Project of Models for Risk and Forecasting

Sapienza Università di Roma

Quantitative risk management in univariate and multivariate framework

Stricelli Alfredo, 1882000
Di Bella Giuseppe, 1621685

December 2021

Contents

1 Introduction

Market risk is defined as the chance of unfavorable market variation in the value of a financial instrument or a portfolio. Banks started to apply a methodology for the estimation of market risk only in the late 80', in particular one of the first models proposed is the well known Value at risk(VaR) introduced by JP Morgan in 1994 with the document risk metrics. Value-at-Risk is a measure of the maximum potential change in value of a portfolio of financial instruments with a given probability over a pre-set horizon. VaR answers the question: how much can is it possible to lose with x% probability over a given time horizon? The aim of this project is to present an analysis of the volatility estimation and risk modelling for the FTSE Mib time series, considering different approaches for the VaR estimation:

- parametric approach: with this approach, the VaR estimation is obtained starting from the assumption that the time series of the returns follows a distribution
- historical simulation: contrary to the parametric approach, the non-parametric technique does not make any distribution assumption. The HS calculates the VaR based on the quantile for a pre-fixed rolling window of data. concerning the daily returns
- semiparametric: it refers to the quantile regression and also in this case there is no need a distribution assumption

The second risk measure proposed in this project is the Expected Shortfall considered as a more reliable measure of risk, obtained as the the mean of the most sever losses, inspecting the maximum unexpected losses over a determined time period. After the estimation of this risk measures, there is the assessment of the performances of the model specified with back testing procedures and model confidence set procedures, in order to determine which is the model that can better forecast volatility.
The last part of the work refers to the multivariate analysis, in fact, the quantitative risk management analysis concerns risky positions whose losses are composed of multiple losses. In order to define risk measurements about these groups of risks, an evaluation of their probabilistic joint behaviour is needed. To build a joint distribution function it is possible to use copula modelling.

2 Risk Measures

Let $L^0(\Omega, F, P)$ the space of all measurable functions on the probability space and let M be the set of measurable functions over the measurable space. A risk measure $\rho(X) : M \to [R]$ where X is a r.v. which describes the loss. There are two ways to identify a risk measure:

- It is possible to define the risk measure in terms of future value of risky position $\rho(V(t+1))$ which represents the capital to be added to the current risky position V(t)
- The risk measure can be also denoted in this way: $\rho(L_{t,t+1})$, which represents the amount of capital necessary to back a risky position with a loss L

2.1 Axioms of coherence of risk measures

A good risky measure is defined as "coherent". In order to judge the coherence of a general risk measure ρ there is the need to define the axioms of coherence of risk measure:

- Axiom1 Monotonicity: Given two lossess $L_1, L_2 \in M$, $L_1 \leq L_2$ a.s.;this imply that $\rho(L_1) \leq \rho(L_2)$. Financially speaking this imply that a risk measure must be able to assign a lesser capital amount to hold, necessary to take position in lesser risky position with respect to riskier positions.
- Axiom2 Translation invariance: $\forall L \in M, x \in R$ where x is a constant amount, than it holds that $\rho(L+x) = \rho(L) + x$. By a financial perspective, adding (or deducting) a risk-free amount to (from) a portfolio and investing it in the reference instrument, results in a decrease (increase) of the risk of the position by exactly the same amount. From this property we can derive a well-known fact in finance: if $x = \rho(L)$, then

  $$\rho(L + x) = \rho(L + \rho(L)) = \rho(L) - \rho(L) = 0$$

  Hence, it is possible to hedge an underwritten risky position, by simply adding a certain amount of risk-free instruments in the portfolio.

- Axiom3 Posiitve homogeneity: $\forall L \in M, \lambda > 0$, it holds that $\rho(\lambda L) = \lambda\rho(L)$, meaning that the variation over scale of the loss is not important
- Axiom4 Sub-additivity: $\forall L_1, L_2 \in M$ it holds that $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$; a risk measure that is sub-additive reflects the diversification over the risky positions in the diversification over losses

A risk measure is coherent if all the axioms are satisfied

2.2 Generalized inverse

Let $T : R \to [R]$ a non decreasing function, that is

$$T(x) \leq T(y) \forall x, y \in R$$

The generalized inverse function $T^{-} : R \to [R] = [-\inf, +\inf]$ such that

$$T^{-}(y) = inf x \in R : T(x) \geq y \qquad y \in R$$

with

$$T^{-}(-\inf) = lim_{(x \to -\inf)} T(x)$$

$$T^{-}(+\inf) = lim_{(x \to +\inf)} T(x)$$

Let T be a distribution function

$$T = F : [0, 1] \to [R]$$

the generalized inverse function of T is the so called quantile function of the distribution F

$$T^{-}(y) = Q(y) = \inf\{x \in R : F(x) \geq y\}$$

where x is a realization of a r.v. X and y as the confidence level $\alpha \in (0, 1)$ If T is:

- increasing: that is $\geq y \in R$ that is $T(x) \geq T(y)$
- is continuous

The generalized inverse function coincides with the ordinary inverse function, that is:

$$T^{-}(y) = T^{-1}(y)$$

2.3 Value at Risk

Value at Risk is probably the most widely used risk measure in finance. It has become the classic measure that financial executives use to quantify market risk. Indeed, when RiskMetrics announced Value at Risk as its measure of risk in 1996, the Basel Committee on Banking Supervision enforced financial institutions to meet capital requirements based on VaR estimates. In order to give a formal definition it is possible to consider a portfolio of risky assets and a fixed time horizon .
Let L be the estimated loss distribution L associated to this portfolio and $F_L(l) = P(L \leq l)$ its distribution function. We would like to evaluate the level of risk associated to the holding of our portfolio over the time period Δ. The $VaR_\alpha(L)$ is the α- quantile function of the distribution function F_L:

$$Var(L) = inf\{l \in R : P(L > l)\} \leq 1 - \alpha = inf\{l \in R : F_L(l) > \alpha\} \tag{1}$$

In other words, given some confidence level $\alpha \in$ (0, 1), the VaR of our portfolio at the confidence level α is given by the smallest number l such that the probability that the loss L exceeds l is no larger than (1 - α)
Typical values for α are α =0.95, α =0.99 or higher.
In market risk management the time horizon Δ is usually 1 or 10 days, while in credit risk management Δ is usually one year.
VaR has the advantage of being very intuitive and the great popularity this instrument has reached, is essentially due to its conceptual clarity. However, VaR has two important drawbacks:

- it does not fulfill (in general) the property of sub-additivity, meaning that it does not award diversification benefits;
- it is tail insensitive. It tells us that in the α100% of the cases the loss will not be greater than a certain level, but it does not give us any clue about the size of the loss in the remaining $(1 \quad \alpha) \cdot$ 100% of the cases.

By definition of VaR at confidence level α, we have that the violation probability of the VaR number equals to 1 - α:

$$P(L > VaR_{\alpha}(L)) = 1 - \alpha. \tag{2}$$

Hence, it is defined as the $1 - \alpha$-th quantile of the conditional distribution of returns.
It has recently been questioned as a measure of risk and consideration is being given to using the Expected Shortfall for the calculation of the minimum reserve capital requirement.

2.4 Expected Shortfall

One of the main weaknesses of VaR is that it does not measure the potential loss in the event of a VaR violation. To overcome this important limitation, the following has been proposed ES: this measures the expected value of the loss once a VaR violation has occurred.
Expected Shortfall, also called Conditional Value-at-Risk (CVaR) measures what happens, in average, when the VaR is violated ($L > VaR_{\alpha}$). Thus, for an integrable loss L with continuous distribution function FL and any$\alpha \in (0, 1)$, we have that

$$ES_{\alpha}(L) = \mathbb{E}(L|L > VaR_{\alpha}) \tag{3}$$

Expected Shortfall is thus related to VaR by this formula

$$ES_{\alpha}(L) = \frac{1}{1 - \alpha} \int_{\alpha}^{1} VaR_u(L)du \tag{4}$$

With respect to VaR,which is a frequency based measure, the ES ia a severity based measure; moreover, ES is a coherent risk measure.

2.5 Econometric approaches to VaR calculation

A general approach to calculating VaR is to use time series models. In this framework the VaR can be simply defined as the $\tau - th$ quantile of the conditional distribution of the returns r_t. The one-step-ahead VaR at τ level, and the one-step-ahead CVaR (or ES) are respectively defined via :

$$Pr(r_t < VaR_{(\tau),t}|\mathcal{F}_{t-1}) = \tau \tag{5}$$

and

$$ES_{\tau} = E[r_t|r_t < VaR_{(\tau),t}, \mathcal{F}_{t-1}], \tag{6}$$

where α is the quantile level and $\mathcal{F}_t$ is the information set at time t.
By imposing $VaR_{\tau,t} = Q_{r_t}(\tau)$ be the VaR at time t for the daily returns r_t at τ level, the formula (5) tells us that there is a probability equal to τ of losing more than the value predicted by $VaR_{\tau,t}$. Similarly the (6) measures what happen on average when VaR is violated, that is when r_t is smaller than $VaR_{\tau,t}$.
In risk management framework, we are interested in estimating the VaR ex-ante (that is the out-of-sample). So we should say we are forecasting VaR. The methodology used to obtain VaR measures can be broadly divided into three main categories:parametric,non parametric and semi-parametric approach.

2.6 Parametric approach

The **Parametric** approach, which requires the estimation of the volatility of the asset under investigation as a primer step and the VaR measures are indirectly obtained by considering these volatility estimates and the quantile at a fixed level of the presumed distribution of the asset.
Several models have been developed to analyze volatility in financial time series. These are fundamental tools to make the best volatility estimates, useful for the calculation of VaR. In this analysis, GARCH, GARCH-MIDAS, DAGM and CAViaR models. The GARCH models were introduced by Bollerslev (1986) as extensions of the AutoRegressive Conditional Heteroscedasticity (ARCH) models. Empirical evidence showed that the (1,1) specification adapts well with financial data. In particular, the equation of GARCH(1,1) model is given by:

$$\sigma_t^2 = w + \alpha\epsilon_{t-1} + \beta\sigma_{t-1}^2$$

with σ_t^2 denoting the conditional variance, the intercept and ϵ_t^2 the residuals. The process is stationary when > 0 and $\alpha + \beta < 1$. To allow for asymmetric effects between positive and negative returns, Nelson (1991) extended the GARCH framework to derive the exponential GARCH model. In particular, an E-GARCH(1,1) model can be written as:

$$ln(\sigma_t^2) = w + \alpha(|z_{t-1}| - E[|z_{t-1}|]) + \gamma z_{t-1} + \beta ln(\sigma_{t-1}^2)$$

where is the intercept, α the sign effect, β the persistence, γ the leverage effect and z_t is a standard Gaussian. The so-called GJR-GARCH model assumes a specific parametric form for the conditional heteroscedasticity and it models positive and negative shocks on the conditional variance asymmetrically via the use of the indicator function I. A GJR-GARCH(1,1) is given by:

$$\sigma_t^2 = w + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\epsilon_{t-1}^2 I(\epsilon_{t-1} < 0)$$

where I_{t1} is the indicator function: it is equal to zero if $\epsilon_{t1}\mu$, and equal to 1 if $\epsilon_{t1} < \mu$.
Another important issue to be managed, is that macroeconomic variables have a lot of influence on volatility because of expectations. Models trying to catch the effect generated by a macroeconomic variable tend to describe the dynamics of the volatility decomposing it into two terms: a short component, with high frequency and a long component, with low frequency. Engle (2013) developed the GARCH-MIDAS model, where MIDAS stands for “Mixed Data Sampling”, in order to take into account the macroeconomic variable (MV) and at the same time to solve the main issue of finding daily returns compared to at least monthly observation of the MV. The model is described by the following:

$$r_{i,t} = \sqrt{\tau_t \times g_{i,t} z_{i,t}}$$

$$g_{i,t} = (1 - \alpha - \beta) + \alpha\frac{(r_{i-1,t}^2}{\tau_t} + \beta g_{i-1,t}$$

$$\tau_t = \exp(m + \theta \textstyle\sum_{k=1}^{K} \delta_k(w) X_{t-k})$$

$$\delta_k(w) = \frac{(k/K)^{w_1-1}(1-k/K)^{w_2-1}}{\sum_{j=1}^{K}(j/K)^{w_1-1}(1-j/K)^{w_2-1}}$$

where $r_{i,t}$ are the log-returns for day i in period t, $z_{i,t}$ are standard gaussian innovations, $g_{i,t}$ is the short-run component following a unit-mean reverting GARCH(1,1), τ_t is the long-run component and $\delta_k(w)$ is the Beta function, which allows giving equal, increasing or decreasing weights to the past realizations of the strictly stationary variable X_t. The Double Asymmetric GARCH-MIDAS (DAGM) model proposed by Amendola (2019) allows specifying three additional sources of asymmetry associated with negative daily returns (γ), to positive and negative 2 values in the macroeconomic variable ($\theta + and\theta$), and to the different system of weights for the macroeconomic variable realizations ($\delta_k(w)+$ and $\delta_k(w)$). In such a model the short-run component and the long-run component are given respectively by:

$$g_{i,t} = (1 - \alpha - \beta - \tfrac{\gamma}{2}) + (\alpha + \gamma I(r_{i-1,t<0})\frac{(r_{i-1,t}^2)}{\tau_t} + \beta g_{i-1,t}$$

$$\tau_t = exp(m + \theta^+ \textstyle\sum_{k=1}^{K} \delta_k(w)^+ X_{t-k} I_{(X_{t-k}>0} + \theta^- \sum_{k=1}^{K}(w)^- X_{t-k} I_{(X_{t-k}>0)})$$

Moreover, the skewed version of the GARCH-MIDAS and DAGM models can be developed, but they won’t be introduced in this paper work.

2.7 Non parametric approach

The non- parametric approach does not make any distribution assumption concerning the daily returns. One of the most prominent examples of this class of models is the Historical simulation (HS, Hendricks(1996)) which is one used in this paper. The HS approach calculates the $VaR_{\tau,t}$ as the sample quantile of past returns computed over a moving-window of length w, that is:

$$VaR_{\tau,t} = \hat{Q}_{r_{(t-w):(t-1)}}(\tau) \tag{7}$$

where $r_{(t-w)} : (t-1) = (r_{t-w}, r_{t-w+1}, .., r_{t-1})$ and $\hat{Q}_{(.)}(\tau)$ represents the sample quantile of . at τ level. The HS approach is distribution free but not assumption free, because it requires returns to be *iid* within the window w. Moreover, for an accurate VaR estimation a long window (large w) is required.

2.8 Semi-parametric approach

Considering **Semi-parametric** approaches, QR (quantile regression) tools are used to model the quantiles of the distribution instead of taking into account the whole distribution. In this work we use the quantile estimation of linear ARCH model (Q-L-ARCH).
The linear ARCH representation of daily returns is defined by:

$$r_t = (\beta_0 + \beta_1|r_{t-1}| + ... + \beta_q|r_{t-q}|)z_t \tag{8}$$

where $z_t \backsim (0, 1)$, whose distribution is not specified. The quantile linear ARCH (Q-L-ARCH) calculates the VaR as:

$$\hat{Q}_{r_t}(\tau|\mathcal{F}_{t-1}) = x_t'\hat{\beta}(\tau), \tag{9}$$

where $x_t = (1, |r_{t-1}|, ..., |r_{t-q}|)'$, and $\hat{Q}_{r_t}$ is the VaR at time t and at the $\tau - th$ level, given information set $\mathcal{F}_{t-1}$.
As in section **??** has been discussed, the MIDAS componens allows to filter the information coming from variables observed at lower frequencies (monthly, quarterly) in context where the dependent variable is usually observed daily.
Recently Amendola(2019) proposed the Q-LARCH-MIDAS model (quantile linear ARCH MIDAS) model highlighting that macroeconomic variables are driving forces of daily assets variability. In this context, the daily log-returns $r_{i,t}$ for the day i and the period t are specified as equation (8), but adding the macroeconomic component:

$$r_{i,t} = (\beta_0 + \beta_1|r_{i-1,t}| + ... + \beta_q|r_{i-q,t}| + \theta \sum_{j=1}^{K} d_k(\omega)|MV_{t-j}|)z_{i,t} \tag{10}$$

where the coefficient θ signals the impact of the weighted summation of the K realizations of the additional stationary variable MV_t. $\delta_k(\omega)$ is already specified in equation **??**.
The VaR estimated using Q-L-ARCH-Midas model is defined as:

$$\hat{Q}_{r_{t,i}}(\tau|\mathcal{F}_{i-1,t}) = x_{i,t}'\hat{\beta}(\tau), \tag{11}$$

where $x_t = (1, |r_{i-1,t}|, ..., |r_{i-q,t}|, WS_{i-1,t})'$, with $WS_{i-1,t} = \sum_{j=1}^{K} d_k(\omega)|MV_{t-j}|$.

The CAViaR-SAV or Symmetric Absolute Value is given by:

$$Q_{r_t}(\tau) = \beta_0 + \beta_1 Q_{r_{t-1}}(\tau) + \beta_2 |r_{t-1}|$$

This model specifications will be correct if data are generated by GARCH process with iid errors, and this model responds symmetrically to past returns. The CAViaR-AS or Asymmetric Slope is described as it follows:

$$Q_{rt}(\tau) = \beta_0 + \beta_1 Q_{r_{t-1}}(\tau) + (\beta_2 I_{rt-1>0} + \beta_3 I_{rt-1>0})|r_{t-1}|$$

Also here the specifications will be correct if data are generated by GARCH process with iid errors, but in this case, the model responds asymmetrically to past returns. The CAViaR-IG or Indirect GARCH is represented by:

$$Q_{rt}(\tau) = -\sqrt{\beta_0 + \beta_1 Q^2_{rt-1}(\tau) + \beta_2 r^2_{t-1}}$$

This model specifications are correct if the underlying data were truly a GARCH(1, 1) with an iid error distribution and as the CAViaR-SAV it responds symmetrically to past returns.

3 Backtesting

The accuracy of VaR predictions is examined using the Backtesting procedures, a set of statistical procedures designed to check if the actual losses are in line with the estimated risk measure.
In this work, Kupiec, Christoffersen and Conditional Coverage tests will be used to check if the models are reliable in estimating VaR.
The Kupiec test, also known as Proportion of Failure test (POF-Test), tests the null hypothesis that the observed rate of violations observed ($x/N = \pi_0$) is statistically equal to the expected violation rate, α, where x is the number of days on which violations occur, and N is the total number of predictions. The likelihood ratio is:

$$LR_{UC}(\alpha) = -2xln\frac{\alpha^x(1-\alpha)^{n-x}}{\pi^x(1-\pi)^{n-x}} \backsim \chi^2_1 \tag{12}$$

The null hypothesis will be rejected if the observed number of violations is statistically different from the expected number of violations indicated by the VaR significance level.
The Christoffersen test tests the null hypothesis according to which VaR violations occur independently, against the alternative hypothesis according to which violations are grouped together.
It is possible to figure the dependence structure of the hit sequence as a first-order Markov's chain with the following transition probability matrix:

$$\Pi_1 = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix} \tag{13}$$

The two conditions for testing serial independence of exceptions are $\pi_{1,1} = \pi_{0,1} = \pi_1$ and $\pi_{0,0} = \pi_{1,0} = \pi_0$. The likelihood ratio needed to test the null hypothesis of independence can be built as:

$$LR_{IND} = -2ln\frac{L(\hat{\pi})}{L(\widehat{\Pi_1})} \backsim \chi^2_1 \tag{14}$$

Given a sample of T observations, we can write the likelihood function of the first-order Markov process as:

$$L(\Pi_1) = (1-\pi_{01})^{T_{00}}\pi_{01}^{T_{01}}(1-\pi_{11})^{T_{10}}\pi_{11}^{T_{11}} \tag{15}$$

The conditional coverage test jointly verify the hypothesis that the number of exceptions is correct and that they are independent.
This corresponds to test that $\pi_{0,1} = \pi_{1,1} = \alpha$. It is possible to write LR_{CC} as:

$$LR_{CC} = LR_{IND} + LR_{UC} \backsim \chi^2_2 \tag{16}$$

Thus, So $H_0 : \pi_{01} = \pi_{11}$ is verified against $H_1 : \pi_{01} \neq \pi_{11}$. If the null hypotheses of both tests are not rejected for a particular model then we can say that that model creates a quantity of expected VaR violations that occur independently of each other.

Backtesting ES involves to check the goodness of the estimated loss distribution, that is to benchmark the ES realizations with respect to the loss realizations. First of all the Expected Shortfall is not a quantile and secondly, it occurs when a VaR violation has occured.
For a one period risky position the ES_α violation is the quantity $\Gamma_{t+1} = (L_{t+1} - ES^t_\alpha(L))\mathbb{1}_t(\alpha)$, where $\mathbb{1}_t(\alpha) = L_{t+1} > VaR^t_\alpha$ and $t+1$ is the forecasted period. In a multiperiod risky position the series of Expected Shortfall violations become:

$$\Gamma_\tau(\alpha), \tau \in [t+1, ..., t+n] \tag{17}$$

To ensure goodness of ES estimation, the (17) has to satisfy the martingale difference property:

$$\mathbf{E}[\Gamma_{t+1}|\mathcal{F}_t] = 0, where \mathcal{F}_t = \sigma(L_t : t \leq s) \tag{18}$$

Thus, one strategy to test ES, its goodness, consists in checking if its violations distribution series Γ_{t+1} has 0 mean.

4 MCS procedure

The MCS consists of a sequence of statistical tests which permits to construct the "Superior Set of Models" (SSM) where the null hypothesis of equal predictive ability (EPA) is not rejected at a certain confidence level. The EPA statistic tests are calculated for an arbitrary loss function. In this case the asymmetric loss function is used, and is defined as

$$\ell\left(y_t, \widehat{\mathrm{VaR}}^\tau_{i,t}\right) = \left(\tau - d^\tau_{i,t}\right)\left(y_t - \widehat{\mathrm{VaR}}^\tau_{i,t}\right), \tag{19}$$

where $\widehat{\mathrm{VaR}}^\tau_{i,t}$ is the VaR estimate of model i at time t at confidence level τ and $d^\tau_{i,t}$, that is 1 if $y_t < \widehat{\mathrm{VaR}}^\tau_{i,t}$ and 0 otherwise, is the τ-quantile loss function.
The procedure starts from an initial set of models M^0 of dimension m encompassing all the model specifications and delivers, for a given confidence level $1-\alpha$, a smaller set, the superior set of models, SSM, $M^*1\alpha$ of dimension $m^* \leq m$. The best scenario is when the final set consists of a single mode, i.e. $m^* = 1$.
Let $d_{ij,t}$ denotes the loss differential between model i and j

$$d_{ij,t} = \ell_{i,t} - \ell_{j,t}, \quad i,j \in M, \quad t = 1, \ldots, n \tag{20}$$

and let

$$d_{i\cdot,t} = (m-1)^{-1} \sum_{j \in M} d_{ij,t} \quad i \in M \tag{21}$$

be the simple loss of model i relative to any other model j at time t. The EPA hypothesis for a given set of models M that are being tested can be formulated in two alternative ways:

$$\begin{aligned} &\mathrm{H}_{0,M} : \mu_{ij} = 0, \text{ for all } i,j \in M \\ &\mathrm{H}_{\mathrm{A},M} : \mu_{ij} \neq 0, \text{ for some } i,j \in M, \end{aligned} \tag{22}$$

or

$$\begin{aligned} &\mathrm{H}_{0,M} : \mu_{i\cdot} = 0, \text{ for all } i,j \in M \\ &\mathrm{H}_{\mathrm{A},M} : \mu_{i\cdot} \neq 0, \text{ for some } i,j \in M, \end{aligned} \tag{23}$$

where $M \subset M^0$ and μ_{ij}=E(d_{ij}) and $\mu_{i\cdot}$=E($d_{i\cdot}$) are assumed to be finite and do not depend on t for all i, $j \in M^0$.

In order to test the two hypothesis above, the following two statistics are constructed:

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{\mathrm{var}}\left(\bar{d}_{ij}\right)}} \quad \text{and} \quad t_{i\cdot} = \frac{\bar{d}_{i,\cdot}}{\sqrt{\widehat{\mathrm{var}}\left(\bar{d}_{i,\cdot}\right)}} \text{ for } i,j \in M, \tag{24}$$

where $\bar{d}_{i,\cdot} = (m-1)^{-1}\sum_{j\in M}\bar{d}_{ij}$ is the simple loss of the i-th model relative to the averages losses across models in the set M, and $\bar{d}_{ij} = m^{-1}\sum_{t=1}^{m} d_{ij,t}$ measures the relative sample loss between the i-th and j-th models, while $\widehat{\text{var}}\left(\bar{d}_{ij}\right)$ and $\widehat{\text{var}}\left(\bar{d}_{i\cdot}\right)$ are bootstrapped estimates of $\text{var}\left(\bar{d}_{ij}\right)$ and $\text{var}\left(\bar{d}_{i\cdot}\right)$ respectively. The two EPA null hypothesis presented in equation (22) and (23) map naturally into the two test statistics

$$\mathrm{T}_{\mathrm{R},M} = \max_{i,j\in M}|t_{ij}| \quad \text{and} \quad \mathrm{T}_{\max,M} = \max_{i\in M} t_{i\cdot}. \tag{25}$$

The MCS procedure computes the statistics in (25). Intuitively, a great value of t_{ij} or $t_{i\cdot}$ reveals that estimates of model i are distant from actual realizations compared to those of any other model $j \in M$ and thus the i-th model may be discarded from M by means of an elimination rule, which are

$$e_{\mathrm{R},M} = \arg\max_i \left\{ \sup_{j\in M} \frac{\bar{d}_{ij}}{\sqrt{\widehat{\text{var}}\left(\bar{d}_{ij}\right)}} \right\}, \quad e_{\max,M} = \arg\max_{i\in M} \frac{\bar{d}_{i,\cdot}}{\sqrt{\widehat{\text{var}}\left(\bar{d}_{i,\cdot}\right)}} \tag{26}$$

respectively. At each iteration, if the null hypothesis in (22) or (23) is rejected at the fixed significance level $1-\alpha$, the elimination rules in (26) cancels out the worst model. The procedure recomputes the statistics in (24) now for $i, j \in M \subset M^0$. The iterations stop when the null hypothesis of EPA in (22) or (23) is accepted and finally set the SSM $\hat{M}^*_\alpha$.

5 Data analysis

The empirical analysis is performed on FTSE MIB index. The period of time of the analysis covers about fourteen years starting on 01-01-2007 and ending on 01-05-2021, so that, it takes into consideration the 2008 financial crisis, the european debt crisis and the beginning of the covid-19 pandemic.
The FTSE MIB (Milano Indice di Borsa) is the benchmark stock market index for the Borsa Italiana, the Italian national stock exchange, which superseded the MIB-30 in September 2004. The index consists of the 40 most-traded stock classes on the exchange. The index was administered by Standard Poor's from its inception until June 2009, when this responsibility was passed to FTSE Group, which is 100% owned by the Borsa Italiana's parent company London Stock Exchange Group.
In Figure 1 it is possible to see the close prices of the index with reference to the observed period of time. The first empirical feature emerging from the historical series is the non-stationarity of the process, in fact, a trend is present for the series. Stationarity is fundamental for modelling purposes.
The plots are realised on the index closing prices transformed in a xts object:

Figure 1: Closing Price Time Series

In financial applications, the quantity of interest is given by the variation of the price levels rather than just their levels because variations give information on profits and losses. Usually, people look at the log variations which are called in finance returns. In fact, the returns make it possible to summarise the performance of an asset for the investor. By applying the following formula we have calculated the log-returns

$$r_t = log\left(\frac{P_t}{P_{t-1}}\right) = log(P_t) - log(P_{t-1}) \tag{27}$$

Figure 2 reports the Log-Returns.

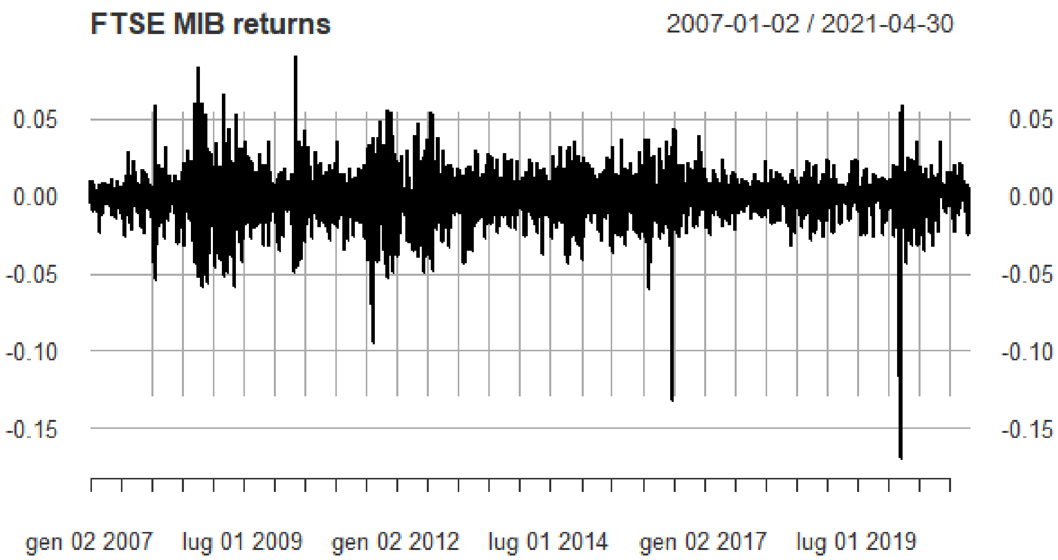


Figure 2: Log-Returns

The calculation of log-returns performs the logarithm of the first-order differentiation that eliminates non-stationarity in the mean. In fact, we see that from the analysis of ACF and PACF (Figures 3) of the log-returns the outcoming process looks being stationary and in particular behaving like a White Noise.

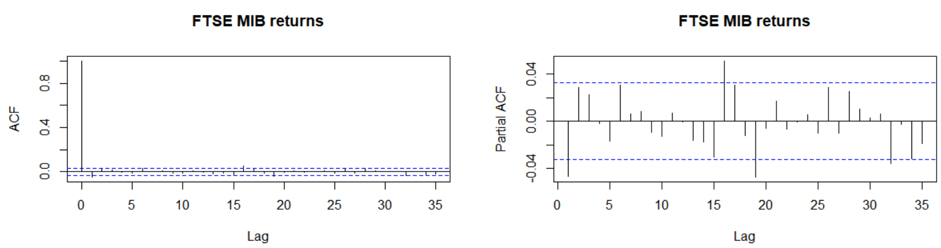


Figure 3: ACF and PACF of Log-Returns

From figure 3, watching at the bars which for lag $h \neq 0$ fall mostly within the confidence interval, we can state the absence of correlation for the values of all the three processes.
If we consider Squared returns, or in absolute terms (Figure 4), we can see that there is a dependency in the data, meaning that they are correlated. This feature will be confirmed by performing the Ljung-Box Test, which test for the Null that returns are uncorrelated.

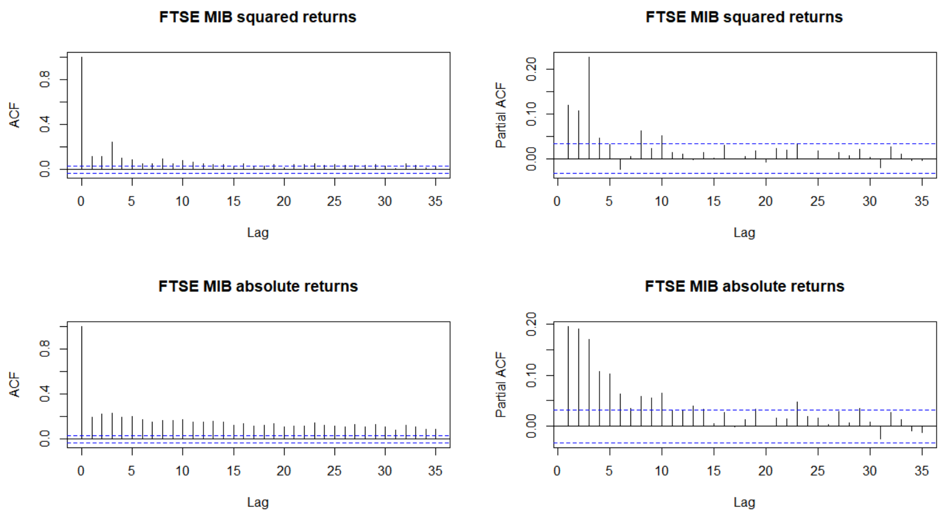


Figure 4: PACF of Log-Returns

In the Table 1 it is possible to observe some descriptive statistic performed on the Log-Returns.
The skewness index measure the asymmetry of a probability distribution with respect to its mean. A probability distribution measures can be positively (right) skewed -fat tails on the right-, negatively (left) skewed - fat tails on the left - or have zero skewness - symmetric distibution. The Kurtosis is a measure of the heaviness of the tails of a distribution i.e. how peaked or how flat a distribution is. The normal distribution has a Kurtosis equal to 3. If it is greater than 3 the Kurtosis is said to be *leptokurtic*.
In our data distributions it is possible to observe both asymmetry and presence of leptokurtosis (heavy tails), which represent two main properties (*stylised facts*) of financial time series.

Table 1: Summary statistics

	N	Min	1st Qu.	Mean	3rd Qu.	Max	Skewness	Kurtosis
FTSE MIB	3486	-0.17039	-7.3983e-03	-6.3912e-04	6.8359e-03	9.062896e-02	-0.88837	10.831

The following figures (5 and 6) represent a graphical support to what has emerged from the analysis on asymmetry and the heavy tails.

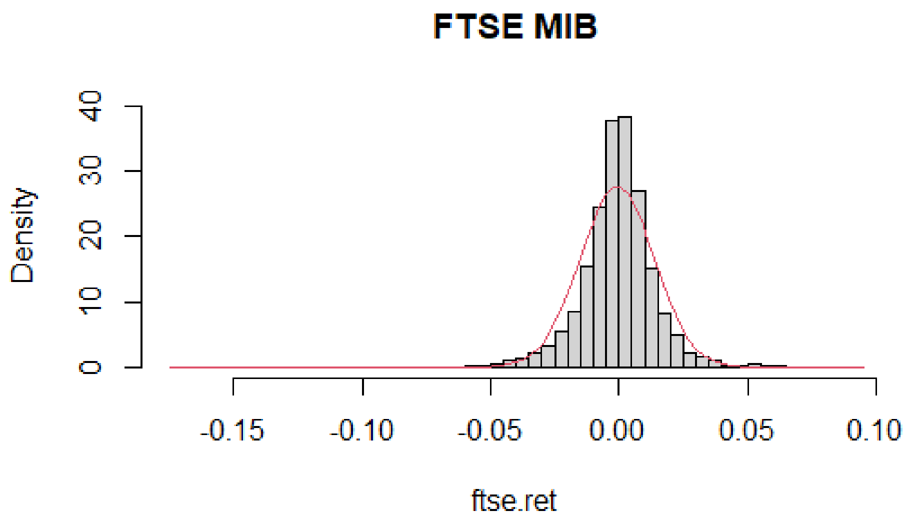


Figure 5: Returns vs Normal density distribution

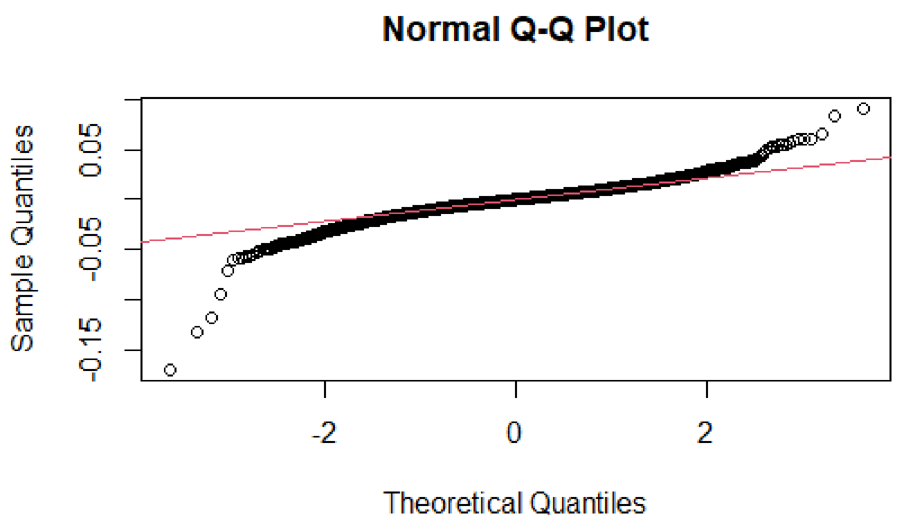


Figure 6: Returns Normal QQ-Plot

The plots in figure 5 shows the density functions of a normal distribution compared to the density function of our returns, and one can immediately see the heavy tails feature, the asymmetry and also a more pointed shape of the returns distribution.
Both figures show contrasting behaviour with respect the normal distribution, since the latter has zero skewness and zero excess kurtosis. This empirical evidence will lead us to consider other distributions to better fit r_t other than the normal one.
All the evidence that have been analysed so far represent characteristics that the financial series have in

common. For the sake of clarity, some tests have been carried out to confirm the *stylised facts*. Table 2 shows the results of tests calculated on our data:

- the Augmented Dickey Fuller test (`adf.test`) is calculated on closing prices, checks the stationarity of the distribution: the alternative hypothesis of the test is the stationarity and it confirm the stationarity;
- Ljung-Box test (`Box.test`) The test is performed on the squared returns with lag equal to 5,10,20 for robustness purposes. The null is rejected in all three cases so there is evidence of heteroskedasticity.
- Arch-LM test (`ArchTest`) checks for the presence of Arch effect, meaning the presence of structure in the series.
- Jarque-Bera (`jarque.bera.test`) checks the normality of the distribution, having as null hypothesis that data are normally distributed. The null is strongly rejected, and this confirms the results of figure 5 and 6 ;

Table 2: Tests

	lag 5		lag 10		lag20	
	Statistic	p.value	Statistic	p.value	Statistic	p.value
LB test	387.09	<2.2e-16	469.69	<2.2e-16	538.31	<2.2e-16

	Statistic	p.value
ARCH LM	283.69	< 2.2e-16
JB test	18273	< 2.2e-16
ADF test	-14.964	0.01

*the null hypothesis is rejected at the 5% significance level.

*the null hypothesis is rejected at the 5% significance level.

5.1 Macrovariables

In this project 5 macrovariables have been choosen and fitted in the class of Garch-midas, Dagm and Linear Arch midas models:

- Gross domestic product for Euro area measured as the growth rate of the same period over the previous year for all the series; the reference year is the 2015.The source for this index is Organization for Economic Co-operation and Development, its frequency is monthly and it is seasonally adjusted.
- Real Gross domestic product for Italy measured as millions of chained 2010 euro, so accounting also the inflation over time. Chained euro is a method of adjusting real euro amounts for inflation over time, to allow the comparison of figures from different years. It generally reflects euro figures computed with 2010 as the base year. The source for this index is the Eurostat, its frequency is quarterly and it is seasonally adjusted.
- Production of total industry Italy which is an index measured taking into consideration that 2010=100. Also in this case,the source for this index is Organization for Economic Co-operation and Development, its frequency is monthly and it is seasonally adjusted.
- Production of total industry Euro area is an index measured fixing 2015=100, the source is Organization for Economic Co-operation and Development, its frequency is quarterly and it is seasonally adjusted
- Long term government bond yield 10 years for Italy whose source is Organization for Economic Co-operation and Development. It is measured as a percentage, its frequency is mounthly and it is not seasonally adjusted.

The macroeconomic variables needs to meet stationarity condition and so 2 test have benn performed to check stationarity: the ADF test and KPSS test.
For all macrovariables ADF test suggests that data are stationary, since the null hypothesis of non-stationarity has been rejected. KPSS test confirms the previous result, since we do not reject the null hypothesis that data is derived from a stationary process.

Table 3: Composite Lending Indicator (CPI) Stationarity test

	ADF Statistic	ADF p.value	KPSS Statistic	KPSS p.value
GDP EU	-6.6118	0.01	0.046139	0.1
GDP IT	-3.5863	0.0396	0.11856	0.1
PTI IT	-5.0431	0.01	0.16146	0.1
PTI EU	-4.5393	0.01	0.03379	0.1
Bond IT	-5.6341	0.01	0.16146	0.1

*the null hypothesis is rejected at the 5% significance level.

6 Model estimation

In order to calculate risk measures with parametric approach, the model estimation phase is important especially to check the significance of the models' parameters. For the class of vanilla and gjr garch the parameters are significant for both the distribution Normal and t-Student. For the class of midas models the θ parameters plays a crucial role because if it is significant, it means that the lagged realizations of the macrovariable have an impact on the long run and so on the volatility of the log returns. In the class of garch midas, there are only 2 models with θ significant: the garch midas with the 10 years long term government bond yield (but with w_2 not significant, but it isn't a big problem)and Student-t distribution for the innovations; the other garch midas with θ significant is the one with production of total industry for Italy as macrovariable and Student-t distribution for the innovations. The class of double asymmetric garch midas models considers two θ parameters: θ^+ and θ^-. Among all the dagm estimated for each macrovariable, the models which have both the positive and negative θ significant are: dagm with government bond yield and Normal distribution for innovations and the dagm with Gross domestic product for Euro area, also in this case with the assumption of Normal distribution for innovation terms. The other models have at least one θ not significant, but the worse cases are those are of dagm with gross domestic product for Italy and production of total industry in Italy as macrovariables for both distributions which registered both θ not significant.

6.1 VaR estimation

After having calculated the VaR with the 3 different approaches described in the previous sections, in this section it is possible to appreciate the performances of the models in the in-sample period. The following table shows the number of failures for each model implemented in this paper, with the percentage of failure for each confidence level providing an initial idea about the models' performances. For the 95% confidence level, the models which seem to perform better are: linear-arch,linear-arch midas with gross domestic product for Euro area,production of total industry for Italy and Euro area and pti eu macrovariables, the SAV model and the AS models reporting a rate of violation exactly of of 5%. In general the semi-paramteric models, a part in same cases, are the models which perform better; the majority of the parametric models registered a percentage in between 3 and 4% meaning that the models have overestimated the VaR. There are few exception in which the models underestimate the VaR reporting a percentage of failure greater slightly greater than 5%. In the more restrcitve 99% scenario, the percentage of failure expected is 1%, thus the best models are: garch-midas with normal distribution for both bond yield curve and gross domestic product Euro area, dagm with production of total industry as macrovariable, the linear arch, the linear- arch midas with production of total industry for Italy and Euro as macrovariables and finally SAV and AS garch. In most of the cases the models overestimate the VaR with the worse one, the gjr- garch with t-student distribution wich reported only 17 VaR violations. The rest of the models underestimate the VaR. This is primary analysis is not enough the decide which models is the best one to forecast VaR because it only takes into account the number of violation but not the magnitude of each violation, for this reason other analysis must be performed.

Table 4: FTSE VaR violations

confidence level	95%	95%	99%	99%
	VaR violations	viol %	VaR violations	viol %
GARCH-norm	112	3.21%	32	0.91%
GARCH-std	123	3.52%	25	0.71%
GJR-GARCH-norm	105	3.01%	32	0.91%
GJR-GARCH-std	119	3.41%	17	0.48%
GARCH-MIDAS-n Bond Yield	114	3.27%	35	1%
GARCH-MIDAS-std Bond Yield	139	3.98%	24	0.68%
GARCH-MIDAS-n GDP EU	113	3.24%	35	1%
GARCH-MIDAS-std GDP EU	127	3.64%	26	0.74%
GARCH-MIDAS-n GDP IT	182	5.22%	64	1.83%
GARCH-MIDAS-std GDP IT	141	4.04%	54	1.54%
GARCH-MIDAS-n PTI	116	3.32%	32	0.91%
GARCH-MIDAS-std PTI	116	3.32%	32	0.91%
GARCH-MIDAS-n PTI EU	130	3.72%	41	1.17%
GARCH-MIDAS-std PTI EU	116	3.32%	32	0.91%
DAGM-n Bond Yield	114	3.27%	37	1.06%
DAGM-std Bond Yield	126	3.61%	22	0.63%
DAGM-n GDP EU	152	4.36%	49	1.40%
DAGM-std GDP EU	128	3.67%	26	0.74%
DAGM-n GDP IT	191	5.47%	46	1.32%
DAGM-std GDP IT	191	5.47%	46	1.32%
DAGM-n PTI	114	3.27%	32	0.91%
DAGM-std PTI	128	3.67%	24	0.68%
DAGM-n PTI EU	112	3.21%	36	1.03%
DAGM-std PTI EU	126	3.61%	25	0.71%
l-ARCH	175	5%	35	1%
l-ARCH-MIDAS Bond Yield	179	5.13%	32	0.91%
l-ARCH-MIDAS GDP EU	173	4.96%	33	0.94%
l-ARCH-MIDAS GDP IT	225	6.45%	66	1.89%
l-ARCH-MIDAS PTI	175	5%	35	1%
l-ARCH-MIDAS PTI EU	174	4.99%	36	1.03%
VaR HS 200	137	3.77%	33	0.9%
SAV garch	173	4.99%	35	1%
AS garch	174	5%	36	1.03%

6.2 Backtesting VaR

In the next table are reported the backtesting results for all the models used in this project. The best results come from the semi parametric models for both 95% and 99% confidence level. For what concern the vanilla garch, gjr garch, garch midas and dagm at 95% the Kupiec test is rejected very often with respect to Christoffersen meaning that there are some volatility clustering in the return; this is not a surprise because volatility clustering is one of the stylized facts of returns. At 99% it seems that normal distribution works better for vanilla garch, gjr garch, garch midas(except for garch midas with GDP Italy macrovariable), and for damg models.

Table 5: FTSE Backtesting results

LR_{UC} and LR_{IND} Critical value: 3.841 at 95% conf level and 6.634 at 99% conf level
LR_{CC} Critical value: 5.991 at 95% conf level and 9.210 at 99% conf level

confidence level	95%			99%		
	LR.uc	LR.ind	LR.cc	LR.uc	LR.ind	LR.cc
GARCH-norm	26.69	0.89	27.59	0.24	0.59	0.83
GARCH-std	17.63	1.67	19.31	3.12	0.36	3.48
GJR-GARCH-norm	33.60	0.51	34.12	0.24	0.59	0.83
GJR-GARCH-std	20.68	0.001	20.68	11.39	0.16	11.56
GARCH-MIDAS-n Bond Yield	24.88	1.02	25.90	0	0.71	0.71
GARCH-MIDAS-std Bond Yield	8.06	3.20	11.26	3.83	0.33	4.16
GARCH-MIDAS-n GDP EU	25.78	0.95	26.73	0	0.71	0.71
GARCH-MIDAS-std GDP EU	14.85	2.01	16.87	2.49	0.39	2.88
GARCH-MIDAS-n GDP IT	0.35	0.03	0.38	19.73	2.39	22.12
GARCH-MIDAS-std GDP IT	7.14	0.61	7.75	9.09	1.70	10.79
GARCH-MIDAS-n PTI	23.15	1.15	24.30	0.24	0.59	0.83
GARCH-MIDAS-std PTI	23.15	1.15	24.30	0.24	0.59	0.83
GARCH-MIDAS-n PTI EU	12.94	2.28	15.23	1.03	0.97	2.01
GARCH-MIDAS-std PTI EU	23.15	1.15	24.30	0.24	0.59	0.83
DAGM-n Bond Yield	24.88	1.02	25.90	0.13	0.79	0.92
DAGM-std Bond Yield	15.52	1.92	17.45	5.51	0.28	5.79
DAGM-n GDP EU	3.13	1.36	4.49	5.14	1.39	6.54
DAGM-std GDP EU	14.20	2.10	16.30	2.49	0.39	2.88
DAGM-n GDP IT	1.60	0.24	1.85	3.26	1.23	4.49
DAGM-std GDP IT	1.60	0.24	1.85	3.26	1.23	4.49
DAGM-n PTI	24.88	1.02	25.90	0.24	0.59	0.83
DAGM-std PTI	14.20	2.10	16.30	3.83	0.33	4.16
DAGM-n PTI EU	26.69	0.89	27.59	0.03	0.75	0.78
DAGM-std PTI EU	15.52	1.92	17.45	3.12	0.36	3.48
l-ARCH	-0.059	0	-0.059	-0.01	0	-0.01
l-ARCH-MIDAS Bond Yield	-0.28	0	-0.28	0.48	0	0.48
l-ARCH-MIDAS GDP EU	0.12	0	0.12	0.26	0	0.26
l-ARCH-MIDAS GDP IT	0.00054	0	0.00054	19.59	0	19.59
l-ARCH-MIDAS PTI	-0.059	0	-0.059	-0.01	0	-0.01
l-ARCH-MIDAS PTI EU	0.02	0	0.02	-0.059	0	-0.059
VaR HS 200	8.800	0	8.800	0.046	0	0.046
SAV garch	0.01	0	0.01	0	0	0
AS garch	0	0	0	0.037	0	0.037

*the null hypothesis is rejected.

6.3 ES estimation

In this section, we presents the results of the Expected Shortfall for both confidence levels. Each model satisfies the Sanity Check condition (ES violations < VaR violations) and, by looking at the Gamma Mean, satisfies the Martingale Condition that is: the violations must be independent and identically distributed with zero mean. All the models show good performances

Table 6: FTSE Martingale condition

confidence level	95%		99%	
	ES violations	Gamma mean	ES violations	Gamma mean
GARCH-norm	103	-0.0009	50	-0.0002
GARCH-std	87	-0.001	21	-0.0002
GJR-GARCH-norm	98	-0.0008	55	-0.0003
GJR-GARCH-std	81	-0.001	26	-0.0001
GARCH-MIDAS-n Bond Yield	108	-0.0009	51	-0.0003
GARCH-MIDAS-std Bond Yield	65	-0.001	17	-0.0002
GARCH-MIDAS-n GDP EU	107	-0.0009	51	-0.0003
GARCH-MIDAS-std GDP EU	51	-0.001	12	-0.0003
GARCH-MIDAS-n GDP IT	61	-0.001	20	-0.0006
GARCH-MIDAS-std GDP IT	25	-0.001	3	-0.0006
GARCH-MIDAS-n PTI	106	-0.0009	53	-0.0002
GARCH-MIDAS-std PTI	53	-0.001	11	-0.0004
GARCH-MIDAS-n PTI EU	123	-0.0009	64	-0.0003
GARCH-MIDAS-std PTI EU	50	-0.001	11	-0.0004
DAGM-n Bond Yield	106	-0.0009	51	-0.0003
DAGM-std Bond Yield	51	-0.001	13	-0.0002
DAGM-n GDP EU	139	-0.001	74	-0.0003
DAGM-std GDP EU	50	-0.001	12	-0.0003
DAGM-n GDP IT	60	-0.001	20	-0.0004
DAGM-std GDP IT	19	-0.001	1	-0.0006
DAGM-n PTI	106	-0.0009	51	-0.0002
DAGM-std PTI	49	-0.001	11	-0.0003
DAGM-n PTI EU	107	-0.0009	48	-0.0003
DAGM-std PTI EU	50	-0.001	12	-0.0003

*the null hypothesis is rejected.

6.4 MCS

Table 7: FTSE Superior Set of Models

Models	$Rank_{max,M}$	v_M	$MCS_{max,M}$	$Rank_R$	v_R	$MCS_{max,R}$	Loss
confidence level 95%							
Linear Arch	4	0.16839153	0.956	3	-1.209637	0.994	- 9.649023e-05
L-Arch midas GDP EU	3	0.03127431	0.998	2	0.07010698	0.999	-1.001660e-04
Asymmetric slope	2	-0.10424302	1.000	4	6.14208569	0.000	-1.036085e-04
SAV	1	-0.10912698	1.000	1	-007010698	1.000	-1.037306e-04
confidence level 99%							
Dagm N GDP EU	1	-1.79998	1	1	-1.79998	1	-0.0004427871

From the MCS table it is possible to see that at 95% level the class of model selected by the procedure is the semi parametric one, with the SAV as the best model. This result is not a surprise and it is in line with the analysis performed before in terms of VaR violations and backtesting results. At 99% level, the model coming from the MCS procedure is Dagm with Normal distribution and Gross domestic product for Europe as macrovariable.
From the results obtained by the backtesting procedure and MCS, we present the plot of the models which perform the best. At 95%, since there are more the one model in the SSM, then the plot in figure 7 shows the result for the VaR at average level giving to all the models equal weights.

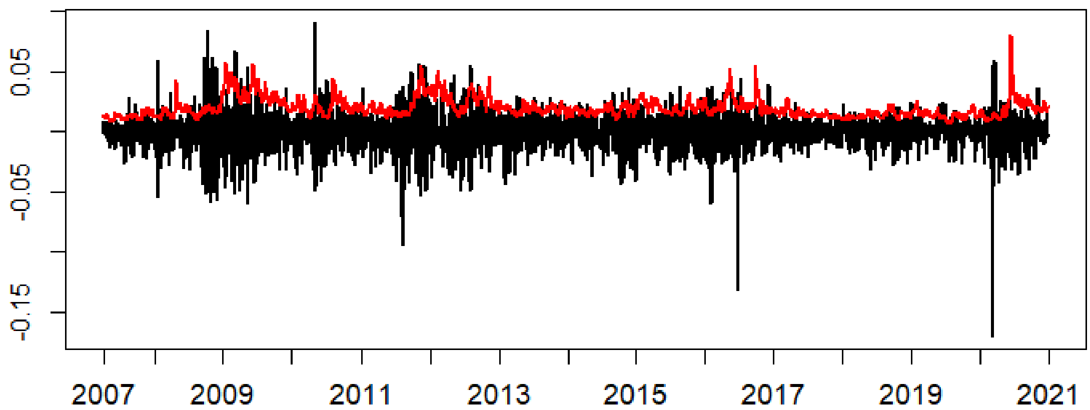


Figure 7: VaR Aggregation 95%

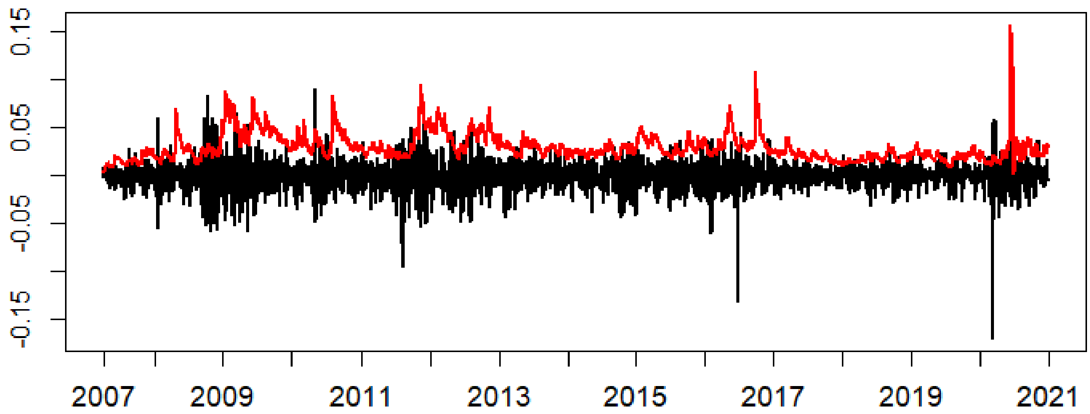


Figure 8: Dagm Normal GDP EU 99%,

7 Copula models

This part of the project focuses the multivariate framework considering jointly different risks, and since the risks are mathematically represented by random variables, this part is about the modelling of the multivariate distribution of these risks. The fundamental tool to identify the joint distribution about a random vector is the so called copula. In the univariate framework the loss distribution F_L is determined by the distribution of risk factor changes, so how is it possible to define the joint distribution of risk factor changes, given that the different risk factor changes could be dependent? Copula models define a joint distribution function for a random vector where the components of this random vector are dependent.
A d-dimensional copula $C : [0, 1]^d \to [0, 1]$ is a cumulative distribution function (CDF) with uniform marginal.

The concept of Copulas is based on Sklar's theory:
Sklar's Theorem: Consider a d-dimensional CDF, F, with marginals $F_1, ..., F_d$. Then there exist a copula C, such that

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d))$$

for all $x_i \in [-\infty, \infty]$ and $i = 1, ..., d$.

There are three main families of copulas characterized by different feature: Fundamental, Elliptical and Archimedean.
The Fundamental family represents the special cases of dependence. In there, we find models in which the marginal distributions have particular dependence structures, such as the independence, the comonotonicity and the countermonotonicity copulas (positive and negative dependence respectively).
The second group comprehends the Gaussian copula and the Tcopula, which are implicit and extrapolated from a multivariate normal distribution and multivariate student-t distribution.
The third group is the Archimedean family of copulas. Those copulas are explicit and no multivariate distribution are known. Gumbel, Clayton and Frank copulas are part of this kind of models.

7.1 Fundamental Copula

The fundamental copulas describe the fundamental dependence structure of r.v. which compose a random vector. The simplest dependence structure of r.v. is the non-dependence that is the stochastic independence where the joint distribution function is defined as the multiplication of the univariate distribution function. It is possible to define the Independent copula as:

$$\prod(u_1, ..., u_d) = \prod_{i=1}^{d} u_i$$

The others fundamental dependence structure are the perfect positive dependence and perfect negative dependence represented by respectively the comonotonicity and counter comonotonicity copulas. In order to introduce these kind of copulas it must be introduced the concept of bound of a distribution function, in particular the so called Fréchet-Hoeffding bounds.

For any C, the Fréchet-Hoeffding Bounds are:

$$max\{1 - d + \sum_{i=1}^{d} u_i; 0\} \leq C(u_1, - - -, u_d) \leq min\{u_1, ..., u_d\}$$

Thanks to this theorem it is possible to define the comonotonicity copula as the upper bound of the Fréchet-Hoeffding

$$U(u_1, ..., u_d) = min\{(u_1, ..., u_d)\}$$

The comonotonicity copula can be seen as the distribution function for the standard uniform vector $U = (U, ..., U)$, where each U is a uniform distribution. On the other hand, the counter comonotonicity copula identifies the distribution function of the random vector whose r.v. are perfectly negative dependent. This copula is weel defined only for d=2
The counter comonotonicity copula is the Fréchet-Hoeffding lower bound for d=2, that is:

$$W(u_1, u_2) = max\{u_1 + u_2 - 1; o\}$$

This copula identifies the joint probability distribution of the random vector U=(U,1-U) where U is an uniform.

7.2 Elliptical Copula

Let $X \sim E_d(\mu, \Sigma, \Psi)$ with $X_1 \sim F_{X_1}, ..., X_d \sim F_{X_d}$. The unique copula C for X is the Elliptical Copula.

7.2.1 Gaussian Copula

Let $X \sim N_d(\mu, \Sigma)$ be a multivariate normal distribution with $X \sim \Phi$ and $X_1 \sim \Phi_1, ..., X_d \sim \Phi_d$, where Φ_i are also normal. From the Sklar's Theorem, the Gaussian copula is:

$$C^{GA}(u_1, ..., u_d) = \Phi(\Phi_1^{-1}(u_1), ..., \Phi_d^{-1}(u_d))$$

It is possible write the Gaussian copula in the explicit form with a generic d:

$$C^{GA}(u_1, ..., u_d) = \int_{-\infty}^{\Phi_1^{-1}(u_1)} ... \int_{-\infty}^{\Phi_d^{-1}(u_d)} \frac{1}{\sqrt{2\pi^d det(\Sigma)}} exp\{-\frac{1}{2}(s-\mu)^T \Sigma^{-1}(s-\mu)\} ds_1 ... ds_d$$

7.2.2 Tcopula

As for the Gaussian case, the T copula is directly extracted from a multivariate t-student distribution. Let $X \sim t_d(\nu, 0, \Psi)$ be a multivariate t distribution with ν degrees of freedom (d.o.f.). The d-dimensional t-copula is then defined as

$$C^t_{\nu,\Sigma}(u_1, ..., u_d) = t_{\nu,\Sigma}(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_d))$$

where again Σ a correlation matrix, $t_{\nu,\Sigma}$ is the joint CDF of X.
As for the Gaussian, we could write the Tcopula in the explicit form; setting $d = 2$ the Tcopula is:

$$C^t_{\nu,\Sigma}(u_1, u_2) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \{1 + \frac{s_1^2 - 2\rho s_1 s_2 + s_2^2}{\nu(1-\rho^2}\} ds_1 ds_2$$

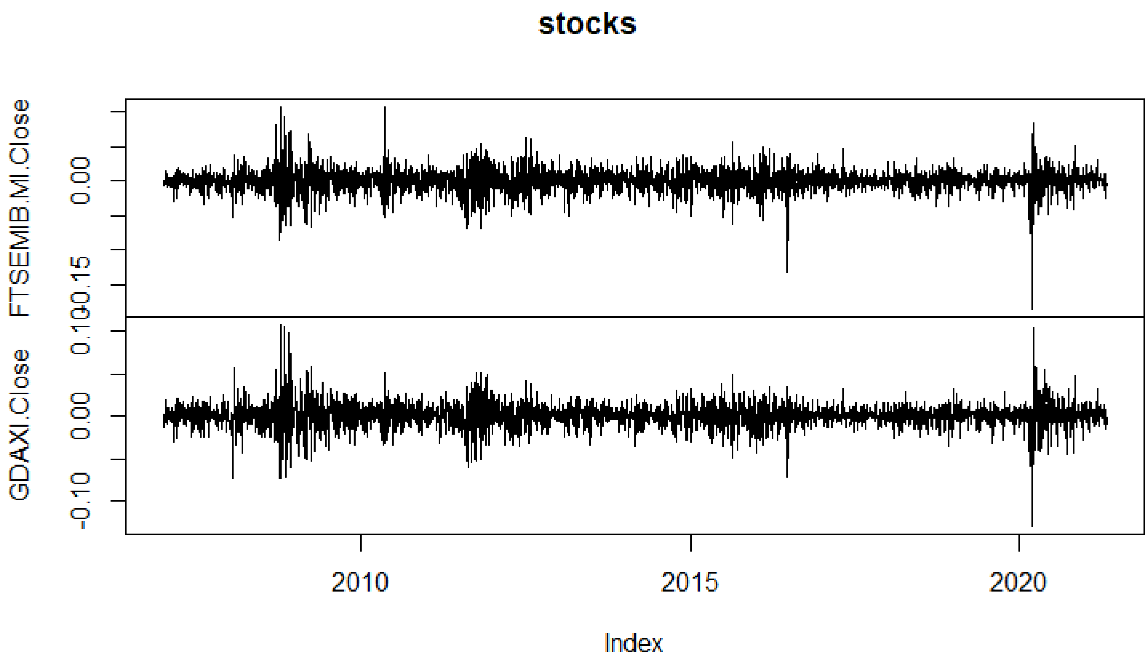
7.3 Archimedean Copula

The three main Archimedean Copulas are:

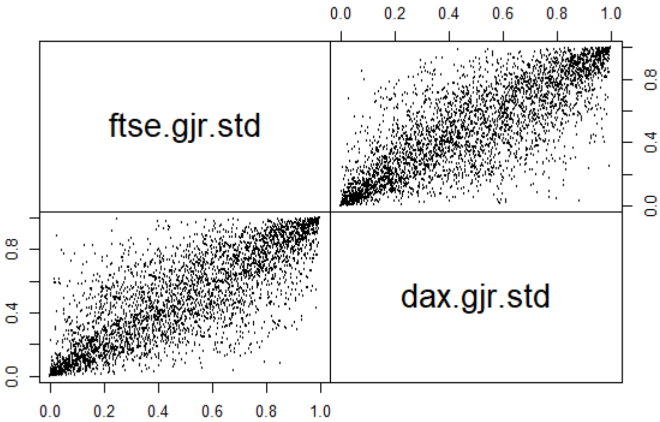
- Gubler Copula: $C^{GU}(u_1, ..., u_d) = exp\{-[\sum_{i=1}^{d}(-ln(u_i))^{\alpha}]^{\frac{1}{\alpha}}\}$, $\alpha > 1$
- Clayton Copula: $C^{CL}(u_1, ..., u_d) = [\sum_{i=1}^{d} u_i^{-\alpha} - d + 1]^{\frac{1}{\alpha}}$, $\alpha > 0$
- Frank Copula: $C^{F}(u_1, ..., u_d) = -\frac{1}{\alpha}ln[1 + \frac{\prod_{i=1}^{d}(e^{-\alpha u_i}-1)}{(e^{-\alpha}-1)^{d-1}}]$, $\alpha > 0$

8 Estimation results

In this part we will focus on the multivariate framework, composing a portfolio with 2 indexes: the ftse(used in the univariate analysis) and the German index Dax.



We want to model the aggregate loss,so we must aggregate the univariate distribution functions. In order to do that we need a copula and it must be fitted to our observed data. To do this we know that we can model the univariate time series and we know that for each sequence of univariate distribution function we can associate one and only one copula model. The first step is to model each marginal time series with the parametric approach, using GARCH-models. In particular we have used the GJR-GARCH model. We created a list of GARCH specifications and we fit each time series in a separate way, so the residuals of each model are independent. Now we are interested in the residuals because they represent the only stochastic part of the GARCH-models. The joint distribution of the time series is related to the joint distribution of the residuals. So in our case the joint distribution function of the log-returns is related to the joint distribution of the residuals. We composed a matrix formed by the standardized residuals series for each index. In our garch modelling, we have considered t-student specification. In the second step, we have build a set of [0,1] values for the residuals in order to fit the copula. Since this new set is not observed, we must build it. We did that with the pseudo observations. To build pseudo observation starting from a set of real observations we must calculate the rank of the real observation. The rank is not intended in a mathematical sense but in statistical sense since is related to the position of each observation within a series. Each pseudo observation is equal to the rank(position) over the dimension of the entire horizon n+1. Since the rank is [0,1] we obtain the pseudo-standard-uniform realization.



As it is possible to observe, for the fitted models there is the presence of weak positive dependence,so we can exclude the family of fundamental copulas from our analysis. Now on we will consider elliptical and archemedean copulas

Table 8: Kendall's Tau

	Gaussian	Tcopula	Gumbel	Clayton	Frank
GJR-STD	0.6161975	0.6179882	0.5920039	0.5278733	0.617257

Kendall's tau is the most important rank correlation estimators. As Kendall's tau is a measure with possible values in the interval [-1,1], when it takes a value of 0, this means that variables are independent. When it takes a value of 1, variables are co-monotonic: perfect positive dependence; while it is equal to -1 in case of perfect negative dependence: variables are counter-monotonic. Our five copulas have positive tau, and then positive dependence.
Next step is the analysis of the main criterion to evaluate the best copula for our series: the Akaike and log-likelihood criterion.

Table 9: Akaike

	Gaussian	Tcopula	Gumbel	Clayton	Frank
GJR-STD	-4027.452	-4206.394	-3833.427	-3450.427	-3749.012

Table 10: Log Likelihood

	Gaussian	Tcopula	Gumbel	Clayton	Frank
GJR-STD	2014.726	2105.197	1917.713	1726.214	1875.506

In terms of those criterion the best copula is the t-copula which entails that higher and lower tail dependence are symmetric in our data. This is in line with the graph seen before

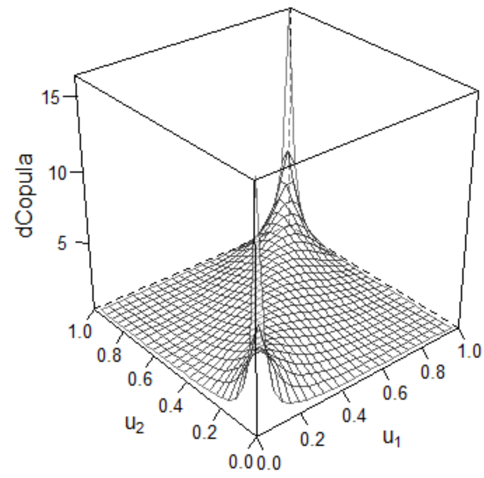
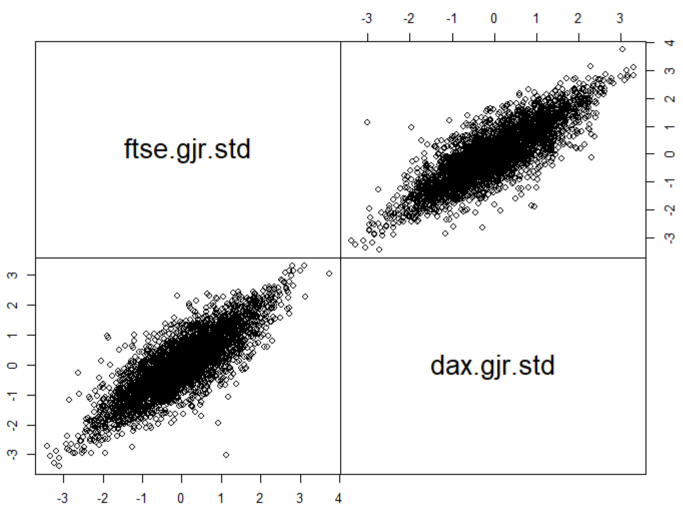


Figure 9: Tcopula density



Since there is not a closed form for the t-copula(as in the case of the Gaussian one) we must perform our analysis by a numerical perspective, so we can build distributions by simulations. We have to simulate jointly the realizations for each single series and these realizations must take into account the dependence between these shares. For each trajectory we must simulate the realization of the t-copula, given our fitted model for each time;we are generating the uniform realization given the t-copula, so the realizations are in [0,1], but they are sampled by a t-copula. Then we pass to the quantile transform, so we pass from the U distributed values to the t distributed values. We have to apply the quantile function to each time series. Now, starting from the realization of the t-student distribution multivariate, given the measure of dependence given by the copula, we want to calculate the loss and the VaR. When we work on shares is very simple to obtain the aggregated loss, since, for a unit value of the portfolio is equal to the sum of risk factor changes(we have log-returns and so we consider the sum of log-returns). So we build a vector that for each time we have the sum of the log-returns for each series and, we do this also with the predicted realizations. Now the simulated aggregated loss is constituted by multiple patterns for each time, so we obtain a matrix. Starting from this matrix, for each row, so for each time, we apply a quantile function and in this way we obtain the multivariate VaR with a non parametric approach

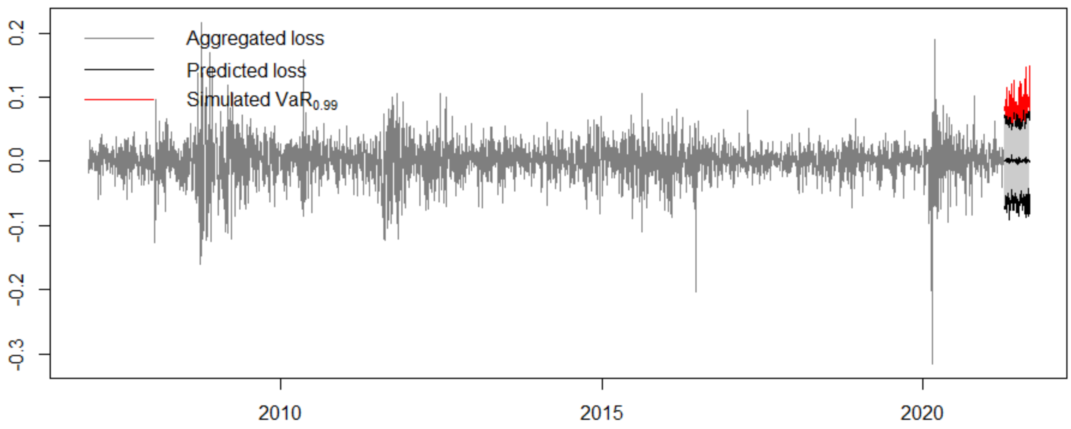


Figure 10: Simulated VaR of Tcopula

The plot below (Figure 10) shows the calculation of the VaR based on the simulation of 200 paths, of 100 trading days. Considering the mean of these different paths observations, the predicted loss, the aggregated loss, we compute the simulated VaR at 99% confidence level.

9 Conclusion

In this paper we have shown the most used risk measures in risk management: Value at risk and Expected shortfall. The VaR has been calculated using different approaches with different kind o models, and through the backtesting and MCS procedure we have selected a set of best models at 95% coming from the family of semiparamtric models and one best model in the more restrictive case of 99% coming from tha midas family with Gross Domestic Product in EU area as macrovariable. For the ES only a first step backtesting has been performed with good results for all the models implemented in the analysis The second part of the paper refers to the risk aggregation and the most common way to model risk in a multivariate framework with copula models. Through the use of these models we performed the evaluation of the VaR at portfolio level considering a 99% confidence interval