1. **Introduction**

The objective of our analysis is to investigate the Swiss Market Index behavior in the last 20 years in relationship with 3 macro-variables (real GDP, unemployment, and 10-years bond yields). We tested it with the univariates models 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.

Doing our analysis, we decided to expand our evaluation and we did a little refer in the effects of the 2007-2008 financial global crisis and in the COVID-19 crisis.

1. **Data description**

To developed our analysis, we considered a time period of 8 years, from 01/01/2000 to 31/12/2020, having 5280 daily observations.

Our analysis is about the relation of the Swiss Market Index (SMI) 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. This index is updated no more than once per second.

Macro-variables:

The **Gross Domestic Product (GDP)** is the most common tool to evaluate the development and economy of a country, evaluating the productivity of the residents of a country in a time period. We can see that the Swiss Economy is one of the most advanced and high-developed free-market economics.

We were interested in the relationship between the most relevant Index of the Swiss stock market and the economy of Switzerland to understand if the GDP have some effect in the Stock Market Index of that important and developed country.

We were also interested in the **Real Gross Domestic Production (RGDP)** that evaluate the Nominal GDP considering the inflation of the country. We added this variable to our analysis to have an extended view of the effect of this tool.

**The unemployment rate** is probably the best-known labour market measure. The unemployment rate is a useful measure of the underutilization of the labour supply. It is thus seen as an indicator of the efficiency and effectiveness of an economy to absorb its labour force and of the performance of the labour market.

[**Bond yield**](https://www.investopedia.com/articles/investing/022516/understanding-different-types-bond-yields.asp) is the [return](https://www.investopedia.com/terms/r/return.asp) an investor realizes on a [bond](https://www.investopedia.com/terms/b/bond.asp). The bond yield can be defined in different ways. Setting the bond yield equal to its [coupon rate](https://www.investopedia.com/terms/c/coupon-rate.asp) is the simplest definition. The [current yield](https://www.investopedia.com/terms/c/currentyield.asp) is a function of the bond's price and its coupon or [interest](https://www.investopedia.com/terms/i/interest.asp) payment, which will be more accurate than the coupon yield if the price of the bond is different than its face value.

2007-2008 Financial crisis (global financial crisis):

Inside of the period of consideration of our dataset we can observe the impact of the 2007-2008 financial crisis, we were interested to do a focus in that financial shock.

**Great Recession**, economic [recession](https://www.britannica.com/topic/recession) that was [precipitated](https://www.britannica.com/dictionary/precipitated) in the [United States](https://www.britannica.com/place/United-States) by the [financial crisis of 2007–08](https://www.britannica.com/event/financial-crisis-of-2007-2008) and quickly spread to other countries. Beginning in late 2007 and lasting until mid-2009, it was the longest and deepest economic downturn in many countries.

The financial crisis, a severe contraction of [liquidity](https://www.britannica.com/topic/liquid-asset) in global financial markets, began in 2007 as a result of the bursting of the U.S. housing bubble.

To evaluate the shock of the financial market on the Swiss Market Index we took the information of the returns from 03-01-2006 up to 30-12-2011 to see the before and the after of the crisis.

Text

Description automatically generated with medium confidenceThe table below show us the descriptive statistics of the dataset.

Chart, line chart

Description automatically generatedChart, line chart

Description automatically generated

We can observe that the SMI was increasing before the crisis, that begin at the middle of 2007, and was followed from a long-term decreasing until the beginning of 2009 that arrive to the lowest point of the graph and then start to increase again.

**Covid-19 crisis:**

We can also observe in our dataset a second big world crisis that come because of the covid-19 economic effects, to do this analysis we took into consideration a period of our dataset from 03-01-2019 up to 30-12-2020, again to understand how was going the market before of the shock.

The 2020 COVID-19 crisis start in March 2020 when a big part of the world governments announces the quarantine because of the pandemic, this created with the fear and the confusion a big financial crisis putting in shock a big portion of the financial markets in the world, followed by an inactive period of production, supply and provisions.

In the table below we can find the descriptive statistics.

Text

Description automatically generated with medium confidence

Chart, bar chart, line chart

Description automatically generatedChart, bar chart

Description automatically generated

In the upstanding plots we can see the effect of the crisis in the SMI, that was having a big increase 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. **Data analysis**

First, we evaluate the autocorrelogram function (ACF). In this test, we want to see whether previous observations are correlated with today’s observation. 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 lagged return is significant to explain today’s returns and we should take it into account in our model.

In this test, all those lags that exceed the nullity band mean that, according to our sample, there is evidence suggesting the correlation is significant.

Afterwards, we test the existence of partial autocorrelogram. In this test we see the direct effect of the lag return respect today’s return, eliminating the influence it contains from other lags (we do not consider the dependency created by the lags between them).

Chart

Description automatically generated with medium confidence

according to the obtained results there is no evidence suggesting dependence between rt and rt-j. In consequence, we could assume that the expected value of the returns is equal to zero.

However, results vary a lot if we transform the variable under analysis to squared returns. If we follow the same procedure than before, we see that dependency with recent lags are significantly different from zero, and then fall inside the nullity band.

A picture containing text, device, caliper

Description automatically generated

This result suggest that shocks are not permanent, that are absorbed through time, and that we can model as there is no white noise process.

A picture containing text, caliper, device

Description automatically generated

Later, we evaluate if the time series could be said to follow a normal distribution. For this, we must calculate its kurtosis and skewness to be able to carry on with the Jarque-Bera test.

We calculate the kurtosis, which is 7.640, much larger than 3 (kurtosis for a normal distribution).

The interpretation of this result of fat tails is that there is a bigger probability of observing large returns and losses.

Then, we calculated the skewness from our time series. We got that it is equal to -0.261 (SIDE skewness), which means that number of negative returns are higher than the number of positive returns.

After obtaining these values, we were able to test the Jarque-Bera to evaluate if the time series is normally distributed. The null hypothesis is that the series is normally distributed, and the test statistic is:

A picture containing diagram

Description automatically generated

Basically, the test is based in the difference between the sample skewness and kurtosis and the values of a normal distribution. In our case, the test statistic took the value of 12,917 (extremely large) with a p-value below ???? meaning we should reject/accept H0, the evidence suggests the time series does not follow a normal distribution.

Chart, histogram

Description automatically generated

Moreover, we tested if the time series could be associated with a white noise.

One possible test is the Box-Pierce Test, where the null hypothesis is that the autocorrelation of the variable against the same lagged variable is equal to zero, meaning there is no relation between past returns. So, if we don’t reject H0 is because evidence suggests the series is a white noise process.

In our SMI time series (for r\_t) the Box-Pierce test statistic took a value of 123.28 with its corresponding p-value of 0.000… that led us to reject our null hypothesis (there is at least one significant lagged variable) and the series is not a white noise.

Using lag=50 in our SMI time series for r\_t, we reject H0 as took a value of 0.000… (Box-Pierce=123,28), but we reject when we test for r\_t^2 (p-value=0.00…). These results are consistent with what we observed from the ACF and PACF, where the time series of r\_t could be associated with a white noise process while it could be said the same for the squared returns. If we reject the null hypothesis, the evidence is suggesting there is at least one significant lagged variable (there is correlation), so we can model. But if we do not reject, we are saying that the correlation coefficients are not statistically different from zero, so the series is a white noise.

We also test this hypothesis with the Ljung-Box test. The results and conclusions are the same that with the Box-Pierce test: accept the null hypothesis for the case of squared returns (there is a correlation that can be modeled) and we do not reject in the case of the absolute returns.

Also consistent with the interpretation of ACF and PACF graphs.

Finally, regarding the tests that could be done for checking the stylized facts of the 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 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.

1. **Models and Risk measures**

The goal of this study 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 95 and 99 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.

**Parametric Approach**

**GARCH MODELS**

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.

**GARCH-MIDAS and DAGM models**

As for the GARCH models, returns 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 a 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) variable on the conditional volatility.

**Semi-parametric Approach**

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 returns 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 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.

The main difference between the two is that the latter consider the possibility of asymmetric effect on the quantile of the distribution depending on the sign of the past value of the returns, meanwhile the first take the absolute value of the past return.

**Non-parametric Approach**

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.

The problem is that the sample quantile is a consistent and asymptotically normal estimator under the i.i.d. assumption of the returns, which is a strong assumption. For an accurate estimation of the VaR a long window is needed. On the other hand, a long estimation window increases the probability of structural breaks

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, semi-parametric and non parametric.

Referring to the parametric one, we consider the following GARCH (1,1) 8 class of models with two different distributions (N for normal and Std for Student-t) for the innovation term: S-GARCH-N, S-GARCH- STD, GJR-GARCH-N, GJR-GARCH-STD, E-GARCH-N, E-GARCH-STD, CS-GARCH-N, CS-GARCH-STD.

To model the variance involving also the MVs as additional volatility determinant, we use (for the 3 different MVs) GM-N-noSkew-MVs, GM-N-Skew-MVs, GM-T-noSkew-MVs, GM-T- Skew-MVs, DAGM-N-noSkew-MVs, DAGM-N-Skew-MVs.

Referring to the semi-parametric approach, we use CAViaR-IG, CAViaR-SAV, CAViaR-AS, L- ARCH models.

The non-parametric approach includes Historical Simulation (HS) with window size: w=250.

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 1, 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 theta, 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 (model the volatility), 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. (Chiedere a qualcuno il perchè)

With regards to the semi-parametric approach, we can observe that all the models perform well for the considered series. These results are actually expected, since we do not assume any kind of distribution for both CAViaR and L-ARCH/L-ARCH-MIDAS models.

Table

Description automatically generatedAn 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).

From these results it is evident how quantile models produce better forecasts. This may be due to several reasons: firstly, the quantile regression approach is free of any restrictive hypothesis on the error term distribution. Secondly, this procedure is a more robust tool with respect to the tails and outliers of the data. In the end, CAViaR and L-ARCH directly model the quantiles of the return distributions that, in a quantile estimation problem, may be more reasonable.

Table

Description automatically generated

We also consider the Expected Shortfall, where we control if the ES is always less than VaR at each time t, we verify if the martingale difference property is satisfied or not. As observed in ES columns, this happens for all series and for each model in both confidence levels, since the values are approximately equal to zero. This suggest that the distribution of ES violations has zero mean. It is worth noting that, in parametric approach, when we assume a Student-t distribution for the innovation term, the ES violations result to be less than those observed for the normal one. It can be due to the presence of fat tails typical of the Student-t, suggesting that extreme events are more likely than those expected assuming a Gaussian distribution.

1. **MODEL CONFIDENCE SET (MCS)**

The semi-parametric approach contains the same best model both for 95% and 99%: being this the Asymmetric Slope (AS) model.

Despite the DAGM-*skew* N covers the second position at 95% of confidence level, it is discarded at 99%. This behavior can be associated to the fact that, being these quantile models, fat tails can strongly affect the results.

As seen before in the Table 3, 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 each confidence interval.

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.

the first table is the MCS for the VaR at 95%, and it reports the inclusion in the superior set of models (SSM) according to the MCS procedure. Column QL represent the vaerages of the loss function.

The second one represents the VaR and ES at the 95%. Column FL represents the averages of the loss function.

the first table is the MCS for the VaR at 99%, and it reports the inclusion in the superior set of models (SSM) according to the MCS procedure. Column QL represent the vaerages of the loss function.

The second one represents the VaR and ES at the 99%. Column FL represents the averages of the loss function.



****

1. **Conclusion**

For our analysis, as you can see, we used the most popular risk measures that are the VaR and ES. For the VaR we used different kind of models; for the backtesting we selected a set of the best models, at 95% level we can concluded that come from the semi-parametric group and the same count for the 99% level, both of them give us “AS model” as the best model considering our 3 macrovariables.