THE ESTIMATE AND EVALUATION OF THE FORECASTS IN THE GARCH MODELS: APPLICATION TO STOCK INDEX

Relatore: Ch.mo Prof. Alessandra Amendola
Candidato: Maddalena Masi
matr. 0222200910

Anno Accademico 2018/2019
a Babbo,
che tra urla e silenzi resterà per sempre il mio primo amore
INDEX

INTRODUCTION 1

CHAPTER 1: ARCH & GARCH MODELS

1. Volatility 5
   1.1 Introduction 5
   1.2 Empirical regularities of time series of returns 7
   1.3 Volatility analysis 10
   1.4 Stylized facts of volatility 11
   1.5 Proxy of volatility 13
      1.5.1 Exploratory techniques 14
      1.5.2 Comparison of exploratory techniques 16

2. ARCH Model 17
   2.1 Specification of an ARCH model 19
   2.2 ARMA representation of an ARCH model 20
   2.3 Test for the presence of heteroskedasticity 22
   2.4 Limitations of an ARCH model 24

3. GARCH Model 25
   3.1 Specification of a GARCH model 25
   3.2 ARMA representation of a GARCH model 27
   3.3 Estimate of the GARCH(p,q) model 28
   3.4 GARCH-in-mean Model 31
   3.5 Standardized residue tests 32
   3.6 Limitations of a GARCH model 33


1. Introduction 35
   1.1 Leverage effect 35
   1.2 Verification of asymmetry 37
      1.2.1 Sign Bias Test 37
      1.2.2 Negative Size Bias Test 38
      1.2.3 Positive Size Bias Test 38
      1.2.4 Joint test 39
CHAPTER 3: EMPIRICAL ANALYSIS

1. Introduction
   1.1 Preliminary analysis of prices
   1.2 Preliminary statistical analysis of returns
   1.3 Analysis of the conditional expected value: the ARMA models
      1.3.1 Analysis of the residues of the GARCH(1,1) model

2. Analysis of volatility forecasts
   2.1 The volatility estimates of the S&P500 index with the GARCH(1,1) model
   2.2 The volatility estimates of the S&P500 index with the TGARCH model

3. Analysis of the volatility expectations of the NASDAQ stock index
   3.1 Forecasts for the volatility of the NASDAQ index with the EGARCH model

CONCLUSIONS

BIBLIOGRAPHY
The focus of the analysis in this thesis paper is the study of volatility forecasting. The latter is intended as a measure of the intensity of the variations suffered by the price of a financial asset in a given period. In other words, volatility is a financial index able to measure the aptitude of returns to take on extreme values; or more precisely, defines the lesser or greater probability that this price varies as the required return varies too.

Since, in a financial series, volatility is a fundamental parameter in the determination of the optimal portfolio, its forecast implies a great interest from part of the practitioners to determine the investment strategies.

In the first chapter, starting from the returns calculated as raw logarithmic differences of prices, the focus was on empirical regularities, or on those recurrent behaviors that are found in the historical series. Graphically it is possible to see that they have been observed the main stylized facts of returns (average level close to zero, weak or negative asymmetry, marked kurtosis and lack of correlation).

The uncertainty of returns implies the study of their volatility. Once defined what is meant by conditional volatility (non-constant volatility compared to time parameter), attention was focused on the stylized facts of volatility (volatility) clustering, thick tails and leverage effect). Subsequently, because the volatility is a latent variable and therefore not directly observable, have been defined proxy of volatility. In fact, to estimate it, we use proxies like: daily returns squared, daily returns in absolute value and realized volatility.

Finally, the conditional heteroscedasticity models are introduced: ARCH and GARCH.

The ARCH model, introduced by Engle in 1982, has as its main objective that of provide an estimate of the volatility as a standard deviation, given by the square root of the variance. The standard deviation is used in financial decisions related to the analysis of risks. The use of this model occurs if there is the presence of autocorrelation of positive transformations (square or absolute value) of returns.

Defined the structure of the ARCH model and observed some tests for the presence of heteroskedasticity (the variance between the innumerable observations that make up the sample is not constant), the limits of this model have been presented. The main limitation of the ARCH model is given by the too high number of yields necessary to adapt it to the
observed data. To overcome this limit, Bollerslev (1986) proposed to use a more general class of models, called GARCH, Generalized Autoregressive Conditional Heteroskedasticity. Main objective of such generalization is to reduce the parameters to be estimated with respect to the structure of the ARCH model; furthermore, the main idea of this model is to reproduce the model's parsimony ARMA in terms of parameters used, introducing delayed values in the ARCH model of conditional variance.

GARCH models consider relevant information on past values of variance conditioning; the model is such that the conditional variance at time \( t \) is expressed through a linear combination of \( q \) delays of itself, synthesis of information past, and of \( p \) squared residual delays, representative of market news and ability to vary conditional variance over time.

In concluding the first chapter the limitations of the GARCH model have been specified. The main limitation of this model is the assumption that the positive error terms and negatives have a symmetrical effect on volatility; in other words, defining the model GARCH as symmetrical, it does not consider an empirical evidence of historical series already noted in the 1970s and that is that negative news have a greater impact on the volatility than positive news.

In the second chapter, once defined the leverage effect (inverse relationship between shock of price and volatility) and once the asymmetry verification tests were introduced, they were identified the variants of the GARCH model that foresee the presence of asymmetry.

The first model of the GARCH family that foresees the presence of asymmetry and the model IGARCH (integrated GARCH) in which the volatility of a period affects all forecasts relating to any subsequent period. In this model, a given shock on variance conditioning is persistent for any future time horizon, becoming one relevant component in the long term.

Subsequently we observed the EGARCH model, exponential GARCH, proposed by Nelson in 1991. The model considers not only the asymmetric effect of the innovations, but also the proportionality of the effect itself with respect to the intensity of the innovations that determine it.

Like the EGARCH model, the Threshold model GARCH, TGARCH, has as main objective to capture the different effects asymmetric that positive and negative innovations have on the conditional volatility of returns. The aim of the TGARCH model is to capture the different behavior at the crossing by delayed innovation of a specific threshold value, which
is generally set equal to zero. As a result, the positive innovations have less impact than negative peer innovations intensity.

In formulating the TGARCH (1,1) model, note the conditional standard deviation present the same functional form as the GJR model (1,1). The GJR model proposed by Glosten, Jagannathan and Runkle in 1993 represents an extension of the GARCH in model which includes a term whose main function is to capture asymmetric evolution of conditional variance.

Finally, we have observed the model AGARCH, Absolute GARCH, in which the volatility comes identified based on $\sigma^{1/2}$ rather than the standard deviation itself.

The models for the study of conditional heteroskedasticity, belonging to the family GARCH, are common tools in many economic and financial applications. In assessing the leverage effect allowed by these models, a useful tool is constituted from the news impact curve (NIC).

The News Impact Curve characterizes the impact of past yield shocks on implied volatility in a volatility model. It reflects, in a GARCH context, the asymmetry and leverage effects of volatility, it also measures how information is incorporated into volatility estimates. Once this impact curve has been formulated for all the models belonging to the family GARCH, attention was focused on the forecast of volatility.

The wording of the prediction is based on the knowledge and re-elaboration of information processed by past manifestations of the historical series under examination, it is an operation of "Link between past and future".

It is possible to classify forecasts by: nature, object or their temporal horizon.

Attention was drawn to the distinction between ex post forecasts and ex-ante forecasts. Ex-post forecasts concern the reconstruction of the historical series of interest on the past trend; while the ex-ante forecasts take into consideration the period that begins later the last known instant and is prolonged indefinitely in the future.

The various forecasting schemes are then defined (fixed scheme, recursive scheme and rolling scheme) have been analytically formulated, for the models belonging to the family GARCH, forecasts of volatility and forecast errors have been defined. In fact, in forecasting future values of volatility it is possible to incur errors defined as “errors of forecast”.

3
The latter take on a central role and represent the difference between the series of real and expected values.

Conditional heteroscedasticity models, hence the GARCH model family, it applied to series with non-normal data and negligible autocorrelation; otherwise, then if the series shows a non-negligible linear dynamic structure, it is possible to build first an ARMA model and then adapt a GARCH model on the residues.

Given the ARMA models for the stock indexes of reference, they have been estimated for both GARCH models, considered the starting points in the simulation of the forecasts of the volatility of the historical series of interest.

Subsequently, considering the same information set of the stock index Standard & Poor 500 has been possible, to make a comparison, to simulate data forecasts with asymmetric variants of the GARCH model. It was taken the Threshold-GARCH model is being examined. While, for the NASDAQ stock index the comparison occurred with another variant of the GARCH model, i.e. it is not considered plus the TGARCH model but the Exponential-GARCH model was considered.
CHAPTER 1: ARCH & GARCH MODEL

1. Volatility

The focus of the analysis of the financial series is the study of return, as a measure of profitability of a financial asset.

The uncertainty of returns implies the study of the volatility of the same returns. The latter is a risk measure that specifies the degree of uncertainty associated with a determined phenomenon in the financial sphere. In other words, from an economic point of view volatility is a statistical measure of the dispersion of the returns for a given market index.

In finance, volatility measure the rate of change in price over a given period. Expressed often as a percentage, it is computed as the annualized standard deviation of the percentage change in the daily price.

1.1 Introduction

One of the main statistical tools which can describe the evolution of the historical series of interest, is the price.

Prices have a slow and persistent trend: to what the price is, in a given day, tends to follow a very similar price the next day.

There is a large part of the literature concerning the adoption, or better to say the trasformation of past series to determine whether the evolution of prices is increasing or decreasing to identify the moment in which to invest and obtain profits, not certain but with a high probability of certainty. However, the future evolution of the performance of the prices is uncertain, in fact, there is no model to forecast them.

The price of an asset at time $t$ is linked to the price of the asset thereof at time $t-1$ through the yield passed between $t-1$:

$$P_t = P_{t-1}(1 + r_t)$$

Which in logarithmic terms can be expressed as:

$$p_t = p_{t-1} + \log(1 + r_t)$$
For $r_t$ (simple net return) sufficiently small, which is certainly due to daily price changes, through the Taylor expansion of the term $\log(1 + r_t)$ around 1, the previous equation it becomes:

$$p_t \approx p_{t-1} + r_t$$

And from which it turns out that:

$$r_t \approx \Delta p_t \equiv p_t - p_{t-1}$$

The returns, calculated as the first logarithmic differences in the prices of a financial asset, define a percentage change in prices. One of the advantages in the adoption of percentage variations, in their approximation of differences logarithmic, is identified by the property of additivity. According to this property, returns can be expressed as the sum of returns identified in intervals of smaller time.

The time series of returns be a finite realization of a variable random\(^1\), therefore as a representation of uncertainty. It is essential to put in relate the elements of the temporal sequence to these random variables.

If the returns, considered as random variables, are independent of each other, the passage of time provides no element of interest and the likelihood of two joint events it will be equal to the product of the single probabilities.

$$\Pr(AB) = \Pr(A) \Pr(B)$$

If, on the other hand, they are linked together, the temporal sequence can be exploited through one conditioned distribution, that is through information on all that is relevant and irrelevant to the entire distribution.

$$\Pr(AB) = \Pr(A|B) \Pr(B)$$

In this sense, we can take advantage of the conditional probability, $\Pr(A \mid B)$, where what comes in aid is the sequence of information on what happens; in other words, the prediction of future returns is based on information derived from past returns.

\(^1\)A random variable is a rule that associates with an elementary event, one and only one real number. Often a random variable is nothing but the immediate numerical translation of elementary events, but it can be also a more complex function defined on the events of the set. Think for example of the “launch of a dice “where, by convention, a number is attributed to each face of the dice.
Based on the econometric approach of San Diego School (Engle, Granger\(^2\)) the point of departure of the analysis of a phenomenon is represented by empirical regularities, that is represented by the recurring characteristics underlying the phenomenon itself.

### 1.2 Empirical regularities of the time series of returns

To analyze a given financial phenomenon, one must start from the stylized facts, that is, from the empirical regularities found in the historical series with the help of elementary statistical tools applied to data over a sample period. The instruments Statistical referenced are:

- Descriptive statistics (average, standard deviation, asymmetry and kurtosis)
- Charts
- Correlogram (correlation between values distant from each other k periods, k = 1, ..., K)

Empirical regularities of the time series of returns are:

- Medium level close to zero;
- Asymmetry is weak and often negative;
- Marker kurtosis;
- Absence of correlation

Referring to the Standard & Poor's 500\(^3\) index as it is the most well-known of the indices shares of the New York Stock Exchange and its variations are a main parameter for to judge the trend of a stock exchange day, we note that the average is close to zero (picture 1).

![Figure 1: Standard & Poor's 500 index](image)
The S&P 500 index is calculated based on the shares listed on the New York Stock Exchange (NYSE) and the American Stock Exchange (ASE), of 500 titles of the most important American companies with high capitalization.

Around zero, yields fluctuate both in positive and in negative, alternating a lot frequent with respect to each other.

There is no trend that can be defined, in fact, without considering the reticence of the markets, the information available is absorbed at the very price level quickly and the possibility of exploiting this information, even in the short / medium term, does not exist to make profits.

Asymmetry and kurtosis can be observed on the histogram. In general, the data financial institutions have a weak and often negative asymmetry: the effect of bad news tends to be larger than the effect of good news. It follows that, the histogram will not take the form of a normal distribution; the distribution will tend towards the left side (picture 2), in which the losses are represented, behaving a different market reaction to earnings.

Figure: 2 NORMALIZED DISTRIBUTION HISTOGRAMS
Finally, we consider the absence of correlation. To observe if there is a linear link between the values of the same variable, detected in different time instants, recourse is made to "Correlation of returns".

Given the series of returns for $t = 1, ..., N$ of an asset, it is possible to derive the series of autocorrelations $\rho(h)$ which allows to investigate and describe the present linear bonds within the series.

The autocorrelation function is given by:

$$\rho(h) = \frac{\sum_{t=1}^{N-h} (r_t - \bar{r})(r_{t+h} - \bar{r})}{\sqrt{\sum_{t=1}^{N-h} (r_t - \bar{r})^2 \sum_{t=1}^{N-h} (r_{t+h} - \bar{r})^2}}$$

Empirical evidence shows that correlation of returns is often close to zero implying that returns observed at different times may have no connection whatsoever of linear type.

![Figure 3: Correlation between returns](image_url)

As previously stated, prices are slow and persistent, however the correlation of returns is often close to zero. Therefore, since there is not linear bonding, information at time $t$ is not enough to predict what will happen at the time $(t-1)$. Looking at picture 3, note that the blue lines represent the confidence interval with respect to the percentile of the normal random
variable standardized ($\alpha=5\%$); while the red bands are inside the blue bands, we accept the hypothesis that the correlation is close to zero.

1.3 Volatility analysis

Volatility analysis meets the needs of an uncertainty assessment of future returns.

Specifically, volatility is a financial index calculated in a given interval of time and can be defined as the deviation of returns with respect to their average value. In other words, it measures the ability of the same returns to take on values extremes; this volatility is called “conditional volatility” and represent a relevant information for the listing of shares and for the management of risk.

Analytically it can be defined as:

$$\text{var}(r_{t+k} | I_{t-1}) = \sigma_{t+k | t}^2,$$

where $k>0$ and $I_{t-1}$ represents the set of information available at time $(t-1)$.

Against the conditional volatility there is the unconditional volatility of the series of returns, unconditional volatility means volatility that does not change.

Analytically:

$$\sigma^2 = \frac{\sum_{t=1}^{N} (r_t - \bar{r}_t)^2}{N-1},$$

where $\bar{r}_t = \frac{1}{N} \sum_{t=1}^{N} r_t$, the time parameter changes and is therefore constant throughout the period considered.

The unconditional volatility imposes a constraint of unrealistic homogeneity with financial returns.

From an economic point of view, volatility can be interpreted as a measure of market risk\(^4\) related to future values of returns, given the information at time $t$.

While, from a statistical point of view the variance indicates the dispersion or distance of the welcome returns of a financial asset with respect to the average of the same returns. Basically, it measures the extent of the variations undergone by the price of an asset financial and therefore defines the lesser or greater probability that this price changes to vary the required returns. As volatility increases, the probability that of wider oscillations increases too, whether they increase or decrease.

---

\(^4\) Market risk: risk represented by the probability that financial assets, traded on a market sufficiently liquid, are subjected to significant fluctuations in their listing, due to of the unpredictability of factors capable of influencing it. In other words, market risk is the risk related to any portfolio losses determined by the market.
Volatility is directly proportional to the return; higher returns of financial asset, the greater risk associated with it, therefore the higher the volatility associated with a financial asset with a low level of return will present a low level of variability; a financial asset is all the riskier the greater variability of its performance around the expected value.

1.4 Stylized facts of volatility

The analysis of the volatility of the volatility of the returns has identified three main characteristics, defined stylized facts, which are:

- volatility clustering;
- thick tails;
- leverage effects.

To explain the concept of volatility clustering it must be kept in mind that volatility is a persistent phenomenon, i.e. the values it assumes are closely related to values assumed in the previous period, but also with the future values that it will assume in the period after the analysis. Or rather, any value assumed by the conditional variance is related to the values assumed by it in the previous period of observation.

To emphasize this phenomenon Mandelbrot (1963) states:

“Big changes tend to be followed by big changes, while small changes tent to be followed by small changes”

This statement contains the concept of volatility clustering which is identified in the alternation of large oscillations with small fluctuations around an average value.

As for Thick tails, the empirical distributions of the returns are generally characterized by a marked deviation from normality. These distributions are defined leptokurtic whose peculiarity consists in the presence of tails that contain more information than a normal distribution. In other words, leptokurtic distributions have as their main characteristic the assignment of a greater probability of events far removed from the average value of the distribution in relationship to the probabilities that would be assigned to such events in a normal distribution, for this reason we speak of thick tails.

---

5 Benoit Mandelbrot, 1963. “The Variation of Certain Speculative Prices”.
From a probabilistic point of view, the presence of heavy tails in the returns implies the presence of numerous extreme values.

The statistical index capable of providing information regarding the information contained in the distribution tails is kurtosis. The latter can be estimated by:

$$\hat{\gamma}_2 = \frac{1}{N} \sum_{t=1}^{N} \left( \frac{r_t - \bar{r}}{S} \right)^4$$

Where:

- $r_t$ is the value of the single observation that is considered;
- $\bar{r}$ is the sample mean;
- $S$ is the sample variance.

Kurtosis, given by the fourth moment of the standardized random variable, will be equal to 3 for normal distribution, less than 3 for platykurtic distributions and greater than 3 for leptokurtic distributions.

To be clearer, we speak of leptokurtic distribution if it is more “pointed”, as can be seen from Picture 4, of the normal and has heavy tails; we speak, instead, of platykurtic distribution if it is more flattened than normal and with thin tails.
Finally, consider the leverage effect. According to this effect the volatility at the time current tends to be negatively correlated with past returns. In other words, there is a negative linear link between the returns of today and the returns of yesterday.

The definition “leverage effect” is due to Black (1976), who noted that there is one tendency of price shock changes to be negatively correlated with change in volatility. In this sense, volatility tends to react asymmetrically compared to the sign of shock.

For investors an increase in the leverage effect corresponds to an increase in the risk that is incorporated in the volatility of the risk itself.

1.5 Proxy of volatility

It is important to remember that volatility is an unobservable variable and therefore not directly observable. To estimate it, it is necessary to use the proxies of the measure.
Proxy of the measurement of volatility can be:

- The squared daily returns: \( r_t^2 \) an undistorted measure of volatility, even if is affected by strong imprecision;
- Daily returns in absolute value: \( |r_t| \)
- Realized volatility, which is based on the sums of information realized with less than a day. It can be understood as a non-parametric measure of volatility.

The concept of realized volatility reported per day, is nothing but the sum of returns calculated in smaller time intervals, such as percentage changes in the price.

Other volatility measures can be based on some known values of the distribution of returns (High-Low, Opening and Closing Prices, Extreme-value, etc.).

Estimators of Extreme Value volatility are based on intra-daily information. Defined \( H_t \) and \( L_t \) as, respectively, the highest and lowest price in a generic day \( t \), it is possible to introduce another volatility proxy:

\[
\sigma_t^2(HL) = \log\left(\frac{H_t}{L_t}\right)^2
\]

Parkinson, in 1980, introduces a volatility estimator if price intra-daily follows a geometric Brownian generating process. Analytically:

\[
\hat{\sigma}_t^2(HL)_{c} = \frac{\sigma_t^2(HL) + (\log(H_t) - \log(L_t))^2}{4\log(2)}
\]

Garman and Klass, in 1980, modify this proxy by presenting a more efficient estimator, also based on a Brownian generating process, which considers besides the logarithm of the ratio between the extreme value also the logarithm of the relationship between the closing prices:

\[
\hat{\sigma}_{(G\&K),t}^2 = 0.5(\ln\left(\frac{H_t}{L_t}\right))^2 - 0.39(\ln\left(\frac{P_t}{P_{t-1}}\right))^2
\]

The estimator can be formulated, in terms of standard deviation as follows:

\[
\hat{\sigma}_t = \sqrt{0.511 (h_t - l_t)^2 - 0.019[c_t(h_t + l_t)] - 2h_t l_t - 0.383 c_t^2}
\]

In which: \( h_t = \ln\left(\frac{H_t}{O_t}\right), \ l_t = \ln\left(\frac{L_t}{O_t}\right) \) and \( c_t = \ln\left(\frac{C_t}{O_t}\right) \). Specifically, \( O_t, C_t \) are the opening price and daily closing.
Note that, a Brownian process is a mathematical tool developed in the field of probability theory. Bachelier\(^6\) was the first to realize that it is possible apply probability theory to financial markets; according to this scholar, investors develop price forecasts based on one independence feature between them.

State that \(r_t^2\) and \(\sigma_t^2(HL)c\) are proxies of undistorted volatility implies that their expected value is equal to the variable of interest. Analytically we have:

\[
E(\sigma_t c_t) = E(r_t^2) = \sigma_t^2
\]

Different approach for estimating the volatility of returns concern the use of techniques exploratory, such as moving variance or exponential smoothing, and the use of models stochastic, like the GARCH model or STOCHASTIC VOLATILITY MODELS.

1.5.1 Exploratory techniques

Exploratory techniques to predict volatility are linear filters that can be applied to the original data of the series we are considering; they are:

1) Moving variance;
2) Exponential smoothing.

In the moving variance, the variance is calculated over a mobile time window:

\[
\hat{\sigma}_{t+1}^2 = \frac{1}{m} \sum_{j=1}^{m} (r_{t-j+1} - \bar{r}_t)^2
\]

Where \(m\) is the width of the windows and where \(\bar{r}_t\) is the moving average of returns. Average mobile returns are as follows:

\[
\bar{r}_t = \frac{1}{m} \sum_{j=1}^{m} r_{t-j+1}
\]

It is calculated from time to time by eliminating the oldest observation and adding the most recent observation.

The weakness of this method lies in the arbitrary choice of window width.

---

\(^6\) Bachelier: French statistician and mathematician. In his thesis Bachelier anticipated many concepts today commonly accepted by financial scholars: "random walk" movement of stock prices, motorcycles Brownian and martingale process, these latter concepts even earlier than the later theories by Einstein and Wiener.
Now we consider Exponential Smoothing, also called Risk-metrics variance. In such method volatility is recursively estimated as a combination of the last volatility expected and the standard deviation of the last deviation of the last observed performance.

Analytically, we have:

$$\hat{\sigma}^2_{t+1|t} = \lambda \hat{\sigma}^2_t + (1 - \lambda)(r_t - \hat{r}_t)^2$$

Where $\hat{r}_{t+1} = \lambda_m \hat{r}_t + (1 - \lambda_m)r_t$ and the parameter $\lambda$, $0 \leq \lambda \leq 1$, is defined as smoothing.

At the ends of the smoothing parameter we have:

- $\lambda = 1$: constant volatility;
- $\lambda = 0$: the most recent deviation from the mean is obtained as an estimate of the variance square

The value of the smoothing parameter is related to the process memory and affects the influence of past values of returns current volatility. A low value of the parameter $\lambda$ causes the term $r^2_{t-1}$ has more weight; this lead, consequently, to obtain a model that is very sensitive to pass returns, or to market innovations. To tend of $\lambda$ towards 1, increases the persistence of volatility. A high lead of $\lambda$ to an answer slow to innovations.

J.P. Morgan suggests, based on experience, to impose $\lambda = 0.94$

The risk-metrics database, created by J.P. Morgan and made available in 1994, is based on an EWMA model (Exponential Weighted Moving Average) with $\lambda = 0.94$ to estimate the daily volatility.

J.P. Morgan showed that, in a wide range of different market variable, the value of $\lambda$ provides forecasts of the variance approximated to the true value of the realized variance.

The notion $\hat{\sigma}^2_{t+1|t}$ emphasizes that the volatility estimated in a given period ($t$) is indeed used as a prediction of volatility in the subsequent period ($t+1$). For completeness, the daily VAR\(^7\) at a confidence level $p$ (for example 95%) can be calculated, under the assumption of normality, multiplying $\hat{\sigma}^2_{t+1|t}$ for the quantile ($1-p$) of the standard distribution.

\(^7\) Var: Value at Risk, a measure of market risk whose disclosure dates to the first inclusion of that indicator in the Basel agreement (1996). In a probabilistic context, the VaR is connected to occurrence of rare events in the left tail of the yield density function.
1.5.2 Comparison of exploratory techniques

In the comparison between the exploratory techniques, it can be deduced that the Exponential smoothing is from consider the best approach for estimating the volatility of returns.

In fact, few data are required in this exploratory technique, given the volatility of the period current depends only on the data of the previous period. Furthermore, it monitors the changes in volatility (if the previous period was very volatile, then the estimate current account will take this event into consideration). Finally, we underline that the parameter establishes how much the estimate is sensitive to the change in volatility in a single period or establishes its sensitivity to innovations.

On the other hand, exponential smoothing does not consider any representative elements of volatility in the long term, moreover, there is no forecast element and consequently the forecast of volatility in the subsequent period is equal to the volatility of the current period.

About the moving variance, or risk-metrics, this method was used to measure market risk through Value at Risk. The model is based exclusively on the assumption of normality of returns and ignores the presence of asymmetry and tick tails, which as previously stated are features fundamentals of financial assets.

In general, the performance in forecasting the volatility of this exploratory technique is considered good because it is based on a close time horizon, in fact it is based on an only period.

The moving variance incorporates the phenomenon of volatility clustering and is based on a rather strong assumption, that the returns are conditionally normal, meaning that they are conditional on the set of information available at time T, which normally consists of the series of returns passed at time T itself.

The methodology of Risk-metrics has allowed the diffusion of numerous methods of statistical measurement of market risk. The first institution to develop a complex methodology for managing market risk is JP Morgan. The database of Risk-metrics, created by this institution was based on a moving average model exponentially weighted; in 2006 the risk-metrics models were updated and the volatility estimate, currently, is no longer based on EWMA models but on models ARCH.
Wanting to represent the typical behaviors of the volatility of returns, such as the clustering of the volatility and the presence of tick tails, we instead resort to use of family of ARCH models.

2. ARCH model

The ARCH model was introduced by Engle\(^8\) in 1982. The economist's main idea was that of a model that could estimate the volatility of the Friedman hypothesis (1997) that the unpredictability of inflation was the primary cause of economic cycles\(^9\).

The acronym ARCH stands for "Auto-Regressive Conditional Heteroskedasticity", i.e. models in which volatility does not is a constant function of the time parameter.

In econometrics the term heteroskedasticity indicates that the variance between the countless observations that make up the sample is not constant; opposed to that concept is the concept of homoskedasticity, with which we want to represent the set of distributions where volatility is constant.

With the term Auto-Regressive, instead, we want to indicate the dependence of the present values from the past of the variable being considered.

The main objective of the model, or more precisely of the model family, is ARCH that of providing an estimate of the volatility as standard deviation (or quadratic deviation medium, is nothing but the square root of the variance). The latter can be used in financial decisions relating to risk analysis.

In having to define the structure of an ARCH model, we consider the latter as linear combination of past returns, or we assume that returns are generated by a stochastic process\(^10\) in which volatility is conditioned by the change of time.

---

\(^8\) Robert F. Engle is an American statistician. In 2003, Engle, along with econometrician Clive Granger, received the Nobel Memorial Prize in Economic Sciences for "methods of analyzing economic time series with time-varying volatility (ARCH)"

\(^9\) Friedman, an American economist with the Nobel prize for economics in 1976, considered money as a pivot of the economic system and the main cause of inflation. Friedman's main thesis is that there is a close relationship between the cyclical movements of quantities of money and those of real income. The autonomous variations of the quantity of money, monetary shocks are the main causes of fluctuations.

\(^10\) Stochastic process: set of random variables ordered according to a parameter \(t \in T\), where \(T\) indicates the parametric space. The historical series of returns can be understood as a finite part of the endless traces that make up a stochastic process.
Note that the number of quadratic returns involved is reported through the notion ARCH (p), where p is the estimated parameter in the model itself.

2.1 Specification of the ARCH model

The use of the ARCH model occurs if the presence of autocorrelation of positive transformations of the returns (square or absolute value).

Defining \( r_t = \mu_t + \varepsilon_t \) as the returns, given by the sum of an average level \( \mu_t \) and an error term \( \varepsilon_t \).

Empirical evidence suggests that in most cases the average of returns is next to zero, it follows that the attention is shifted to the error term, \( \varepsilon_t \); such term has a conditional variance component and a standardized return: \( \varepsilon_t = \sigma_t z_t \).

In particular:

- The conditional variance is: \( \sigma_t^2 = \text{var}(r_t | I^{t-1}) \);
- \( I^{t-1} \) is the set or set of information on the time series at time (t-1);
- \( z_t \sim D(0,1) \), where D (0,1) is a random variable with mean 0 and variance 1.

The structure of the model is given by the linear combination of past returns, analytically:

1. \( r_t = \sqrt{\sigma_t^2} z_t \)
2. \( \sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 \)

The coefficients \( \alpha_i \) (\( i = 1, \ldots, p \)) are the parameters of the model that need to be estimated.

Note that equation (1) implies a link between returns and volatility, moreover, identifies the idiosyncratic component or innovation; while equation (2) identifies the structure of variability.

In identifying a model for variance, it is necessary to ascertain that it does not generate negative values; to guarantee the positivity of conditional variance, \( \sigma_t^2 \), the constraints of non-negativity that must be met by the parameters are: \( w > 0 \) and \( \alpha_i \geq 0 \).

A feature of ARCH models is that they can be written as an equation that expresses the return \( r_t \) in a way dependent only on volatility.
In fact, writing the returns as:

\[ r_t = \sigma_t^2 + v_t \]

it is possible to obtain \( v_t \):

\[ v_t = r_t - \sigma_t^2 \]

This error term is given by a sequence of serially uncorrelated random variables but not independent.

### 2.2 ARMA representation of an ARCH model

To obtain a more general version of the ARCH model, the use of a model ARMA\(^{11}\) (Moving Average Auto-Regressive). This autoregressive model associates, in fact, an ARMA(p,q) model for the historical series of returns to a model ARCH(p) for innovations.

In other words, it is also possible to identify the ARCH (p) model as a model Auto-regressive order p, AR (p), for squared returns, \( r_t^2 \):

\[ r_t^2 = w + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + v_t = \Phi_0 + \Phi_p(B) r_t^2 + v_t \]

Being a linear model, we can observe the correlogram of \( r_t^2 \), in such a way that identify the order of the model more easily.

This process cannot be done on \( r_t \) as it has no component on average, there is no linear part. The AR (p) model, therefore considering only the component autoregressive, it can determine the variations in the amplitude of the oscillations of the historical series.

As we have already said, the ARCH models serve for the estimation of conditional variance. An ARCH model is stationary if the p roots of the polynomial \( (\sum_{i=1}^{p} \alpha_i B^i) \) reside at outside the unit circle (since the roots of the delay polynomial can be provided either from real numbers, or from complex numbers, this characteristic is identified in the conditions in which their form is greater than 1).

---

\(^{11}\) ARMA Model (Auto Regressive to Moving Average): linear mathematical model which provides instant for instant an output value based on the previous input and output values.
If it imposes $\alpha_i \geq 0$ to guarantee the positivity of $\sigma_t^2$ the stationarity in variance is had when:

$$\sum_{i=1}^{p} \alpha_i < 1$$

Therefore, starting from the condition that the parameters $\alpha_i$ are all positive, the stationarity invariance is guaranteed when the sum of all the above parameters is less than 1.

Generalizing the model under condition of stationarity in variance, the variance does not conditioned performance is given by:

$$
\left(1 - \sum_{i=1}^{p} \alpha_i\right) E(r_t^2) = w \rightarrow E(r_t^2) = \frac{w}{1 - \sum_{i=1}^{p} \alpha_i}
$$

The condition $w > 0$ is also necessary to ensure the non-positive variance conditioning of returns. The conditions sufficiently relevant for the non-obligation negativity becomes:

$w > 0, \alpha_i > 0$ and $(\alpha_1 + \ldots + \alpha_p) < 1$.

In the simplest case of a model ARCH (1), where, we remind that $p = 1$ implies that there is only one quadratic yield involved, we have that (under the conditions of stationarity in variance):

- The conditional variance is equal to: $\sigma_t^2 = w + \alpha r_{t-1}^2$; it depends only on a delay previous one.
- The unconditional variance is equal to: $E(r_t^2) = \frac{w}{1 - \alpha}$; it is given from the moment according to the entire distribution.
- Kurtosis is equal to: $\gamma_2 = \mu_4 \frac{1 - \alpha^2}{1 - \mu_2 \alpha^2}$, in which $\mu_4$ is the fourth moment of $z_t$ and in the case of normality is equal to 3.

To understand the relationship between conditional variance and unconditional variance, it must be bore in mind that for the law of expected values iterated, the unconditional variance, $\sigma^2$, can be written as:

$$\sigma^2 = E(r_t^2) = E[\text{var}(r_t^2 | I^{t-1})] = E(\sigma^2)$$
Which in the specific case of an ARCH (1) becomes:

\[ E(r_t^2) = E(\sigma^2) = E(w + \alpha_t r_{t-1}^2) = w + \alpha_t E(r_{t-1}^2) \]

Under the hypotheses of stationary nature of the process, this equation becomes:

\[ E(\sigma^2) = E(\sigma_{t-1}^2) = \sigma^2 \]

In conclusion, we will therefore have: \( \sigma^2 = w + \alpha_t \sigma^2 \) from which it is possible to obtain:

\[ \sigma^2 = E(r_t^2) = \frac{w}{1 - \alpha} \]

In ARCH models, the conditional variance is given by the sum of the non-conditional variance and a fraction of the difference between the most recent squared innovation and the its unconditional expected value.

In fact, an interesting interpretation of the conditional variance equation is given from: \( \sigma_t^2 = w + \alpha_t r_{t-1}^2 \) which can be written as:

\[ \sigma_t^2 = \sigma^2(1 - \alpha_t) + \alpha_t \sigma_{t-1}^2 = \sigma^2 + \alpha_t (\sigma_{t-1}^2 - \sigma^2) \]

This equation demonstrates some relevant properties of the model:

- The coefficient \( w \) must be strictly positive;
- The coefficient \( 0 \leq \alpha_t < 1 \);
- The conditional variance may be greater or less than the non-conditional variance;
- The unconditional distribution of returns is leptokurtic.

In a nutshell, the ARCH model and its generalizations are, therefore, applied to the modeling, among other things, interest rates, exchange rates and stock returns.

### 2.3 Test for the presence of heteroskedasticity

To determine the presence of autocorrelation in the residues, tests are carried out for the presence of heteroskedasticity. The ARCH test proposed by Engle himself foresees how alternative hypothesis:

\[ H_1: r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_m r_{t-m}^2 + u_t \]
where \( u_t \) is a White Noise stochastic process, meaning a stationary process in the sense weak (or covariance) identified as \( u_t \sim WN(O, \sigma^2) \).

Opposed to the alternative hypothesis is the null hypothesis, defined as:

\[
H_0 : \alpha_0 = \cdots = \alpha_m = 0
\]

Expressing the temporal link between returns (innovations) squared as a function of delays:

\[
r_t^2 = f(r_{t-1}^2, r_{t-2}^2, \ldots)
\]

and resorting to the use of an Auto-Regressive Media Mobile model (ARMA, model) specified for the time series average) it is possible to verify the actual presence of autocorrelation. Based on the ARMA model, the estimate of returns, \( \hat{r}_t \), is given by:

\[
r_t - (\mu + \Phi_1 \hat{r}_{t-1} + \cdots + \Phi_p \hat{r}_{t-p} + \Theta_1 \epsilon_{t-1} - \cdots - \Theta_q \epsilon_{t-q})
\]

The resulting squares of the residuals regressed to a constant and a regression auxiliary of past p values:

\[
\hat{r}_t^2 = \alpha_0 + \alpha_1 \hat{r}_{t-1}^2 + \cdots + \alpha_p \hat{r}_{t-p}^2 + \epsilon_t
\]

Under the null hypothesis, Engle's test statistic, obtained as a product of the numerosity of the sample N for \( R^2 \) of the auxiliary regression, it is distributed as a chi-square with p degrees of freedom.

\[
NR^2 \overset{d}{\rightarrow} \chi^2_p
\]

Different to the ARCH test are the Liung-Box Q tests and the Box-Pierce Q tests. Both tests consider the following null hypothesis: \( H_0: u_t \sim WN(O, \sigma^2) \), which implies that the correlations they are different from zero. Consequently, the alternative hypothesis provides for null correlations.

Before specifying the test statistics, we specify that the two types of tests differ simply because of the different weighting system used, but asymptotically they have the same distribution.
The test statistic of the Liung-Box Q Test is:

\[ Q = N(n + 2) \sum_{k=1}^{m} \frac{\hat{\rho}_a^2}{N - k} \]

While, the Box-Pierce Test Q test statistic is:

\[ Q = n \sum_{k=1}^{m} \hat{\rho}_a^2 \]

In both statistics, wherein \( \hat{\rho}_a(k) \) is the estimate of the autocorrelation function of residuals, under the null hypothesis the Q statistic is distributed as a chi-square with \( g = m - (\hat{p} + \hat{q} + 1) \) degrees of freedom.

In general, the test results are identified by the p-value. The information provided by probability value indicates the minimum significance value for which the null hypothesis comes we refused.

\[ P\left( \chi^2_p > \chi^2_\alpha \right) = \lambda \]

In the case of Engle's ARCH Test, the null hypothesis is rejected if it is verified that \( \lambda < \alpha \), otherwise the hypothesis is accepted.

In the specific cases of Q tests, the null hypothesis is rejected if \( Q > \chi^2_p(\alpha) \). The rejection of the hypothesis nothing indicates the inadequacy of the model.

To perform the test for the presence of ARCH it is not necessary to specify the above real ARCH model. To conclude, in choosing the number of delays there is a trade-off between test accuracy and significance of results: accuracy in capturing the effect of persistence is given by the auxiliary regression, the latter will be better if there is the addition of delays in the same regression; however, the addition of delays could attribute too much weight to distant events and leads the test statistic to be non-significant.

2.4 Limitations of the ARCH model

ARCH is a process with zero mean, constant variance and conditional variance dependent on the squares of the idiosyncratic components (or innovations). It succeeds in capture the
phenomenon of fluctuations in the historical series relative to the returns of a financial asset and, therefore, can represent volatility clustering.

However, the limitations to this model concern:

1. The model assumes that shocks, positive or negative, we have the same price effect on volatility; the prices of a financial asset are so responsive different to shocks.
2. The model provides only a mechanism to describe the volatility trend, of facts, it does not provide nor any indication on the causes of such behavior nor new information to understand the origin of variance.
3. The model is likely to overestimate the volatility because it responds slowly to big isolated shocks for the return series.

3. GARCH model

A limit of ARCH models can be given by the too high number of yields necessary to adapt it to the observed data. Moreover, the empirical evidence shows that these models rarely consider the persistence that characterizes them the trend of conditional variance of returns. Which is why, to overcome this limit, Bollerslev (1986) proposed to use a more general class of models, defined GARCH, Generalized Autoregressive Conditional Heteroskedasticity.

The $GARCH (p,q)$ models introduce the delayed values of the conditional variance $\varepsilon$ consider relevant information on past values of conditional variance. In other words, the model is such that the conditional variance at time $t$ is expressed through a linear combination of $q$ delays of itself, synthesis of the past information, and of $p$ squared residual delays, representative of market news and capacity of variation over time of conditional variance.

The main objective of this generalization is to reduce the parameters to be estimated with respect to ARCH model structure.

3.1 Specification of the GARCH model

In analogy with the ARCH model, in the GARCH model the variance conditional on time $t$ is given by a linear combination of $p$ delays of squared residuals and of $q$ delays of the conditional variance, analytically we have:
\[ \sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i r_{t-1}^2 + \sum_{i=1}^{q} \beta_i h_{t-i}^2 \]

Which in the case of a GARCH (1,1) becomes:

\[ r_t = \sqrt{\sigma_t^2} z_t \]

and

\[ \sigma_t^2 = w + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1}^2 \]

The non-negativity conditions of the variance, in this case, are: \( w > 0, \alpha_i \geq 0, \beta_j \geq 0 \) e \( \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1 \). Such conditions, as demonstrated by Nelson and Cao (1992), may be less restrictive for higher-order models.

The parameters of the model must satisfy, in addition to the conditions of non-negativity of the variance, some regularity conditions. To this end it is assumed that \( z_t \) is distributed as a Normal Standard distribution or a standardized Student t distribution.

We have:

- \( z_t | I_{t-1} \sim N(0; 1) \)

then: \( r_t | I_{t-1} \sim N(0; \sigma_t^2) \), we will have that innovation, given the information set at time \( t-1 \), is a normal random variable with zero mean and variance that depends on time.

Therefore, as the time parameter changes, probability distributions will be realized normal of a different type, dispersed compared to their center of symmetry.

The index \( r_t^2 \) covers the role of the observable variable in the ARMA model and the role of innovation for the process of variance is played by the difference \( r_t^2 - \sigma_t^2 \):

\[ \sigma_t^2 = w + (\alpha_1 + \beta_1) r_{t-1}^2 + (r_t^2 - \sigma_t^2) - \beta_1 (r_{t-1}^2 - \sigma_{t-1}^2) \]

From this equation it is easy to deduce the autoregressive nature of the coefficient relative to \( r_{t-1}^2 \), while the coefficient 1 refers to a moving average component.
3.2 ARMA representation of a GARCH model

In analogy with what we have seen for the ARCH model, a GARCH model can be written as an ARMA model \((p, q)\) for \(r_t^2\):

\[
r_t^2 = w + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i)r_{t-1}^2 + v_t - \sum_{i=1}^{q} \beta_i v_{t-i}
\]

From this transformation it is immediate to better understand the reasons for the constraints, previously defined. The unconditioned average of the ARMA model is defined by:

\[
E(r_t^2) = \frac{w}{1 - \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i)}
\]

\(\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1\) is necessary for the denominator to be positive.

In a GARCH model, the conditional variance reacts to the delayed value of the term of squared noise (innovation for returns) in an amount equal to \(\alpha_1\). In fact, the variance conditioning \(\sigma_t^2\) will be greater than \(\sigma_{t-1}^2\) itself:

\[
w + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 > \sigma_{t-1}^2
\]

Through simple algebraic steps, you will have:

\[
r_{t-1}^2 > \frac{(1 - \beta_1) \sigma_{t-1}^2 - w}{\alpha_1}
\]

In the stationary conditions in variance, the moments of a GARCH \((1,1)\) model are:

- **Conditional variance:**
  \[
  \sigma_t^2 = w + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1}^2
  \]

- **Unconditional variance:**
  \[
  E(r_t^2) = \frac{w}{1 - \alpha_1 - \beta_1}
  \]

- **Kurtosis:**
  \[
  \gamma_2(r_t) = v_4 \frac{1 - (\alpha_1 + \beta_1)^2}{(1 - \alpha_1^2 v_4 - \beta_1^2 - 2\alpha_1\beta_1)}
  \]
- ACF:

\[ \rho(k) = (\alpha_1 + \frac{\alpha_1^2 \beta_1}{1 - 2\alpha_1 \beta_1 - \beta_1^2})(\alpha_1 + \beta_1)^{k-1} \]

About the unconditional variance, we have:

\[ E(r_t^2) = \sigma^2 = w + \alpha_1 E(\sigma^2_{t-1}) + \beta_1 E(\sigma^2_{t-1}) = w + \alpha_1 \sigma^2 + \beta_1 \sigma^2 \]

From which \( w = \sigma^2(1 - \alpha_1 - \beta_1) \) with \( (\alpha_1 + \beta_1) < 1 \).

Replacing this result in the expression of conditional variance we obtain:

\[ \sigma_t^2 = \sigma^2(1 - \alpha_1 - \beta_1) + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 = \sigma^2 + \alpha_1 (r_{t-1}^2 - \sigma^2) + \beta_1 (\sigma_{t-1}^2 - \sigma^2) \]

The persistence of conditional variance is a clear intuition of the equation.

**3.3 Estimate of the GARCH(p,q) model**

The estimation phase of a GARCH model is obtain by the maximum likelihood function.

The likelihood function, under the normality hypotheses of the distribution conditioning, is the following:

\[ \mathcal{L}(w, \alpha_1, \beta_1 | r_2, ..., r_T; r_1) = \prod_{t=2}^{T} \mathcal{L}_t (w, \alpha_1, \beta_1 | r_2, ..., r_T; r_1) \]

in which \( r_1 \), the first value of the time series of returns, is considered as a value initial and therefore constant; this term is used to start the estimation procedure.

Considering exclusively the observation at time \( t \), the function \( \mathcal{L} \) becomes:

\[ \mathcal{L}_t = \frac{1}{\sqrt{2\pi \sigma_t^2}} e^{(-\frac{1}{2\sigma_t^2} r_t^2)} \]

Note that supposing \( r_t = \mu + \epsilon_t \), where \( \epsilon_t = \sigma_t v_t \) is the error term that a has component in conditional variance and in which \( v_t \) is an independent random variable eidentically distributed (with zero mean and unit variance), and assuming that \( \mu_t = 0 \), yes will have \( r_t = \epsilon_t = \sigma_t v_t \).
It follows that the function \( \mathcal{L} \) can be written as follows:

\[
\mathcal{L}_t = \frac{1}{\sqrt{2\pi \sigma_t^2}} e^{\left(-\frac{1}{2\sigma_t^2} \epsilon_t\right)}
\]

Having to obtain the parameter estimation, we maximize the likelihood function, or, to simplify this estimate, the likelihood log function:

\[
l(\epsilon_{p+1}, \epsilon_T | w, \alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \epsilon_1, \ldots, \epsilon_p) = \sum_{t=p+1}^{T} -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{1}{2} \epsilon_t^2 / \sigma_t^2
\]

By isolating the only terms that depend on the parameter of the model we obtain:

\[
l(\epsilon_{p+1}, \epsilon_T | w, \alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \epsilon_1, \ldots, \epsilon_p) = -\sum_{t=p+1}^{T} \frac{1}{2} \ln(\sigma_t^2) - \frac{1}{2} \epsilon_t^2 / \sigma_t^2
\]

The maximization problem will be the following:

\[
\begin{align*}
\max_{\theta} \{ l(\theta) \} &= \max_{\theta} \left\{ -\sum_{t=p+1}^{T} \frac{1}{2} \ln(\sigma_t^2) - \frac{1}{2} \epsilon_t^2 / \sigma_t^2 \right\} \\
\text{s.t.} \quad \sigma_t^2 &= w + \sum_{i=1}^{p} \alpha_i r_{t-1}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2
\end{align*}
\]

The constraint of the problem is defined precisely by the conditional variance of the model GARCH (p, q).

Additional constraints are given by:

\[
\begin{align*}
\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) &< 1 \\
w &> 0 \\
\alpha_i &> 0 \quad \forall i \\
\beta_i &> 0 \quad \forall i
\end{align*}
\]

To maximize the objective function, and therefore obtain the solution, the derivative with respect to the parameters \( w, \alpha_i, \beta_i, \) (sinteticamente \( \Theta \)) equals zero.

First order condition:
$$\frac{\partial l(\theta)}{\partial \theta} = - \frac{1}{2} \sum_{t=1}^{T} \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} + \frac{1}{2} \sum_{t=1}^{T} \frac{\epsilon_t^2}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} = \sum_{t=1}^{T} \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} \left[ \frac{\epsilon_t^2}{\sigma_t^2} - 1 \right] = 0$$

In the specific case of a GARCH (1,1), the conditional likelihood function becomes:

$$L_t = \frac{1}{\sqrt{2\pi(w + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2)}} e^{(-\frac{1}{2(w + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2)} r_{t-1}^2)}$$

And in the specific case of \( r_t = \epsilon_t \) with innovation whose conditional variance follows a GARCH (1,1) you will have:

$$\frac{\partial \sigma_t^2}{\partial w} = 1$$

$$\frac{\partial \sigma_t^2}{\partial \alpha_1} = r_{t-1}^2$$

$$\frac{\partial \sigma_t^2}{\partial \beta_1} = \sigma_{t-1}^2$$

The three conditions that jointly produce the maximum likelihood estimators \( \hat{\theta}, \hat{\alpha}_1, \hat{\beta}_1 \) are:

$$\sum_{t=2}^{T} \frac{1}{\sigma_t^2} \left[ \frac{r_{t-1}^2}{\sigma_{t-1}^2} - 1 \right] = 0$$

$$\sum_{t=2}^{T} \frac{r_{t-1}^2}{\sigma_{t-1}^2} \left[ \frac{r_{t-1}^2}{\sigma_{t-1}^2} - 1 \right] = 0$$

$$\sum_{t=2}^{T} \frac{\sigma_{t-1}^2}{\sigma_t^2} \left[ \frac{r_{t-1}^2}{\sigma_{t-1}^2} - 1 \right] = 0$$

These conditions cannot be expressed in a closed form in terms of parameters unknown; this operation, in fact, is often performed not analytically, but through of the appropriate numerical algorithms. The latter consist of a set of procedures iterative that suggest partial solutions.
3.4 GARCH- in mean model

In the ARCH and GARCH models the empirical evidence suggests that the average of the returns it is conventionally set equal to zero, implying that the conditional average is completely independent of conditional volatility.

In fact, it has been assumed that the historical series of returns is generated by:

\[ r_t = \mu_t + \varepsilon_t = \varepsilon_t = \sqrt{\sigma_t^2} z_t \]

In which:

- \( \varepsilon_t = \sigma_t z_t \);
- \( \sigma_t^2 = \text{var} (r_t | I_t^{-1}) \);
- \( z_t \sim D(0,1) \), where \( D(0,1) \) is a random variable with mean 0 and variance 1;
- \( I_t^{-1} \) is the set or set of information on the time series at time \( t-1 \);

These models, therefore, do not meet the need to obtain feedback between the media and volatility; remember that, according to the classical portfolio theory, in front of one higher volatility expects a higher return. Synthetically, since the conditional variance is a measure of risk, a financial asset with a high risk would pay on average a higher yield; therefore, the average return depends on the riskiness of the asset and defines the risk premium (risk premium).

To keep this need in mind, Engle, Lilien and Robins (1987) introduced an important extension of the GARCH model, the GARCH-in mean model (below GARCH-M). The returns, in this case, will be defined as follows:

\[ r_t = \mu_t + \varepsilon_t \]

What changes is the average component of facts: \( \mu_t = \delta \sigma_t \).

As for returns, \( r_t | I_t^{-1} \sim D(\delta \sigma_t, \sigma_t^2) \): conditional variance also determines the conditioned average, based on the previous linear relationship (\( \mu_t = \delta \sigma_t \)). The model has a conditional mean and variance, that vary in non-linear manner in relation to past values of the time series of returns, \( r_t \).
The GARCH-M model is based on the GARCH (1,1) model, introduced by Bollorselv (1986) and consists of two equations, one for the average and one for the volatility of the time series of the returns:

1) The equation for the mean: \( r_t = \mu + \beta_1 \sigma^2_t + \epsilon_1 \)

2) The equation for variance: \( \sigma^2_t = w + \alpha \epsilon^2_{t-1} + \beta \sigma^2_{t-1} \)

In general, the GARCH-M model can be defined more adequately by

\[
\begin{align*}
    r_t &= X_t' b + \delta \sigma^2_t + u_t \\
    \sigma^2_t &= w + \alpha_1 u^2_{t-1} + \beta_1 \sigma^2_{t-1}
\end{align*}
\]

where:

- \( \delta \) is the volatility coefficient (risk premium) for the average;
- \( u_t = \sigma_t z_t \) represents the residual of the model at time \( t \);
- \( \alpha_1 \) ARCH component of the model;
- \( \beta_1 \) GARCH component of the model

Please note that a positive value of this coefficient indicates that the return is positively linked to your past volatility.

- \( u_t = \sigma_t z_t \) represents the residual of the model at time \( t \);
- \( \alpha_1 \) ARCH component of the model;
- \( \beta_1 \) GARCH component of the model

Given the presence of \( \sigma^2_t \) in the structure of the media, to obtain consistent estimates of the model parameters, the parameters for the average and for the conditional variance must be estimated simultaneously.

### 3.5 Testing standardized residues

Pivotal hypothesis of a GARCH(p,q) model is on the conditional distribution of the innovations, \( r_t = \sqrt{\sigma^2_t} z_t \), in which \( z_t | I_{t-1} \sim N(0; 1) \) implies that innovations standardized are not autocorrelated.

The goodness of adaptation of a GARCH model can be evaluated in relation to capacity of conditional variance to make standardized residues as close as possible to be normally distributed. The tools used for this purpose are the tests for autocorrelation, ARCH effects test, Jarque and Bera test.
Focusing only on the Jarque and Bera test (1980), it is defined as a Normality Test.

The main purpose of the test is to check if the interest series is distributed as one normal. The hypothesis system is:

\[ H_0: \text{the series is distributed as a Normal distribution} \]

\[ H_1: \text{the series is not distributed as a Normal distribution} \]

The test statistic, instead, is the following:

\[ JB = \frac{N - g}{6} \left( S^2 + \frac{(K - 3)^2}{4} \right) \]

Where: \( S \) (skewness) indicates the asymmetry of the interest series, \( K \) the kurtosis and the number of estimated coefficients.

In other words, the test statistic measures the difference between the values of asymmetry and kurtosis on the series with the known ones of the Normal random variable. A high-test value indicates that the joint deviation of asymmetry and kurtosis of the reference values of the Gaussian is statistically significant.

The value of the test statistic is then compared with a chi-square with 2 degrees of freedom; of facts, the Normality Test is distributed as:

\[ JB \overset{d}{\rightarrow} \chi^2_2 \]

### 3.6 Limitations of the GARCH model

As previously stated, the GARCH model is a first extension of the model ARCH. The main idea of this model is to reproduce the parsimony of the model ARMA in terms of parameters used, introducing delayed values in the ARCH model of conditional variance.

In many cases, the conditional variance equation based on the GARCH, under hypothesis of normality, provides a reasonably good model for time series analysis financial and the conditional volatility estimate. However, in some cases there are aspects of the model that can be improved so that you can capture the features and features dynamics of a historical series.
The GARCH model, in other words, allows us to capture the effects of volatility clustering but it does not consider the dependence of volatility on the sign of past returns.

Empirical evidence has shown that daily returns have very little correlation: $\text{Corr}(r_t; r_{t-1}) \approx 0$, and that squared returns and returns in value absolute show a positive correlation with past values:

$$\text{Corr}(r_t^2; r_{t-1}^2) > 0 \quad \text{Corr}(|r_t|; |r_{t-1}|) > 0.$$ 

These results are included in the GARCH model, but what is left out is the possibility of finding a negative correlation between past returns and variance:

$$\text{Corr}(\sigma_t; r_{t-1}) < 0.$$ 

This last effect is recurrent in the historical financial series, specifically, volatility reacts differently to an abrupt increase in price rather than an equally high price sharp price collapse. This asymmetry was highlighted for the first time by Black (1976) and is more commonly referred to as Leverage Effect.

As for the leverage effect, when a negative yield shock occurs volatility records higher values than those obtained following one positive shock of the same amplitude. The asymmetry is therefore configured in the fact that the volatility tends to increase in correspondence with the "bad news", i.e. in those periods in where the level of returns is less than expected and tends to decrease when one is in presence of "good news".

The GARCH model, in fact, assumes that the positive and negative error terms have a symmetrical effect on volatility, that is good and bad news have the same effect on volatility.

Since the GARCH model is a symmetrical model, to consider, the different effects, can be modified by generating:

- Asymmetric GARCH;
- Extended GARCH.
CHAPTER 2: VARIANTS OF THE GARCH MODEL

1. Introduction

In the GARCH model, a big negative shock has the same impact on the future value of the volatility of a great positive shock of the same dimension. This model does not consider empirical evidence of historical series already noted in the 1970s and that is that negative news has an impact on the greater volatility than positive news.

In other words, the model, as illustrated above, tends to underestimate the amount of volatility when there is "bad news" and, consequently, it tends to overestimate the amount of volatility when one is in presence of "good news"; furthermore, it underestimates the value of volatility after a shock high and overestimate it after a contained shock.

What is therefore not taken into consideration is the concept of asymmetry: when a financial asset has movements that tend to fall, then its volatility is greater than the movements in which it has movements that tend to rise: an investor, who is supposed to be risk averse, perceives downward movements as a danger to himself.

In other words, bad news produces greater volatility than those produced from the good news. This phenomenon is defined, in the financial sphere, as leverage effect and not is captured by the GARCH models so far examined.

1.1 Leverage effect

One of the most important empirical regularities is the inverse relationship between shock price and volatility. F. Black (1976) was the first to document the phenomenon, called "Leverage effect", in the financial sphere.

Black exposed this phenomenon in terms of a company's financial leverage: a return negative implies a decrease in the value of the shares of the company itself, it reduces the share capital of the company and subsequently this increases the leverage. For the investors, an increase in leverage corresponds to an increase in the risk that is incorporated in the volatility of the security.
The economist to study the relationship between the volatility and returns of a single security or of a portfolio composed of these, in his paper\textsuperscript{12}, uses daily data, from 1964 to 1975, of a sample consisting of 30 titles. For each title summarizes and estimates the volatility, in 21-day intervals, such as the square root of squared daily returns. Similarly, for portfolios it estimates, as defined by him, the "market return summed up "or" estimate of market volatility ".

Black defines the "variance of volatility" as the difference between the estimate of the current volatility and the volatility of the previous period, divided by the volatility estimate of the previous period.

The results, to which he arrives, suggest a strong inverse relationship between performance and stock volatility: the 1% decrease, in terms of yield, implies an increase of equal volatility.

The economist proposes two plausible explanations for this phenomenon:

1) The first explanation, defined as "direct causality", refers to the causal relationship between the returns of a security and volatility.

A decrease in the value of the company's shares causes a return of the security and increases the financial leverage of the security, the debt / capital ratio; an increase in this ratio implies an increase in volatility.

2) The second explanation, defined as "indirect causality", refers to the relationship causal between the change in volatility and stock returns.

Changes in technology lead to uncertainty about investment profits, consequently, for the investor to continue to hold the security, yes will have to check, against an increase in volatility, a reduction in the price of the title itself.

The leverage effect and the consequent explanations have been empirically confirmed by numerous studies, such as those of Christie (1982), Cheung and Ng (1992) and Duffee (1995).

\textsuperscript{12} Black, F., 1976, "Studies of Stock Price Volatility Changes"
1.2 Verification of asymmetry

To verify the presence of the leverage effect they have been developed and proposed by Engle and Ng (1993) different tests. The incorrect model specification tests are:

- Sign Bias Test (test for distortion due to the sign);
- Negative Size Bias Test (negative impact distortion test);
- Positive Size Bias Test (positive impact distortion test).

The logic behind the tests is to verify whether the standardized squared residues can be explained by some variable observed in the past and linked to the effect of asymmetry.

1.2.1 Sign Bias Test

The test verifies the impact that shocks, both positive and negative, have on volatility. Of facts, it is based on the result of a regression of standardized squared residuals on one constant and on a Dummy variable that takes the value of 1 at values negative of the residual estimate delayed by a period.

$$\frac{\hat{r}_t^2}{\sigma_t^2} = \alpha + \beta D_{t-1} - 1 + u_t$$

Where: $D_{t-1} = \begin{cases} 1 & \text{if there are negative shocks} \\ 0 & \text{if there are no negative shocks} \end{cases}$

As far as the parameter $\beta$ is concerned, it takes a positive value if there is a shock negative, otherwise this parameter will not give any contribution since it is canceled.

The logic of the test is to verify if the average of the standardized squared residuals is significantly different depending on whether the immediately preceding residues are positive or negative.

The test statistic, given by the relationship between $\hat{\beta}$ and its standard error, refers to a hypothesis in the absence of difference in the mean and is distributed as a t of Student with

---

13 Dummy variable: it is a variable that takes the value 0 or 1, depending on whether a date is satisfied condition. Its use in a regression is intended to capture the effect of a qualitative variable on average value of the dependent variable; in this sense it allows to improve the adaptation of the regression, since it allows to capture and insert in the system of variables also extra-statistical factors
(t-2) degrees of freedom. The alternative hypothesis defines the presence of the leverage effect ($\beta > 0$).

### 1.2.2 Negative Size Bias Test

The test analyzes the influence of the size of the shock; it predicts that $r_{t-1}$ has an effect not only on the average linked to the sign, but also an effect tied to its own dimension. The equation of the regression referred to is:

$$\frac{\hat{r}_t^2}{\sigma_t^2} = \alpha + \gamma D_{t-1} \hat{r}_{t-1} + u_t$$

Where: $D_{t-1} = \begin{cases} 1 \text{ se } \hat{r}_{t-1} < 0 & \text{will have: } \frac{\hat{r}_t^2}{\sigma_t^2} = \alpha + \gamma \hat{r}_{t-1} \\ 0 \text{ se } \hat{r}_{t-1} > 0 & \text{will have: } \frac{\hat{r}_t^2}{\sigma_t^2} = \alpha \end{cases}$

Use the t statistic on the single coefficient $\gamma$. In other words, the test statistic to which one does reference is based on this parameter.

The hypotheses of the test will be:

$$\begin{cases} H_0: \gamma = 0 \\ H_1: \gamma < 0 \end{cases}$$

Under the hypothesis of leverage effect $\gamma$ is negative, or the leverage effect occurs when this parameter is statistically negative since it leads to a reduction of the volatility.

### 1.2.3 Positive Size Bias Test

The test, shown below, is constructed starting from the Dummy variable $D_{t-1}^+ = 1 - D_{t-1}^-$. You will then have:

$$\begin{cases} D_{t-1}^- = 0 \\ D_{t-1}^+ = 1 \end{cases}$$

The regression to which this test refers is like that of the Negative Size Bias Test, yes differs for the sign positive rather than negative:

$$\frac{\hat{r}_t^2}{\sigma_t^2} = \alpha + \delta D_{t-1}^+ \hat{r}_{t-1}$$
\( \delta \) measures the differential effect relative to positive innovations. In this test, the system of hypothesis is given by: \( H_0: \delta = 0 \) \( H_1: \delta > 0 \), so if we accept the alternative hypothesis we are in presence of leverage effect.

1.2.4 Joint test

The hypotheses of the tests, previously exposed, can be submitted together with the order to verify the impact and influence of shocks. The reference model uses the following regression equation:

\[
\frac{\hat{\sigma}_t^2}{\sigma_t^2} = \alpha + \beta D_{t-1}^- + \gamma D_{t-1}^- \hat{r}_{t-1} + \delta D_{t-1}^+ \hat{r}_{t-1} + u_t
\]

In which the Dummy variables are as previously formulated. The hypothesis system is given by:

\[
\begin{cases}
H_0: \beta = \gamma = \delta = 0 \\
H_1: \text{at least one of the parameters is different from 0}
\end{cases}
\]

The null hypothesis of absence of leverage effect corresponds to the hypothesis that there are no effects differentials of any kind, neither on the average nor on the angular coefficients. This hypothesis can be checked with a regression F test.

2. Variants of the GARCH model

We can move on to the introduction of the models that provide for presence of asymmetry.

2.1 IGARCH model

An IGARCH model, Integrated GARCH, is obtained if the self-regressive polynomial of the GARCH component, \( \sum_{t=1}^{q} \beta_t h_{t-1}^2 \), has at least one-unit root. In other words, the IGARCH model is non-stationary in variance so it proves invaluable in the case where the conditional variance is strongly autocorrelated, i.e. in cases where shock suffered variance in the past have repercussions on its future values.

The IGARCH(1,1) model turns out to be:

\[
\sigma_t^2 = w + \beta_1 \sigma_{t-1}^2 + \alpha_1 r_{t-1}^2
\]
In this model, the impact of the squares of the delayed shocks, $\eta_{t-i} = \epsilon_{t-i}^2 - \sigma_{t-i}^2$, on $\epsilon_t^2$ with $i>0$ is persistent. In practice you will have:

$$\alpha_i, \beta_i > 0$$

and

$$\sum \alpha_i + \sum \beta_j = 1$$

Consequently, the IGARCH model (1.1) can be formulated as:

$$\sigma_t^2 = w + \beta_1 \sigma_{t-1}^2 + (1 - \beta_1) r_{t-1}^2$$

Where precisely, note that:

$$\alpha_1 + \beta_1 = 1$$

From which it is possible to obtains parameter: $\alpha_1 = 1 - \beta_1$

In an IGARCH model, the volatility of a period affects all forecasts relating to any subsequent period. A given shock on conditional variance is persistent for any future time horizon, becoming a significant component in the long-term period.

### 2.2 EGARCH model

The Exponential GARCH, EGARCH model proposed by Nelson (1991) is applied to rates of return on equities, for which the variance logarithm is assumed conditioning depends on a parametric set that is not necessarily and strictly positive. In other words, the model considers not only the effect asymmetric innovation, but also the proportionality of the effect itself compared to the intensity of the innovations that determine it.

In the formulation of an EGARCH(p,q) model one must bear in mind that Nelson considered weighted innovation:

$$g(u_t) = \alpha u_t + \gamma [u_t - E(|u_t|)]$$

In which $\alpha e \gamma$ are real constant and where $[u_t - E(|u_t|)]$ are random variables with zero mean if the standardized innovations are normally distributed. It is supposed, moreover, that $E[g(u_t)] = 0$.

The model is characterized by the fact that the conditional variance looks like one asymmetric function of past values of returns; this aspect made it possible to deal with the
modeling of returns on securities for which one was recognized negative correlation
between returns and volatility.

The asymmetric effect of the function $g(u_t)$, emerges from the fact that:

$$g(u_t) = \begin{cases} (\alpha + \gamma)u_t - \gamma E(|u_t|) \text{ se } u_t \geq 0 \\ (\alpha - \gamma)u_t - \gamma E(|u_t|) \text{ se } u_t \geq 0 \end{cases}$$

In general, an EGARCH(p;q) model can be written as:

$$\ln(\sigma_t^2) = w + \sum_{i=1}^{q} g(u_t) + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2)$$

The EGARCH(p,q) model does not require any restrictions to ensure the positivity of the
conditional variance since the logarithmic function removes all non-constraints negativity
on constraints. Ergo, not imposing constraints of non-negativity on the parameters, yes
verify the convergence problems that could arise in the GARCH models.

In the specific case of an EGARCH(1,1) model, the model equation becomes:

$$\ln(\sigma_t^2) = w + \beta \ln(\sigma_{t-1}^2) + \alpha \left( \frac{|r_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \frac{2}{\sqrt{\pi}} \right) + \gamma \frac{r_{t-1}}{\sqrt{\sigma_{t-1}^2}}$$

Note that, any value assumed by the parameters, the exponential function ensures the non-
negativity of variance.

We have that:

- $\beta \ln(\sigma_{t-1}^2)$ captures the effect of persistence in volatility and stationarity is secured
  by condition $0 < \beta < 1$
- The term:
  $$\left( \frac{|r_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \frac{2}{\sqrt{\pi}} \right)$$
  is a zero-average random variable that allows for taking the possibility into account
  of an asymmetrical reaction proportional to the innovations.
- The asymmetric effect is highlighted by the term:
  $$\gamma \frac{r_{t-1}}{\sqrt{\sigma_{t-1}^2}}$$
In which $r_{t-1}$ can take any sign while the parameter $\gamma$ is negative for consider the leverage effect.

The estimated conditional variance is obtained by using an exponential function of the values provided in the model:

$$\hat{\sigma}_t^2 = \hat{w} \hat{\beta} e \left( \frac{\hat{\alpha} |r_{t-1}| + \hat{\gamma} r_{t-1}}{\hat{\sigma}_{t-1}^2} \right)$$

In which: $\hat{w} = \left( w - \alpha \frac{2}{\sqrt{\pi}} \right)$

The coefficients $(\alpha + \gamma)$ and $(\alpha - \gamma)$ show the leverage effect related to past performance, $r_{t-1}$. Assuming that $\gamma < 0$, two cases can occur:

1. Positive shock: in this case the past performance will have an overall impact equal to $(\alpha + \gamma) < \alpha$.
2. Negative shock: the effect will be $(\alpha - \gamma) > \alpha$. In such a situation there will be an effect amplified.

The EGARCH model presents strong complications in the estimation phase given that the procedure requires numerous iterations before achieving convergence. Further limit of this model is represented by the fact that:

$$E[\ln(\sigma_t^2)] \neq \ln(\sigma_t^2)$$

that is, the limit lies in the forecast of conditional variance. The calculation of forecasting of conditional variance is not immediate as it is structured by analyze the logarithm of the conditional variance and not $\sigma_t^2$ same; note that in making the estimate of variance is using the conditional mean, however the EGARCH model provides, at most, the expected value of $\ln(\sigma_t^2)$.

### 2.3 TGARCH model

The model with asymmetrical effects proposed by Glosten, Jagannathan and Runkle, in 1993, e from Zakoian, in 1994, it is the TGARCH or the Theshold GARCH.

The main objective of this model is to capture the different asymmetrical effects that it has positive and negative innovations have on the conditional volatility of returns. The goal is
to capture the different behavior in correspondence of the crossing by delayed innovation of a certain threshold value, which in general it is set equal to zero. As a result, positive innovations have a less impact compared to negative innovations of equal intensity.

The main feature of the model is the fact that it is not modeled variance, but the conditioned standard deviation, \( \sigma_t = \sqrt{\sigma_t^2} \).

The model can be formulated as follows:
\[
\sigma_t^2 = w + \alpha r_{t-1}^2 + \beta \sigma_{t-1} + \gamma D_{t-1} r_{t-1}^2
\]
in this formulation, the Dummy variable takes value 1 in correspondence with values negative. You will have:
\[
\sigma_t^2 = \begin{cases} 
    w + (\alpha + \gamma) r_{t-1}^2 + \beta \sigma_{t-1} & \text{se } r_{t-1} < 0 \\
    w + \alpha r_{t-1}^2 + \beta \sigma_{t-1} & \text{se } r_{t-1} \geq 0
\end{cases}
\]
The coefficient \( \gamma \) measures the differential effect that negative shocks have with respect to those positive. It is assumed that this coefficient is positive.

In the TGARCH(1,1) model, the model takes the following form:
\[
\sigma_t = w + \alpha^+ r_{t-1}^+ + \alpha^- r_{t-1}^- + \beta \sigma_{t-1}
\]
Where: \( r_{t}^+ = \max\{r_t; 0\} \) and \( r_{t}^- = \min\{r_t; 0\} \). The equation can be parametrized in the following way:
\[
\sigma_t = w + \alpha |r_{t-1}| + \beta \sigma_{t-1} + \gamma D_{t-1} r_{t-1}
\]
In this way, in the TGARCH(1,1) model the presence-conditioned standard deviation the same functional form as the GJR(1,1) model.

The GJR model proposed by Glosten, Jagannathan and Runkle in 1993 represents an extension of the GARCH model in which a term whose function is included the main one is to capture the asymmetric evolution of conditional variance.

In a GJR(p,q) model the dynamic equation of conditional variance, given p=q=1, is given by:
\[
\sigma_t^2 = w + \alpha r_{t-1}^2 + \gamma D_{t-1} r_{t-1}^2 + \beta \sigma_{t-1}^2
\]
Again, the coefficient \( \gamma \) measures the different impact that negative shocks have on conditioned variance. To be linear with Black's considerations (1976), the expected sign of
this coefficient is positive: volatility is amplified by negative shocks. Finally, a coefficient $\gamma$ statistically non-zero indicates the absence of effects.

2.4 AGARCH model

In the GARCH model the predictions on future values of variance are in linear relation with those relating to the present and the past, while the squares of innovations serve as consideration for the forecasts themselves.

The Absolute GARCH, AGARCH model, the volatility is identified based on power $\sigma^\lambda$ rather than the standard deviation itself. This family is defined by the following formulation, in which the conditional standard deviation is raised to the power $\lambda$.

$$\frac{\sigma^\lambda_t - 1}{\lambda} = w + \beta \left( \frac{\sigma^\lambda_{t-1} - 1}{\lambda} \right) + \alpha \sigma^\lambda_{t-1} [f(u_{t-1})]^\nu$$

where $\sigma_t = \sigma_t^{1/2}$. When $\lambda < 1$ the representation of this model is given by a function concave, otherwise the function is convex.

The function $f(u_t)$ contained in the model is the curve that shows the impact that each new one event has on volatility: it always has positive value to prevent volatility from being less than zero. The function $f(u_t)$ raised to a generic power ensures that the volatility is identified based on a power of $\sigma_t^{1/2}$ rather than from $\sigma_t^{1/2}$ same. Pagan and Schwert (1990) introduce, for this equation, the following expression:

$$f(u_t) = |u_t - b| + c(u_t - b)$$

The coefficients $b$ (displacement parameter) and $c$ (angular coefficient) show two effects distinct asymmetry and cannot be null at the same time. In this context, it is important to point out that while $b \in [-\infty; +\infty]$, it must be $|c| \leq 1$, to make sure that $f(u_t)$ always has a positive value.

From the formulation of the AGARCH model it is possible to derive, through appropriate hypotheses on the parameters, other types of the ARCH family that contemplate a delay both for $\sigma^2_t$ both for $\sigma^2_t$; indeed:

- For $\lambda = v = 2$, $b = c = 0$ e $\beta = 0$ we obtain the formulation of the ARCH model;
- For $\lambda = v = 2$, $b = c = 0$ we obtain the formulation of the GARCH model;
- For $v = 1$, $b = 0$ e $\lambda \to 0$ we obtain the formulation of the EGARCH model;
For \( \lambda = v = 1, b = 0 \) and \( |c| \leq 1 \) we obtain the formulation of the TGARCH model;

For \( \lambda = v = 1 \) and \( |c| \leq 1 \) we obtain the formulation of the AGARCH model;

For \( \lambda = v = 2 \) and \( b = 0 \) we obtain the formulation of the GJR model.

For \( \lambda = v, b = 0 \) and \( |c| \leq 1 \) we obtain the formulation of the APARCH model, that is, the formulation of an Asymmetric Power model. The formulation of the model, proposed by Ding, Engle and Granger (1993), is the following:

\[
\sigma_t^2 = w + \sum_{i=1}^{q} \alpha_i (|u_{t-i}| - \gamma u_{t-i})^\nu + \sum_{j=1}^{p} \beta_j \sigma_t^{\nu_j}
\]

3. News Impact Curve

The models for the study of conditional heteroskedasticity, belonging to the family GARCH, are common tools in many economic and financial applications. In to assess the leverage effect allowed by these models, a useful tool is the news impact curve (NIC).

Engle and Ng characterize a range of alternative models for the volatility conditioned by a so-called news impact curve, which describes the impact of the latest shock of returns, or news, on current volatility, maintaining all information dated to (t-2). In other words, the News Impact Curve produces a certain value, \( \sigma_t^2 \), set the other variables of the model to the stationary values. Therefore, measuring the relationship of volatility to achievements of innovations, one can graphically represent how innovations translate into volatility.

3.1 Formulazione della NIC per le varianti del modello GARCH

The News Impact Curve characterizes the impact of past performance shocks on implied volatility in a volatility model. It reflects, in a GARCH context, the asymmetry and leverage effects of volatility, also measures how information is incorporated in volatility estimates.

Using the news impact curve, Engle and Ng compare the various models of asymmetric volatility and conclude that the models differ in the way they accommodate the asymmetry. To explain this phenomenon, they categorize return shocks or the term error of the average equation by collectively measuring news. So, a positive return - an unexpected increase in the price - suggests the arrival of some good news, while a negative return - an unexpected drop in prices - denotes the arrival of some bad one’s news. Similarly, a great value of returns in absolute value implies that the news is "significant" or "large" in the sense that it produces a great change unexpected in the price.
These curves are used to visually examine if there are effects of asymmetry in the volatility for a set of data.

We specify that the curves of impact of the news, subsequently obtained in the models, are defined *Theoretical NICs* as they are derivable as model properties for variance conditioning. They differ from the *Empirical NICs* which, instead, are estimated based on the estimated coefficients.

Referring to the theoretical news impact curve and considering a GARCH(1,1) model, the formulation of the same, defined by $NIC: \mathbb{R} \rightarrow \mathbb{R}^+$, will be:

$$NIC_{(GARCH)}(r_{t-1})\sigma_t^2 = A + \alpha_1 r_{t-1}^2$$

In which $A = w_0 + \beta_1 \bar{\sigma}$

$\bar{\sigma}$ represents the unconditional standard deviation of the historical series of returns that they are analyzed.

Graphically you will have:

![Figure 5: News Impact Curve for the GARCH (1,1) model](image)

Note that, given the limitations of the GARCH model, the representation of the NIC (figure 5) will be a perfectly symmetrical curve. This implies that, to assess the impact of past performance shock on implied volatility, the model is inefficient.
If we consider an EGARCH (1,1) model, the NIC formulation will be defined in following way:

\[ NIC_{EGARCH}(r_{t-1}): \sigma_t^2 = \begin{cases} 
Ae^{\frac{\alpha+\gamma}{\sigma}e_{t-1} \sigma_{t-1}^2} & \text{if } e_{t-1} \sigma_{t-1}^2 > 0 \\
Ae^{\frac{\alpha-\gamma}{\sigma}e_{t-1} \sigma_{t-1}^2} & \text{if } e_{t-1} \sigma_{t-1}^2 \leq 0 
\end{cases} \]

In which \( \sigma^2 = \beta_1 e^{\alpha w + \frac{\alpha^2}{\sqrt{2\pi}}} \)

An alternative version:

\[ NIC_{EGARCH}(r_{t-1}): \ln(\sigma_t^2) = w + \alpha \frac{|r_t|}{\sigma} + \gamma \frac{r_t}{\sigma} + \beta \ln(\sigma) \]

The presence of the leverage effect, in this case, implies that the parameter is negative, \( \gamma < 0 \); more generally, the asymmetry in the conditional variance will occur if \( \gamma \neq 0 \).

Graphically we have:

![Figure 6: News Impact Curve for the EGARCH(1,1) model](image)

As can be seen from Figure 6, the substantial equality of \( \hat{\alpha} \) and of \( \hat{\gamma} \) makes the constant branch of the NIC corresponding to positive innovations, while the reactivity of the function is greater at negative innovations.
The News Impact Curve can be examined for many other models. If we consider the GJR(1,1) asymmetric mode, the news impact curve is defined as follows:

$$NIC_{GJR}(r_{t-1}): \sigma_t^2 = w + \alpha_1 r_{t-1}^2 + \beta_1 \bar{\sigma} + \gamma r_{t-1}^2 I(r_{t-1} < 0)$$

Graphically:

![Figure 7: News Impact Curve for the GJR(1,1) model](image)

While, for the AGARCH(1,1) model the formulation of the News Impact Curve will be defined as follows:

$$NIC_{AGARCH}(r_{t-1}): \sigma_t^2 = A + \alpha (r_{t-1} + \gamma)^2$$

Where, remember that $\sigma_t^2$ is the conditional variance at time $t$, $r_{t-1}$ is the unpredictable return at time $t-1$ and finally $A = \omega + \beta \sigma^2$ ($\omega$ is a constant parameter, $\beta$ is the parameter of the conditional variance, $\alpha$ e $\gamma$ are parameters of the squared term). The loss of the curve is indicative of the cases with:

$$\omega > 0, 0 \leq \beta < 1, \sigma > 0, 0 \leq \alpha < 1, (\alpha + \beta) < 1 \text{ e } \gamma > 0$$

Finally, considering the TGARCH(1,1) model, the formulation of the News Impact Curve will be defined by:
\[ NIG_{TGARCH}(r_{t-i}) : \sigma_t^2 = \begin{cases} A + \alpha_1 r_i & \text{if } r_i \geq 0 \\ A + \alpha_2 r_i^2 + \gamma r_i^2 & \text{if } r_i < 0 \end{cases} \]

In which \( A = w + \beta_1 \sigma^2 \)

Graphically:

**Figure 8: News Impact Curve for the TGARCH(1,1) model**

In the graphic comparison between the news impact curves, we note that the GARCH model and the EGARCH model differ in two fundamental aspects:

1. The EGARCH model allows good news and bad news to have a different impact on volatility; the standard GARCH model does not allow this differentiation.
2. Unlike the standard GARCH model, the EGARCH model allows the great news to have a different impact on volatility.
Figure 9 compares the two News Impact Curves identified in the models; in model GARCH this curve is quadratic and centered on $r_{t-1} = 0$, while in the model EGARCH the NIC has its minimum in $r_{t-1} = 0$ and increases exponentially in both directions but with different parameters. It follows that, the returns, $r_t$, have an impact quadratic in the GARCH model and exponential in the EGARCH model; this consequence shows that for larger shocks the response of the EGARCH model is greater than the response of the GARCH model as the exponential function dominates the quadratic one.

In the comparison between NICs identified in the GJR model and in the AGARCH model one can note that the news impact curves identified in the GJR model capture the effects asymmetric volatility more efficiently.
From figure 10 it is noted that the News Impact Curve of the GJR model is characterized by a greater slope on its negative side compared to the positive side; on the other hand, the model AGARCH allows the NIC to be centered on a positive value of $r_{t-1}$.

These differences between the impact curves of the news in the different models have important implications for portfolio selection and asset price. Since the volatility of the market is linked to the risk premium, the two models imply premiums for the risk of all different. The differences in the impact curves of the news given by the models imply important implications for option prices. A significant difference in the Expected volatility after the arrival of some important news brings a difference significant in the current price of the financial asset.
Figure 11: News Impact Curve

Figure 11 shows the different impact curves of the news obtained in the various models of the GARCH family. From the figure it can be seen how the NICs formulated for the ARCH models and GARCHs are symmetrical, while the NICs formulated in the other models show substantial asymmetries with respect to negative news.

It should of course be emphasized that news affects volatility in the same way both in long term than in the short term. With a wide range of models for volatility, each with a different specification for the conditional variance dynamics, can be difficult to determine the precise effect of a shock on the variance itself.

News impact curves measure the effect of a shock in the current period on the conditional variance in the subsequent period and facilitate model comparison.

The setting of variance in the current period on unconditional variance has two consequences:

1- First, it ensures that the NIC does not depend on the level of variance;
2- Improve the comparison between the linear and non-linear specifications of the different models.

It should also be emphasized that the impact curves of the news introduced so far are a way convenient to summarize the effect of the news on the conditional volatility of the model; by comparing these curves with the volatility forecasting models one can highlight the difference between the same models. In other words, testing whether the news impact curve of a model offers good adaptability to the data, on can understand the quality of the template. In conclusion, all models detect that negative shocks introduce greater volatility compared to positive shocks, with particularly evident effects if such shocks are larger, however the
model to be considered best for observing these effects is the one proposed by Glosten, Jagannathan and Runkle (GJR).

4. The prediction of variance

The variance of a financial series is a fundamental parameter in determining the optimal portfolio of the investor: in the definition of the latter it is necessary to find the right compromise between expected average return and riskiness, measured precisely by the variance. The forecasts of the volatility of financial instruments, in other words, are of great interest for operators as they are the basis of investment strategies.

The present value of volatility can be calculated using two alternative methods:

- Subjective probability measure: standard deviation of historical observations;
- Risk-neutral probability measure: value of the volatility that equals risk-neutral price at the market price.

The future value of volatility can instead be calculated using forecasts.

The modeling and forecasting problem have been addressed in the literature with one multiplicity of tools and different purposes. As for the purposes, a first one classification may be as follows:

1. Forecasting models of the price;
2. Descriptive models;
3. Forecasting models of volatility.

By forecasting models, we mean the models whose main purpose is precisely the forecast of the price, while with descriptive models we mean models that aim at a better one description of the behavior of volatility, which does not necessarily lead to one best forecast. For example, knowing that in addition to the "normal" component of the volatility dynamics there is a discontinuous component, "of leap", with certain statistical properties, it does not necessarily improve the forecast but can be useful for purposes such as the evaluation of as many derivative instruments.
The formulation of the forecast is based on the knowledge and reworking of information processed by the past events of the historical series under examination, it is an operation of "connection between past and future"\textsuperscript{14}.

4.1 Classification of forecasts

The classification of the forecasts can be made based on the main one’s characteristics of the same as the nature, the object or the horizon of the same.

Based on the nature of the forecasts, these are distinguished in qualitative or forecasts quantitative; however, both aspects often occur in complex phenomena.

The object of the forecast can be the future value of a phenomenon that manifests itself with continuity or can be detected at regular intervals; or the time in which a phenomenon, or the modalities and characteristics of an event that will occur in the future.

The horizon of the forecast, that is the distance in the future of the events that are expected, it allows to classify the forecasts in short, medium and long term. However, the distinction is not clear and precise. We talk about short-term forecasts when the structural conditions remain unchanged, vice versa we speak of a forecast a long horizon (or long term) if the fundamental conditions that determine the event to be expected are still substantially uncertain.

A further classification concerns the dimensions, for this purpose the forecasts are divided into unit-varied forecasts and multivariate forecasts. The first concern only one phenomenon, while the latter can simultaneously concern several connected phenomena.

A further distinction of the forecasts is given by: static forecasts and forecasts dynamics. In the first case the information is enriched over time and is used to predict conditional variance a forward period; in the second case, yes hypothesizes that the available information runs out with the final period T and a forecast for a generic value $\tau \geq 2$ must be conditioned by the information set in T.

About static forecasts, note that they are applied to phenomena conceptually defined to be subject to objective measurement. In static forecasts, since phenomena that are developed are taken into consideration over time, the data present precise relationships of dynamic dependence, which involve not only their value, but also their mutual position over time.

\textsuperscript{14}Cipolletta, 1992
The purpose of the static forecasts is to study the future manifestations of a phenomenon determined by the permanence of a stability in the general structural properties that have existed verified in past. Note that, with structural properties we mean the characteristics of the statistical and probabilistic mechanism that generates the phenomenon; it requires that these are static to predict the phenomenon. The request for stability between the past and future is minimal to make sense of the statistical forecast, because if it came to miss then the knowledge of the past would be insufficient to formulate guesses about the future.

Unlike the static forecasts, in the dynamic forecasts the information set is not updates adding the sample observations, but the forecasts are used as these are derived. In other words, in the dynamic forecast the values provided are based on elements subject to uncertainty, which leads to greater probability of forecast error. In fact, through dynamic forecasting they are obtained "less precise" forecasts.

Finally, let us consider the distinction between ex post forecasts and ex ante forecasts.

The ex post forecasts concern the reconstruction of the historical series of interest on past performance. That is, they are inferences on the values assumed by the variables of interest in the sample range, therefore on already known quantities.

In the ex-post forecasts, the observations are divided into two sets:

1- Estimation set, indicated with "T-H", used in the specification and estimate of parameters of the forecaster;
2- Forecast set, indicated with "H", used to judge its capacity ex post forecast.

Note that with "T" the information set available is indicated while with "H" yes represents the set of forecasts made. In the case of ex-post precision, it is advisable to calculate dynamic forecasts; therefore, for each forecast a model must be estimated in which a new observation is inserted and calculate the forecast of the next day: if the set H consists of 30 days, they will get 30 models for the 30 forecasts.

In the ex-ante forecasts, instead, future values are generated, not immediately observable.
The ex-ante forecasts take into consideration the period starting after the last known instant and is prolonged indefinitely in the future, and they are intended to describe the properties and characteristics that the studied phenomenon presents in its unfolding in time.

The key assumption is that past and future events are believed to be generated by a unique stochastic process, whose main properties remain stable and therefore apply both to the past and to the future.

The study and inference of these properties are carried out in based on the set of known information, and therefore mainly using the manifestations already detected of the phenomenon, or the data already known at the time of the forecast.

To generate future values, involving the use of an estimator: $\hat{Z}_t(h)$. In symbols:

$\hat{Z}_t(h) = E[Z_{t+h}|I^t]$  

This estimator is defined as the conditional expected value of the reference time series given the set of information available at time $t$. The forecaster $\hat{Z}_t(h)$ is a linear function of the random variables contained in the information set $I^t$; in general, for processes Gaussians, linear forecasters are optimal because they minimize the mean square error.

### 4.2 Forecasting schemes

The out of sample forecasts to be used in the models to make comparisons are typically calculated through forecasting schemes.

It should be noted that in the out-of-sample forecasts the series of a forecast obtained afterwards of the estimation of a model, consists of several successive forecasts which represent the forecasting horizon.

The most commonly used forecasting schemes are:

1) **Fixed scheme**: fixed forecast scheme that considers only the information from 1 to T without updating the information set without using a scroll window.
Specifically, this forecasting scheme estimates the parameters of the model only once per the entire forecast period. For the subsequent forecasts, we continue to use them estimated parameters on information from 1 to T.

2) **Recursive scheme**: in this scheme the forecaster uses all the information available at time T to predict observations at time T + 1. Once the forecast at time T + 1 was generated, to generate the forecast for the period next T + 2 the information set will be expanded with information, from 1 to T + 1.

In this case the parameters are re-estimated each time a new one is added information and the breadth of the information set increases with each step.

3) **Rolling scheme**: the scheme uses a sliding window of width equal to w starting from T - w + 1 up to T. In other words, we start with an initial data sample from \( t = 1, ..., T \) used to identify a time window of width T, per estimate the model. The next steps are predicted starting from the same T. The next step the time window moves forward one step and the model comes again estimated using data from \( t = 2, ..., T + 1 \); consequently, the subsequent steps they are produced starting from the time T + 1.

Also, in this case the parameters are estimated always using the same amplitude window sample. The use of a fixed sample size allows for more effective evaluation of the forecasts in comparative terms.

5. **Volatility forecast with stochastic model**

The problem of forecasts can be formalized in the context of process theory stochastic studying the relations of a set of variables that refer to the past and present with a process variable referring to the future.

Models of the GARCH family can be used to predict variance conditioned based on the information set available at a given moment, assuming let T be the last instant used for the estimate.

For simplicity we consider a model GARCH (1,1):

\[
\sigma_t^2 = w + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2
\]

In general, the best forecast is given by the conditional expected value of the observation: \( E[x_{T+h} | I^T] \); if we suppose that we want to make forecasts one- step- forward in values
future volatility ($\sigma_{T+1}^2$), calculate its expected value $\hat{\sigma}_{T,1}^2 = E(\sigma_{T+1}^2 | I^T)$, that in the model considered it is given by:

$$\hat{\sigma}_{T,1}^2 = w + \alpha_1 E(r_1^2 | I^T) + \beta_1 E(\sigma_1^2 | I^T) = w + \alpha_1 \hat{r}_T^2 + \beta_1 \hat{\sigma}_T^2$$

In which $\hat{r}_T$ represents the residue of the equation estimation of the average and $\hat{\sigma}_T$ represents the value estimated conditional variance at time $T$.

The information is enriched over time and is used to predict the variance conditioned a forward period. In a broader context, for a forecast h- pass-forward you will have:

$$\hat{\sigma}_{T,s}^2 = w \sum_{i=1}^{s-1} (\alpha_1 + \beta_1)^{i-1} + (\alpha_1 + \beta_1)^{s-1} \hat{\sigma}_{T,1}^2$$

Similarly, for the TGARCH (1,1) model, if the parameter estimation is assumed carried out only once on the original sample, you will have:

$$E(r_{T+\tau+1}^2 | I_{T+\tau-1}) = \bar{\omega} + \alpha_1 \hat{r}_{T+\tau-1}^2 + \beta_1 \hat{\sigma}_{T+\tau-1}^2 | I_{T+\tau-2} + \gamma D_{T+\tau-1} \hat{\sigma}_{T+\tau-1}$$

While for the EGARCH (1,1) model we will have:

$$E(r_{T+\tau+1}^2 | I_{T+\tau-1}) = \bar{\omega} \sigma_{T+\tau-1}^2 | I_{T+\tau-2}^2 \exp \left( \frac{\hat{\alpha} |\hat{r}_{T+\tau-1}| + \hat{\gamma} |\hat{r}_{T+\tau-1}|}{\sqrt{\hat{\sigma}_{T+\tau-1}^2 | I_{T+\tau-2}^2}} \right)$$

In a context of dynamic forecasting, so if we assume that the information a disposition ends with the period $T$, the forecast for a generic horizon $\tau \geq 2$ must be conditioned to the information set in $T$.

For $\tau = 2$ we will have:

$$E(r_{T+2}^2 | I_T) = \bar{\omega} + \alpha_1 E(r_{T+1}^2 | I_T) + \beta_1 \hat{\sigma}_{T+1}^2 | I_T = \bar{\omega} + (\alpha_1 + \beta_1) \hat{\sigma}_{T+1}^2 | I_T = \sigma_{T+2}^2 | I_T$$

Similarly, for $\tau = 3$ we can write:

$$E(r_{T+3}^2 | I_T) = \bar{\omega} + \alpha_1 E(r_{T+2}^2 | I_T) + \beta_1 \hat{\sigma}_{T+2}^2 | I_T = \bar{\omega} + (\alpha_1 + \beta_1) \hat{\sigma}_{T+2}^2 | I_T$$

$$= \bar{\omega} + (\alpha_1 + \beta_1)(\bar{\omega} + (\alpha_1 + \beta_1) \hat{\sigma}_{T+1}^2 | I_T)$$

$$= \bar{\omega} + (1 + (\alpha_1 + \beta_1))(\alpha_1 + \beta_1) \hat{\sigma}_{T+2}^2 | I_T$$

$$= \bar{\omega} + (1 + (\alpha_1 + \beta_1))(\alpha_1 + \beta_1) \hat{\sigma}_{T+2}^2 | I_T$$
For a generic horizon $T - \tau$, the forecast of conditional variance is equal to:

$$E(r_{T+\tau}^2|I_T) = \tilde{\omega}(1 + (\tilde{\alpha}_1 + \tilde{\beta}_1) + \cdots + (\tilde{\alpha}_1 + \tilde{\beta}_1)^{\tau-2} + (\tilde{\alpha}_1 + \tilde{\beta}_1)^{\tau-1} \tilde{\sigma}_{T+1|T}^2$$

From which we note that the forecasts are a function of the information available and of a coefficient, successive powers of $(\tilde{\alpha}_1 + \tilde{\beta}_1)$, which decreases exponentially as the forecast horizon increases. In fact, for $\tau \to \infty$ we will have it $(\tilde{\alpha}_1 + \tilde{\beta}_1) \to 0$; given that $(\tilde{\alpha}_1 + \tilde{\beta}_1) < 1$ the prediction converges to:

$$\lim_{\tau \to \infty} E(r_{T+\tau}^2|I_T) = \lim_{\tau \to \infty} \sum_{i=0}^{\tau-2} (\tilde{\alpha}_1 + \tilde{\beta}_1)^i = \sum_{i=0}^{\infty} \tilde{\omega}(\tilde{\alpha}_1 + \tilde{\beta}_1)^i = \frac{\tilde{\omega}}{1 - \tilde{\alpha}_1 - \tilde{\beta}_1}$$

About the dynamic forecast for the TGARCH model, note that the presence of the asymmetric component modifies the profile of the forecasts, mainly for the forecast a period ahead. In fact, the forecast, if the last estimated innovation turns out to be positive, will be:

$$E(r_{T+\tau}^2|I_T) = \tilde{\omega} + \tilde{\alpha}_1 r_{T}^2 + \tilde{\beta}_1 \tilde{\sigma}_{T}^2$$

In contrast, if the last estimated innovation turns out to be negative, the forecast will be:

$$E(r_{T+\tau}^2|I_T) = \tilde{\omega} + \tilde{\alpha}_1 r_{T}^2 - \tilde{\gamma} r_{T}^2 + \tilde{\beta}_1 \tilde{\sigma}_{T}^2$$

Once the forecast has been identified, a forward period, starting from time $T + 1$, is the sign innovation is not known and the probability of obtaining a positive sign is the same as that to get a negative sign: $E(D_{T+\tau}^-) = \frac{1}{2}$, $\forall \tau \geq 2$, we will have:

$$E(r_{T+\tau}^2|I_T) = \tilde{\omega} + \tilde{\alpha}_1 E(r_{T+1}^2|I_T) + \tilde{\gamma} E(r_{T+1}^-|I_T)E(D_{T+\tau}^-|I_T) + \tilde{\beta}_1 \tilde{\sigma}_{T+1|T}^2$$

$$= \tilde{\omega} + (\tilde{\alpha}_1 + \frac{\tilde{\gamma}}{2} + \tilde{\beta}_1) \tilde{\sigma}_{T+1|T}^2$$

And so on for a generic horizon $\tau > 2$. For a growing forecast horizon to the ha convergence to an estimate of the non-conditional variance different from the GARCH case:

$$\frac{\tilde{\omega}}{1 - \tilde{\alpha}_1 - \frac{\tilde{\gamma}}{2} - \tilde{\beta}_1}$$

Finally, we consider the dynamic forecast for the EGARCH model. In this context, the forecast is based on the following expression referring to the period $T + 1$:
\[ \hat{\sigma}_{T+1|T}^2 = \hat{\omega}^* \hat{\sigma}_T^\beta \exp \left( \frac{\hat{\alpha}|\hat{T}_T| + \hat{\gamma}_h}{\sqrt{\hat{\sigma}_T^2}} \right) \]

For the forecast at the time T + 2 conditional on the information at time T, we will have:

\[ E \left( \exp \left( \frac{\hat{\alpha}|\hat{T}_{T+1}| + \hat{\gamma}_{T+1}}{\sqrt{\hat{\sigma}_{T+1}^2}} \right) \right) = \hat{K}^* \]

The result \( \hat{K}^* \) is a constant, this is since standardized random variables they have hypotheses for moments that are independent of time. Therefore, for T + 2 the forecast can be formulated as follows:

\[ \hat{\sigma}_{T+2|T}^2 = \hat{\omega}^* \hat{K}^* \hat{\sigma}_{T+1|T} = \hat{R} \hat{\sigma}_{T+1|T}^\beta \]

Similarly, for T + 3 you will have:

\[ \hat{\sigma}_{T+3|T}^2 = \hat{K} \hat{\sigma}_{T+2|T} = \hat{R} (K \hat{\sigma}_{T+1|T})^\beta = \hat{R} (K \hat{\sigma}_{T+1|T})^\beta \]

Finally, for a generic horizon \( \tau \) we get

\[ \hat{\sigma}_{T+\tau|T}^2 = \hat{K} \hat{\sigma}_{T+\tau-1|T} = \hat{R} (K \hat{\sigma}_{T+\tau-1|T})^{\beta} = \hat{R} (1 + \beta + \cdots + \beta^{\tau-2}) \hat{\sigma}_{T+1|T}^{\beta-1} \]

This forecast converges, for \( \tau \to \infty \), to: \( \hat{K} \frac{1}{1-\beta} \) since \( \beta < 1 \).

### 5.1 Forecast errors

From a statistical point of view, to predict means to determine with the least possible error the realization of a random variable through the realization of other variables random. Therefore, it is necessary to choose a loss function and determine the forecaster excellent, that is the function of the observable variables, which minimizes the expected loss.

In fact, in predicting future values of volatility you can incur errors defined as "errors foresight". The latter take on a central role and represent the difference between the series of real and expected values. Analytically

\[ e_t = z_{t+h} - \hat{z}_t(h) \]

The values, associated with an optimal forecaster, should respect certain properties to generate errors that:

- Have an average of zero and consequently they can be defined as correct;
- Do not underestimate or overestimate the future values of the variable of interest;
- They are compatible with that observed in the estimation interval.

The estimator for the conditional variance $\hat{\sigma}_t(h)$ cannot calculate the forecast error since the real value is not directly observable (the variance is a latent variable, not directly observable. To this end, volatility is used as a proxy for returns daily squared or in absolute value).

In predicting, there are situations in which an error by default involves much higher costs than an error in excess. Note that a cost function generally involves a forecaster optimal distorted. In other words, the expected value of the forecast error is not zero, because the loss function implies more frequent negative values of the forecast error.

Often, however, there is no explicit cost function, especially when they produce forecasts that must be used by a large mass of users. In these cases, we use a symmetrical loss function, so that the forecast is not distorted (yes will have an expected error of null forecast). For reasons linked above all to simplicity mathematics the most used loss function is the quadratic error.

The accuracy of the forecasts in terms of accuracy takes place using functions of loss such as: mean square error and mean absolute error.

$$MSE = \frac{1}{k} \sum_{\sigma=1}^{k} (r_{T+H}^2 - \hat{\sigma}_{T;H}^2)^2$$

$$MAE = \frac{1}{k} \sum_{\sigma=1}^{k} |r_{T+H}^2 - \hat{\sigma}_{T;H}^2|$$

Note that any volatility proxy can be replaced at $r_{T+H}^2$ in order to calculate the loss functions.

From the mean square error loss function it is possible to obtain another loss function, or given the historical series of interest, composed of $T$ observations, and given the sequence of forecasts having the same size, the quadratic average of the forecast error defines the Root Mean Square Error (RMSE). Analytically we have:

$$RMSE = \sqrt{\frac{1}{k} \sum_{\sigma=1}^{k} (r_{T+H}^2 - \hat{\sigma}_{T;H}^2)^2}$$
This indicator can only take positive values as it is constructed as an average of squares. Its theoretical minimum value is the zero that would be configured if the forecasts are traced perfectly the observations about the dependent variable.

A further synthetic measure for the evaluation of the forecast goodness is the Theil index. Analytically this coefficient is given by:

\[
THEIL = \sqrt{\frac{1}{T-t} \sum_{t=t+1}^{T} (r_t - \hat{\delta}_t)^2}
\]

\[
\sqrt{\frac{1}{T-t} \sum_{t=t+1}^{T} (r_t)^2 + \frac{1}{T-t} \sum_{t=t+1}^{T} (\hat{\delta}_t)^2}
\]

The smaller the calculated error measure is, the greater will be considered the predictive ability of the model. The Theil index takes values between zero and one, thus signaling the two extreme situations of adaptation: very bad (value index equal to one) or perfect (index value of zero).

The Theil index is, of course, limited by 0 and 1, where the lower bound is the case where \( r_t = \hat{\delta}_t \), which is the event that you get perfect forecasts, while the upper limit is has if the index of \( THEIL = 1 \), that is, if there is the maximum inequality, such situation occurs when there is a negative proportionality or if one of the variables is identically zero.

The Theil index is closely related to the components of the average forecast error. The latter are defined as:

- **Bias proportion**: which indicates the distance of the average of the forecasts from the forecasts of the average of the real values;
- **Variance proportion**: which indicates the difference between the variability of the forecasts e that of real values;
- **Covariance proportion**: which measures the remaining non-systematic forecast error.

The main feature of these components is that they add up one; moreover, a good one forecast provides that the first two components are close to zero while the last one component is very close to one.

With attention to the bias component, it is possible to ascertain that it represents the error in the estimates caused by systematic errors that lead to consistently high results o low, compared to actual values.
In general, it is possible to state that the model that produces the smallest value of these statistics is judged to be the best. Certainly, the evaluation statistics of the prediction are random variables and a formal statistical procedure should be used to determine whether a model has a higher predictive performance.
CHAPTER 3: EMPIRICAL ANALYSIS OF THE FORECAST OF VOLATILITY OF STOCK INDICES

1. Introduction

In this chapter we will present an empirical analysis of the quotations regarding Standard & Poor's 500 (GSPC) and NASDAQ Composite (IXIC). The data-set of the series historical reference, having as source the site http://www.oxford-man.ox.ac.uk/, concern daily data from 3 January 2000 to 3 June 2019, for a total of 4778 observations.

The Standard & Poor's 500, or simply S&P500, is the most important stock index American; it contains 500 shares of companies listed in New York, representative about 80% of market capitalization. It is in fact the main stock benchmark relating to securities listed on Wall Street and the underlying for an incredibly large range of derivative products, such as futures, options and certificates.

The S&P500 can be considered the index that par excellence groups the most capitalized titles on the US market; the methodology by which companies are capitalized to be included in the basket represents one of the main differences with other stock indices. The 10 titles that currently have a greater weight in the basket and that together reach about 21% of the total, are: Apple, Microsoft Corp, Amazon, Berkshire Hathaway, Johnson & Johnson, JP Morgan Chase, Facebook, Exxon Mobil, Alphabet C and Alphabet A. As far as the individual sectors are concerned, the most represented are the IT sector (Information Technology) with 20.7%, Health Care (15.0%) and Financials (13.6%).

About NASDAQ Composite (IXIC) is concerned, it is possible to state that this acronym stands for National Association of Securities Dealers Automated Quotation. It is the first example to the world of electronic stock market, a market made up of a computer network.

The NASDAQ Composite is an index based on floating capitalization; in that index the quotations of shares of IT giants like Microsoft, Cisco are grouped Systems, IBM, Apple, Google, Yahoo and Facebook.
1.1 Preliminary analysis of prices

The historical series of reference include the adjusted closing prices (AdjClose) of the stock indices\textsuperscript{15}. The adjusted closing price, in general, changes the closing price of an action to accurately reflect the value of such stock after accounting for any corporate actions. It is considered the true price of a security and is used when the returns are examined, or a detailed analysis is carried out.

The Time Plot graphs to the stock indices shown in figures 12 and 13 show the slow and persistent of prices, in fact it can be seen that with the succession of days prices tend to be followed by similar prices. Moreover, from the price trend of both stock index it is possible to note periods of financial crisis over the years 2000/2004 and in the years 2007/2010.

The securitization process, together with the adoption of an organizational structure complex, it is one of the key factors of the financial and economic crisis that has struck before the United States and later Europe.

This process is an operation financial by which one or more illiquid and undivided assets that are present in the financial statements of the financial intermediary or of a generic transferor company and who can generate cash flows, they are converted into divided and salable assets or securities bonds denominated Asset Backed Securities (ABS).

The trend of historical price series clearly shows the various historical periods and gives an idea of the intrinsic variability of the phenomenon.

\textsuperscript{15} The indices are the basket values (portfolios) of financial securities, generally used to evaluate performance of a certain market. We distinguish stock market indices from indices that refer to bonds. The first (like S&P500 and DJI) are built by choosing the securities of the company with the highest capitalization or belonging to some specific industrial sectors, while the latter (like the EMBI and the MAE) measure the value of the debt of the emerging countries or include government bond issues.

The indices are not real financial securities, being the virtual portfolios of the latter.
In these cases, the dynamic trend under study shows structural features which exercise a similar influence throughout the survey and are then the subject of greater interest for forecasting purposes, given that we can legitimately wait for that affect similar ways in the future.

From the preliminary analysis of the historical series we essentially find the evidence of a non-stationary trend; the time plot, shown in the previous figures, shows how the evolution of prices does not fluctuate around a constant value, highlighting the presence of a (non-
linear) trend, understood as the underlying trend of the historical series of reference, characterized by a relatively slow variation over time.

It is possible, at this point, to proceed to an in-depth analysis of the returns, calculated, as observed in the first chapter, as raw logarithmic differences of the prices of the time series of interest: \( r_t \approx \Delta p_t \equiv p_t - p_{t-1} \)

### 1.2 Preliminary statistical analysis of returns

The preliminary step of any statistical analysis is to analyze the data in terms purely descriptive, with the aim of obtaining the necessary information use of more sophisticated techniques.

It is first opportune to observe the time graph of the returns, from which it is possible to draw some preliminary indications on the possible presence and on the typology of time dependence.

*Figure 14: daily returns of the S&P500 index (22 May 2000-20 May 2019)*
Figures 14 and 15 show the historical series of returns. The graphs show immediately a typical characteristic of financial returns analyzed in the first chapter, the so-called aggregation of volatility (volatility clustering). In intuitive terms, the observations adjacent tend to have a similar level of volatility, and therefore follow periods of high and low volatility. This feature, as we will see later, indicates a structure of non-linear dependence and therefore to be modeled through a self-variance approach regressive.

A numerical summary of the characteristics of the series in analysis can be obtained by the command of Gretl "Descriptive statistics".

![Figure 15: daily returns on the NASDAQ index (22 May 2000-20 May 2019)](image)

![Figure 16: descriptive statistics of the historical series of interest](image)
The results, shown in Figure 16, indicate that the distributions have a weak asymmetry and, in the case of the S&P500 negative share index, they also show a much more kurtosis wide than normal and therefore have thicker tails than normal.

Values of asymmetry and empirical kurtosis provide information not only about non-normality of the data, but also on the reason why they are not normal.

*Figure 17: frequency distribution of daily returns of the S&P500 index*

*Figure 18: frequency distribution of daily returns of the NASDAQ index*
Figures 17 and 18 show the frequency distributions of the daily yields of the historical series under examination, from which we note that the distributions are not normal, not so much because asymmetric (asymmetry does not differ much from zero) but because leptokurtic (kurtosis, in the two distributions, is much greater than three).

To verify the normality of the distributions it is also possible to make the tests of Normality. The Jarque-Bera test is based on the simultaneous evaluation of the asymmetry \( e \) of kurtosis; the system of hypotheses, explained in paragraph 3.5 of the first chapter, provides as null hypothesis that the series of distributes as a Normal while the alternative hypothesis assumes that the reference time series is not distributed as a Normal one.

The results of the test statistic obtained for the two share indices are:

- Jarque-Bera test (\(^{^\wedge} \) S&P500) = 15812, with p-value 0;
- Jarque-Bera test (\(^{^\wedge} \) NASDAQ) = 7359.49, with p-value 0;

The high values of the tests imply that the joint deviation of the asymmetry and of the kurtosis is statistically significant; moreover, observing the p-value, null in both cases, the null hypothesis is rejected (this was predictable given the high values of kurtosis).

For what concerns the study of the dependence of ordered observations in time, this can be carried out using the correlogram. The graphical representations of the functions of empirical autocorrelations, that is the correlogram, are reported in the following figures:

![Correlogram of daily returns of the S&P500 index (22 May 2000-20 May 2019)](image-url)

*Figure 19: correlogram of daily returns of the S&P500 index (22 May 2000-20 May 2019)*
Figure 20: correlogram of daily return of the NASDAQ index (22 May 2000 - 20 May 2019)

As can be seen, from figures 19 and 20, various values are outside the confidence interval; although in most cases they are not very far from the confidence limits, it seems unlikely that the time series of interest are not autocorrelated.

Out of a total of 36 global autocorrelation coefficients, for the index S&P 500, there are 7 extreme values at the 99% confidence interval and 4 extreme values at the 95% confidence interval; the percentages are respectively 7/36 = 19.44% and 4/36 = 11.11%. While on a total of 36 coefficients of partial autocorrelations, the values extremes at the 99% confidence interval represent 19.44% and extreme values at the 95% confidence interval they are 5.56%.

Similarly, for the NASDAQ index, there are 8 extreme values at the interval of 99% confidence and 4 extreme values at the 95% confidence interval; the percentages are respectively 8/36 = 22.22% and 4/36 = 11.11%. While on a total of 36 coefficients of partial autocorrelations, the extreme values at a significance level of 1% are the 19.44% and the extreme values at a significance level of 5% are 5.56%.
1.3 Analysis of the conditional expected value: the ARMA models

An examination of the range of returns of a stock index is often observed behavior characterized by constant mean but by alternating period with variance moderate and others with high variance. In this case, it is advisable to identify models that they can represent these changes in the characteristics of variability.

Conditional heteroscedasticity models, hence the GARCH model family, are applied to series with non-normal data and negligible autocorrelation; otherwise, then if the series shows a non-negligible linear dynamic structure, it is possible to build first an ARMA model and then adapt a GARCH model on the residues.

Let us first proceed with the identification of an ARMA model. An important aspect to be considered in this phase is that of the possible inclusion of a constant; note that the forecasts starting from an ARMA model are not distorted, the average of the residuals it must not be significantly different from zero. When this does not occur, it is possible introduce a constant to ensure that the average of the residues is zero. The best constant estimate is the average of the differentiated and transformed series.

In general, it is a good idea to start with very simple models such as AR (1) and MA (1) ed then add one order at a time until the estimated residues are not without autocorrelation. Experience suggests that, for data relating to stock indices, the model ARMA(1,1) is often able to capture all the autocorrelation present in the returns.

With reference to the historical series of returns of the S&P500 stock index, through the GRETL statistical software, three ARMA models have been estimated:

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Standard Error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>phi_1</td>
<td>0.418420</td>
<td>0.146902</td>
<td>2.848</td>
</tr>
<tr>
<td>theta_1</td>
<td>-0.502266</td>
<td>0.140054</td>
<td>-3.585</td>
</tr>
</tbody>
</table>

Figure 21: ARMA(1,1) model estimated for the S&P500 index (22 May 2000-20 May 2019)
In the estimation of the three ARMA models, shown in figures 20, 21 and 22, it was noted how the inclusion of a constant was not significant. The significance of the coefficients is evaluated by observing the p-values of the same; the p-value indicates the probability of obtaining an equal or "more extreme" result than the one observed. With reference to significance only of the estimated parameters in the ARMA models, it is possible to state that in the ARMA(1,1) model both parameters are significant at 99%, in the ARMA(2,2) model only the parameter phi_1 is 99% significant, finally in the ARMA(3,3) model the parameters phi_1 and theta_1 they are not at all significant.

In practice the identification of ARMA models is rather difficult and consequently Numerous automatic identification procedures are developed.
The information criteria, useful to assess the soundness of the model estimate, most frequently applied are:

- Akaike's asymptotic information criterion, defined as follows:
  \[ AIC(k) = -2 \log(\hat{\sigma}^2_k) + 2k \]
  Where \( k \) is the number of model parameters and \( \hat{\sigma}^2_k \) corresponds to the variance of residuals calculated after estimating the ARMA model \((p, q)\);

- The Bayesian information criterion (BIC) proposed in 1978, defined in the following so:
  \[ BIC(k) = -2 \log(\hat{\sigma}^2_k) + k \log(n) \]
  The term \((k \log(n))\) indicates that the penalty for adding parameters additional increases with the sample size.

According to these criteria, the parity of the parameters being the same, the preferred model is that which corresponds to the lowest value; of facts, in the ARMA models the estimated variance of the residues always decreases with the introduction of a new parameter, for which the criteria are a compromise between a function (decreasing) of the variance of the residues and a function (increasing) of the parameters.

Observing the criteria, among the three proposed models, the one to be preferred is the ARMA\((1,1)\) model. As can be seen, the model is characterized only by statistically parameters significant. By accepting that returns follow (in expected value) a process ARMA\((1,1)\), the conditional variance analysis can start from the identification of a GARCH model.

Using the Gretl statistical software, it was possible to identify the following GARCH\((1,1)\) model whose parameters are significant, and which represents the point of departure in the estimate of subsequent forecasts:
Figure 24: GARCH(1,1) model estimated for the S&P500 index

Looking at Figure 24, the GARCH model can be defined as follows:

$$\sigma_t^2 = 0.00000173457 + 0.0940596\sigma_{t-1}^2 + 0.892966\sigma_{t-1}^2$$

In continuing the empirical analysis of the S&P500 stock index it was noted that the models GARCH with parameters p and q, GARCH (1,1), GARCH (1,2), GARCH (2,1), even considering a low order of parameter content, seem to be enough to model the dynamics of variance even for large sampled periods.

Through the estimation of the parameters, first of the ARMA model and then of the GARCH model, we can manage to get better and more reliable estimates that allow more identification quick model that is best suited to describe the phenomenon in question.

1.3.1 Analysis of the residues of the GARCH(1,1) model

In defining which model GARCH is considered the best, in addition to the significance of the estimated parameters, an analysis of the residues of the same model was performed.

First, the correlogram was observed.
Figure 25: correlogram of the residuals of the GARCH (1,1) model estimated for the S&P500 index

The correlogram, represented in the figure, does not give a completely satisfactory result since the main hypothesis of the GARCH model on conditional distribution implies that the residues are not auto-correlated.

The goodness of adaptation of a GARCH model can be evaluated in relation to capacity of conditional variance of making standardized residues as close as possible to be normally distributed. To this end, one of the most commonly used tools is the residue normality test.

Figure 26: Normality test of the residual GARCH (1,1) model estimated for the S&P500 index

Figure 26 shows that the residuals of the GARCH(1,1) model are not distributed as one normal distribution. However, it should be borne in mind that the historical reference series does not only presents an excessive number of observations, as far as it shows an excess of kurtosis.
2. Analysis of volatility forecast

The GARCH models, and their asymmetrical variants, are used for forecasting purposes, taking advantage of the ARMA model local forecasting methodology. It should be kept in mind but that the forecasts will not concern the future values of the series, but their variability, whose forecast is of special interest in the study of financial markets.

To carry out an analysis efficient of the forecasts of volatility, the reference sample was reduced, which previously included observations concerning the period from 3 January 2000 to June 3, 2019; in fact, observations regarding the period are considered from 3 January 2000 to 31 December 2014. In addition, for this purpose the forecasts will be compared estimated with the realized volatility contained in the Oxford Man dataset.

The following figure shows part of the reference data set available:

<table>
<thead>
<tr>
<th>Date</th>
<th>Close</th>
<th>Adj Close</th>
<th>Log of Adj Close</th>
<th>rendimenti</th>
<th>r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>31/12/2014</td>
<td>2058</td>
<td>2058.899502</td>
<td>7.6299279095944700</td>
<td>-0.0012703264391784</td>
<td>0.0001470620047226</td>
</tr>
<tr>
<td>02/01/2015</td>
<td>2058.199551</td>
<td>2058.199551</td>
<td>7.6295876969594700</td>
<td>0.0052031378609851</td>
<td>0.001031126929295</td>
</tr>
<tr>
<td>05/01/2015</td>
<td>2020.799560</td>
<td>2020.179560</td>
<td>7.62511935861546400</td>
<td>-0.01647712347561444</td>
<td>0.0009643580011112</td>
</tr>
<tr>
<td>06/01/2015</td>
<td>2002.099500</td>
<td>2002.099500</td>
<td>7.62206601279431000</td>
<td>-0.0069323288391784</td>
<td>0.0008595952351916</td>
</tr>
<tr>
<td>07/01/2015</td>
<td>2022.099520</td>
<td>2025.909200</td>
<td>7.62517893707345000</td>
<td>0.0126724345044404</td>
<td>0.000912838954794</td>
</tr>
<tr>
<td>08/01/2015</td>
<td>2062.193930</td>
<td>2062.193930</td>
<td>7.62149995060116100</td>
<td>0.01773014636161213</td>
<td>0.0006104664607579</td>
</tr>
<tr>
<td>09/01/2015</td>
<td>2044.310670</td>
<td>2044.310670</td>
<td>7.62200101847936000</td>
<td>-0.0094395212340362</td>
<td>0.0008299574937339</td>
</tr>
<tr>
<td>12/01/2015</td>
<td>2028.099100</td>
<td>2028.099100</td>
<td>7.61493910063294000</td>
<td>-0.0091329930293970</td>
<td>0.000949734340392</td>
</tr>
<tr>
<td>13/01/2015</td>
<td>2023.099290</td>
<td>2023.099290</td>
<td>7.61523188032839000</td>
<td>-0.028188359681188</td>
<td>0.0010259831877560</td>
</tr>
<tr>
<td>14/01/2015</td>
<td>2011.270600</td>
<td>2011.270600</td>
<td>7.60652165226750000</td>
<td>-0.0583026586161861</td>
<td>0.0009383674248676</td>
</tr>
<tr>
<td>15/01/2015</td>
<td>1992.570640</td>
<td>1992.570640</td>
<td>7.59723074826553700</td>
<td>-0.0925093532103649</td>
<td>0.000876009993241</td>
</tr>
<tr>
<td>16/01/2015</td>
<td>2013.420640</td>
<td>2013.420640</td>
<td>7.61050594223417000</td>
<td>0.0133483318939934</td>
<td>0.000978775344959</td>
</tr>
<tr>
<td>20/01/2015</td>
<td>2032.550640</td>
<td>2032.550640</td>
<td>7.61141394732779000</td>
<td>0.0154875294125994</td>
<td>0.0018172533513232</td>
</tr>
<tr>
<td>21/01/2015</td>
<td>2031.199500</td>
<td>2031.199500</td>
<td>7.61483469516231000</td>
<td>0.0475240447956427</td>
<td>0.0008589027763636</td>
</tr>
<tr>
<td>22/01/2015</td>
<td>2061.499500</td>
<td>2061.499500</td>
<td>7.63189917257808000</td>
<td>0.0152534391083180</td>
<td>0.000850512799802</td>
</tr>
<tr>
<td>23/01/2015</td>
<td>2051.420640</td>
<td>2051.420640</td>
<td>7.62648216386181000</td>
<td>-0.00596032893330</td>
<td>0.0009392906517999</td>
</tr>
<tr>
<td>26/01/2015</td>
<td>2057.096680</td>
<td>2057.096680</td>
<td>7.62590748513161000</td>
<td>0.020651862629423</td>
<td>0.0009421898070568</td>
</tr>
<tr>
<td>27/01/2015</td>
<td>2029.506490</td>
<td>2029.506490</td>
<td>7.61565996726806000</td>
<td>-0.0134782377484405</td>
<td>0.00036165191956</td>
</tr>
<tr>
<td>28/01/2015</td>
<td>2002.160040</td>
<td>2002.160040</td>
<td>7.60738197543031000</td>
<td>-0.015759752905459</td>
<td>0.0018278846866296</td>
</tr>
<tr>
<td>29/01/2015</td>
<td>2021.260000</td>
<td>2021.260000</td>
<td>7.61247410891231000</td>
<td>0.009485917480102</td>
<td>0.000928815513141</td>
</tr>
<tr>
<td>30/01/2015</td>
<td>1994.369900</td>
<td>1994.369900</td>
<td>7.59839431176751000</td>
<td>-0.0130770912379398</td>
<td>0.00062127285009</td>
</tr>
<tr>
<td>02/02/2015</td>
<td>2020.493760</td>
<td>2020.493760</td>
<td>7.6127348290101000</td>
<td>0.0124791732321593</td>
<td>0.0001294872531188</td>
</tr>
</tbody>
</table>

Figure 27: Reference Data-Set of the S&P500 stock index

The different columns represent the date of the quotations, the closing price of the day, the closing price adjusted (as already stated, the adjustment of the quotations are necessary when dividends are paid), the natural logarithms of the AdjClose column (starting point in the models for risk assessment and for predict the future market trend), the returns calculated as the first difference of the "ln (AdjClose)" and finally the realized volatility (non-parametric measure of volatility, it considers the sum of returns realizations with respect to length intervals fixed, in this case this length is fixed every 5 minutes).
2.1 The volatility estimates of the S&P 500 index with the GARCH(1,1) model

Once the process and its parameters have been defined, forecasts can be simulated by using the following analytical formula:

\[ \hat{\sigma}_t^2 = w + \alpha_1 E(\hat{r}_t^2 | I_t) + \beta_1 E(\hat{\sigma}_t^2 | I_t) = w + \alpha_1 \hat{r}_t^2 + \beta_1 \hat{\sigma}_t^2 \]

As previously stated, \( \hat{r}_t^2 \) and \( \hat{\sigma}_t^2 \) represent, respectively, the residual and the estimated value of conditional variance in the previous period.

The first observations are:

1) Return = 0.0031108326
2) Volatility estimated in the previous period = 0.0000211494010

It is possible, using the parameters estimated in the GARCH model, to predict the volatility a step forward:

\[ \hat{\sigma}_{5\text{gen}}{15,1}^2 = 0.000125 + 0.08145796 \hat{r}_{2015,1}^2 + 0.910196 \hat{\sigma}_{2015,1}^2 = 0.000030185182 \]

Taking advantage of this information it is possible to derive the forecast of the volatility of the following day, replacing the just predicted volatility as the volatility of the previous day:

\[ \hat{\sigma}_{6\text{gen}}{15,1}^2 = 0.000125 + 0.08145796 \hat{r}_{2015,1}^2 + 0.910196(0.00000129541771808) = 0.00000129541771808 \]

This procedure was carried out for the subsequent provisions. Follow the figure, part of the subsequent forecasts:

![Figure 28: Forecasts made with the GARCH model](image-url)
You can create a new data set by making forecasts through the calculation methodology "one step forward".

Once all the forecasts have been identified, until 4 June 2019, it is possible first observe the comparison between rv5 and the forecasts obtained from the following graph:

![Figure 29: comparison between rv5 and the forecasts obtained with the GARCH model](image)

From figure 29 it is possible to observe that the expected volatility trend, through this model, it follows in principle that of the realized volatility, however determined peaks reached by the latter fail to be captured by forecasts.

The main property of stock market volatility is the persistence that connects the persisting shocks of variance, whose perturbations tend to generate effects persistent in the long run; which is why the forecasts, in the phases in which volatility does not undergoes disturbances, they succeed in describing the true in the best way realized volatility.

Graphically the distance between the two trends defines the forecast error, given in fact by from the difference between the actual value and the expected value.

In continuing the analysis, the forecast errors of the first two are reported observations and later a part of the relative errors will be illustrated in a figure to the observations contained in the data set:
\[ e_1 = z_{t+h} - \hat{z}_t(h) = 0.00006435800 - 0.000012954177 = 0.0000630625833 \]
\[ e_2 = z_{t+h} - \hat{z}_t(h) = 0.0000836950629 - 0.000030185181 = 0.00005350999 \]

![Figure 30: forecast errors](image)

Finally, the indexes of measurement of the goodness of adaptation of the forecasts have been calculated to determine their accuracy.

\[
MSE = \frac{1}{k} \sum_{\sigma=1}^{k} (r_{T+H}^2 - \hat{\sigma}_{T,H}^2)^2 = 0.00008335631027
\]
\[
MAE = \frac{1}{k} \sum_{\sigma=1}^{k} |r_{T+H}^2 - \hat{\sigma}_{T,H}^2| = 0.0000275278884
\]
\[
RMSE = \sqrt{\frac{1}{k} \sum_{\sigma=1}^{k} (r_{T+H}^2 - \hat{\sigma}_{T,H}^2)^2} = 0.000912997
\]

Finally, the U-di-Theil index was calculated, which is very close to zero; in fact:

\[
THEIL = \sqrt{\frac{1}{T-t} \sum_{t=t+1}^{T} (r_t - \hat{\sigma}_t)^2} = 0.009373348
\]

The indicators assume values close to zero, which implies that, as already stated, the model GARCH basically follows the observations about the variable depend.
2.2 The volatility estimates of the S&P500 index with the TGARCH model

If we consider the same information set as the Standard & Poor share index 500 it is possible, to make a comparison, to calculate forecasts with the variants asymmetric of the GARCH model.

Examining the Threshold-GARCH, TGARCH model, an analysis was performed quite like the one previously shown.

The figure 31 shows the forecast of TGARCH(1,1) model parameters used to estimated forecasts of the volatility. So, considering the first one’s observations available in the data set and using the analytical form of the calculation of the forecasts of volatility in this context:

\[ E(r_{T+\tau+\tau-1}^2 | r_{T+\tau-1}) = \hat{\omega} + \hat{\alpha}_1 \hat{r}_{T+\tau-1}^2 + \hat{\beta}_1 \hat{\sigma}_{T+\tau-1}^2 + \hat{\gamma} D_{T+\tau-1} \hat{r}_{T+\tau-1} \]

forecasts have been identified for the entire reference period.

The substantial difference with the GARCH model is the presence of the Dummy variable. this variable will assume value 1 if the returns in the previous period turn out to be negative, otherwise it will assume a value of 0. In other words, the asymmetric component changes the forecast profile for the one-step-ahead.

Analytically two different formulations will be obtained: the first one already exposed, while in the second case you will have:
\[ E(r_{T+1}^2|I_{T+\tau-1}) = \hat{\omega} + \hat{\alpha}_1 r_{T+\tau-1}^2 + \hat{\beta}_1 \hat{\sigma}_{T+\tau-1}^2 | r_{T+\tau-2} \]

For forecast horizons greater than a period, however, the sign of innovation is not known.

<table>
<thead>
<tr>
<th>rendimento</th>
<th>VARIABILE DUMMY</th>
<th>rv5</th>
<th>Previsioni modello TGARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>0,0031108325860672</td>
<td>0,0000000000000000</td>
<td>0,000032043635349</td>
<td>0,0000211494010224</td>
</tr>
<tr>
<td>0,0005453717845443</td>
<td>0,0000000000000000</td>
<td>0,0000836525972747</td>
<td>0,0000200234133702</td>
</tr>
<tr>
<td>0,039932178895576</td>
<td>0,0000000000000000</td>
<td>0,000156671588534</td>
<td>0,000189177501686</td>
</tr>
<tr>
<td>0,028775830215331</td>
<td>0,0000000000000000</td>
<td>0,000128300830438</td>
<td>0,000197881823069</td>
</tr>
<tr>
<td>0,017452316292985</td>
<td>0,0000000000000000</td>
<td>0,000132319390912</td>
<td>0,00020754634270</td>
</tr>
<tr>
<td>-0,0092545457869802</td>
<td>0,0000000000000000</td>
<td>0,000190818255718</td>
<td>0,000191973720762</td>
</tr>
<tr>
<td>0,082914563243834</td>
<td>0,0000000000000000</td>
<td>0,000102424298664</td>
<td>0,0002728896153286</td>
</tr>
<tr>
<td>0,0024234862235293</td>
<td>0,0000000000000000</td>
<td>0,000178617981878</td>
<td>0,000261053063465</td>
</tr>
<tr>
<td>-0,010882126616956</td>
<td>0,0000000000000000</td>
<td>0,000142163032803</td>
<td>0,000246085641916</td>
</tr>
<tr>
<td>0,0124221150101506</td>
<td>0,0000000000000000</td>
<td>0,000125996330693</td>
<td>0,000390879791298</td>
</tr>
<tr>
<td>-0,0105647307543287</td>
<td>0,0000000000000000</td>
<td>0,000150402872313</td>
<td>0,000376163748505</td>
</tr>
<tr>
<td>-0,01912667576379148</td>
<td>0,0000000000000000</td>
<td>0,00050402872313</td>
<td>0,000305230311290</td>
</tr>
<tr>
<td>-0,0223897789555867</td>
<td>0,0000000000000000</td>
<td>0,000207234681741</td>
<td>0,000396245008797</td>
</tr>
<tr>
<td>0,004587558584412</td>
<td>0,0000000000000000</td>
<td>0,000251317986085</td>
<td>0,001575067073195</td>
</tr>
<tr>
<td>-0,004120575270346</td>
<td>0,0000000000000000</td>
<td>0,000101809409019</td>
<td>0,001467102184192</td>
</tr>
<tr>
<td>0,004868228183939</td>
<td>0,0000000000000000</td>
<td>0,0002620908002481</td>
<td>0,001901067900858</td>
</tr>
<tr>
<td>-0,011886125800155</td>
<td>0,0000000000000000</td>
<td>0,00098677842122</td>
<td>0,001298032920286</td>
</tr>
<tr>
<td>-0,009877798573724</td>
<td>0,0000000000000000</td>
<td>0,0001009635902149</td>
<td>0,001399813754312</td>
</tr>
<tr>
<td>0,0141653052825124</td>
<td>0,0000000000000000</td>
<td>0,0001830830701244</td>
<td>0,001435808659134</td>
</tr>
<tr>
<td>0,012895205561932</td>
<td>0,0000000000000000</td>
<td>0,000061195490671</td>
<td>0,001350783145249</td>
</tr>
<tr>
<td>-0,0054893503911009</td>
<td>0,0000000000000000</td>
<td>0,0000557079317717</td>
<td>0,0012695412614433</td>
</tr>
</tbody>
</table>

**Figure 32: Forecasts made with the TGARCH model**

Figure 32 shows the estimate of the forecasts made. Reporting this result to a graph is obtained:
From the graph this model is not able to "capture" the perturbations that volatility suffers over time. This result can also be obtained in analytical way. In fact, considering the U-d\-Theil index:

\[
THEIL = \frac{\sqrt{\frac{1}{T-t} \sum_{t=t+1}^{T} (r_t - \hat{\sigma}_t)^2}}{\sqrt{\frac{1}{T-t} \sum_{t=t+1}^{T} (r_t)^2} + \sqrt{\frac{1}{T-t} \sum_{t=t+1}^{T} (\hat{\sigma}_t)^2}} = 1
\]

It is equal to 1, which indicates that the model considered is not capable of adapt to the values assumed by the volatility of the historical series taken in consideration.

From the comparison between the two models used it is certainly possible to conclude that the model GARCH is more efficient for estimating volatility forecasts for the stock index S&P 500. Below a similar analysis will be performed for the stock index NASDAQ; the GARCH model will be compared with its other asymmetrical variants.

3. Analysis of the volatility expectations of the NASDAQ stock index

To make a comparison of the estimated volatility forecasts for the index NASDAQ shareholder proceeds with an analysis very similar to that previously observed; the only difference consists in comparing the estimates made with the model GARCH with its other asymmetrical variant, the TGARCH model is no longer considered but the Exponential-GARCH model will be considered.
Once the index time series was imported, the reference sample was reduced (the starting date-set included observations from 3 January 2000 to 3 June 2019, the new data-set includes the observations from 3 January 2000 to 31 December 2014) and, finally, the ARMA model has been estimated.

According to figure 34, the ARMA model can be defined as follow:

\[ r_t = \Phi_0 + \Phi_1 r_{t-1} - \Theta_1 \varepsilon_{t-1} = 0.0000946131 + 0.500284r_{t-1} - (-0.555953)\varepsilon_{t-1} \]

Once the ARMA(1,1) model was identified it was possible to identify the model GARCH; the latter represents the starting point for estimating the forecasts of the volatility for this historical series.

The GARCH model calculated for 3675 observations is considered the best according to Akaike information criterion and according to the Bayesian information criterion is reported in the following figure:
It should be kept in mind that in the GARCH models the conditional volatility is calculated based on a long-term average variance rate. The term (1,1) of the GARCH (1,1) indicates that conditional volatility is based on the most recent observation of yields and on most recent estimate of the variance rate.

At this point the forecast is estimated. This part of the analysis follows the same trend of that provided in the previous paragraphs. Once the estimates have been identified e reported the latter in a graph it is possible to make the comparison between the same and those considered as realized volatility.

---

**Figure 35: estimated GARCH(1,1) model for the NASDAQ index**

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Error Std.</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>0.000639408</td>
<td>0.000178639</td>
<td>3.579</td>
</tr>
<tr>
<td>alpha(0)</td>
<td>1.76619e-06</td>
<td>4.06341e-07</td>
<td>4.354</td>
</tr>
<tr>
<td>alpha(1)</td>
<td>0.0795197</td>
<td>0.00873796</td>
<td>9.101</td>
</tr>
<tr>
<td>beta(1)</td>
<td>0.912222</td>
<td>0.00922673</td>
<td>98.87</td>
</tr>
</tbody>
</table>

Mean var. del dipendente -1.30e-06  SQM var. dipendente 0.016402
Log-veroimiglianza 10768.56  Criterio di Akaike -21507.13
Criterio di Schwarz -21496.08  Hannan-Quinn -21496.08

Note: SQM = scarto quadratico medio; E.S. = errore standard

Varianza dell'errore non condizionale = 0.000214233
Test del rapporto di veroimiglianza per i termini (C)ARCH: Chi-quadrato(2) = 1734.45 [0]
As shown in Figure 36, for the prediction of volatility generally the best result is achieved by forecasts made a step ahead and forecasting capacity decreases to growing forecast horizon. An obvious feature given by the indicators in the case of volatility is to show good performance also for horizons of further forecast.

Furthermore, the graph shows how the expected trend does not take the same large swings of the actual trend, this agrees with known results in the theory of stochastic processes; in fact, the series actually observed can always be interpreted by a predictable process but with a forecast error that has non-zero variance even when the forecaster uses all the information contained in the data observed in the periods prior to that of forecast. To confirm this, forecasting errors have been identified e subsequently estimated the indexes of measurement of the goodness of adaptation; the latter are reported in the following table:

<table>
<thead>
<tr>
<th>Means Square Error</th>
<th>Means Absolute Error</th>
<th>Root Means Square Error</th>
<th>U-di Theil</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00000372</td>
<td>0.000058150</td>
<td>0.001929478</td>
<td>0.000202756</td>
</tr>
</tbody>
</table>

It is specified that in evaluating the statistical loss functions it was used, as a proxy of volatility, the volatility achieved as it is not based on daily returns but on the sum is information collected at a lower rate.
As in the previous case, the GARCH model follows suitably the trend of the historical reference series; this implies that in predicting the volatility is determined on as little error as possible.

### 3.1 Forecasts for the volatility of the NASDAQ index with the EGARCH model

In making volatility forecasts with the Exponential GARCH, EGARCH, reference is made, for a generic time horizon $\tau$, to the following formula:

$$\hat{\sigma}^2_{T+\tau|T} = R \hat{\sigma}^\beta_{T+\tau-1|T} = R^{(1+\beta+\cdots+\beta^{\tau-2})} \hat{\sigma}^{\beta^{\tau-1}}_{T+1|T}$$

The parameters necessary to predict volatility have been obtained by estimation of the EGARCH model with statistical software.

<table>
<thead>
<tr>
<th>coefficiente</th>
<th>errore std.</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>0.000150612</td>
<td>0.000166890</td>
<td>0.9025</td>
</tr>
</tbody>
</table>

**Conditional variance equation**

<table>
<thead>
<tr>
<th>coefficiente</th>
<th>errore std.</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>omega</td>
<td>-0.214602</td>
<td>0.0355970</td>
<td>-6.034</td>
</tr>
<tr>
<td>alpha</td>
<td>0.114322</td>
<td>0.0168906</td>
<td>6.768</td>
</tr>
<tr>
<td>gamma</td>
<td>-0.0906073</td>
<td>0.0107259</td>
<td>-8.448</td>
</tr>
<tr>
<td>beta</td>
<td>0.985705</td>
<td>0.00306526</td>
<td>321.6</td>
</tr>
</tbody>
</table>

**Figure 37**: estimated EGARCH(1,1) model for NASDAQ index (3 June 2000-31 December 2014)

Retracing what was described in the second chapter, with reference to the model EGARCH, the expected sign for the parameter $\gamma$ is negative, as an effect is expected amplifier on volatility in the case of negative returns and a reduced impact in the case of positive returns.

Having available information concerning returns and past volatility is make the dynamic forecast for this model one step forward.

<table>
<thead>
<tr>
<th>Data</th>
<th>Rendimenti</th>
<th>Volatilità Realizzata</th>
</tr>
</thead>
<tbody>
<tr>
<td>31/12/2014</td>
<td>-0.0087014120491595</td>
<td>0.00003174021792655670</td>
</tr>
</tbody>
</table>
With reference to the data reported in this table and applying the following formula:

$$\hat{\sigma}_{t+1|T}^2 = \hat{\omega}^* \hat{\sigma}_T^2 \exp \left( \frac{\hat{\alpha}|\hat{r}_T| + \hat{\gamma} \hat{r}_T}{\sqrt{\hat{\sigma}_T^2}} \right)$$

The volatility forecast for January 2, 2015 is derived.

Analytically we will have:

$$\hat{\sigma}_{gennaio 2015|31 dicembre 2014}^2 = (-0.214802)(0.000031740)^{0.000031740} \exp \left( \frac{0.114322\left[(-0.0087014120) + (-0.0906073)\right]}{\sqrt{0.000031740}} \right)$$

$$= 0.0000108507$$

Considering this forecast as volatility at the previous time it is possible to identify the volatility for the next day; in other words, it is possible to carry out one further forecast a step forward. Analytically we will have:

$$\hat{\sigma}_{gennaio 2015|31 dicembre 2014}^2 = (-0.214802)(0.000108507)^{0.0000108507} \exp \left( \frac{0.114322\left[(-0.001952844926) + (-0.0906073)\right]}{\sqrt{0.0000108507}} \right)$$

$$= 0.0000219625$$

Once all the volatility forecasts have been estimated, it is possible to make the graphical comparison with the true volatility achieved:

![Figure 38: comparison between rv5 and the forecasts obtained with the EGARCH model](image-url)
Already from the graphical comparison it is possible to observe how the EGARCH model is more suitable in forecasting the volatility of the historical series taken into consideration. Indeed, this model manages to describe its progress more efficiently, managing to capture, unlike the GARCH model, determined peaks of realized volatility.

Calculating the indices of the measure of goodness of adaptation, we obtain:

- MSE: 0.000000009892551187648
- MAE: 0.00004154791237015580000
- RMSE: 0.00009946130497659888000
- THEIL: 0.0081204670741579600000

The comparison of these indices with those previously calculated for the GARCH model confirm what the graphical comparison has shown.

The variance of a financial series is a fundamental parameter in determining the excellent portfolio of the investor: in the definition of the latter we need to find the right compromise between the expected average return and the riskiness, measured precisely by the variance.

The empirical analysis carried out shows that between the GARCH model and the two variants observed of the same, the best to predict volatility is the exponential model; indeed, in forecasting the volatility carried out using a GARCH model, there is no informative content able to linearly explain the volatility of the following day.
CONCLUSION

The analysis carried out in this thesis paper had as its main objective that to investigate the different models belonging to the Generalized Autoregressive Conditional Heteroskedasticity family, it was observed the application of the model GARCH and two of its asymmetrical variants in the context of the forecasts of the volatility of the stock indices.

The empirical analysis, presented in the third chapter, concerns the quotations concerning two stock indices: The Standard & Poor's 500 (GSPC) and the NASDAQ Composite (IXIC). The date set of historical series of reference, having as source the site http://www.oxford-man.ox.ac.uk/, concern daily data from 3 January 2000 to 3 June 2019, for a total of 4778 observations.

It has been observed, using time plots, the slow and persistent trend of the historical reference prices. Moreover, once the returns were identified, it was possible to draw some preliminary indications on the possible presence and on the type of time dependence. Indeed, a typical characteristic of returns is evident financial, volatility clustering: adjacent observations tend to have an analogous one level of volatility, and therefore periods of high and low volatility follow.

To efficiently analyze the time series, have been identified empirical regularities, i.e. the recurrent characteristics underlying the phenomenon itself were analyzed. Through the study of descriptive statistics, it was noted how the series stock indices show a weak asymmetry and a marked kurtosis; to verify the normality of the distributions was, subsequently, carried out the Test of Jarque-Bera's normality whose result implies that the joint deviation of asymmetry and kurtosis is statistically significant.

For what concerns the study of the dependence of ordered observations in time, this was carried out using the correlogram showing that various values are outside confidence interval; although in most cases they are not very far apart from the confidence limits.

Next step of the empirical analysis concerns the identification of models capable of to represent the changes characterizing the volatility of the series of returns.

Conditional heteroscedasticity models, hence the GARCH model family, are applied to series with non-normal data and negligible autocorrelation; otherwise, then if the series
shows a non-negligible linear dynamic structure, it is possible to build first an ARMA model and then adapt a GARCH model on the residues.

First, an ARMA model was identified for the series of index returns stock S&P 500, whose parameters were statistically significant and considered the best, observing the information criteria useful for assessing the overall goodness of the model estimation.

By accepting that returns follow (in expected value) an ARMA (1,1) process, the analysis of conditional variance can start from the identification of a GARCH model.

Using the Gretl statistical software, it was possible to identify the model GARCH (1,1) whose parameters are significant, and which represents the starting point in the estimate of subsequent forecasts.

To perform an analysis from the estimate of volatility forecasts in an efficient way the reference sample was reduced, which previously included observations concerning the period from 3 January 2000 to 3 June 2019; in fact, the observations are considered concerning the period from January 3, 2000 to December 31, 2014. In addition, the forecasts estimated with the realized volatility contained in the Oxford-Man dataset.

The fulcrum of the empirical analysis carried out was the estimate of the volatility forecast of the two stock indices, S&P 500 and NASDAQ.

The forecasts of volatility have been estimated for both historical series of returns considering as a starting point the model GARCH (1,1). In both cases, from graphical comparison with the true volatility achieved, it was possible to observe how estimated forecasts generally follow the trend of realized volatility, however certain peaks reached by the latter fail to be captured.

Indeed, the main property of stock market volatility is the persistence that it links the persistence of shocks of variance, which perturbations tend to generate long-term effects; that is why the forecasts, in the phases in which the volatility does not undergo disturbances, they succeed in describing the true in the best way realized volatility.

In continuing the analysis, forecast errors were reported and calculated indexes of measurement of the goodness of adaptation of the forecasts to determine the accuracy of the
same. The indicators, both for the S&P 500 index and for the stock index NASDAQ, take values close to zero, which leads to a good estimate of the volatility forecasts.

Subsequently, considering the same information set of the share index Standard & Poor 500 has been possible, to make a comparison, to simulate data forecasts with asymmetric variants of the GARCH model. It was taken the Threshold-GARCH model is being examined, estimated using the statistical software Gretl. While, for the NASDAQ stock index the comparison happened with another variant of the GARCH model, i.e. the TGARCH model is no longer considered but has been considered the Exponential-GARCH model, also estimated using the Gretl statistical software.

From the analysis of the estimate of volatility forecasts for the S&P 500 index through the TGARCH model and from the graphical comparison it was possible to observe how this model is not able to capture the perturbations that volatility undergoes over time. This result was also obtained analytically by calculating the U-index di-Theil.

From the comparison between the two models used, GARCH and TGARCH, it was certainly possible to conclude that the GARCH model is more efficient for estimating volatility for the S&P 500 stock index.

Last part of the empirical analysis carried out concerns the estimation of the forecasts of volatility for the NASDAQ index. This estimate was made using the EGARCH model.

Through the aid of the graph it was possible to observe how the EGARCH model is more appropriate in forecasting the volatility of the historical series taken into consideration. this model can more efficiently describe the progress of the same, managing to capture, unlike the GARCH model, certain peaks of realized volatility.

The empirical analysis carried out shows that between the GARCH model and the two variants observed of the same, the best to predict volatility is the exponential model.
BIBLIOGRAPHY

- Appunti di analisi delle serie storiche, Riccardo ‘Jack’ Lucchetti;
- Asymptotic Theory for GARCH-in-mean Models, Weiwei Liu, The University of Western Ontario;
- Exact Maximum Likelihood estimation for the BL-GARCH model under elliptical distributed innovations, Abdou Kâ Diongue, Dominique Guegan, Rodney C. Wolff;
- Black’s Leverage Effect Is Not Due to Leverage, Jasmina Hasanhodzic and Andrew W. Lo;
- Modelli econometrici per Fenomeni Soggetti a Smooth Transition: Lo studio della volatilità sui mercati finanziari, Massimiliano Cecconi e Giampiero M. Gallo;
- Conditional Heteroskedasticity in Asset Returns: A New Approach, Daniel B. Nelson;
- Metodi di previsione statistica, Francesco Battaglia;
- Evaluating exponential GARCH models, Hans Malmsten;
- Finanza Quantitativa con R, Marco Bee & Flavio Santi;
- Guida ai comandi di gretl;
- Introduzione alla Stima della Volatilità con modelli Statistici, Dott. Luigi Piva, Quantlab Limited, Bridgewater Road;
- CONTRIBUTIONS ON TIME SERIES ECONOMETRICS;
- Measuring and Testing the Impact of News on Volatility, ROBERT F. ENGLE and VICTOR K. NG;
- Time series analysis; prof. Massimo Guidolin;
- Finanza Quantitativa. Risk Management e Statistica dei Mercati Monetari e Finanziari, Gianluca Cassese e Matteo Pelagatti;
- PREVISIONE DEI RENDIMENTI MINIMI E MASSIMI DI UN TITOLO IN BORSA MEDIANTE UN MODELLO MULTIVARIATO DI VOLATILITÀ;
- On the asymmetric impact of macro–variables on volatility, Alessandra Amendola a, Vincenzo Candila, Giampiero M. Gallo.