1

Literature review

1.1 Stylized facts of returns

When analyzing returns as a time-series, we look at log returns. The log returns are similar to simple returns so the stylized facts of returns apply to both. One assumption that is made often in financial applications is that returns are iid, or independently and identically distributed, another is that they are normally distribution. Are these valid assumptions? Below the stylized facts[1] following Annaert [1] for returns are given.

- Returns are *small and volatile* (with the standard deviation being larger than the mean on average)
- Returns have very little serial correlation as mentioned by for example Bollerslev [2].
- Returns exhibit conditional heteroskedasticity, or *volatility clustering*. There is no constant variance, but it is time-varying (homoskedasticity). Bollerslev [2] describes it as “rates of return data are characterized by volatile and tranquil periods”.

[1]Stylized facts are the statistical properties that appear to be present in many empirical asset returns (across time and markets)

- Returns also exhibit *asymmetric volatility*, in that sense volatility increases more after a negative return shock than after a large positive return shock. This is also called the *leverage effect.*

- Returns are *not normally distributed* which is also one of the conclusions by Fama [3]. Returns have tails fatter than a normal distribution (leptokurtosis) and thus are riskier than under the normal distribution. Log returns **can** be assumed to be normally distributed. However, this will be examined in our empirical analysis if this is appropriate. This makes that simple returns follow a log-normal distribution, which is a skewed density distribution.

Firms holding a portfolio have a lot of things to consider: expected return of a portfolio, the probability to get a return lower than some threshold, the probability that an asset in the portfolio drops in value when the market crashes. All the previous requires information about the return distribution or the density function. What we know from the stylized facts of returns that the normal distribution is not appropriate for returns. Below we summarize some alternative distributions that could be a better approximation of returns than the normal one.

1.1.1 Alternative distributions than the normal

Student's t-distribution

A common alternative for the normal distribution is the Student t distribution. Similarly to the normal distribution, it is also symmetric (skewness is equal to zero). The probability density function (pdf), again following Annaert [1], is given by equation (1.1). As will be seen in 1.2, GARCH models are used for volatility modeling in practice. Bollerslev [2] examined the use of the GARCH-Student or GARCH-t model as an alternative to the standard Normal distribution, which relaxes the assumption of conditional normality by assuming the standardized innovation to follow a standardized Student t-distribution [4].

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n}}(1+\frac{x^2}{n})^{-(n+1)/2} \tag{1.1}$$

As can be seen the pdf depends on the degrees of freedom n. To be consistent with Ghalanos [5], the following general equation is used for the pdf (1.2).

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\beta\pi\nu}}\left(1+\frac{(x-\alpha)^2}{\beta\nu}\right)^{-(\nu+1)/2} \tag{1.2}$$

where α, β and ν are respectively the location, scale and shape (tail-thickness) parameters. The symbol Γ is the Gamma function.

Unlike the normal distribution, which depends entirely on two moments only, the student t distribution has fatter tails (thus it has a kurtosis coefficient), if the degrees of freedom are finite. This kurtosis coefficient is given by equation (1.3). This is useful while as already mentioned, the standardized residuals appear to have fatter tails than the normal distribution following Bollerslev [4].

$$kurt = 3 + \frac{6}{n-4} \tag{1.3}$$

Generalized Error Distribution

The GED distribution is nested in the generalized t distribution by McDonald and Newey [6] is used in the GED-GARCH model by Nelson [7] to model stock market returns. This model replaced the assumption of conditional normally distributed error terms by standardized innovations that following a generalized error distribution. It is a symmetric, unimodal distribution (location parameter is the mode, median and mean). This is also sometimes called the exponential power distribution [4]. The conditional density (pdf) is given by equation (1.4) following Ghalanos [5].

$$f(x) = \frac{\kappa e^{\left|\frac{x-\alpha}{\beta}\right|^{\kappa}}}{2^{1+\kappa^{(-1)}}\beta\Gamma(\kappa^{-1})} \tag{1.4}$$

where α, β and κ are respectively the location, scale and shapeparameters .

Skewed t-distribution

The density function can be derived following Fernández and Steel [8] who showed how to introduce skewness into uni-modal standardized distributions [9]. The first equation from Trottier and Ardia [9], here equation (1.5) presents the skewed t-distribution.

$$f_\xi(z) \equiv \frac{2\sigma_\xi}{\xi + \xi^{-1}} f_1\left(z_\xi\right), \quad z_\xi \equiv \begin{cases} \xi^{-1}\left(\sigma_\xi z + \mu_\xi\right) & \text{if } z \geq -\mu_\xi/\sigma_\xi \\ \xi\left(\sigma_\xi z + \mu_\xi\right) & \text{if } z < -\mu_\xi/\sigma_\xi \end{cases} \tag{1.5}$$

where $\mu_\xi \equiv M_1\left(\xi - \xi^{-1}\right)$, $\sigma_\xi^2 \equiv (1 - M_1^2)\left(\xi^2 + \xi^{-2}\right) + 2M_1^2 - 1$, $M_1 \equiv 2\int_0^\infty u f_1(u)du$ and ξ between 0 and ∞. $f_1(\cdot)$ is in this case equation (1.1), the pdf of the student t distribution.

According to Giot and Laurent [10]; Giot and Laurent [11], the skewed t-distribution outperforms the symmetric density distributions.

Skewed Generalized Error Distribution

What also will be interesting to examine is the SGED distribution of Theodossiou [12] in GARCH models, as in the work of Lee, Su, and Liu [13]. The SGED distribution extends the Generalized Error Distribution (GED) to allow for skewness and leptokurtosis. The density function can be derived following Fernández and Steel [8] who showed how to introduce skewness into uni-modal standardized distributions [9]. It can also be found in Theodossiou [12]. The pdf is then given by the same equation (1.5) as the skewed t-distribution but with $f_1(\cdot)$ equal to equation (1.4).

Skewed Generalized t-distribution

The SGT distribution of introduced by Theodossiou [14] and applied by Bali and Theodossiou [15] and Bali, Mo, and Tang [16]. According to Bali, Mo, and Tang [16] the proposed solutions (use of historical simulation, student's t-distribution, generalized error distribution or a mixture of two normal distributions) to the non-normality of standardized financial returns only partially solved the issues of skewness and leptokurtosis. The density of the generalized t-distribution of McDonald and Newey [6] is given by equation (1.6) [17].

$$f\left[\varepsilon_t \sigma_t^{-1}; \kappa, \psi\right] = \frac{\kappa}{2\sigma_t \ \cdot \psi^{1/\kappa} B(1/\kappa, \psi) \cdot \left[1 + |\varepsilon_t|^{\kappa} / \left(\psi b^{\kappa} \sigma_t^{\kappa}\right)\right]^{\psi + 1/\kappa}} \tag{1.6}$$

where $B(1/\eta, \psi)$ is the beta function ($=\Gamma(1/\eta)\Gamma(\psi)\Gamma(1/\eta + \psi)$), $\psi\eta > 2,\ \eta > 0\ \mathit{and}\ \psi > 0,\ \beta = [\Gamma(\psi)\Gamma(1/\eta)/\Gamma(3/\eta)\Gamma(\psi - 2/\eta)]^{1/2}$), the scale factor and one shape parameter κ.

Again the skewed variant is given by equation (1.5) but with $f_1(\cdot)$ equal to equation (1.6) following Trottier and Ardia [9].

1.2 Volatility modeling

1.2.1 Rolling volatility

When volatility needs to be estimated on a specific trading day, the method used as a descriptive tool would be to use rolling standard deviations. Engle [18] explains the calculation of rolling standard deviations, as the standard deviation over a fixed number of the most recent observations. For example, for the past month it would then be calculated as the equally weighted average of the squared deviations from the mean (i.e. residuals) from the last 22 observations (the average amount of trading or business days in a month). All these deviations are thus given an equal weight. Also, only a fixed number of past recent observations is examined. Engle regards this formulation as the first ARCH model.

1.2.2 ARCH model

Autoregressive Conditional Heteroscedasticity (ARCH) models, proposed by Engle [19], was in the first case not used in financial markets but on inflation. Since then, it has been used as one of the workhorses of volatility modeling. To fully capture the logic behind GARCH models, the building blocks are examined in the first place. There are three building blocks of the ARCH model: returns, the innovation process and the variance process (or volatility function), written out in respectively equation (1.7), (1.8) and (1.9). Returns are written as a constant part (μ) and an unexpected part, called noise or the innovation process. The innovation process is the volatility (σ_t) times z_t, which is an independent identically distributed random variable with a mean of 0 (zero-mean) and a variance of 1 (unit-variance). The independent from iid, notes the fact that the z-values are not correlated, but completely independent of each other. The distribution is not yet assumed. The third component is the variance process or the expression for the volatility. The variance is given by a constant ω, plus the random part which depends on the return shock of the previous period squared (ε_{t-1}^2). In that sense when the uncertainty or surprise in the last period increases, then the variance becomes larger in the next period. The element σ_t^2 is thus known at time $t-1$, while it is a deterministic function of a random variable observed at time $t-1$ (i.e. ε_{t-1}^2).

$$R_t = \mu + \varepsilon_t \tag{1.7}$$

$$\varepsilon_t = \sigma_t * z_t,\ where\ z_t \overset{iid}{\sim} (0, 1) \tag{1.8}$$

$$\sigma_t^2 = \omega + \alpha_1 * \varepsilon_{t-1}^2 \tag{1.9}$$

From these components we could look at the conditional moments (or expected returns and variance). We can plug in the component σ_t into the conditional mean innovation ε_t and use the conditional mean innovation to examine the conditional mean return. In equation (1.10) and (1.11) they are derived. Because the random variable z_t is distributed with a zero-mean, the conditional expectation is 0. As a consequence, the conditional mean return in equation (1.11) is equal to the unconditional mean in the most simple case. But variations are possible using ARMA (eg. AR(1)) processes.

$$\mathbb{E}_{t-1}(\varepsilon_t) = \mathbb{E}_{t-1}(\sqrt{\omega + \alpha_1 * \varepsilon_{t-1}^2} * z_t) = \sigma_t \mathbb{E}_{t-1}(z_t) = 0 \tag{1.10}$$

$$\mathbb{E}_{t-1}(R_t) = \mu + \mathbb{E}_{t-1}(\varepsilon_t) = \mu \tag{1.11}$$

For the conditional variance, knowing everything that happened until and including period $t-1$ the conditional innovation variance is given by equation (1.12). This is equal to σ_t^2, while the variance of z_t is equal to 1. Then it is easy to derive the conditional variance of returns in equation (1.13), that is why equation (1.9) is called the variance equation.

$$var_{t-1}(\varepsilon_t) = \mathbb{E}_{t-1}(\varepsilon_t^2) = \mathbb{E}_{t-1}(\sigma_t^2 * z_t^2) = \sigma_t^2 \mathbb{E}_{t-1}(z_t^2) = \sigma_t^2 \tag{1.12}$$

$$var_{t-1}(R_t) = var_{t-1}(\varepsilon_t) = \sigma_t^2 \tag{1.13}$$

The unconditional variance is also interesting to derive, while this is the long-run variance, which will be derived in (1.17). After deriving this using the law of iterated expectations and assuming stationarity for the variance process, one would get (1.14) for the unconditional variance, equal to the constant c and divided by $1-\alpha_1$, the slope of the variance equation.

$$\sigma^2 = \frac{\omega}{1-\alpha_1} \tag{1.14}$$

This leads to the properties of ARCH models.

- Stationarity condition for variance: $\omega > 0$ and $0 \leq \alpha_1 < 1$.
- Zero-mean innovations
- Uncorrelated innovations

Thus a weak white noise process ε_t

Stationarity implies that the series on which the ARCH model is used does not have any trend and has a constant expected mean. Only the conditional variance is changing.

The unconditional 4th moment, kurtosis $\mathbb{E}(\varepsilon_t^4)/\sigma^4$ of an ARCH model is given by equation (1.15). This term is larger than 3, which implicates that the fat-tails (a stylised fact of returns).

$$3\frac{1-\alpha_1^2}{1-3\alpha_1^2} \tag{1.15}$$

Another property of ARCH models is that it takes into account volatility clustering. Because we know that $var(\varepsilon_t) = \mathbb{E}(\varepsilon_t^2) = \sigma^2 = \omega/(1-\alpha_1)$, we can plug in ω for the conditional variance $var_t(\varepsilon_{t+1}) = \mathbb{E}(\varepsilon_{t+1}^2) = \sigma_{t+1}^2 = c + \alpha_1 * \varepsilon_t^2$. Thus it follows that equation (1.16) displays volatility clustering. If we examine the RHS, as $\alpha_1 > 0$ (condition for stationarity), when shock ε_t^2 is larger than what you expect it to be on average σ^2 the LHS will also be positive. Then the conditional variance will be larger than the unconditional variance. Briefly, large shocks will be followed by more large shocks.

$$\sigma_{t+1}^2 - \sigma^2 = \alpha_1 * (\varepsilon_t^2 - \sigma^2) \tag{1.16}$$

Excess kurtosis can be modeled, even when the conditional distribution is assumed to be normally distributed. The third moment, skewness, can be introduced using a skewed conditional distribution as we saw in part 1.1.1. The serial correlation for squared innovations is positive if fourth moment exists (equation (1.15), this is volatility clustering once again.

The estimation of ARCH model and in a next step GARCH models will be explained in the methodology. However how will then the variance be forecasted? Well, the conditional variance for the k-periods ahead , denoted as period $T+k$, is given by equation (1.17). This can already be simplified, while we know that $\sigma^2_{T+1} = \omega + \alpha_1 * \varepsilon^2_T$ from equation (1.9).

$$\begin{aligned} \mathbb{E}_T(\varepsilon^2_{T+k}) &= \omega * (1 + \alpha_1 + ... + \alpha^{k-2}) + \alpha^{k-1} * \sigma^2_{T+1} \\ &= \omega * (1 + \alpha_1 + ... + \alpha^{k-1}) + \alpha^k * \sigma^2_T \end{aligned} \tag{1.17}$$

It can be shown that then the conditional variance in period $T+k$ is equal to equation (1.18). The LHS is the predicted conditional variance k-periods ahead above its unconditional variance, σ^2. The RHS is the difference current last-observed return residual ε^2_T above the unconditional average multiplied by α^k_1, a decreasing function of k (given that $0 \leq \alpha_1 < 1$). The further ahead predicting the variance, the closer α^k_1 comes to zero, the closer to the unconditional variance, i.e. the long-run variance.

$$\mathbb{E}_T(\varepsilon^2_{T+k}) - \sigma^2 = \alpha^k_1 * (\varepsilon^2_T - \sigma^2) \tag{1.18}$$

1.2.3 Univariate GARCH models

An improvement on the ARCH model is the Generalised Autoregressive Conditional Heteroscedasticity (GARCH). This model and its variants come in to play because of the fact that calculating standard deviations through rolling periods, gives an equal weight to distant and nearby periods, by such not taking into account empirical evidence of volatility clustering, which can be identified as positive autocorrelation in the absolute returns. GARCH models are an extension to ARCH models, as they incorporate both a novel moving average term (not included in ARCH) and the autoregressive component.

All the GARCH models below are estimated using the package rugarch by Alexios Ghalanos [20]. We use specifications similar to Ghalanos [5]. Parameters have to be restricted so that the variance output always is positive, except for the eGARCH model, as this model does not mathematically allow for a negative output.

GARCH model

The standard GARCH model [21] is written consistent with Alexios Ghalanos [5] as in equation (1.19) without external regressors.

$$\sigma_t^2 = \omega + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{1.19}$$

where σ_t^2 denotes the conditional variance, ω the intercept and ε_t^2 the residuals from the used mean process. The GARCH order is defined by (q, p) (ARCH, GARCH). As Ghalanos [5] describes: “one of the key features of the observed behavior of financial data which GARCH models capture is volatility clustering which may be quantified in the persistence parameter $\hat{P}$” specified as in equation (1.20).

$$\hat{P} = \sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j. \tag{1.20}$$

The unconditional variance of the standard GARCH model of Bollerslev is very similar to the ARCH model, but with the Garch parameters (β's) included as in equation (1.21).

$$\begin{aligned}\hat{\sigma}^2 &= \frac{\hat{\omega}}{1-\hat{P}} \\ &= \frac{\hat{\omega}}{1-\alpha-\beta}\end{aligned} \tag{1.21}$$

IGARCH model

Following Alexios Ghalanos [5], the integrated GARCH model [21] can also be estimated. This model assumes the persistence $\hat{P} = 1$. This is done by Ghalanos, by setting the sum of the ARCH and GARCH parameters to 1. Because of this unit-persistence, the unconditional variance cannot be calculated.

EGARCH model

The EGARCH model or exponential GARCH model [7] is defined as in equation (1.22). The advantage of the EGARCH model is that there are no parameter restrictions, since the output is log variance (which cannot be negative mathematically), instead of variance.

$$\log_e(\sigma_t^2) = \omega + \sum_{j=1}^{q}(\alpha_j z_{t-j} + \gamma_j(|z_{t-j}| - E|z_{t-j}|)) + \sum_{j=1}^{p}\beta_j \log_e(\sigma_{t-j}^2) \tag{1.22}$$

where α_j captures the sign effect and γ_j the size effect.

GJRGARCH model

The GJRGARCH model [22] models both positive as negative shocks on the conditional variance asymmetrically by using an indicator variable I, it is specified as in equation (1.23).

$$\sigma_t^2 = \omega + \sum_{j=1}^{q}(\alpha_j \varepsilon_{t-j}^2 + \gamma_j I_{t-j}\varepsilon_{t-j}^2) + \sum_{j=1}^{p}\beta_j \sigma_{t-j}^2 \tag{1.23}$$

where γ_j represents the *leverage* term. The indicator function I takes on value 1 for $\varepsilon \leq 0$, 0 otherwise. Because of the indicator function, persistence of the model now crucially depends on the asymmetry of the conditional distribution used according to Ghalanos [5].

NAGARCH model

The NAGARCH or nonlinear asymmetric model [23]. It is specified as in equation (1.24). The model is *asymmetric* as it allows for positive and negative shocks to differently affect conditional variance and *nonlinear* because a large shock is not a linear transformation of a small shock.

$$\sigma_t^2 = \omega + \sum_{j=1}^{q} \alpha_j (\varepsilon_{t-j} + \gamma_j \sqrt{\sigma_{t-j}})^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{1.24}$$

As before, γ_j represents the *leverage* term.

TGARCH model

The TGarch or threshold model [24] also models assymetries in volatility depending on the sign of the shock, but contrary to the GJRGARCH model it uses the conditional standard deviation instead of conditional variance. It is specified as in (1.25).

$$\sigma_t = \omega + \sum_{j=1}^{q} (\alpha_j^+ \varepsilon_{t-j}^+ \alpha_j^- + \varepsilon_{t-j}^-) + \sum_{j=1}^{p} \beta_j \sigma_{t-j} \tag{1.25}$$

where ε_{t-j}^+ is equal to ε_{t-j} if the term is negative and equal to 0 if the term is positive. The reverse applies to ε_{t-j}^-. They cite Davidian and Carroll [25] who find that using volatility instead of variance as scaling input variable gives better variance estimates. This is due to absolute residuals (contrary to squared residuals with variance) more closely predicting variance for non-normal distributions.

TSGARCH model

The absolute value Garch model or TS-Garch model, as named after Taylor [26] and Schwert [27], models the conditional standard deviation and is intuitively specified like a normal GARCH model, but with the absolute value of the shock term. It is specified as in (1.26).

$$\sigma_t = \omega + \sum_{j=1}^{q}(\alpha_j \left|\varepsilon_{t-j}\right|) + \sum_{j=1}^{p} \beta_j \sigma_{t-j} \tag{1.26}$$

EWMA

A alternative to the series of GARCH models is the exponentially weighted moving average or EWMA model. This model calculates conditional variance based on the shocks from previous periods. The idea is that by including a smoothing parameter λ more weight is assigned to recent periods than distant periods. The λ must be less than 1. It is specified as in (1.27).

$$\sigma_t^2 = (1-\lambda)\sum_{j=1}^{\infty}(\lambda^j \varepsilon_{t-j}^2) \tag{1.27}$$

In practice a λ of 0.94 is often used, such as by the financial risk management company RiskMetrics^{TM} model of J.P. Morgan.

1.3 Value at Risk

Value-at-Risk (VaR) is a risk metric developed to calculate how much money an investment, portfolio, department or institution such as a bank could lose in a market downturn. According to VaR was adopted in 1998 when financial institutions started using it to determine their regulatory capital requirements. A VaR_{99} finds the amount that would be the greatest possible loss in 99% of cases. It can be defined as the threshold value θ_t. Put differently, in 1% of cases the loss would be greater than this amount.It is specified as in (1.28).

$$Pr(R_t \leq \theta_t | \Omega_{t-1}) \equiv \phi \tag{1.28}$$

With R_t expected returns in period t, Ω_{t-1} the information set available in the previous period and ϕ the chosen confidence level.

1.4 Conditional Value at Risk

One major shortcoming of the VaR is that it does not provide information on the probability distribution of losses beyond the threshold amount. This is problematic, as losses beyond this amount would be more problematic if there is a large probability distribution of extreme losses, than if losses follow say a normal distribution. To solve this issue, the conditional VaR (cVaR) quantifies the average loss one would expect if the threshold is breached, thereby taking the distribution of the tail into account. Mathematically, a $cVaR_{99}$ is the average of all the VaR with a confidence level equal to or higher than 99. It is commonly referred to as expected shortfall (ES) sometimes and was first introduced by [28]. It is specified as in (1.29).

To calculate θ_t, VaR and cVaR require information on the expected distribution mean, variance and other parameters, to be calculated using the previously discussed GARCH models and distributions.

$$Pr(R_t \leq \theta_t | \Omega_{t-1}) \equiv \int_{-\infty}^{\theta_t} f(R_t | \Omega_{t-1}) \, \mathrm{d}R_t = \phi \tag{1.29}$$

With the same notations as before, and f the (conditional) probability density function of R_t.

According to the BIS framework, banks need to calculate both VaR_{99} and $VaR_{97.5}$ daily to determine capital requirements for equity, using a minimum of one year of daily observations [29]. Whenever a daily loss is recorded, this has to be registered as an exception. Banks can use an internal model to calculate their VaRs, but if they have more than 12 exceptions for their VaR_{99} or 30 exceptions for their $VaR_{97.5}$ they have to follow a standardized approach.Similarly, banks must calculate $cVaR_{97.5}$.

References

[1] Jan Annaert. *Quantitative Methods in Finance*. Version 0.2.1. Antwerp Management School, Jan. 2021.

[2] Tim Bollerslev. “A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return”. In: *The Review of Economics and Statistics* 69.3 (1987). Publisher: The MIT Press, pp. 542–547. URL: `https://www.jstor.org/stable/1925546`.

[3] Eugene F. Fama. “The Behavior of Stock-Market Prices”. In: *The Journal of Business* 38.1 (1965), pp. 34–105. URL: `http://www.jstor.org/stable/2350752`.

[4] Tim Bollerslev. “Glossary to ARCH (GARCH)”. In: (Sept. 4, 2008), p. 46. URL: `Available%20at%20SSRN:%20https://ssrn.com/abstract=1263250`.

[5] Alexios Ghalanos. *Introduction to the rugarch package. (Version 1.4-3)*. Tech. rep. 2020. URL: `http://cran.r-project.org/web/packages/`.

[6] James B. McDonald and Whitney K. Newey. “Partially Adaptive Estimation of Regression Models via the Generalized T Distribution”. In: *Econometric Theory* 4.3 (Dec. 1988), pp. 428–457. URL: `https://www.cambridge.org/core/product/identifier/S0266466600013384/type/journal_article`.

[7] Daniel B. Nelson. “Conditional Heteroskedasticity in Asset Returns: A New Approach”. In: *Econometrica* 59.2 (Mar. 1991). Publisher: JSTOR, pp. 347–347.

[8] Carmen Fernández and Mark F. J. Steel. “On Bayesian Modeling of Fat Tails and Skewness”. In: *Journal of the American Statistical Association* 93.441 (Mar. 1998), pp. 359–371. URL: `http://www.tandfonline.com/doi/abs/10.1080/01621459.1998.10474117`.

[9] Denis-Alexandre Trottier and David Ardia. “Moments of standardized Fernandez-Steel skewed distributions: Applications to the estimation of GARCH-type models”. In: *Finance Research Letters* 18 (Sept. 4, 2015), pp. 311–316. URL: `https://ssrn.com/abstract=2656377`.

[10] Pierre Giot and Sébastien Laurent. “Value-at-risk for long and short trading positions”. In: *Journal of Applied Econometrics* 18.6 (Nov. 2003), pp. 641–663. URL: `http://doi.wiley.com/10.1002/jae.710`.

[11] Pierre Giot and Sébastien Laurent. “Modelling daily Value-at-Risk using realized volatility and ARCH type models”. In: *Journal of Empirical Finance* 11.3 (June 1, 2004), pp. 379–398. URL: `https://www.sciencedirect.com/science/article/pii/S092753980400012X`.

[12] Peter Theodossiou. “Skewed Generalized Error Distribution of Financial Assets and Option Pricing”. In: *SSRN Electronic Journal* (2000). URL: `http://www.ssrn.com/abstract=219679`.

[13] Ming-Chih Lee, Jung-Bin Su, and Hung-Chun Liu. “Value-at-risk in US stock indices with skewed generalized error distribution”. In: *Applied Financial Economics Letters* 4.6 (Oct. 22, 2008), pp. 425–431. URL: http://www.tandfonline.com/doi/abs/10.1080/17446540701765274.

[14] Panayiotis Theodossiou. “Financial data and the skewed generalized t distribution”. In: *Management Science* 44.12 PART 1 (1998). Publisher: INFORMS Inst.for Operations Res.and the Management Sciences, pp. 1650–1661. URL: http://pubsonline.informs.org.https//doi.org/10.1287/mnsc.44.12.1650http://www.informs.org0025-1909/98/4412/1650$05.00.

[15] Turan G. Bali and Panayiotis Theodossiou. “A conditional-SGT-VaR approach with alternative GARCH models”. In: *Annals of Operations Research* 151.1 (Feb. 22, 2007), pp. 241–267. URL: http://link.springer.com/10.1007/s10479-006-0118-4.

[16] Turan G. Bali, Hengyong Mo, and Yi Tang. “The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR”. In: *Journal of Banking and Finance* 32.2 (Feb. 2008). Publisher: North-Holland, pp. 269–282.

[17] Tim Bollerslev, Robert F. Engle, and Daniel B. Nelson. “Chapter 49 Arch models”. In: *Handbook of Econometrics* 4 (Jan. 1994). Publisher: Elsevier, pp. 2959–3038.

[18] Robert Engle. “GARCH 101: The use of ARCH/GARCH models in applied econometrics”. In: *Journal of Economic Perspectives* (2001).

[19] R. F. Engle. “Autoregressive Conditional Heteroscedacity with Estimates of variance of United Kingdom Inflation,journal of Econometrica, Volume 50, Issue 4 (Jul., 1982),987-1008.” In: *Econometrica* 50.4 (1982), pp. 987–1008.

[20] Alexios Ghalanos. *rugarch: Univariate GARCH models.* R package version 1.4-4. 2020.

[21] Tim Bollerslev. “Generalized Autoregressive Conditional Heteroskedasticity”. In: *Journal of Econometrics* 31 (1986), pp. 307–327.

[22] Lawrence R. Glosten, Ravi Jagannathan, and David E. Runkle. “On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks”. In: *The Journal of Finance* 48.5 (Dec. 1993). Publisher: John Wiley and Sons, Ltd, pp. 1779–1801. URL: http://doi.wiley.com/10.1111/j.1540-6261.1993.tb05128.x.

[23] Robert F. Engle and Victor K. Ng. “Measuring and Testing the Impact of News on Volatility”. In: *The Journal of Finance* 48.5 (Dec. 1993). Publisher: John Wiley and Sons, Ltd, pp. 1749–1778. URL: http://doi.wiley.com/10.1111/j.1540-6261.1993.tb05127.x.

[24] Jean Michel Zakoian. “Threshold heteroskedastic models”. In: *Journal of Economic Dynamics and Control* 18.5 (1994), pp. 931–955.

[25] M. Davidian and R. J. Carroll. “Variance Function Estimation”. In: *Journal of the American Statistical Association* 82.400 (Dec. 1987). Publisher: JSTOR, pp. 1079–1079.

[26] Stephen J. Taylor. *Modelling financial time series.* Chichester: John Wiley and Sons, Ltd., 1986.

[27] G. William Schwert. “Why Does Stock Market Volatility Change Over Time?” In: *The Journal of Finance* 44.5 (1989), pp. 1115–1153.

[28] Dimitris Bertsimas, Geoorey J Lauprete, and Alexander Samarov. “Shortfall as a risk measure: properties, optimization and applications”. In: *Journal of Economic Dynamics and Control* 28 (2004), pp. 1353–1381. URL: `www.elsevier.com/locate/econbase`.

[29] Basel Committee on Banking Supervision. *Minimum capital requirements for market risk*. Tech. rep. Issue: January Publication Title: Bank for International Settlements. 2016, pp. 92–92. URL: `https://www.bis.org/basel_framework/`.