**Abstract**

**Al final de todos los modelos explicar el backteting y del msc que nos hace falta.**

1. **Introduction**

The stock market is one of the financial markets with the most risk, and volatility is a standard metric for measuring risk. Stock market collapses, wars, natural disasters, and commodities crises all seem to coincide with periods of extreme volatility.

The stock market crisis in 1987, was the catalyst for the creation of value-at-risk (VaR) as a risk indicator. Later, the conditional Value-at-Risk, known as the Expected Shortfall (ES), was included.

Value-at-Risk is the greatest loss on an investment over a certain time period, whereas the Expected Shortfall is the average of the losses that exceed the Value-at-Risk [2].

Creating proper volatility estimates is an important aspect of evaluating risk. As a result, volatility modeling may be likened to calculating the risk of investing in a certain asset, portfolio, or market.

Increased market risk is caused by unexpected changes in market pricing. This indicates that the higher the amount of volatility, the higher the level of risk. Volatility must be assessed since it is not readily apparent in the market [3].

The standard deviation of returns is the most basic measure of volatility. Financial returns, on the other hand, are known to have certain characteristics, known as stylized facts. Volatility clustering, asymmetry, and leptokurtosis are among the stylized facts, according to Bollerslev, Engle, and Nelson [4].

The objective of our analysis is to investigate the Swiss Stock Market Index behavior in the last 20 years. We test different models in the univariate case and two risk measures: Value at Risk (VaR) and Expected Shortfall (ES). We work on the Profit & Loss distribution represented by the daily returns of the stocks.

While doing our analysis, we decided to expand our evaluation and include the effects of the Financial Global Crisis (2007-2008) and the COVID-19. We also include three macro-variables (Real GDP, Unemployment, and 10-years Bond Yields).

Several models are used for this purpose; we can split them out into three main approaches: parametric, semi-parametric and non-parametric one.

The models are evaluated in-sample for this purpose. For the in-sample study, backward-looking assessment methods are applied, and these estimates are subsequently used in VaR and ES risk measures; Backtesting is then used to assess the risk measures.

The last step of our analysis is the Model Confidence Set (MCS) procedure, and results in a smaller set of superior models, also called, Superior Set of Models (SSM).

The rest of the paper is organized as follows. Section 2 contains the data description of our research. Section 3 demonstrates the methodology. The empirical analysis is covered in Section 4. Finally, Section 5 contains the conclusion.

* 1. **Data description**

The analysis considers a time period of 20 years, from 01/01/2000 to 31/12/2020, having 5280 daily observations. Our analysis is about the relation of the Swiss Stock Market Index (SSMI) that is the most important index in the Swiss Stock Market, the index is composed by the 20 largest and most liquid stocks in the market, which 19 are large-caps and one is middle-cap. We consider such a long time period in order to also evaluate the effects of the *Global Financial Crisis* and *COVID-19* on the SSMI.

In order to understand the relationship between the Swiss Economy and the SSMI, we consider three different macro-variables, and it effect in the index of that important and developed country.

Firstly, we consider the *Real Gross Domestic Production*that evaluate the Nominal GDP considering the inflation and deflation of the country. The*Unemployment rate* is a useful measure of the underutilization of the labor supply. Is seen as an indicator of the efficiency and effectiveness of an economy to absorb its labor force and of the performance of the labor market. *Bond Yield (10 Years)* is the return that an investor realizes on a bond. The simplest definition can be setting the bond yield equal to its coupon rate.

Within the considered period of our dataset, we can observe the *Financial Global Crisis*, as we were interested to focus on that financial shock, as it was the longest and deepest economic downturn in many countries. To evaluate the shock in the SSMI, we took the information of the returns from 03-01-2006 up to 30-12-2011 to evaluate before and after of the crisis.

**Table**

The table show us the descriptive statistics of the dataset.

**Plot**

We can observe that the SSMI was increasing before the crisis, that begin, as said before, at the middle of 2007, and was followed from a long-term decreasing until the beginning of 2009 when arrive to the lowest point of the plot and then started to increase again.

We can also observe in our dataset another big world crisis, the *COVID-19*. To analyze this crisis, we took into consideration a period from 03-01-2019 up to 30-12-2020. Once again, to also understand how the situation was before of the shock. The economic shock started when a big part of the world governments announces the quarantine due the pandemic, this created a shock in most of the financial markets in the world, followed by an inactive period of production, supply and provisions.

**Table**

In the table we can find the descriptive statistics.

**Plot**

In the plot we can see the effect of the crisis in the SSMI, which was increasing until the end of march of 2020, when the covid crisis had the biggest effect in the financial markets, followed by a fast increase of the market.

1. **Methodology**

*Parametric approach*

In the project several univariate models were used to specify the volatility dynamics of SMI returns series.

The GARCH models were thought as an extension of ARCH models (Bollerslev 1986). The standard GARCH (1,1) was one of the models used in this project. In this model, the conditional variance of returns is constructed as:

Text

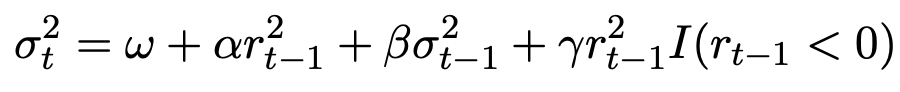
Description automatically generated

It depends on a constant term, last observation of squared returns and previous conditional variance. It is easily to see that this model is able to take into account the clustering phenomenon, where high (low) volatility tends to be followed by high (low) volatility in the next period. Empirical evidence suggests that the specification GARCH (1,1) is suitable for explaining financial data series. It can be shown that the ARCH () can be replaced with a GARCH (1,1), respecting the parsimonious principle.

It is covariance stationarity if and to assure ,

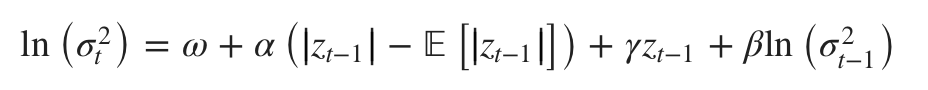
However, this model is uncapable of taking into consideration the leverage effect on today’s volatility, as empirical evidence suggests volatility increases when past returns are negative. In Standard GARCH, good and bad news have the same effect on volatility.

This inconvenient was solved by the GJR-GARCH (1,1) model. In this specification, a new term is added to explain returns volatility where it will enter in the equation to increase volatility only when past returns are negative, assigning asymmetrical treatment to the sign of previous returns. The dynamic equation of a GJR (1,1) for the conditional variance is given by:



The last term becomes operative only when past returns are negative (, adding extra volatility to the original specification in the Standard GARCH. For covariance stationarity, is needed: (.

Considering a return time series , where  is the expected return and  is a zero-mean white noise. Despite of being serially uncorrelated, the series  does not need to be serially independent. For instance, it can present conditional heteroskedasticity. The Exponential GARCH (EGARCH) model assumes a specific parametric form for this conditional heteroskedasticity. More specifically, we say that ~EGARCH if we can write , where  is standard Gaussian and:



Then, the CSGARCH model of Lee and Engle (1999) decomposes the conditional variance into a permanent and transitory component so as to investigate the long- and short-run movements of volatility affecting securities. Letting qt represent the permanent component of the conditional variance, the component model can then be written as:

Text, letter

Description automatically generated

where effectively the intercept of the GARCH model is now time-varying following first order

autoregressive type dynamics. The difference between the conditional variance and its trend, is the transitory component of the conditional variance.

In addition, there were used another set of models which include macroeconomic variables in the dynamic equation of conditional variance. The idea behind is that macroeconomic variables indicators can influence financial asset prices because of expectations or macro analysis. The problem of including MV (macroeconomic variables) is that these are observed in a lower frequency than daily returns.

The univariate GARCH-MIDAS (MIxed DAta Sampling) model includes macroeconomic variables. The problem of different frequency on observations is solved by decomposing the dynamic equation into a short-term component depending on past squared returns observations and past value of the short-term component, and a long-term component depending on a constant variable and the macroeconomic variable. The model is defined as:



Where:

The short run component is given by:Diagram

Description automatically generated

While the long run component is defined as:Schematic

Description automatically generated with low confidence

Where:

m is the intercept

the coefficient to estimate

: function that weighs past K realizations of

The Beta function is defined by:

Text

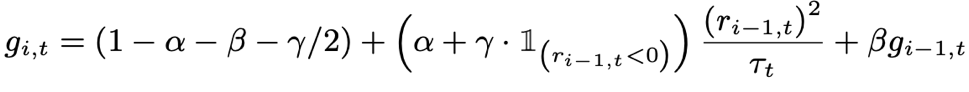
Description automatically generated with medium confidence

If recent observations weigh more in the long-term component. We assume is independent of and that is strictly stationary. The drawback of this model is that is does not considers the leverage effect for past negative returns and bad macroeconomic news.

In our project, we estimated this model with three different macroeconomic variables: Real GDP, Unemployment Rate and 10-years Bond Yields.

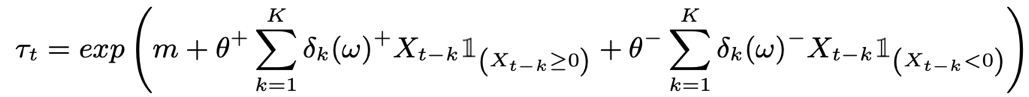
The Double Asymmetric GARCH MIDAS (DAGM), proposed by Amendola (2019), is the model that includes macroeconomic variables as an explanatory determinant and the effect of past negative returns and macroeconomic indicators.

Short-run component is given by:



Where the term determined by the coefficient become operative in case past returns are negative (increasing volatility).

Long-run component is defined as:



Where different weighs are given to past macroeconomic variable whether they took negative or positive values.

*Non-parametric approach*

In the non-parametric models, we do not make any assumption on the distribution of daily returns, and we do not have to estimate any parameter.

The most known method of this approach is the Historical Simulation (Hendricks (1996)) method, which was implemented in the project for calculating the based on the sample quantile for a fixed rolling window of data. This means calculating using a fixed number of previous observations and using it as a forecast. The window length (w) is fixed and is updated from the newest observation.

Even though this model does not need assumptions on daily returns distribution, it needs that returns entering the moving window must be independent and identically distributed (iid).

VaR is estimated as:

Where represents the daily returns that are included in the length for the sample used in the calculations. And is the sample quantile at level.

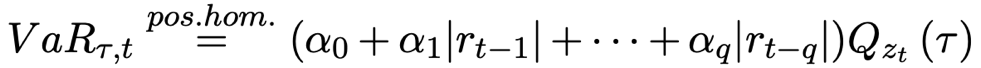
The problem is that this model needs a large window to get an accurate estimation, but it increases the chances of returns not respecting the previous assumption (less likely to be iid) and the estimator would be bad as structural breaks could happen.

In the project, the length of the rolling window was set at 250 observations.

*Semi-parametric approach*

This kind of approach also uses quantiles regression for the estimation as we directly estimate it by using the -th sample quantile ()

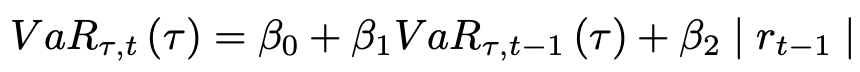
The Quantile Linear ARCH model (Koenker and Zhao-1996) uses this logic of approach, where the is directly calculated and is defined as:



One difference from the original ARCH model is that, instead of lagged squared returns, the semi-parametric approach uses past absolute returns.

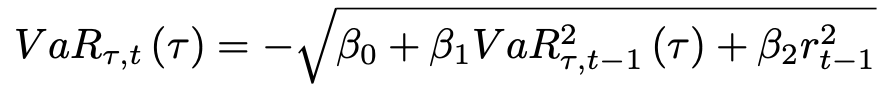
The rest of the semi-parametric approach models are different specifications of the CAViaR model (Engle and Manganelli-2004), where is estimated using previous estimations of it () and past returns observations introduced using its absolute value. Three different specifications were used for our project.

The CAViaR-Symmetric Absolute Value (CAViaR-SAV) is defined as:



The VaR at level depends on most recent VaR estimation and recent observation of absolute returns.

The CAViaR-Indirect GARCH (CAViaR-IG) is similar to the specification from above: instead of past absolute returns, it uses past squared returns, and squared instead. It follows the next specification:



One drawback from these two previous models is that positive and negative returns have the same treatment. This means that they are not able to model the Leverage Effect, which can be solved with the CAViaR-Assymetric Slope (CAViaR-AS), defined as:



This model allows us to give different relevance to past returns depending on whether they took positive or negative values.

A picture containing text, clock, watch

Description automatically generatedParameters from CAViaR models are estimated by minimizing the following quantile loss function:

1. **Empirical Analysis**

Once we transformed our data set into a time series, we evaluated the characteristics of the Swiss Stock Market Index (SSMI) series from 01/01/2000 to 31/12/2020. We focus our data analysis using the main tests for evaluating the stylized facts.

*Auto-correlograms*

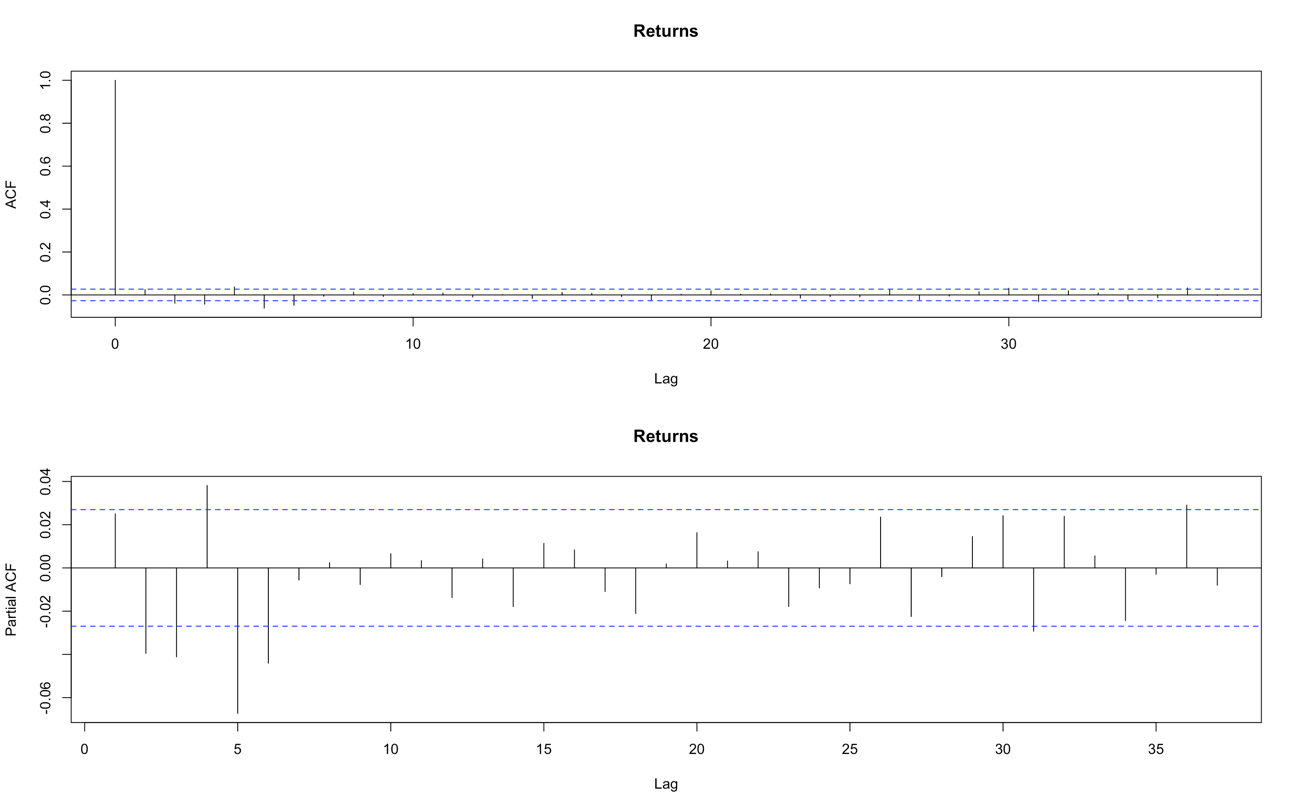
We started by evaluating the autocorrelation function (ACF). The ACF is a function where daily returns depend on past returns observations. In this test, we want to see whether previous observations are correlated with today’s observation, and we test for each individual correlation coefficient. So, our null hypothesis is that there is no correlation with past returns , coefficients accompanying lagged returns are equal to zero. But, if we reject H0, is because the lagged return is statistically significant and different to zero. Meaning that, the lagged return is significant to explain daily returns and we should take it into account in our autoregressive model.

By looking at the auto-correlogram, all those lags that exceed the nullity band mean that, according to our sample, there is evidence suggesting the autocorrelation is significant. We found no significant lags for today’s returns as every test statistic fell inside the nullity band. This result suggest that the series can be associated with a white noise process because if there are no significant past observations, daily returns only depend on the error term (characteristic of a white noise process).

Chart

Description automatically generated with medium confidence

Later, we tested the partial autocorrelation function. In this test we see if the direct effect/influence of the lag return respect to daily returns, eliminating the influence it contains from other lags (we do not consider the dependency created by the lags between them).



Once again, the obtained results suggest there is no dependency between daily returns and past returns. In consequence, we could assume that the expected value of daily returns is equal to zero.

However, results varied a lot when we transformed the time series into squared and absolute returns.

A picture containing text, device, caliper

Description automatically generatedA picture containing text, caliper, device

Description automatically generated

Dependency of daily returns with recent past observations seemed to be significantly different from zero, and then started falling into the nullity band. We can interpret that, shocks are not permanent, they are absorbed through time, and that they should not be associated with a white noise process as there are relevant lags. The effect of lagged variables losses relevance.

These results are important because let us use the squared returns as a proxy of conditional variance, due to the fact that daily returns could be thought as a zero mean process and squared returns cannot.

*Normality test*

Furthermore, we evaluated if the daily returns series could be considered to follow a normal distribution. This null hypothesis was tested with the Jarque-Bera test.

Firstly, we needed to calculate the kurtosis and the skewness of the series. The kurtosis measures the heaviness of the tails of the distribution. According to our sample, the series had a kurtosis equal to 7.64, much larger than 3 (value for a normal distribution), being a case of leptokurtic distribution. The interpretation of the fat tails result is that there is a higher probability of observing large returns and losses.

The estimated skewness took a negative value of -0.261, which means that number of observations for negative daily returns are greater than the number of positive returns.

These values led us to reject the null hypothesis, the evidence suggests the daily returns do not follow a normal distribution as we can inferred from the histogram below.

Chart, histogram

Description automatically generated

It would be wrong to consider daily returns series to follow a normal distribution because we found asymmetry and heavy tails.

*White Noise tests*

To evaluate if the series could be associated with a white noise process, we used the Box-Pierce test and the Ljung-Box test. For both, the null hypothesis is that there is no correlation between the dependent variable and the past observations of it. In this case, it is a global test, meaning that rejecting the null hypothesis suggest there is at least one relevant lag, not specifying which one. These tests were done for the daily returns and the squared returns. In both cases we rejected the null hypothesis using five lags, which is inconsistent with our findings of the auto-correlograms analysis. However, as we increased lags, the p-value associated with daily returns also increased while it was constant for the squared value case. This means that squared return results are more robust than those obtained for the daily returns.

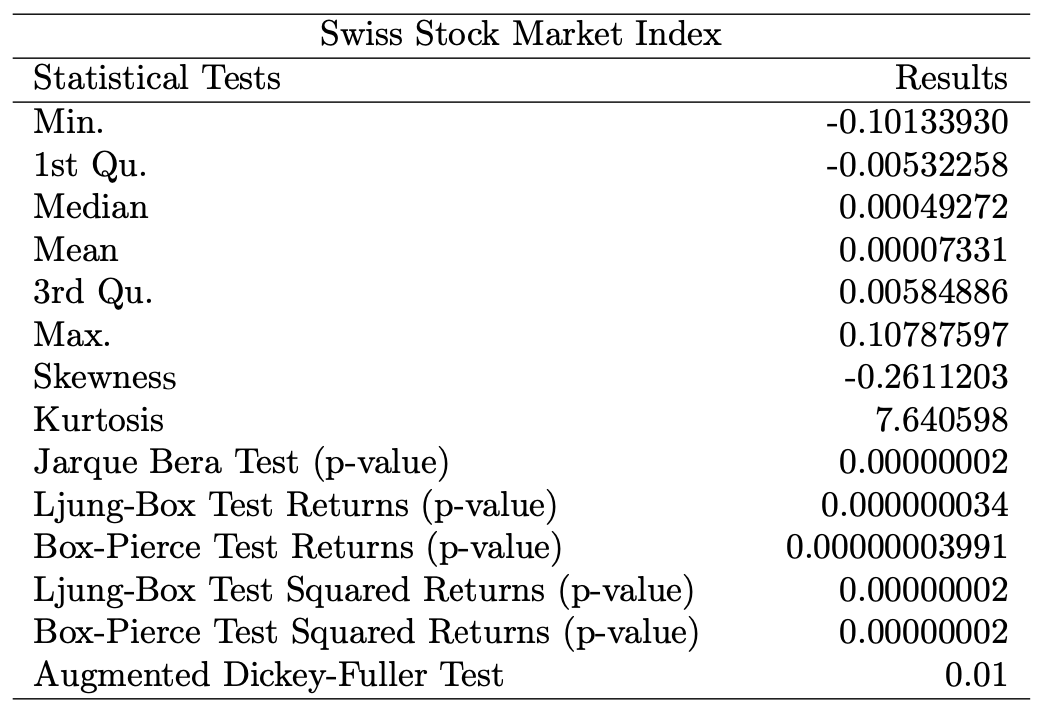
*Stationarity Test*

Finally, regarding the tests that could be done for checking the stylized facts of daily returns time series, we ran the Augmented Dickey Fuller Test. This test evaluates the existence of unit roots that implies that the process is no stationary because the influenced of lagged variables is not absorbed over time, shocks are permanent and cumulative, generating deterministic tendency. The results we obtained indicate that the process is stationary as the evidence suggests there is no presence of unit roots.

Table

Description automatically generatedGiven the p-value=0.01, we reject the null hypothesis of the existence of unit roots.

In the table below we can see the main results of the tests we ran for the evaluation of the series stylized facts:



**3.1) Models and Risk Measures**

The goal is to determine the best estimate of two risk measures: **Value at Risk (VaR)** which is a loss quantile; it represents the highest projected loss over a particular time horizon (typically one day), given a confidence level (in our case 1 and 5 percent); and **Expected Shortfall (ES)** which is the average of Var’s that takes into account the loss distribution's tail that the VaR does not address. Three different approaches are used: the parametric approach, the semi parametric approach and the non- parametric one.

For the parametric approach, we decide to implement various GARCH models, since they are better suited for modelling financial time series characterized by volatility clustering effect. Standard GARCH (SGARCH) models can model the volatility clustering phenomenon. Besides SGARCH models with different distributions (both normal and student-t distributions), we also considered Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) and Exponential GARCH (EGARCH), they take into account a stylized fact that is not contemplated in the SGARCH model, which is the empirically observed fact that negative shocks at time *t−1* have a stronger impact in the variance at time *t* than positive shocks. This asymmetry used to be called leverage effect because the increase in risk was believed to come from the increased leverage induced by a negative shock. In addition, the Component Standard GARCH (CSGARCH) decomposes the conditional variance into a permanent and transitory component so as to investigate the long-and short-run movements of volatility affecting securities.

As for the models that we just mentioned, return distribution assumption has been made normal, student-t distribution and their skewed version.

We consider different Mixing Data Sampling GARCH models (GARCH- MIDAS) and Double Asymmetric GARCH-MIDAS (DAGM); these models take into account Macroeconomic Variables (MVs) to estimate more faithfully the volatility. We compute each -MIDAS including and excluding the skewness parameter (γ) to include in the short-run equation an eventual asymmetric response of conditional variance to returns’ negative changes.

We pair each MIDAS model with three different Macroeconomic Variables, discarding those in which the long-run component parameter θ results to be non-significant, meaning that it does not carry any relevant information for a better estimation of the volatility. We considered the 3 MVs previously mentioned (Real GDP, Unemployment, and 10 years Bond Yields). The main difference between GARCH-MIDAS model and DAGM model is that the latter consider the possibility of asymmetric effect of the returns and the (MIDAS or MV) variable on the conditional volatility.

The second class of models is referred to semi-parametric approach that attempts to model directly quantile through quantile regression. They do not need a specification assumption for the return’s distribution, avoiding the possibility to incur in some misspecification error. We use different type of Conditional Autoregressive Value at Risk (CAViaR) models and Linear ARCH models (L-ARCH). The CAViaR models implemented are: the CAViaR-Symmetric Absolute Value (CAViaR- SAV), CAViaR- Asymmetric Slope (CAViaR-AS) and CAViaR-Indirect GARCH (CAViaR-IG). The main difference between them is that the CAViaR-SAV and CAViaR-IG respond symmetrically to past returns, whereas the CAViaR-AS allows the response to positive and negative returns to be different.

Another possible approach is the Historical Simulation. This is a non- parametric approach because it calculates the VaR as the sample quantile over a moving window of data. This approach does not require a distributional assumption for the errors. We consider just the 250 windows size (w) for the VaR evaluation.

**3.2) Backtesting VaR and ES**

In order to understand which of the aforementioned models better estimate the risk measures, we continue the analysis by fitting them on our financial time series. As said in the previous section, we work on three different approaches: parametric, semiparametric and non-parametric.

These models are used to perform the Backtest, namely a set of statistical procedures designed to check if real losses are in line with estimated risk measures. In particular we focus on three hypothesis tests: the *Proportion Of Failure test* (LRuc), which inspects if the theoretical VaR violations are equal to the estimated ones; the *Conditional Coverage test* (LRcc), composed by the sum of Portion Of Failure test and the independence test (which examines if the VaR violation at time *t* depends on the outcome at time (*t − 1*)); and, the *Dynamic Quantile test* (DQ) which verifies the independence of the VaR violations jointly with the correctness of the number of violations as the CC test, but it has been demonstrated to have more power than this latter (the CC) test. In particular, the DQ test consists of running a linear regression where the dependent variable is the sequence of VaR violations, and the covariates are the past violations and eventually any other explanatory variables.

We have a good model in terms of VaR predictability if it accepts the null hypothesis of the three aforementioned tests. The analysis is conducted at two confidence levels of the risk measures: 95% and 99%.

Looking at table, we can notice another column, the *Actual over Expected* (AE) exceedance ratio, that is the ratio between the actual number of exceptions of the returns respect to the estimated VaR and the theoretical exceptions. The closer the AE ratio is to 1, the more precise is the model; if the ratio is less than 1 the model overestimates the risk, while if it is greater than 1 the model underestimates the risk.

The only excluded models before the Backtesting procedure regards - MIDAS models: when they present a non-significant parameter , they are discarded. Indeed, observing Table 3, we can realize that not all these models are presented.

At the given confidence levels, looking at the parametric approach, the results suggests that, when we assume the normal distribution for MIDAS models, the considered MVs used to for the evaluation, better predict the VaR, when compared with the MIDAS models with student-t distribution. Indeed, in most of the cases of our evaluation the AE values related to the normally distributed models are closer to one and the null hypothesis of UC and CC tests fail to reject (are accepted). Regarding the DQ, we observe that the test works better for the -MIDAS models which include the skew parameter.

With regards to the semi-parametric approach, we can observe that all the models perform well for the considered series. The only exception is for SAV at 99% confidence level, which is for the only model that AE is overestimated; also, CC and DQ are rejected.

An opposite situation occurs in the non-parametric approach, where none of Historical Simulation passes the tests. The only exceptions regard HS (w = 250), in fact is the only HS that we decided to include in the evaluation, and it only pass the POF (UC) test. Also, the AE value is close to one, but it did not pass either the conditional coverage (CC) and the Dynamic Quantile (DQ).

TABLE

Backtesting the ES involves checking the goodness of the estimated loss distribution, that is to benchmark the ES realizations with respect to the loss realizations.

In this work we use the following test: unconditional coverage tests for ES (McNeil and Frey 2000).

The test of McNeil and Frey is based on the size of the discrepancy between the actual return and the estimated ES when a VaR violation occurs.

**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 of models, and through the Backtesting and MCS procedure we have selected a set of best models at 95% and 99% coming from the family of semiparametric models, in particular the “AS model” for both the different confidence levels.