Throughout the paper, the returns are Student's $t$ distributed. As seem in Figure 1b, the returns have heavy tails and the Student's $t$ distribution is better able to capture potential outliers, making all models more robust in the tails of the distribution.




(a) Depicted are the log returns for 2000 to 2021. $r_t$ being the variable for returns at time t.




(b) The unconditional distribution of the returns. The mean of the returns is 0.014. The kurtosis is 13.8, which is about twice as much as that of a N(0,1) normal distribution.

# 3 Methods

## 3.1 RT-GARCH model

The RT-GARCH model below is derived from Smetanina (2017) and the notes on this paper by Lange (2021). The RT-GARCH is an extension of the simple GARCH formulation, meaning the GARCH model is nested within the RT-GARCH. The RT-GARCH model uses current information whereas the simple GARCH model only uses information up to time $t-1$. The following model specification as a joint process of $(r_t; \lambda_t^2)$ is given following Smetanina (2017):

$$\begin{aligned} r_t =& \mu + \lambda_t \varepsilon_t, \\ \lambda_t^2 =& \alpha + \beta \lambda_{t-1}^2 + \gamma r_{t-1}^2 + \phi \varepsilon_t^2. \end{aligned} \tag{3}$$

where $r_t$ is the return series and $\varepsilon$ is i.i.d. random variable with the density function $f_e$, such that $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = 1$. The real time adjustment compared to simple GARCH is represented by $\phi\varepsilon_t^2$. For $\phi = 0$, the model is a simple symmetric GARCH model. The first important take away is that the volatility process is now represented by $\lambda_t^2$ instead of $\sigma_t^2$, since the volatility is now no longer independent of $\varepsilon_t$. That is, $E(r_t^2 - \mu) \neq E(\lambda_t^2)$. The error terms $\varepsilon_t$ are assumed to be Student's $t$ distributed with $v > 0$ degrees of freedom (Lange (2021)). The p.d.f. is as follows:

$$p(\varepsilon_t) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi(\nu-2)}} \left(1 + \frac{\varepsilon_t^2}{\nu-2}\right)^{-(\nu+1)/2}. \tag{4}$$

For the Asymmetric RT-GARCH process Smetanina (2017) uses leverage and feedback effects on both $\phi$ as $\gamma$ respectively, resulting in:

$$\lambda_t^2 = \alpha + \beta\lambda_{t-1}^2 + \gamma_1 r_{t-1}^2 \mathbb{I}_{(r_t \geq 0)} + \gamma_2 r_{t-1}^2 \mathbb{I}_{(r_t < 0)} + \phi_1 \varepsilon_t^2 \mathbb{I}_{(\varepsilon_t \geq 0)} + \phi_2 \varepsilon_t^2 \mathbb{I}_{(\varepsilon_t < 0)}. \tag{5}$$

where $\mathbb{I}$ are indicator functions. Interestingly the indicator function for the feedback effect relies on $r_t$, even though $\gamma$ is supposed to measure effect on lagged returns. For this very reason this research alters this formulation and especially this indicator function to $\mathbb{I}_{(r_{t-1} \geq 0)}$ and $\mathbb{I}_{(r_{t-1} < 0)}$.

To implement the RT-GARCH model the unconditional expectations of $r_t^2$ and $\lambda_t^2$ are needed. Lange (2021) formulates this same process, but splits the process into 2 parts. The part of the process that relies on all information up to time $t-1$ and a part of the process that relies on the real time adjustment based on $\varepsilon_t$. From here on, all derivations of the RT-GARCH process have been taken from Lange (2021).

$$\begin{aligned} r_t &= \mu + \sqrt{h_t}\varepsilon_t, \\ g_t &= \omega + (\alpha_1 \mathbb{I}_{r_{t-1} \leq \mu} + \alpha_2 \mathbb{I}_{r_{t-1} > \mu})(r_{t-1} - \mu)^2 + \beta h_{t-1}, \\ h_t &= g_t + (\phi_1 \mathbb{I}_{\varepsilon_t \leq 0} + \phi_2 \mathbb{I}_{\varepsilon_t > 0})\varepsilon_t^2. \end{aligned} \tag{6}$$

This formulation of the model is not symmetric, but has a leverage factor for both $(r_{t-1})^2$ and $\varepsilon_t^2$. The conditional expectation $E_{t-1}(r_t)$ is then given by:

$$E_{t-1}(r_t) = \mu + E_{t-1}(\sqrt{h_t} \cdot \varepsilon_t) = \mu + E_{t-1}\left(\sqrt{g_t + \phi_1 \mathbb{I}_{\varepsilon_t \leq 0} + \phi_2 \mathbb{I}_{\varepsilon_t > 0} \varepsilon_t^2} \cdot \varepsilon_t\right) = \mu. \tag{7}$$

For the symmetric model, that is $\phi_1 = \phi_2$, in equation 7, the second term vanishes. Since $z_t = \sqrt{g_t + \phi\varepsilon_t^2}$ is an even function and $\varepsilon_t$ an odd function, the expectation becomes an expectation over an odd function which equals zero. For the asymmetric model, this term does not vanish, meaning $\mu$ is the median of the returns and not the mean. The unconditional expectation of $h_t$ in equation 6 can also be calculated as follows:

$$E(h_t) = E(g_t) + \frac{\phi_1}{2}E(\varepsilon_t^2) + \frac{\phi_2}{2}E(\varepsilon_t^2) = E(g_t) + \phi, \tag{8}$$

where $\phi_i E[\mathbb{I}_{\varepsilon_t \leq 0}\varepsilon_t^2] = \frac{\phi_i}{2}E[\varepsilon_t^2]$, $\phi = (\phi_1 + \phi_2)/2$ and $\alpha = (\alpha_1 + \alpha_2)/2$.
Intuitively, this is a logical result, because the process of $g_t$ and $h_t$ are only different by i.i.d. shocks. Using this result, the unconditional expectation over $h_t$ and $(r_t - \mu)^2$ can be determined. Firstly,

$$\begin{aligned} E(h_t) &= \omega + (\alpha_1/2)E[(r_t-\mu)^2|r_{t-1} \leq \mu] + (\alpha_2/2)E[(r_t-\mu)^2|r_{t-1} > \mu] + \beta E(h_{t-1}) + \phi E(\varepsilon_t^2), \\ E[(r_t-\mu)^2|r_t \leq \mu] &= E[h_t\varepsilon_t^2] = E(g_t) + \phi_1 E(\varepsilon_t^4) = E(h_t) - \phi + \phi_1 K_\varepsilon, \\ E[(r_t-\mu)^2|r_t > \mu] &= E[h_t\varepsilon_t^2] = E(g_t) + \phi_2 E(\varepsilon_t^4) = E(h_t) - \phi + \phi_2 K_\varepsilon, \end{aligned} \tag{9}$$

where $K_\varepsilon$ is the kurtosis of $\varepsilon_t$ and equation 8 is used to substitute $E(g_t)$. The two expectations in equation 9 can be combined leading to the following unconditional expectation of $h_t$:

$$h := E(h_t) = \frac{\omega + \phi + \alpha_1(\phi_1 K_\varepsilon - \phi)/2 + \alpha_2(\phi_2 K_\varepsilon - \phi)/2}{1 - \alpha - \beta}, \tag{10}$$

where we make the assumption that $E(h_t) = E(h_{t-1})$ and $0 \leq \alpha + \beta < 1$. When estimating this model, $h_1$ is set equal to $h$ in the above expression to initialise. This means $g_1 = h - \phi$.

When estimating the model, we need filtering equations, to initialise the system and to recursively estimate. For that, a expression for $\varepsilon_t$ needs to be made. For simplicity in the derivations, the indicator functions for $phi_i$ is denoted as

$$F(r_t) := \phi_1 \mathbb{I}_{r_t \leq \mu}\varepsilon_t^2 + \phi_2 \mathbb{I}_{r_t > \mu}\varepsilon_t^2, \tag{11}$$

where $(\varepsilon_t < 0) = (r_t < \mu)$. Following equation 6, we can write:

$$0 = F(r_t)\varepsilon_t^4 + g_t\varepsilon_t^2 - (r_t - \mu)^2.$$

As can be seen, this is a quadratic form equation that can be solved using the following equation:

$$\varepsilon_t^2 = \frac{-g_t + \sqrt{g_t^2 + 4F(r_t)(r_t-\mu)^2}}{2\phi},$$

where, we only evaluate this expression of the abc-formula for the plus sign, since we cannot have a negative value for $\varepsilon_t^2$. Consequently, following equation 6:

$$\begin{aligned} h_t &= g_t + F(r_t)\varepsilon_t^2, \\ &= g_t + \frac{-g_t + \sqrt{g_t^2 + 4F(r_t)(r_t-\mu)^2}}{2}, \\ &= \frac{g_t + \sqrt{g_t^2 + 4F(r_t)(r_t-\mu)^2}}{2} \geq g_t. \end{aligned}$$

This results in two general functions for both $\varepsilon_t$ and $h_t$

$$H(r, \mathcal{I}_{t-1}) = \frac{g_t + \sqrt{g_t^2 + 4F(r_t)(r-\mu)^2}}{2}, \tag{12}$$

$$E(r, \mathcal{I}_{t-1}) = \text{sign}(r-\mu) \cdot \sqrt{\frac{-g_t + \sqrt{g_t^2 + 4F(r_t)(r-\mu)^2}}{2F(r_t)}}, \tag{13}$$

where $\mathcal{I}_{t-1}$ is the information set containing all information up to time $t-1$.

In order to estimate the model by means of Maximum Likelihood (ML), we need a Log-likelihood function. This estimation method is based on the distribution of the returns $r_t$ conditional on information up to time $t-1$.

$$\begin{aligned} f(r|\mathcal{I}_{t-1}) &= \frac{d}{dr} P_\varepsilon \left[E(r, \mathcal{I}_{t-1})\right], \\ &= \frac{dE(r, \mathcal{I}_{t-1})}{dr} \cdot p(E(r, \mathcal{I}_{t-1})), \\ &= \frac{\sqrt{H(r, \mathcal{I}_{t-1})}}{H(r, \mathcal{I}_{t-1}) + F(r_t)E(r, \mathcal{I}_{t-1})^2} \cdot p(E(r, \mathcal{I}_{t-1})), \end{aligned} \tag{14}$$

where $p$ is the p.d.f. of $\varepsilon_t$ given in equation 4. The last line of the derivation is explained in more detail in appendix A. For a realization of $r = r_t$, we get an easier looking function denoted by $h_t$ instead of $H(r, \mathcal{I}_{t-1})$. The log-likelihood function can now be formulated as

$$\mathbf{\Theta}_{ML} = \text{argmax}_{\mathbf{\Theta}} \sum_{t=1}^{T} \left[ \log\left( \frac{\sqrt{h_t}}{h_t + F(r_t)\varepsilon_t^2} \right) + \log p(\varepsilon_t) \right]. \tag{15}$$

The predicted volatility (PVol) over a $d$-day horizon we take as follows:

$$\text{PVol}_t = \sqrt{E_t(\sum_{i=1}^{d} r_{t+i}^2)} = \sqrt{d(\mu^2 + h + \phi(K_\varepsilon - 1)) + \frac{1-(\alpha+\beta)^d}{1-\alpha-\beta}(g_{t+1} + \phi - h)}, \tag{16}$$

where $h$ is defined in equation 10 and $d$ is number of days ahead to forecast.

The NIC is an function that shows how fast the volatility estimate reacts to news represented by shocks. The relation of shocks $\varepsilon_t$ and the volatility estimates $h_t$, is represented by the following equation:

$$
\begin{aligned}
E_t\left[(r_{t+1}^2)\right] =& E_t\left[(\mu + \sqrt{h}_{t+1}\varepsilon_{t+1})^2\right], \\
&= \mu^2 + \phi\kappa_\varepsilon + g_{t+1}, \\
&= \mu^2 + \phi\kappa_\varepsilon + \omega + \alpha(r_t - \mu)^2 + \beta h_t,
\end{aligned} \tag{17}
$$

where $E\left[\varepsilon_t^2\right] = 1$ and $K_\varepsilon = E\left[\varepsilon_t^4\right]$. $h_t$ is defined as in equation 10. The NIC for both GARCH and the RT-GARCH is given in figure 2. Here we see that the shocks in the RT-GARCH have bigger impact on the volatility than with the standard GARCH model. The NIC of the simple GARCH is equal to formulation 17, but for $\phi = 0$. The difference is thus due to the time-varying kurtosis $K_\varepsilon$.

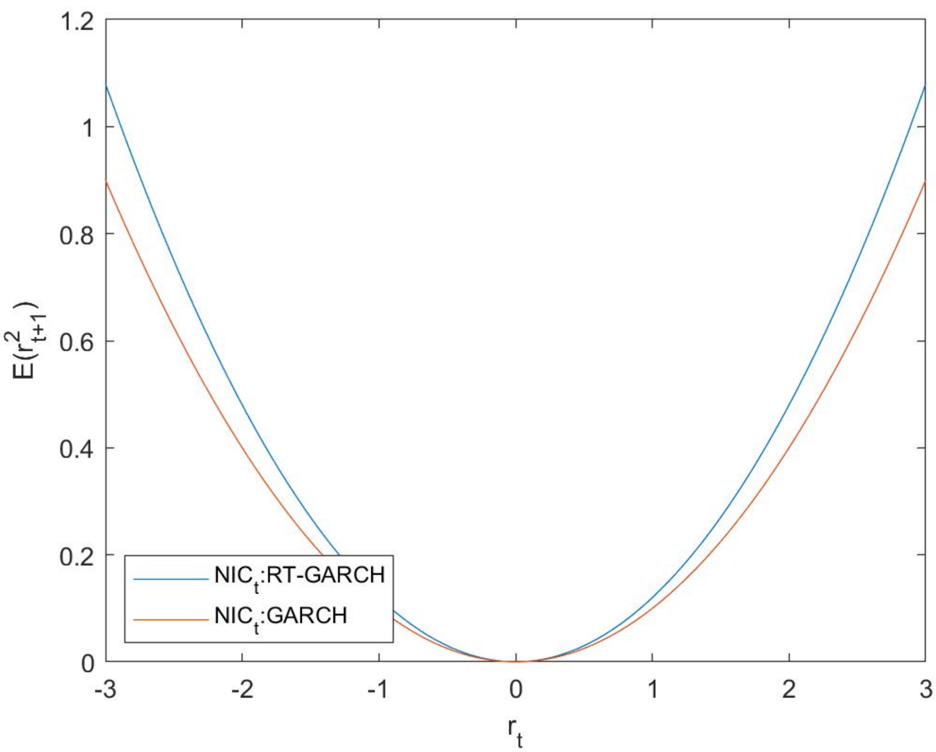



Figure 2: Above are the two estimated NIC's. The parameters for the RT-GARCH are set to $(\mu, \omega, \alpha, \beta, \phi)$ = (0, 0, 0.13, 0.86, 0.01) and for the GARCH model $(\mu, \omega, \alpha, \beta)$ = (0,0,0.11,0.88)

## 3.2 Beta-t-EGARCH

Now on to the Beta-t-EGARCH model described by Harvey and Chakravarty (2008). First of all, the EGARCH part is due to the exponential transformation of the volatility, that is

$$
\sigma_t = \exp(\lambda_t), \tag{18}
$$

where the process of $\lambda_t$ is unrestricted. This class of exponential GARCH models was first introduced by Nelson (1991). The Beta-t-EGARCH formulated by Harvey and Chakravarty is related to the following specification:

$$\begin{aligned} r_t &= \mu + \exp(\lambda_t)\varepsilon_t, \\ \lambda_t &= \lambda(1-\phi) + \phi\lambda_{t-1} + \kappa u_{t-1} + \tilde{\kappa} v_{t-1}, \\ u_t &= \frac{\sqrt{\nu+3}}{\sqrt{2\nu}}\left(\frac{\nu+1}{\nu-2+\varepsilon_t^2}\varepsilon_t^2 - 1\right), \\ v_t &= \frac{\sqrt{(\nu+3)(\nu-2)}}{\sqrt{\nu(\nu+1)}}\frac{\nu+1}{\nu-2+\varepsilon_t^2}\varepsilon_t. \end{aligned} \tag{19}$$

Here, $\varepsilon_t$ follows the same p.d.f. with $\nu > 2$ degrees of freedom. The shocks $\varepsilon_t$ follow a Student's $t$ distribution. For this reason the formulation is dependent on the score of the implied distribution of $r_t$ and a location measure represented by $u_t$ and $v_t$ respectively. The predictive distribution of $r_t$ conditional on information set $\mathcal{I}_{t-1}$ is given as:

$$f(r_t|\mathcal{I}_{t-1}) = \frac{1}{\sigma_t} p\left(\frac{r_t - \mu}{\sigma_t}\right), \tag{20}$$

where $\sigma_t = \exp(\lambda_t)$ and $p(.)$ is given in equation 4 to ensure that $E(\varepsilon_t) = 0$ and $V(\varepsilon_t) = 1$. $u_t$ is then derived by taking the derivative of equation 20, with respect to $\lambda_t$. $v_t$ is the derivative of this same equation with respect to $\mu$.

Now we look at the unconditional expectation of this model:

$$E(r_t) = \mu + \exp(E(\lambda_t))E(\varepsilon_t) = \mu, \tag{21}$$

given $E(\varepsilon_t) = 0$. This means the unconditional expectation of $r_t$ is simply the mean of the return series. The unconditional expectation of $\lambda_t$ is as follows:

$$E(\lambda_t) = \lambda(1-\phi) + \phi E(\lambda_{t-1}) + \kappa E(u_{t-1}) + \hat{\kappa} E(v_{t-1}) = \lambda(1-\phi) + \phi E(\lambda_t) + 0 = \frac{\lambda(1-\phi)}{1-\phi} = \lambda. \tag{22}$$

This results follows from the assumption that $E(\lambda_t) = E(\lambda_{t-1})$ and $E(v_t) = E(u_t) = 0$.

The filtering equations of the Beta-$t$-EGARCH model are more straight forward than that of the RT-GARCH, because the process is now only dependent on past information i.e. information up to time $t-1$. The process needs only to be initialized. $\lambda_1$ will be equal to the unconditional expectation of the process. $\varepsilon_1 = (r_t - \mu)\exp(-\lambda_1)$ and with the residuals, $v_1$ & $u_1$ can be estimated.

For estimation of the model, the maximum likelihood function needs be derived. The probability density function of $r_t$ is given as:

$$f(r_t|\mathcal{I}_{t-1}) = \frac{1}{\exp(\lambda_t)} p(\varepsilon_t), \tag{23}$$

where $p(.)$ is the p.d.f. of the Student's $t$ distribution and $\varepsilon_t = \frac{r_t - \mu}{\exp(\lambda_t)}$ The Log-likelihood function estimator is then:

$$\mathbf{\Theta}_{ML} = \text{argmax}_{\mathbf{\Theta}} \sum_{t=1}^{T} [-\lambda_t + \log(p(\varepsilon_t))] \tag{24}$$

The predicted variance of a an $d$-day horizon can be computed following, this derivation is further explained in appendix B:

$$\text{PVol}_t = \sqrt{d\mu^2 + \sum_{i=1}^{d} \exp\left(2\lambda + 2\phi^{i-1}(\lambda_{t+1} - \lambda) + 2(\kappa^2 + \hat{\kappa}^2)\frac{1-\phi^{2(i-1)}}{1-\phi^2}\right)} \tag{25}$$

The NIC for the Beta-$t$-EGARCH can be viewed as the unexpected part of the difference between $\log(\sigma_{t+1})$ and $\log(\sigma_t)$. The functions reads as follows:

$$NIC_t = \log(\sigma_{t+1}^2) - \log(\sigma_t^2) - E_{t-1}\left[\log(\sigma_{t+1}^2) - \log(\sigma_t^2)\right] = 2\kappa u_t + 2\tilde{\kappa} v_t, \tag{26}$$

where the expectation is taken for information up to $t-1$.

### 3.3 Two-component Beta-$t$-EGARCH

An extension on the general Beta-$t$-EGARCH model is one with two components. One long-run component and one short-run component. The model proposed is described in details by Harvey and Lange (2018) and looks as follows:

$$\begin{aligned} \lambda_t &= \omega + \lambda_{1,t} + \lambda_{2,t}, \\ \lambda_{i,t} &= \phi_i \lambda_{i,t-1} + \kappa_i u_{t-1} + \tilde{\kappa}_i v_{t-1} \quad i = 1, 2. \end{aligned} \tag{27}$$

where $u_t$ and $v_t$ are similar to their definitions in equation 19. If $\phi_1 > \phi_2$, it can be said that $\lambda_{1,t}$ denotes the long-run component.

Just as in the Beta-$t$-EGARCH all parameter estimates follow the log-likelihood function in equation 24. The process assumes the error terms to be Student's $t$ distributed as in all models.

To predict over $d$-day horizon this research uses the following equation:

$$\text{PVol}_t = \sqrt{d\mu^2 + \sum_{i=1}^{d} \exp\left(2\omega + 2\phi_1^{d-1}\lambda_{1,t+1} + 2\phi_2^{d-1}\lambda_{2,t-1}\right) + 2\sum_{i=1}^{2}(\kappa_i^2 + \tilde{\kappa}_i^2)\frac{1-\phi_i^{2(d-1)}}{1-\phi_i^2}}. \tag{28}$$

The derivation of this formula can be found in appendix B.

### 3.4 RT-Beta-$t$-EGARCH

This research designs a new model by the name of RT-Beta-t-EGARCH. This model will have real time features and be driven by robust variance and location measures as in the Beta-t-EGARCH. The model looks as follows:

$$\begin{aligned} r_t &= \mu + \exp(\lambda_t)\varepsilon_t, \\ \lambda_t &= \lambda(1-\phi) + \phi\lambda_{t-1} + \kappa u_{t-1} + \tilde{\kappa} v_{t-1} + \varphi r_t^2, \\ u_t &= \frac{\sqrt{\nu+3}}{\sqrt{2\nu}}\left(\frac{\nu+1}{\nu-2+\varepsilon_t^2}\varepsilon_t^2 - 1\right), \\ v_t &= \frac{\sqrt{(\nu+3)(\nu-2)}}{\sqrt{\nu(\nu+1)}}\frac{\nu+1}{\nu-2+\varepsilon_t^2}\varepsilon_t. \end{aligned} \tag{29}$$

In this model compared to the original Beta-$t$-EGARCH, an extra term $\varphi r_t^2$ is added. The reason to take the squared returns rather than $\varepsilon_t^2$ as in the RT-GARCH of Smetanina (2017), is due to the exponential. When computing a formulation for $\varepsilon_t^2$ in terms of $r_t$ so that the model can be estimated, the need of the lambert W(x) function, an imaginary multidimensional function which is not defined for $\varepsilon_t < -1$, is compulsory.

Compared to the Beta-$t$-EGARCH model, the filtering process does not change to much. The process initialises with starting values of the parameters equal to the estimated parameters in the normal Beta-$t$-EGARCH model and $\varphi = 0$. Then, it will update the $\lambda_t$ series following equation 29.

For estimation of the RT-Beta-$t$-EGARCH, the log-likelihood function is not much different than equation 24. The only main difference is in the definition of $\lambda_t$, that has now an additional parameter $\varphi$.

To predict with the model, the research focuses only on one step ahead forecasts. The one step ahead prediction for volatility is

$$\text{PVol}_t = \sqrt{\mu^2 + 2\mu\exp(\lambda_{t+1})\varepsilon_{t+1} + \exp(\lambda_{t+1})^2\varepsilon_{t+1}^2}. \tag{30}$$

For the RT-Beta-$t$-EGARCH the NIC is defined as follows:

$$\begin{aligned} NIC_t =& 2\lambda_{t+1} - 2\lambda_t - (E[2\lambda_{t+1} - 2\lambda_t]) \\ =& 2\lambda_{t+1} - E_{t-1}[2\lambda_{t+1}] \\ =& 2\kappa u_t + 2\tilde{\kappa} v_t + 2\varphi(r_{t+1}^2 - E_{t-1}[r_{t+1}^2]) \\ =& 2\kappa u_t + 2\tilde{\kappa} v_t \quad \text{(Since, we want to have a function in } r_t \text{, not } r_{t+1}) \end{aligned} \tag{31}$$

This formula is the same as the normal Beta-$t$-EGARCH model, so to see the difference between the two formulas, the difference is determined by parameter estimates of $\kappa$ & $\tilde{\kappa}$. In figure 3, the NIC's are given with parameter estimates later in the research estimated. The NIC behaves differently

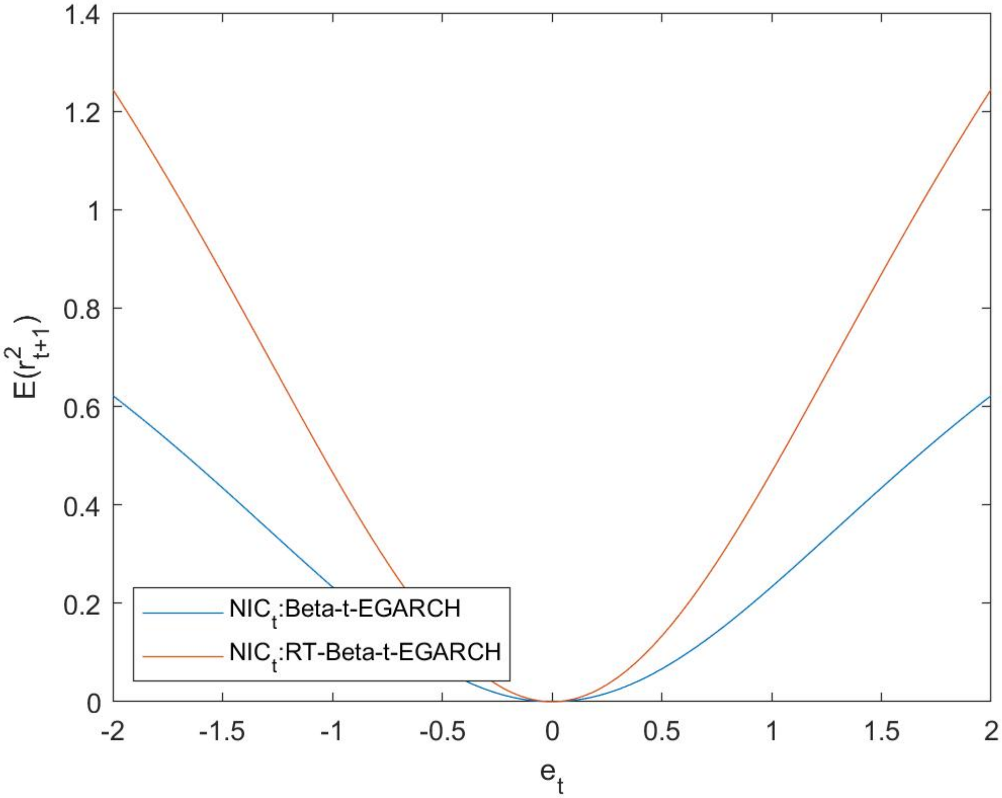



Figure 3: The News Impact Curves for two models where the parameters are Log-likelihood estimators from section 4. For Beta-$t$-EGARCH: $(\kappa, \tilde{\kappa}) = (0.09, -0.01)$. For RT-Beta-$t$-EGARCH: $(\kappa, \tilde{\kappa}) = (0.21, -0.01)$

compared to the GARCH and RT-GARCH. For smaller shocks, the impact on volatility increases quadratic, but as the shocks increase, the effect on volatility slows down. This is a good property, showing robustness of the models. It is desired to have bigger shocks not impact volatility to much since it could potentially be an outlier.

### 3.5 Evaluation

After estimation of the models, evaluation is the next step. The research uses the multi-horizon forecast comparison methods from Quaedvlieg (2021) and alters its loss functions to an quadratic and absolute loss function. The multi-horizon forecast comparison is related to the Diebold-Mariano test. The Diebold-Mariano test takes the difference between an estimated value and the true value denoted by $d_t = y_t - \hat{y}_t$. However, this test only compares model performance for a single forecast horizon. The multi-horizon test maps the forecast errors following a loss function $\mathbf{L} = L(y_t, \hat{y}_{i,t})$ into a H-dimensional vector, with elements $L_{i,t}^h = L(y_t, \hat{y}_{i,t}^h)$. Now the loss differential is denoted by $\mathbf{d}_{ij,t} = \mathbf{L}_{i,t} - \mathbf{L}_{j,t}$, which is an H-dimensional vector. Define $E(\mathbf{d}_{ij,t})= \boldsymbol{\mu}_{ij,t}$. The test statistic this research uses, is the average superior predictive ability (aSPA) discussed by Quaedvlieg (2021) and the. Its definition is as follows:

$$\mu_{ij}^{(Avg)} = \mathbf{w}'\boldsymbol{\mu_{ij}} = \sum_{h=1}^{H} w_h \mu_{ij}^h, \tag{32}$$

where $\mathbf{w}$ is a weight vector.

To show how well a certain model is able to predict for certain lengths of the horizon, the following $\mathbf{w}$ are used:

$$\text{To evaluate long horizon performance: } \mathbf{w} = [0.05; 0.1; 0.2; 0.3; 0.35],$$
$$\text{To evaluate short horizon performance: } \mathbf{w} = [0.35; 0.3; 0.2; 0.1; 0.05],$$
$$\text{To evaluate equally weighted overall performance: } \mathbf{w} = [0.2; 0.2; 0.2; 0.2; 0.2].$$

By adjusting the weights for long horizon and short horizon predictions, a difference in performance for short term and long term can be distinguished.

The used loss functions will be the mean squared error (MSE) and mean absolute error (MAE). In this instance, the H-dimensional vector consists of a 'true' volatility estimate $y$ =RVol and an estimated volatility $\hat{y}$ =PVol for all given horizons.

This research compares the volatility estimates of all against a proxy for the 'true' standard deviation of returns, namely Realized Variance (RV) based on five-minute intraday returns. To compare h-day ahead forecast with h-day ahead RV, the following formula is used to estimate h-day ahead RV:

$$\text{Realized volatility} \;=\; \text{RVol}_t = \sqrt{\frac{4}{3}\sum_{d=1}^{h} rv5_{t+d}} \tag{33}$$

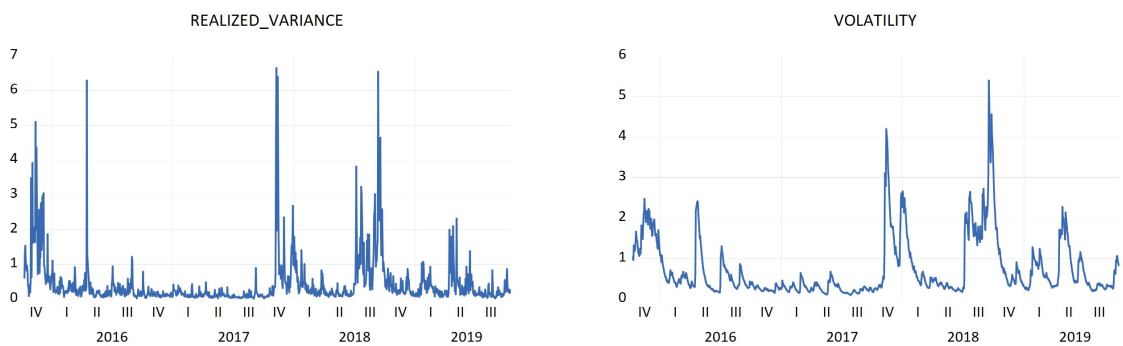
where, $rv5_t$ is the five-minute intraday RV for day $t$. The $\frac{4}{3}$ term is due to overnight volatility that is artificially added via this term Hansen and Lunde (2005). I caculate this $\frac{4}{3}$ factor via the following equation:

$$\hat{\sigma}_t^2 = \hat{c} \cdot RV_t \quad \text{where} \quad \hat{c} = \left( \frac{n^{-1}\sum_{t=1}^{n}(r_t - \hat{\mu}_t)^2}{n^{-1}\sum_{t=1}^{n} RV_t} \right) \approx 1.35 \approx \frac{4}{3} \tag{34}$$

where $RV_t$ are the five-minute intraday $rv5_t$.

# 4 Results

First of all, the estimated volatility and Realized variance for the period October 2015 till December 2019 will be shown. This gives insight in which periods in time volatility is expected to be high and forecasts may be poor.




(a) The day to day Realized Variance for 5-minute intra-day frequency.

(b) The estimated volatility with estimated RT-GARCH model.

Figure 4: Depicted is the volatility and realized variance over the period October 2015 till December 2019.

## 4.1 GARCH and RT-GARCH

### 4.1.1 Parameter estimation of the models

For estimation of all models, the entire dataset of the S&P500 is used. In table 1 the parameter estimates for GARCH, RT-GARCH, Asymmetric RT-GARCH and Smetanina's Asymmetric RT-GARCH where indicator function for feedback effect is different (see equation 5) are given with their likelihood values.

When looking at the Likelihood values, the RT-GARCH has a higher value, meaning the model is a better fit. Since GARCH is nested within we can use the likelihood-ratio test to see whether this difference is significant:

$H_0 : \Theta = \Theta_0$ & $H_1 : \Theta = \Theta_1$.

$\lambda_{LR} = -2\left[l(\Theta_0) - l(\Theta_1)\right]$

Table 1: For all parameters the values have been estimated following the log-likelihood function described in section 3. The values within brackets are the standard deviations. All values in bold are significant at a 1% level. The asymmetric RT-GARCH are different in their indicator function for $\alpha$. Equation 6 uses indicator that relies on the lagged returns, while the equation 5 given by Smetanina uses indicator function for the lagged squared returns that relies on current returns.

| model | $\mu$ | $\omega$ | $\alpha$ | $\beta$ | $\phi$ | | | $\nu$ | Likelihood |
|---|---|---|---|---|---|---|---|---|---|
| GARCH | 0.060 | **0.021** | **0.128** | **0.861** | | | | | -6992 |
| | (0.011) | (0.003) | (0.009) | (0.008) | | | | | |
| RT-GARCH | **0.071** | **0.005** | 0.115 | **0.880** | **0.008** | | | 7.205 | -6868 |
| | (0.010) | (0.006) | (0.010) | (0.004) | (0.007) | | | | |
| | $\mu$ | $\omega$ | $\alpha_1$ | $\alpha_2$ | $\beta$ | $\phi_1$ | $\phi_2$ | $\nu$ | Likelihood |
| As-RT-GARCH | **0.040** | **0.000** | 0.162 | **0.000** | **0.885** | **0.051** | **0.003** | 17.341 | -6750 |
| Indicator equation 6 | ( 0.010) | (0.005) | (0.017) | (0.010) | (0.010) | (0.009) | (0.003) | | |
| As-RT-GARCH | **0.070** | **0.000** | 0.138 | **0.035** | **0.878** | 0.064 | **0.000** | 15.127 | -6793 |
| Indicator equation 5 | (0.001 ) | (0.005) | (0.016) | (0.010) | (0.010) | (0.011) | (0.004) | | |

where, $\Theta_0$ & $\Theta_1$ are the parameter estimates of GARCH and RT-GARCH respectively. The likelihood ratio test gives $\lambda_{LR} = 248 >>> 6$ (critical value). All parameters are significant on a 5% critical value.

Interestingly, the asymmetric RT-GARCH model introduced by Smetanina (2017) has a significantly lower loglikelihood than the asymmetric RT-GARCH in equation 6, with a likelihood-ratio test statistic of $86 >> 6$, meaning that the indicator function by Smetanina is not optimal.

#### 4.1.2 Forecasting performance

To see which model is able to perform better forecasts, a pseudo out-of-sample forecast is done. For 5 different horizons being (1;2;5;10;15), Quaedvlieg's Multi-Horizon forecast is given in table 4.

The difference between multi-horizon test statistics between long term and short term emphasis is within line of expectation. For longer step ahead forecasts, the predictions get less accurate. The research proved the indicator function of Smetanina earlier to be wrong based on the fit and this is also underlined by the predictions with lower test statistic values for both the MSE as the MAE. Adding asymmetry improves the predictive capability of the model with 6% (that is the

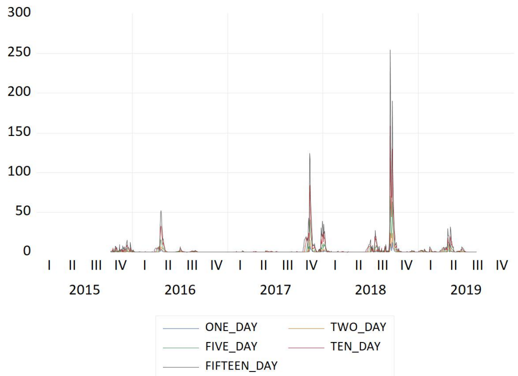
Table 2: Given above are all test statistics for two different loss functions, namely Mean Squared Error (MSE) and Mean Absolute Error (MAE). Long is for weight vector [0.05;0.1;0.2;0.3;0.35], short is for weight vector [0.35;0.3;0.2;0.1;0.05] and mean for equal weight 0.2. The horizons predicted for are $d = [1; 2; 5; 10; 15]$.

| | MSE | | | MAE | | |
|---|---|---|---|---|---|---|
| | Long | Short | Mean | Long | Short | Mean |
| GARCH | 0.586 | 0.221 | 0.403 | 0.188 | 0.108 | 0.148 |
| RT-GARCH | 0.243 | 0.088 | 0.166 | 0.171 | 0.094 | 0.133 |
| As. RT-GARCH | 0.229 | 0.084 | 0.156 | 0.170 | 0.092 | 0.131 |
| As. RT-GARCH (Smetanina's) | 0.259 | 0.093 | 0.176 | 0.187 | 0.101 | 0.144 |
| VIX | 0.248 | 0.098 | 0.173 | 0.193 | 0.115 | 0.154 |

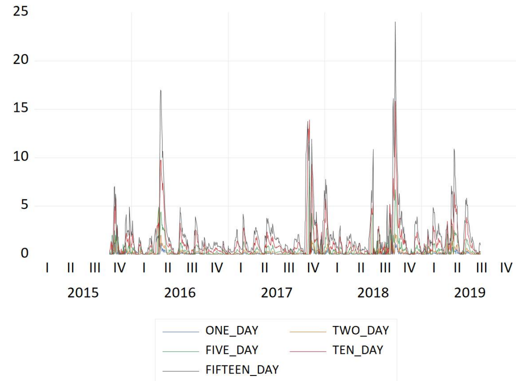
Mean statistic for MAE compared to the RT-GARCH). The VIX, which is an known authority for volatility modelling, is not able to outperform the asymmetric RT-GARCH from this research's indicator function but is equally well performing as the asymmetric RT-GARCH of Smetanina. Also, the asymmetric RT-GARCH following Smetanina is not predicting better than the normal RT-GARCH.

To get a better insight of the pseudo out-of-sample predictions, the difference between volatility forecasts and Realized Variance for the given horizons are depicted in figure 5.

Figure 5 shows the following. First of all, both models seem to perform badly during the same periods in time (take for instance the first and second quarter of 2016). Secondly, the spikes in the graph seem to be higher for the GARCH model. The GARCH model showed to predict worse based on table 4 and the model does have more trouble predicting for volatile periods in time. Another important take-away is the fact that the GARCH model is over estimating volatility more often than the RT-GARCH. This can be concluded from the fact that the RT-GARCH has more negative spikes and GARCH more positive spikes (spikes being big differences between forecast and RV).




(a) For normal GARCH model




(b) For RT-GARCH model with Student's $t$ distribution

Figure 5: Depicted are the squared prediction differences of the two models over the prediction period for all five different horizons. The comparison is towards the Realized Variance.

## 4.2 Beta-$t$-EGARCH, two-component Beta-$t$-EGARCH and RT-Beta-$t$-EGARCH

### 4.2.1 Parameter estimation of the models

Now on to the (RT-)Beta-$t$-EGARCH models. Below, in table 3 we see the estimated parameters for all models with their likelihoods. Based on the latter, the RT-Beta-$t$-EGARCH shows a better fit for the S&P500. Performing a Likelihood ratio test once more gives:

$H_0 : \Theta = \Theta_0$ & $H_1 : \Theta = \Theta_1$.

$\lambda_{LR} = -2\left[l(\Theta_0) - l(\Theta_1)\right]$

where, $\Theta_0$ & $\Theta_1$ are the parameter estimates of Beta-$t$-EGARCH and RT-Beta-$t$-EGARCH respectively. The likelihood ratio test gives $\lambda_{LR} = 1766 >>> 6$ (critical value).

Interestingly, the $\varphi$ in the RT-Beta-$t$-EGARCH seems to absorb a lot of explanatory power. Looking at $\kappa$ and $\tilde{\kappa}$, the 'real time' model is still dependent on past information. Especially the score and location measure. The explanatory power of previous volatility, that is parameter $\phi$, has reduces drastically compared to the normal Beta-$t$-EGARCH.

The Two-component Beta-$t$-EGARCH model has a marginally higher log-likelihood than the Beta-$t$-EGARCH with a likelihood ratio test statistic of $18 > 6$, showing to be significantly higher. Parameter $\tilde{\kappa_i}$, that represents the effect of the location measure is positive for long run volatility measure $\lambda_{1,t}$ and negative for short term volatility measure $\lambda_{2,t}$, meaning that the location of the previous returns, only have short term effects on volatility.

Table 3: B-$t$-E: Beta-$t$-EGARCH, RT-B-$t$-E: RT-Beta-$t$-EGARCH and TC-B-$t$-E: Two-component Beta-$t$-EGARCH. All parameters in bold are significant at 1% level. Parameters estimated following their log-likelihood functions given in section 3

| model | $\mu$ | $\lambda$ | $\phi$ | $\kappa$ | $\tilde{\kappa}$ | $\varphi$ | | | $\nu$ | likelihood |
|---|---|---|---|---|---|---|---|---|---|---|
| B-$t$-E | **0.072** | 0.804 | **0.977** | 0.099 | **-0.023** | | | | 5.77 | -6884 |
| | (0.010) | (0.731) | (0.007) | (0.015) | (0.027) | | | | | |
| RT-B-$t$-E | **0.046** | 0.000 | **0.008** | 0.575 | -0.663 | **0.227** | | | 340.33 | -6001 |
| | (0.008) | (0.058) | (0.010) | (0.048) | (0.062) | (0.007) | | | | |
| TC-B-$t$-E | $\mu$ | $\omega$ | $\phi_1$ | $\phi_2$ | $\kappa_1$ | $\tilde{\kappa}_1$ | $\kappa_2$ | $\tilde{\kappa}_2$ | | |
| | **0.072** | 0.249 | **0.994** | 0.910 | **0.034** | **0.001** | 0.081 | -0.043 | 5.882 | -6875 |
| | (0.008) | (0.153) | (0.002) | (0.023) | (0.008) | (0.002) | (0.013) | (0.026) | | |

#### 4.2.2 Forecasting performance

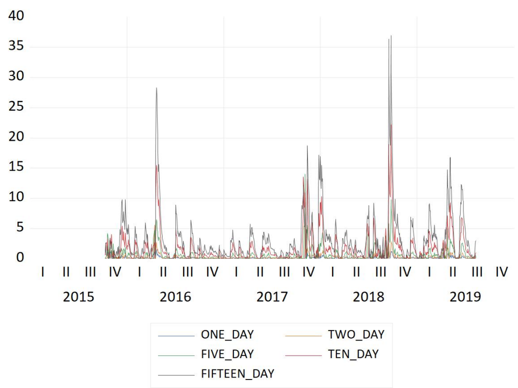
For prediction of volatility using Beta-$t$-EGARCH and the two component variant, we get the following values when it comes to the multi-horizon prediction test of Quaedvlieg (2021):

Table 4: Given above are all test statistics for two different loss functions, namely Mean Squared Error (MSE) and Mean Absolute Error (MAE). Long is for weight vector [0.05;0.1;0.2;0.3;0.35], short is for weight vector [0.35;0.3;0.2;0.1;0.05] and mean for equal weight 0.2. The horizons predicted for are $d = [1; 2; 5; 10; 15]$.

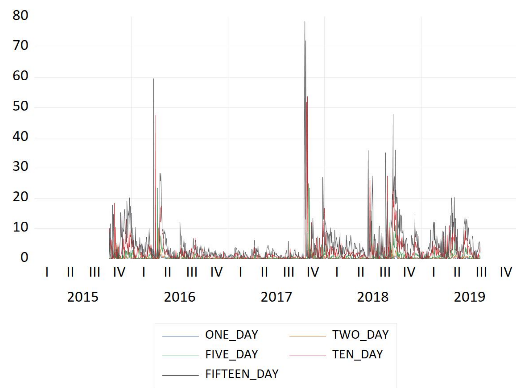
| | MSE | | | MAE | | |
|---|---|---|---|---|---|---|
| | Long | Short | Mean | Long | Short | Mean |
| Beta-$t$-EGARCH | 0.359 | 0.119 | 0.239 | 0.206 | 0.105 | 0.155 |
| two component Beta-$t$-EGARCH | 0.504 | 0.156 | 0.330 | 0.239 | 0.115 | 0.177 |

Based on the Multi-horizon test statistics, the Beta-$t$-EGARCH performs better than the GARCH model, but worse than the RT-GARCH model (see table 4), since all values are in between test statistics of said models. This means the real-time feature has proven to be valuable. The Beta-$t$-EGARCH is however not able to outperform the VIX in table 4. This means there is still ground to win in the Beta-$t$-EGARCH model formulations. The two component variant had proven to be a better fit than the Beta-$t$-EGARCH based on its likelihood value, but is not able to predict better.

All values for both MSE as MAE are higher.




(a) For normal Beta-$t$-EGARCH model




(b) For two component Beta-$t$-EGARCH model

Figure 6: Depicted are the squared prediction differences of the two models over the prediction period for all five different horizons. The comparison is towards the Realized Variance.

Taking a closer look at the mean difference per iteration of every horizon in figure 6, it is clear that during volitile periods in time, the two component Beta-$t$-EGARCH model has an exponential difference almost twice as big.

For the RT-Beta-$t$-EGARCH one step ahead forecasts compared with the one step ahead forecasts of RT-GARCH and Beta-$t$-EGARCH, the following MSE and MAE are found:

Table 5: In the table the Mean Squared Error and Mean Absolute Error are given for the three models. The prediction is for one step ahead forecasts.

| model | MSE | MAE |
|---|---|---|
| RT-GARCH | 0.023 | 0.053 |
| Beta-$t$-EGARCH | 0.027 | 0.056 |
| RT-Beta-$t$-EGARCH | 0.051 | 0.075 |

For the one step ahead forecasts of the RT-beta-$t$-EGARCH the main finding is that it is not able to predict better than its competitors. This means that the 'real time' feature did not add predictive value to the model.

# 5 Conclusion

In the introduction the question was raised which of the RT-GARCH and Beta-$t$-EGARCH models were able to predict volatility best. In section 4 it was shown that the RT-GARCH model has the best fit for the S&P500 data since its log-likelihood value is significantly higher than both the simple GARCH as the Beta-$t$-EGARCH. The asymmetric RT-GARCH as this research describes it with different indicator functions than Smetanina's outperforms all models on a predictive level. This is due to the low values of both the MSE and MAE test statistic. By means of the multihorizon test statistic provided by Quaedvlieg (2021), it is clearly visible the simple GARCH and VIX are no match for models with 'real time' features.

To investigate the robustness of both models, meaning the responsiveness towards returns, the NICs were set up. The NIC of the RT-GARCH proves that the effect of shocks on volatility are bigger compared to the GARCH model. However, the NIC of Beta-$t$-EGARCH shows a feature both GARCH and RT-GARCH were missing. In this NIC it could be seen that for bigger shocks that could potentially be viewed as outliers, the effect on volatility seemed to decline, making the Beta-$t$-EGARCH model potentially more robust.

These very results lead to a combination of the two models by the name of RT-Beta-$t$-EGARCH. After estimation, the log-likelihood value was significantly higher than the Beta-$t$-EGARCH. This result shows lots of promise. The NIC of the RT-Beta-$t$-EGARCH has some matching properties as the NIC of Beta-$t$-EGARCH. The shape is almost equal, meaning that it is less sensitive for outliers than standard GARCH models. The NIC however, was also more steep. This means the effects of shocks are bigger on volatility. However, the one step ahead forecasts of the model are not better than the simple RT-GARCH and Beta-$t$-EGARCH. Therefore in conclusion, the construction of the RT-Beta-$t$-EGARCH is not optimal and the research advises against application of the model.

All these findings lead to a better understanding on RT-GARCH and Beta-$t$-EGARCH models. The main research question that was raised in the introduction can now be answered. Both models are well able to capture volatility. The RT-GARCH model performs better on predictive level, while the Beta-$t$-EGARCH seems to more robust and handle potential outliers in the data. A combination of the both could be further investigated, where the 'real time' feature could be represented by the squared errors as in RT-GARCH instead of squared current returns.

## A Log-likelihood function RT-GARCH

The derivation of function 14:

The final step I describe in this equation, needs use of the chain rule. This derivation is provided by R. Lange in his notes on the RT-GARCH model of Smetanina.

$$\begin{aligned}
\frac{d}{dr}E[r,\mathcal{I}_{t-1}] &= \text{sign}(r-\mu)\cdot\frac{d}{dr}\sqrt{\frac{-g_t+\sqrt{g_t^2+4F(r_t)(r-\mu)^2}}{2F(r_t)}},\\
&= \frac{1}{2E(r,\mathcal{I}_{t-1}}\frac{d}{dr}\left[\frac{-g_t+\sqrt{g_t^2+4F(r)(r-\mu)^2}}{2F(r)}\right],\\
&= \frac{1}{2E(r,\mathcal{I}_{t-1}}\frac{8F(r)(r-\mu)}{4F(r)\sqrt{g_t^2+4F(r)(r-\mu)^2}},\\
&= \frac{1}{E(r,\mathcal{I}_{t-1}}\frac{(r-\mu)}{\sqrt{g_t^2+4F(r)(r-\mu)^2}},\\
&= \frac{1}{E(r,\mathcal{I}_{t-1}}\frac{(r-\mu)}{F(r)E(r,\mathcal{I}_{t-1}+H(r,\mathcal{I}_{t-1})},\\
&= \frac{\sqrt{H(r,\mathcal{I}_{t-1})}}{H(r,\mathcal{I}_{t-1})+F(r)E(r,\mathcal{I}_{t-1})^2}.
\end{aligned} \tag{35}$$

The derivation can only be calculated for $r \neq \mu$ due to the sign function, but the expression we're left with is can now be directly evaluated. This is because $E(r,\mathcal{I}_{t-1})$ and $H(r,\mathcal{I}_{t-1})$ are continuous functions in $r$.

## B two component Beta-$t$-EGARCH prediction formula

$$\begin{aligned}
E_t\left(\sum_{i=1}^{d} r_{t+i}^2\right) &= d\mu^2 + E_t\left(\sum_{i=1}^{d}(r_{t+i}-\mu)^2\right),\\
&= d\mu^2 + E_t\sum_{i=1}^{d}\exp(2\lambda_{t+i|t+i-1})\varepsilon_{t+i}^2,\\
&= d\mu^2 + \sum_{i=1}^{d}E_t\exp(2\lambda_{t+i|t+i-1}), \quad \text{since } \varepsilon_t \text{ i.i.d. with mean zero and variance one},\\
&\approx d\mu^2 + \sum_{i=1}^{d}\exp(2E_t[\lambda_{t+i|t+i-1}]+2V_t[\lambda_{t+i|t+i-1}]) \quad \text{log normal approximation.}
\end{aligned} \tag{36}$$

For the standard Beta-$t$-EGARCH model, I use the derivation by R. Lange, that is:

$$\begin{aligned} E_t[\lambda_{t+i|t+i-1}] &= \lambda + \phi^{i-1}(\lambda_{t+1} - \lambda), \\ V_t[\lambda_{t+i|t+i-1}] &= (\kappa^2 + \tilde{\kappa}^2)\frac{1-\phi^{2(i-1)}}{1-\phi^2}. \end{aligned} \tag{37}$$

For the two component model, the expectation and variance differ however. Firstly,

$$\begin{aligned} E_t[\lambda_{t+i|t+i-1}] &= E_t[\omega + \lambda_{1,t+i|t+i-1} + \lambda_{1,t+i|t+i-1}], \\ &= \omega + \phi_1^{i-1}\lambda_{1,t+1} + \phi_2^{i-1}\lambda_{2,t+1}. \end{aligned} \tag{38}$$

$$\begin{aligned} V_t[\lambda_{t+i|t+i-1}] &= V_t[\lambda_{1,t+i|t+i-1}] + V_t[\lambda_{2,t+i|t+i-1}] + 2cov[\lambda_{1,t+i|t+i-1}, \lambda_{2,t+i|t+i-1}], \\ &= (\kappa_1^2 + \tilde{\kappa_1}^2)\frac{1-\phi_1^{2(i-1)}}{1-\phi_1^2} + (\kappa_2^2 + \tilde{\kappa_2}^2)\frac{1-\phi_2^{2(i-1)}}{1-\phi_2^2} \end{aligned} \tag{39}$$

# C Brief explanation used MATLAB code

The zip file containing all code contains four types of functions for all models. The GarchFilter.m, which describes how to initialise the process. The NegativeLogLikelihood.m, that gives the log likelihood function for the model, where the log likelihood is taken to be negative since MATLAB minimizes rather than maximizes. The predictor-horizons.m function, accepts a vector of five different forecast horizons and using the estimated parameters, it produces forecasts for the five different horizons and calculates the multi-horizon test statistic. The demo.m file is the code that is used to call all functions and estimate the models. For each model in the demo file, starting values of parameters and their lower- and upperbound are given.

Table 6: In this table, the indicator number per function for each model in the zip file can be found. For instance, the Beta-$t$-EGARCH model uses: garchfilter3.m and so on.

| model | GarchFilter | NegativeLogLikelihood | predictor-horizons |
|---|---|---|---|
| GARCH | - | - | - |
| RT-GARCH | 2 | 2 | 2 |
| Beta-$t$-EGARCH | 3 | 3 | 3 |
| RT-Beta-$t$-EGARCH | 4 | 4 | 4 |
| two component Beta-$t$-EGARCH | 5 | 5 | 5 |
| asymmetric RT-GARCH | 6 | 6 | 6 |
| asymmetric RT-GARCH (smetanina) | 7 | 6 | 6 |

## References


Ding, Y. (2021). Augmented Real-Time GARCH: A Joint Model for Returns, Volatility and Volatility of Volatility. (2112). https://ideas.repec.org/p/cam/camdae/2112.html

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), 1–50. https://doi.org/10.1080/07474938608800095

Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a garch(1,1)? *Journal of Applied Econometrics*, *20*(7), 873–889. https://doi.org/https://doi.org/10.1002/jae.800

Harvey, A., & Chakravarty, T. (2008). Beta-t-(e)garch. https://EconPapers.repec.org/RePEc:cam:camdae:0840

Harvey, A., & Lange, R.-J. (2018). Modeling the interactions between volatility and returns using egarch-m. *Journal of Time Series Analysis*, *39*(6), 909–919. https://doi.org/https://doi.org/10.1111/jtsa.12419

Lange, R.-J. (2021). A short explanation of rt-garch (with proofs).

Lunde, A., & Hansen, P. (2004). Realized variance and iid market microstructure noise. *Econometric Society, Econometric Society 2004 North American Summer Meetings*, *24*. https://doi.org/10.2139/ssrn.506542

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, *59*(2), 347–70. https://EconPapers.repec.org/RePEc:ecm:emetrp:v:59:y:1991:i:2:p:347-70

Patton, A. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*(1), 246–256. https://EconPapers.repec.org/RePEc:eee:econom:v:160:y:2011:i:1:p:246-256

Quaedvlieg, R. (2021). Multi-horizon forecast comparison. *Journal of Business & Economic Statistics*, *39*(1), 40–53. https://doi.org/10.1080/07350015.2019.1620074

Smetanina, E. ( (2017). Real-time garch. *Journal of Financial Econometrics*, *15*, 561–601. https://doi.org/10.1093/jjfinec/nbx008