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 very 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 indepently 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 fat tails or show leptokurtosis and thus riskier than under the normal distribution (excess kurtosis that is larger than 3). 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 are log-normally distributed, 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. Well, it all 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. What distribution is then appropriate?

1.1.1 Alternative distributions than the normal

Student t distribution

One, often used alternative for the normal distribution is the Student t distribution. It is also a symmetric distribution, this means 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 model as an alternative to the standard Normal distribution.

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

As can be seen the pdf depends on degree of freedom parameter n. To be consistent with Ghalanos [4], the following general equation is used for the pdf (1.2)

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

where α, β and ν are the location parameters (scale and shape 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 has a kurtosis coefficient). This kurtosis coefficient is given by equation (1.3).

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

General Error Distribution
Skewed-t Distribution
Skewed Generalized Error Distribution
Skewed Generalized t distribution
Standardized Generalized Hyperbolic Distribution

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 [5] 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 [6], 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 (1.4), (1.5) and (1.6). 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).

$$R_t = \mu + \varepsilon_t \tag{1.4}$$

$$\varepsilon_t = \sigma_t * z_t,\ where\ z_t \overset{iid}{\sim} (0,1) \tag{1.5}$$

$$\sigma_t^2 = \omega + \alpha_1 * \varepsilon_{t-1}^2 \tag{1.6}$$

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 (1.7) and (1.8) 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 (1.8) is equal to the unconditional mean in the most simple case. But variations are possible using ARMA (eg. AR(1)) processes.

$$\mathbb{E}_{t-1}(\varepsilon_t) = \mathbb{E}_{t-1}(\sqrt{\omega + \alpha_1 * \varepsilon_{t-1}^2} * z_t) = \sigma_t \mathbb{E}_{t-1}(z_t) = 0 \tag{1.7}$$

$$\mathbb{E}_{t-1}(R_t) = \mu + \mathbb{E}_{t-1}(\varepsilon_t) = \mu \tag{1.8}$$

For the conditional variance, knowing everything that happened until and including period $t-1$ the condiitional innovation variance is given by equation (1.9). 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 (1.10), that is why equation (1.6) is called the variance equation.

$$var_{t-1}(\varepsilon_t) = \mathbb{E}_{t-1}(\varepsilon_t^2) = \mathbb{E}_{t-1}(\sigma_t^2 * z_t^2) = \sigma_t^2 \mathbb{E}_{t-1}(z_t^2) = \sigma_t^2 \tag{1.9}$$

$$var_{t-1}(R_t) = var_{t-1}(\varepsilon_t) = \sigma_t^2 \tag{1.10}$$

The unconditional variance is also interesting to derive, while this is the long-run variance, which will be derived in (1.14). After deriving this using the law of iterated expectations and assuming stationarity for the variance process, one would get (1.11) for the unconditional variance, equal to the constant c and divided by $1 - \alpha_1$, the slope of the variance equation.

$$\sigma^2 = \frac{\omega}{1 - \alpha_1} \tag{1.11}$$

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

The unconditional 4th moment, kurtosis $\mathbb{E}(\varepsilon_t^4)/\sigma^4$ of an ARCH model is given by equation (1.12). This term is larger than 3, which implicates that the fat-tails (a stylised fact of returns).

$$3\frac{1-\alpha_1^2}{1-3\alpha_1^2} \tag{1.12}$$

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 (1.13) 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.

$$\sigma_{t+1}^2 - \sigma^2 = \alpha_1 * (\varepsilon_t^2 - \sigma^2) \tag{1.13}$$

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 will see in part **??**. The serial correlation for squared innovations is positive if fourth moment exists (equation (1.12), 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 (1.14). This can already be simplified, while we know that $\sigma_{T+1}^2 = \omega + \alpha_1 * \varepsilon_T^2$ from equation (1.6).

$$\begin{aligned} \mathbb{E}_T(\varepsilon_{T+k}^2) &= \omega * (1 + \alpha_1 + ... + \alpha^{k-2}) + \alpha^{k-1} * \sigma_{T+1}^2 \\ &= \omega * (1 + \alpha_1 + ... + \alpha^{k-1}) + \alpha^k * \sigma_T^2 \end{aligned} \tag{1.14}$$

It can be shown that then the conditional variance in period $T + k$ is equal to equation (1.15). 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_T^2 above the unconditional average multiplied by α_1^k, a decreasing function of k (given that $0 \leq \alpha_1 < 1$). The further ahead predicting the variance, the closer α_1^k comes to zero, the closer to the unconditional variance, i.e. the long-run variance.

$$\mathbb{E}_T(\varepsilon_{T+k}^2) - \sigma^2 = \alpha_1^k * (\varepsilon_T^2 - \sigma^2) \tag{1.15}$$

1.2.3 Univariate GARCH models

Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models come in because of the fact that rolling period standard deviations give an equal weight to the deviations, by such not taking into account volatility clustering, which can be identified as positive autocorrelation in the absolute returns. All these GARCH models are estimated using the package rugarch by Alexios Ghalanos [7]. We use specifications similar to Ghalanos [4].

sGARCH model

The standard GARCH model [8] is written consistent with Alexios Ghalanos [4] as in equation (1.16) without external regressors.

$$\sigma_t^2 = \omega + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{1.16}$$

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 [4] 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 (1.17).

$$\hat{P} = \sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j. \tag{1.17}$$

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 (1.18).

$$\begin{aligned} \hat{\sigma}^2 &= \frac{\hat{\omega}}{1 - \hat{P}} \\ &= \frac{\hat{\omega}}{1 - \alpha - \beta} \end{aligned} \tag{1.18}$$

iGARCH model

Following Alexios Ghalanos [4], the integrated GARCH model [8] 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 [9] is defined as in equation (1.19),

$$\log_e(\sigma_t^2) = \omega + \sum_{j=1}^{q}(\alpha_j z_{t-j} + \gamma_j(|z_{t-j}| - E|z_{t-j}|)) + \sum_{j=1}^{p} \beta_j \log_e(\sigma_{t-j}^2) \tag{1.19}$$

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

gjrGARCH model

The gjrGARCH model [10] models both positive as negative shocks on the conditional variance asymmetrically by using an indicator variable I, it is specified as in equation (1.20).

$$\sigma_t^2 = \omega + \sum_{j=1}^{q}(\alpha_j \varepsilon_{t-j}^2 + \gamma_j I_{t-j} \varepsilon_{t-j}^2) + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{1.20}$$

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 [4].

naGARCH model (Engle & Ng)
tGARCH model (Zakoian)
avGARCH model (in our paper: TS-GARCH to Taylor and Schwert)

1.3 Value at Risk

1.4 Conditional Value at risk

References

[1] Jan Annaert. *Quantitative Methods in Finance.* Vol. 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] Alexios Ghalanos. *Introduction to the rugarch package. (Version 1.4-3).* Tech. rep. 2020. URL: `http://cran.r-project.org/web/packages/`.

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

[6] 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.

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

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

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

[10] 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 Sons, Ltd, pp. 1779–1801. URL: `http://doi.wiley.com/10.1111/j.1540-6261.1993.tb05128.x`.