Swiss Stock Market Index – Univariate Analysis

Gonzalo Lacunza, Andres Alberto Quintero, and Nagib Stefano Souki

Faculty of Economics

MEMOTEF Department

Sapienza University of Rome

Abstract

A new framework for the joint estimation and forecasting of dynamic Value-at-Risk (VaR) and Expected Shortfall (ES) is proposed by incorporating intraday information into a generalized autoregressive score (GAS) model, introduced by Patton, Ziegel and Chen (2019) to estimate risk measures in a quantile regression setup. We consider four intraday measures: the realized volatility at 5-min and 10-min sampling frequencies, and the overnight return incorporated into these two realized volatilities. In a forecast- ing study, the set of newly proposed semiparametric models is applied to 4 international stock market indices: the S&P 500, the Dow Jones Industrial Average, the NIKKEI 225 and the FTSE 100, and is compared with a range of parametric, nonparametric and semiparametric models including historical simulations, GARCH and the original GAS models. VaR and ES forecasts are backtested individually, and the joint loss function is used for comparisons. Our results show that GAS models, enhanced with the realized volatility measures, outperform the benchmark models consistently across all indices and various probability levels.

Keywords: Swiss Stock Market Index, Univariate Analysis, GARCH, Value-at-Risk, Expected Shortfall, Backtesting.

JEL Classification:

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.

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.

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. [1]

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.

2) Data description

All the mentioned below data was imported from Yahoo Finance database and the Federal Reserve Economic Data database.

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. [2]

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. [3] 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.[4] *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. [5]

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.

Financial Crisis Shock	
Statistical Test	Results
Min.	-0.0810810
1st Qu.	-0.0064311
Median	0.0004966
Mean	-0.0001657
3rd Qu.	0.0066730
Max.	0.1078760
Skewness	0.05437087
Kurtosis	6.731367
Jarque Bera Test (p-value)	0.00000002
Ljung-Box Test Returns (p-value)	0.00000005754
Box-Pierce Test Returns (p-value)	0.0000001026
Ljung-Box Test Squared Returns (p-value)	0.000000002
Box-Pierce Test Squared Returns (p-value)	0.000000002
Augmented Dickey-Fuller Test	-11.495

The table show us the descriptive statistics of the dataset.

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.

COVID-19 CRISIS	
Statistical Test	Results
Min.	-0.1013393
1st Qu.	-0.0036022
Median	0.0008624
Mean	0.0004681
3rd Qu.	0.0059655
Max.	0.0678046
Skewness	-1.434697
Kurtosis	15.4008
Jarque Bera Test (p-value)	0.00000002
Ljung-Box Test Returns (p-value)	0.0004052
Box-Pierce Test Returns (p-value)	0.008021
Ljung-Box Test Squared Returns (p-value)	0.000000002
Box-Pierce Test Squared Returns (p-value)	0.000000002
Augmented Dickey-Fuller Test	0.01

The above table show us the descriptive statistics of the dataset.

In the plot we can see the effect of the crisis in the SSMI, which was increasing until the end of march of 2020, followed by a huge set-back when the COVID-19 crisis had the biggest effect in the financial markets, followed by a fast increase of the market.

3) 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:

$$\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

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 $\alpha + \beta < 1$ and to assure $\sigma_t^2 > 0$, ω; α; $\beta \geq 0$. 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. [6]

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:

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

The last term becomes operative only when past returns are negative ($I = 1\ if\ r_{t-1} < 0; I = 0\ if\ r_{t-1} > 0$) , adding extra volatility to the original specification in the Standard GARCH. For covariance stationarity, is needed: $(\alpha + \beta + \frac{\gamma}{2}) < 1$.

Considering a return time series $rt = \mu + \varepsilon t rt = \mu + \varepsilon t$, where μ is the expected return and εt is a zero-mean white noise.[7] Despite of being serially uncorrelated, the series εt does not need to be serially independent.[8] For instance, it can present conditional heteroskedasticity. The Exponential GARCH (EGARCH) model assumes a specific parametric form for this conditional heteroskedasticity.[9] More specifically, we say that εt~EGARCH if we can write $\varepsilon t = \sigma t z t \varepsilon t = \sigma t z t$, where zt is standard Gaussian and:

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

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: [10]

$$\sigma_t^2 = q_t + \sum_{j=1}^{q}\alpha_j\left(\varepsilon_{t-j}^2 - q_{t-j}\right) + \sum_{j=1}^{p}\beta_j\left(\sigma_{t-j}^2 - q_{t-j}\right)$$
$$q_t = \omega + \rho q_{t-1} + \phi\left(\varepsilon_{t-1}^2 - \sigma_{t-1}^2\right)$$

where effectively the intercept of the GARCH model is now time-varying following first order autoregressive type dynamics. [11] The difference between the conditional variance and its trend, $\sigma_{t-1}^2 - q_{t-j}$ is the transitory component of the conditional variance. [12]

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:

$$r_{i,t} = \sqrt{\tau_t \times g_{i,t}} z_{i,t}, \quad \text{with } i = 1, \cdots, N_t$$

Where:

$$r_{i,t}: represents\ the\ \log return\ for\ day\ i\ of\ the\ period\ t$$
$$N_t: is\ the\ number\ of\ days\ for\ period\ t$$

$$z_{i,t} |\ \mathcal{F}_{i,t} \sim N\ (0,1)\ where\ \mathcal{F}_{i,t}\ denotes\ the\ information\ set\ up\ to\ day\ i-1\ of\ period\ t$$
$$g_{i,t}\ follows\ a\ unit\ mean\ reverting\ GARCH\ (1,1)\ process\ (short\ run\ component)$$
$$\tau_t: provides\ the\ slow\ moving\ average\ level\ of\ volatility\ (long\ run\ component)$$

The short run component is given by:

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

While the long run component is defined as:

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

Where:
m is the intercept
θ: the coefficient to estimate
$\delta_k\ (w)$: function that weighs past K realizations of X_t

The Beta function is defined by:

$$\delta_k(\omega) = \frac{(k/K)^{\omega_1 - 1}(1 - k/K)^{\omega_2 - 1}}{\sum_{j=1}^{K}(j/K)^{\omega_1 - 1}(1 - j/K)^{\omega_2 - 1}}$$

If $\omega_1 < \omega_2$ recent observations weigh more in the long-term component. We assume X_t is independent of $z_{i,t}$ 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:

$$g_{i,t} = (1 - \alpha - \beta - \gamma/2) + \left(\alpha + \gamma \cdot \mathbb{1}_{(r_{i-1,t}<0)}\right) \frac{(r_{i-1,t})^2}{\tau_t} + \beta g_{i-1,t}$$

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

Long-run component is defined as:

$$\tau_t = exp\left(m + \theta^+ \sum_{k=1}^{K} \delta_k(\omega)^+ X_{t-k} \mathbb{1}_{(X_{t-k} \geq 0)} + \theta^- \sum_{k=1}^{K} \delta_k(\omega)^- X_{t-k} \mathbb{1}_{(X_{t-k}<0)}\right)$$

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 $VaR_{\tau,t}$ 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:

$$\widehat{VaR^w}_{\tau,t} = \hat{Q}_{r_{\tau,t}}(\tau)$$

Where $r_{\tau,t}$ represents the daily returns that are included in the length for the sample used in the calculations. And $\hat{Q}_{r_{\tau,t}}(\tau)$ is the sample quantile at τ 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 i.i.d.) 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 $VaR_{\tau,t}$ estimation as we directly estimate it by using the τ-th sample quantile ($\hat{Q}_{r_{\tau,t}}(\tau)$).

The Quantile Linear ARCH model (Koenker and Zhao-1996) uses this logic of approach, where the $VaR_{\tau,t}$ is directly calculated and is defined as:

$$VaR_{\tau,t} \overset{pos.hom.}{=} (\alpha_0 + \alpha_1|r_{t-1}| + \cdots + \alpha_q|r_{t-q}|)Q_{z_t}(\tau)$$

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 $VaR_{\tau,t}$ is estimated using previous estimations of it ($VaR_{\tau,t-i}$) 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:

$$VaR_{\tau,t}(\tau) = \beta_0 + \beta_1 VaR_{\tau,t-1}(\tau) + \beta_2 \mid r_{t-1} \mid$$

The VaR at level τ 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 $VaR_{\tau,t-1}$ instead. It follows the next specification:

$$VaR_{\tau,t}(\tau) = -\sqrt{\beta_0 + \beta_1 VaR^2_{\tau,t-1}(\tau) + \beta_2 r^2_{t-1}}$$

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:

$$VaR_{\tau,t}(\tau) = \beta_0 + \beta_1 VaR_{\tau,t-1}(\tau) + (\beta_2 I(r_{t-1} > 0) + \beta_3 I(r_{t-1} < 0)) \mid r_{t-1} \mid$$

This model allows us to give different relevance to past returns depending on whether they took positive or negative values.
Parameters from CAViaR models are estimated by minimizing the following quantile loss function:

$$\hat{\boldsymbol{\beta}}(\tau) = \arg\min_{\boldsymbol{\beta}} \frac{1}{T}\sum_{t=1}^{T} \rho_\tau(r_t - VaR_{\tau,t}(\boldsymbol{\beta}))$$

3.1) Backtesting and Model Selection

Value at Risk has become one of the most popular risk measurement techniques. Hence, building models able to predict VaR accurately is importance, since they are useful only if they predict future risks accurately. For this reason, it is quite relevant to evaluate the quality of the VaR estimates by performing a set of targeted tests. Nowadays, in this context, Backtesting is the most used test procedure, which is widely adopted to verify the precision of the VaR prediction. The Backtesting technique relies on quantitative tests which scrutinize the model performances in terms of accuracy and efficiency with respect to a defined criterion.

We evaluate the predicting ability of extreme negative returns of each model using:

- the *Proportion Of Failure test* (LRuc) test of Kupiec (1995), which inspects if the theoretical VaR violations are equal to the estimated ones;
- the *Conditional Coverage test* (LRcc) test of Christoffersen (1998), 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 $(t - 1)$);
- the *Dynamic Quantile* (DQ) test of Engle and Manganelli (2004), 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. [13] In this work, the DQ test uses lagged violations at lag $q = 4$;
- and the *AE ratio*, which tracks the actual number of times that the returns have exceeded the estimated VaR over the expected VaR violations. For instance, if daily VaR. estimates are computed at t confidence level, one would expect a percentage of violations of $100(1 - t)\%$. In particular, the closer the ratio gets to one, the better the model estimates VaR. If the ratio is less than 1 the model overestimates the risk, while if it is greater than 1 the model underestimates the risk.

By performing the Backtesting procedure, however, it may not happen that a single model outperforms all the others. For this reason, in order to give a more in-depth analysis in terms of the "best fitting model", we propose the use of the Model Confidence Set procedure which will be explained in the following subsection. [14]

3.2) Model Confidence Set

In order to reduce the set of models considered in the analysis to a smaller set that contains the best model(s) with a given level of confidence, Hansen et al. (2011) introduced the Model Confidence Set (MCS) procedure.

The MCS procedure makes use of a loss function $\ell : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ mapping the distance between the observed returns, r_t, and the estimated VaR from model l. Let $\widehat{VaR^{\tau}}_{i,t}$ be the VaR estimate of the day t obtained from model l at level τ. The loss function associated to model l penalizing more heavily negative returns which overcome VaR is given by:

(QL) $$\ell(r_t, \widehat{VaR}_{l,t,\tau}) = (\tau - I(r_t < \widehat{VaR}_{l,t,\tau}))(r_t - \widehat{VaR}_{l,t,\tau}).$$

The relative performance of the models, say l and k, is evaluated by means of the loss differential $d_{lk,t}$, that is:

$$d_{lk,t} = \ell_{l,t} - \ell_{k,t}, \quad \forall l, k \in \mathcal{M}^0,\, l \neq k.$$

Let $\mu_{lk} \equiv E[d_{lk,t}]$ be the expectation of the loss differential.

We rank alternatives in terms of expected loss μ_{lk}, so that when $\mu_{lk} < 0$, model *l* is preferred (on average) to model *k.*

The Superior Set of Models is formally defined as follows:

$$\mathcal{M}^* = \{i \in \mathcal{M}^0 : \mu_{lk} \geq 0 \quad \text{for all} \quad j \in \mathcal{M}^0\}.$$

The objective of the MCS procedure is to determine M^*.

This is done by a sequence of Equal Predictive Ability (EPA) tests where models that are being found to be significantly inferior to other elements of M^0 are eliminated. The null hypotheses that are being tested take the form:

$$H_{0,\mathcal{M}} : \mu_{lk} = 0, \quad \forall\, l, k \in \mathcal{M},\, l \neq k, \text{ with } \mathcal{M} \subset \mathcal{M}^0.$$

Meanwhile, we denote the alternative hypothesis

$$H_{1,\mathcal{M}} : \mu_{lk} \neq 0 \text{ for some } l, k \in \mathcal{M}$$

The test used in the MCS procedure are based on the following t-statistics:

$$\mathcal{T}_{l\cdot} = \frac{\bar{d}_{l\cdot}}{\sqrt{\widehat{var}(\bar{d}_{l\cdot})}} \quad \text{for} \quad l \in \mathcal{M},$$

where $\overline{d_{l.}} = (m - 1)^{-1} \Sigma_{k\in M} \overline{d_{lk,t}}$ is the loss of model *l* relative to the averages losses across models in the set $\mathcal{M}$, $\overline{d_{lk}} = H^{-1} \Sigma_{t=1}^{H} d_{lk,t}$ measures the relative loss between models *l* and *k*, while $\widehat{var}(\overline{d_{l.}})$ is a bootstrapped estimate of $\widehat{var}(\overline{d_{l.}})$.

Intuitively, large values for $\mathcal{T}_{l.}$ provide evidence that model l has a bad performance relatively to that of the other models in $\mathcal{M}$, such that it should be eliminated from $\mathcal{M}$, according to a defined elimination rule.

The choice of the worst model to be eliminated is made using the following rule:

$$e_{\max,\mathcal{M}} = \arg\max_{l\in\mathcal{M}} \mathcal{T}_{l\cdot} = \arg\max_{l\in\mathcal{M}} \frac{\bar{d}_{l\cdot}}{\sqrt{\widehat{var}(\bar{d}_{l\cdot})}}$$

This rule above mentioned, removes the model that contributes most to the test statistic. This model has the largest standardized excess loss relative to the average across all models in M

In practice, since the asymptotic distribution of T_1· is nonstandard because it depends on nuisance parameters, the relevant distribution under the null hypothesis is estimated using a bootstrap approach.

3.3) A new scoring function for VaR and ES

The class of strictly consistent scoring functions introduced by Fissler and Ziegel (2016), for jointly evaluating VaR and ES forecasts, are of the following form:

$$
\begin{aligned}
S(VaR_{t,\tau}, ES_{t,\tau}, r_t) = {} & (I(r_t \leq VaR_{t,\tau}) - \tau) G_1(VaR_{t,\tau}) \\
& - I(r_t \leq VaR_{t,\tau}) G_1(r_t) \\
& + G_2(ES_{t,\tau})(ES_{t,\tau} - VaR_{t,\tau} + I(r_t \leq VaR_{t,\tau}) \\
& + (VaR_{t,\tau} - r_t)/\tau) - \xi_2(ES_{t,\tau}) + a(r_t),
\end{aligned}
$$

where G_1, G_2, ξ_2 and a are functions satisfying the properties that G_1 is increasing, $G_2 = \xi_2'$ and ξ_2 is increasing and convex. Fissler et al. (2015): $G_1(x) = x,\ G_2(x) = 1/(1+e^{-x})$ and $a = \ln(2)$

(FZ)
$$
f(r_t \mid VaR_{t,\tau}, \sigma_t, \tau) = \frac{\tau - 1}{ES_{t,\tau}} \exp\left(\frac{(r_t - VaR_{t,\tau})(\tau - I(r_t \leq VaR_{t,\tau})}{\tau ES_{t,\tau}}\right)
$$

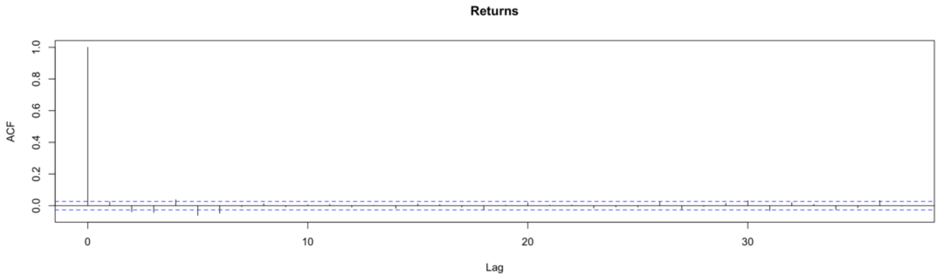
4) Empirical Analysis

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.

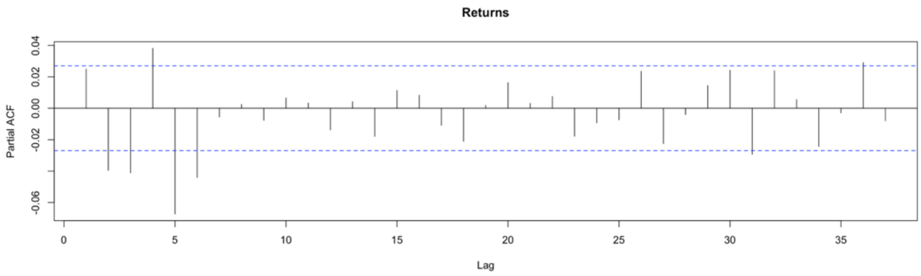
Auto-correlograms

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 ($\rho = 0$), 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).

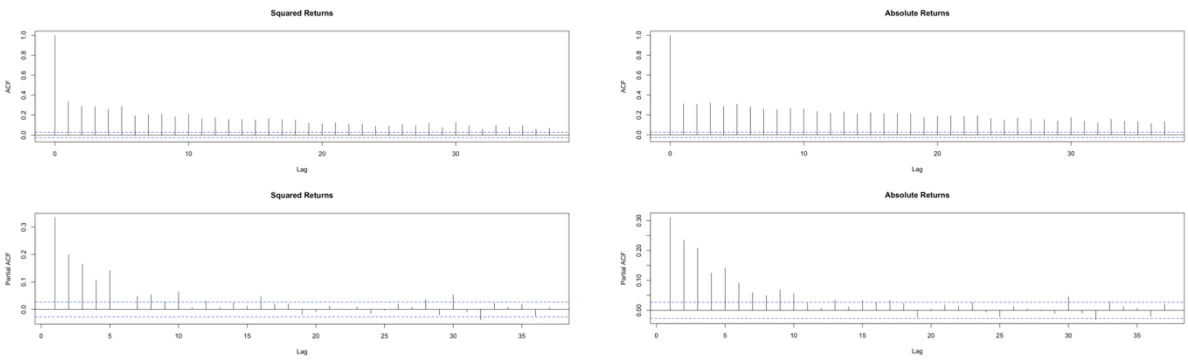


Later, we tested the partial autocorrelation function. In this test we see if the direct effect/influence of the lag return respect to daily returns, eliminating the influence it contains from other lags (we do not consider the dependency created by the lags between them).



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.

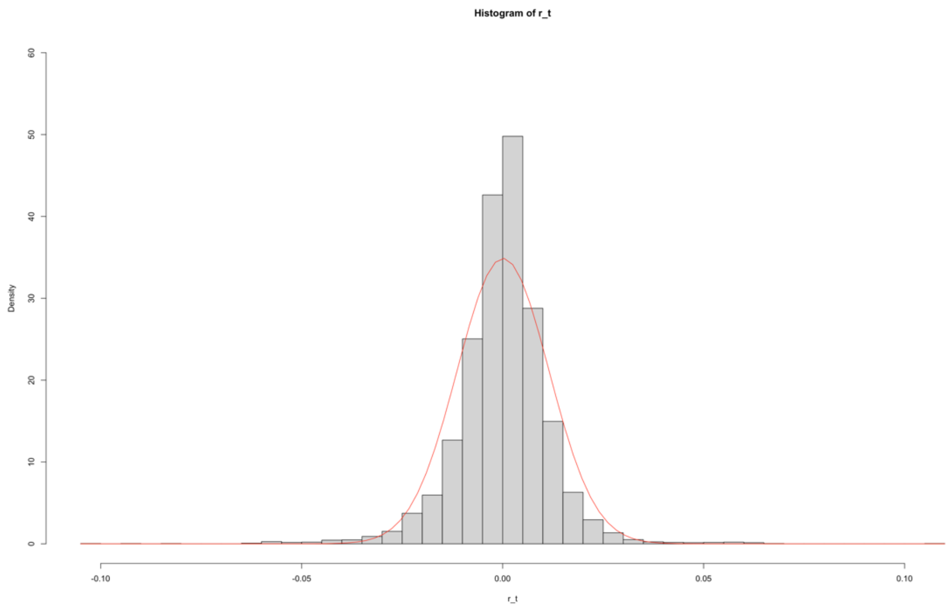


Dependency of daily returns with recent past observations seemed to be significantly different from zero, and then started falling into the nullity band. We can interpret that, shocks are not permanent, they are absorbed through time, and that they should not be associated with a white noise process 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.[15] 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.



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.

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

Stationary Test	Real GDP	Unemployment	Bond Yield 10 Years
ADF Statistic	-3.52	-3.40	-3.02
ADF p-value	0.04	0.05	0.14

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

Swiss Stock Market Index	
Statistical Tests	Results
Min.	-0.10133930
1st Qu.	-0.00532258
Median	0.00049272
Mean	0.00007331
3rd Qu.	0.00584886
Max.	0.10787597
Skewness	-0.2611203
Kurtosis	7.640598
Jarque Bera Test (p-value)	0.00000002
Ljung-Box Test Returns (p-value)	0.000000034
Box-Pierce Test Returns (p-value)	0.00000003991
Ljung-Box Test Squared Returns (p-value)	0.00000002
Box-Pierce Test Squared Returns (p-value)	0.00000002
Augmented Dickey-Fuller Test	0.01

4.1) 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.

We have a good model in terms of VaR predictability if it accepts the null hypothesis of the three aforementioned tests. The significance level used for the analysis are: 1 and 5 %.

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

	5%				1%			
	AE	UC	CC	DQ	AE	UC	CC	DQ
sGARCH N	1.144	0.019*	0.062	0.031*	1.742	0.000*	0.000*	0.000*
gjrGARCH N	1.129	0.035*	0.098	0.255	1.705	0.000*	0.000*	0.000*
eGARCH N	1.110	0.072	0.153	0.190	1.723	0.000*	0.000*	0.000*
csGARCH N	1.125	0.041*	0.087	0.046*	1.705	0.000*	0.000*	0.000*
sGARCH Std	1.193	0.002*	0.006*	0.001*	1.326	0.023	0.030	0.009*
gjrGARCH Std	1.193	0.002*	0.006*	0.020*	1.402	0.006*	0.008*	0.033
eGARCH Std	1.174	0.005*	0.014*	0.056	1.364	0.012	0.016	0.003*
csGARCH Std	1.189	0.002*	0.004*	0.002*	1.307	0.032	0.041	0.012
GM N - U	1.133	0.030*	0.092	0.027*	1.648	0.000*	0.000*	0.000*
GM-skew N - U	1.098	0.106	0.238	0.091	1.648	0.000*	0.000*	0.000*
GM Std - U	1.186	0.003*	0.010*	0.001*	1.856	0.000*	0.000*	0.000*
GM-skew Std - U	1.178	0.004*	0.011*	0.023*	1.402	0.006*	0.022	0.063
DAGM N - U	1.114	0.063	0.161	0.015*	1.742	0.000*	0.000*	0.000*
DAGM-skew N- U	1.110	0.072	0.166	0.144	1.686	0.000*	0.000*	0.000*
GM N - Real GDP	1.121	0.047*	0.130	0.044*	1.648	0.000*	0.000*	0.000*
GM-skew N - Real GDP	1.129	0.035*	0.098	0.067	1.648	0.000*	0.000*	0.000*
GM Std - Real GDP	1.186	0.003*	0.010*	0.001*	1.326	0.023	0.030	0.009*
GM-skew Std - Real GDP	1.178	0.004*	0.011*	0.030*	1.345	0.017	0.022	0.097
DAGM N - Real GDP	1.121	0.047	0.130	0.042	1.705	0.000*	0.000*	0.000*
DAGM N-skew - Real GDP	1.102	0.093	0.212	0.296	1.648	0.000*	0.000*	0.000*
GM N - Bond Yields	1.121	0.047*	0.130	0.031*	1.648	0.000*	0.000*	0.000*
GM-skew N - Bond Yields	1.091	0.135	0.296	0.297	1.686	0.000*	0.000*	0.000*
GM Std - Bond Yields	1.337	0.000*	0.000*	0.000*	1.705	0.000*	0.000*	0.000*
DAGM N - Bond Yields	1.152	0.014*	0.044*	0.009*	1.629	0.000*	0.000*	0.000*
DAGM-skew N - Bond Yields	1.098	0.016	0.238	0.245	1.723	0.000*	0.000*	0.000*
SAV	1.008	0.900	0.438	0.111	0.966	0.802	0.000*	0.000*
AS	1.008	0.900	0.594	0.540	1.023	0.869	0.570	0.952
IG	1.004	0.950	0.579	0.138	1.023	0.869	0.570	0.937
Lin. ARCH	1.004	0.950	0.579	0.585	1.023	0.869	0.299	0.201
HS (w=250)	1.117	0.054	0.000*	0.000*	1.610	0.000*	0.000*	0.000*

(*) denotes that the null hypothesis is rejected at the given significance level.

Note: The table reports the Backtesting results for the models in the first column. Column AE denotes the Actual over Expected exceedance ratio. Columns UC and CC report the p-values of the Unconditional and Conditional coverage tests. Column DQ represents the p-value of the Dynamic Quantile test. UC, CC and DQ tests at significance level $\alpha = 0.05$ and $\alpha = 0.01$.

At the given significance 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 α = 0.01 significance 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).

	5%			1%		
	Expected Exc.	Actual Exc.	P-value	Expected Exc.	Actual Exc.	P-value
sGARCH N	264.000	92.000	0.000	264.000	92.000	0.000
gjrGARCH N	264.000	90.000	0.000	264.000	90.000	0.000
eGARCH N	264.000	91.000	0.001	264.000	91.000	0.001
csGARCH N	264.000	90.000	0.000	264.000	90.000	0.000
sGARCH Std	264.000	70.000	0.095	264.000	70.000	0.095
gjrGARCH Std	264.000	74.000	0.119	264.000	74.000	0.119
eGARCH Std	264.000	72.000	0.110	264.000	72.000	0.110
csGARCH Std	264.000	69.000	0.104	264.000	69.000	0.104
GM N - U	264.000	87.000	0.000	264.000	87.000	0.000
GM-skew N - U	264.000	87.000	0.000	264.000	87.000	0.000
GM Std - U	264.000	98.000	0.028	264.000	98.000	0.028
GM-skew Std - U	264.000	74.000	0.195	264.000	74.000	0.195
DAGM N - U	264.000	92.000	0.000	264.000	92.000	0.000
DAGM-skew N- U	264.000	89.000	0.000	264.000	89.000	0.000
GM N - Real GDP	264.000	87.000	0.000	264.000	87.000	0.000
GM-skew N - Real GDP	264.000	87.000	0.000	264.000	87.000	0.000
GM Std - Real GDP	264.000	70.000	0.165	264.000	70.000	0.165
GM-skew Std - Real GDP	264.000	71.000	0.259	264.000	71.000	0.259
DAGM N - Real GDP	264.000	90.000	0.000	264.000	90.000	0.000
DAGM N-skew - Real GDP	264.000	87.000	0.000	264.000	87.000	0.000
GM N - Bond Yields	264.000	87.000	0.000	264.000	87.000	0.000
GM-skew N - Bond Yields	264.000	89.000	0.000	264.000	89.000	0.000
GM Std - Bond Yields	264.000	90.000	0.035	264.000	90.000	0.035
DAGM N - Bond Yields	264.000	86.000	0.000	264.000	86.000	0.000
DAGM-skew N - Bond Yields	264.000	91.000	0.000	264.000	91.000	0.000
SAV	264.000	51.000	0.449	264.000	51.000	0.449
AS	264.000	54.000	0.535	264.000	54.000	0.535
IG	264.000	54.000	0.478	264.000	54.000	0.478
Lin. ARCH	264.000	54.000	0.470	264.000	54.000	0.470
HS (w=250)	264.000	85.000	0.032	264.000	85.000	0.032

(*) denotes that the null hypothesis is rejected at the given significance level.

Note: The table reports the ES Backtesting results for the models in the first column - according to the McNeil and Frey (2000) test, whose null is: "excess conditional shortfall (excess of the actual series when VaR is violated), is i.i.d. and has zero mean". Column Expected Exc. denotes the expected ES exceedance. Column Actual Exc. reports the actual ES exceedance.

4.2 Model Confidence Set (MCS)

The successive step of our analysis is the Model Confidence Set (MCS) procedure, here we implement the MCS procedure at significance level $\alpha = 0.25$.

a) **For $\tau = 0.05$**

	QL
sGARCH N	0.122
gjrGARCH N	0.118*
eGARCH N	0.117*
csGARCH N	0.121
sGARCH Std	0.122
gjrGARCH Std	0.118*
eGARCH Std	0.117*
csGARCH Std	0.121
GM N - U	0.121
GM-skew N - U	0.117*
GM Std - U	0.124
GM-skew Std - U	0.117*
DAGM N - U	0.121
DAGM-skew N- U	0.117*
GM N - Real GDP	0.121
GM-skew N - Real GDP	0.117*
GM Std - Real GDP	0.122
GM-skew Std - Real GDP	0.117*
DAGM N - Real GDP	0.121
DAGM N-skew - Real GDP	0.117*
GM N - Bond Yields	0.121
GM-skew N - Bond Yields	0.117*
GM Std - Bond Yields	0.124
DAGM N - Bond Yields	0.121
DAGM-skew N - Bond Yields	0.117*
SAV	0.122
AS	0.117*
IG	0.121
Lin. ARCH	0.123
HS (w=250)	0.138

	FZ
sGARCH N	-3.786
gjrGARCH N	-3.821
eGARCH N	-3.827*
csGARCH N	-3.792
sGARCH Std	-3.788
gjrGARCH Std	-3.823
eGARCH Std	-3.828*
csGARCH Std	-3.795
GM N - U	-3.787
GM-skew N - U	-3.825*
GM Std - U	-3.738
GM-skew Std - U	-3.827*
DAGM N - U	-3.79
DAGM-skew N- U	-3.829*
GM N - Real GDP	-3.788
GM-skew N - Real GDP	-3.825*
GM Std - Real GDP	-3.79
GM-skew Std - Real GDP	-3.827*
DAGM N - Real GDP	-3.788
DAGM N-skew - Real GDP	-3.826*
GM N - Bond Yields	-3.787
GM-skew N - Bond Yields	-3.822
GM Std - Bond Yields	-3.411
DAGM N - Bond Yields	-3.788
DAGM-skew N - Bond Yields	-3.823*
SAV	-3.795
AS	-3.839*
IG	-3.798
Lin. ARCH	-3.782
HS (w=250)	-3.608

Note: The tables report the inclusion in the SSM according to the MCS procedure. Column QL represents the averages of the loss function - denoted in the page 10 by (QL). Column FZ represents the averages of the loss function - denoted in the page 11 by (FZ). (*) denote the inclusion in the SSM, at significance level α = 0.25.

b) For $\tau = 0.01$

	QL
sGARCH N	0.036
gjrGARCH N	0.035
eGARCH N	0.034
csGARCH N	0.036
sGARCH Std	0.036
gjrGARCH Std	0.034
eGARCH Std	0.034
csGARCH Std	0.035
GM N - U	0.036
GM-skew N - U	0.035
GM Std - U	0.037
GM-skew Std - U	0.034
DAGM N - U	0.036
DAGM-skew N- U	0.035
GM N - Real GDP	0.036
GM-skew N - Real GDP	0.035
GM Std - Real GDP	0.035
GM-skew Std - Real GDP	0.034
DAGM N - Real GDP	0.036
DAGM N-skew - Real GDP	0.035
GM N - Bond Yields	0.036
GM-skew N - Bond Yields	0.035
GM Std - Bond Yields	0.037
DAGM N - Bond Yields	0.036
DAGM-skew N - Bond Yields	0.035
SAV	0.035
AS	0.033*
IG	0.035
Lin. ARCH	0.035
HS (w=250)	0.045

	FZ
sGARCH N	-3.338
gjrGARCH N	-3.374
eGARCH N	-3.4
csGARCH N	-3.346
sGARCH Std	-3.402
gjrGARCH Std	-3.428
eGARCH Std	-3.446
csGARCH Std	-3.403
GM N - U	-3.344
GM-skew N - U	-3.382
GM Std - U	-3.273
GM-skew Std - U	-3.434
DAGM N - U	-3.35
DAGM-skew N- U	-3.389
GM N - Real GDP	-3.347
GM-skew N - Real GDP	-3.391
GM Std - Real GDP	-3.406
GM-skew Std - Real GDP	-3.442
DAGM N - Real GDP	-3.349
DAGM N-skew - Real GDP	-3.393
GM N - Bond Yields	-3.344
GM-skew N - Bond Yields	-3.378
GM Std - Bond Yields	-2.161
DAGM N - Bond Yields	-3.347
DAGM-skew N - Bond Yields	-3.379
SAV	-3.434
AS	-3.482*
IG	-3.424
Lin. ARCH	-3.441
HS (w=250)	-3.056

Note: The tables report the inclusion in the SSM according to the MCS procedure. Column QL represents the averages of the loss function - denoted in the page 10 by (QL). Column FZ represents the averages of the loss function - denoted in the page 11 by (FZ). (*) denote the inclusion in the SSM, at significance level $\alpha = 0.25$.

The semi-parametric approach contains the same best model both for $\tau = 0.01$ and 0.05, being this the Asymmetric Slope (AS) model.

Despite the DAGM-skew N covers the second position at $\tau = 0.05$, but it is discarded at $\tau = 0.01$. This behavior can be associated to the fact that, being these quantile models, fat tails can strongly affect the results.

The parametric approach does not work as well as the semi-parametric approach.

Even thought, GARCH-MIDAS and DAGM models are spread on all the ranking within the first six positions. Historical Simulations occupy the bottom of the ranking, being the first model to be eliminated for both $\tau = 0.01$ and 0.05.

Model Confidence Set procedure confirms what seen in the Backtesting, with some exceptions: quantile models seem to have best performances in both confidence levels. Non-parametric models remain the worse ones.

Here we have the plots of the AS model which is the best model in both of our confidence levels.

5) 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.

References:

[1] Tim Bollerslev, Robert F. Engle, and Daniel B Nelson. “ARCH models”. In: *Handbook of econometrics 4* (1994), pp. 2959–3038.

[2] Wikipedia contributors. Swiss Market Index. *Wikipedia, The Free Encyclopedia*. (2022).URL: https://en.wikipedia.org/w/index.php?title=Swiss_Market_Index&oldid=1083878004

[3] Akhilesh Ganti. Real Gross Domestic Product. *Investopedia.* (2022). URL: https://www.investopedia.com/terms/r/realgdp.asp

[4] *Top PDF Unemployment and Economic Growth in Ethiopia*, URL: https://1library.net/title/unemployment-and-economic-growth-in-ethiopia.20

[5] *What is a Bond Yield - by Evan Green - eNtEr THe oKeEdoE*, URL: https://entertheokeedoe.substack.com/p/what-is-a-bond-yield.20

[6] *Mexican Stock Market Index Volatility - jfbmnet.com*, URL: http://jfbmnet.com/journals/jfbm/Vol_2_No_3_4_December_2014/1.pdf.20

[7] *Distinguish the indistinguishable: a Deep Reinforcement Learning ...*, URL: https://hal.archives-ouvertes.fr/hal-03202431/document.20

[8] *Abstract - Preprints*, URL: https://www.preprints.org/manuscript/201811.0552/v1/download.20

[9] *V-Lab: Exponential GARCH Volatility Documentation*, URL: https://vlab.stern.nyu.edu/docs/volatility/EGARCH.20

[10] *Read Introduction_to_the_rugarch_package.pdf*, URL: http://www.readbag.com/cran-r-project-web-packages-rugarch-vignettes-introduction-to-the-rugarch-package.20

[11] *The Predictive Performance of Extreme Value Analysis Based-Models in ...*, URL: https://www.scirp.org/journal/paperinformation.aspx?paperid=111091.20

[12] *The full fgarch model of hentschel 1995 when δ λ*, URL: https://www.coursehero.com/file/p1l60t2/The-Full-fGARCH-model-of-Hentschel-1995-when-%CE%B4-%CE%BB-submodel-ALLGARCH-The/.20

[13] *(PDF) Using mixed-frequency and realized measures in ... - ResearchGate*, URL: https://www.researchgate.net/publication/345215064_Using_mixed-frequency_and_realized_measures_in_quantile_regression.20

[14] *Selection of Value at Risk Models for Energy Commodities*, URL: https://www.sciencedirect.com/science/article/pii/S0140988318302548.20

[15] *Stress Intensity Factor Prediction on Offshore Pipelines using ...*, URL: https://www.sciencedirect.com/science/article/pii/S2214509522001772.20