Deepanshu Lakhwan*, Aaradhya Dave**

DETERMINING THE MOST EFFICIENT TECHNICAL INDICATOR OF INVESTING IN FINANCIAL MARKETS BASED ON TRENDS, VOLUME, MOMENTUM AND VOLATILITY

DOI: 10.26399/meip.3(70).2020.17/d.lakhwan/a.dave

Introduction

In economics and finance, stock markets are of utmost importance. There are different practices and techniques used by traders and investors to make gains from stock price movements. Even though the time horizon of the investment along with risk appetite governs most investment decisions, the objective of every investor is to maximise his/her returns. To attain this purpose, traders employ various strategies.

Predicting stock market prices by using historical data would be in contradiction to weak market efficiency, which says that market should reflect all past information in its prices and hence stock market forecasts using technical analysis should not be possible. However, the efficient market and random walk hypotheses contradict this approach by proclaiming that the publicly available information on the market is instantly reflected in terms of prices, and that it is impossible to achieve abnormal returns made on the basis of knowing historical data (Malkiel, and Fama 1970: 383–417).

* Deepanshu Lakhwan – MA in International Business Economics (dual degree), Faculty of Economics and Management, Lazarski University in Warsaw, e-mail: lakhwandeepanshu@gmail.com

** Co-author Aaradhya Dave – master's degree student of Madras School of Economics, e-mail: aaradhyadave@gmail.com

This paper attempts to fit a model of stock market prices to check the accuracy of the forecasts. It also employs a comparative evaluation of traditional indicators such as Bollinger Bands, SMA, EMA, VWAP, MACD and RSI to ascertain the most efficient way to comprehend and forecast future price trajectories.

Forecasting financial time series of the stock market has induced significant attention among applied researchers because of the impact of the stock market on the growth of any nation. The autoregressive integrated moving average (ARIMA) model has been the most widely used time series model for forecasting stock market series. This tool for the stock market prediction has attracted vast literature on empirical analysis because of its importance in the development of national economy.

For the research presented below, we have successfully applied the ARIMA model to the stock prices of 2 major indices namely IXIC (NASDAQ Composite Index) and SPY (SPDR S&P 500 ETF Trust) for the time period of 20 years (Adebayo, Sivasamy, and Shangodoyin 2014: 65–77).

1. Literature Review

The purpose of the analysis of the topic is to understand the working and application of the technical indicators used in trading in various financial instruments on the US stock market. Financial markets mean any marketplace where the trading in securities takes place. This includes forex, bonds, stock, commodities, and derivatives among other financial instruments. For the smooth operation of capitalist economies, financial markets are of vital importance (Taylor, and Allen 1992: 304–314). A technical analysis studies the historical price patterns, trends, and other hints responsible for the indication of price movements in the future; it has been increasingly famous over recent years among the financial practitioners to make investment decisions to maximise profitability. In a survey, it was found that at least 90% of respondents use technical analysis in making decisions about their portfolios on the UK market (Chong, and Ng 2008: 1111–1114). According to the basic definition by Robert D. Edwards, John Magee, W. H.C. Bassetti (2018), fundamental prerequisites of a technical analysis are market prices that reflect all events, repeated historical prices and financial instrument's charts changing trends. It can be further divided into two groups, namely technical indicators and price action. When it comes to trading options, the combination of using Bollinger bands with double deviation forms a variety of option strategies where others are shown to offer a wide range of exotic ones like binary or digital options discussed in the paper by Adrea Kolková (2017: 35–40). Trading on these markets means the purchase and sale of financial instruments to profit from short term gains from price fluctuations for themselves or for their clients. To do so, every trader must have an edge in their trading systems by implication of indicators that are helpful in predicting the price movements of these instruments. Lukas Menkhoff, and Mark Tylor (2007: 936–972) showed that a technical analysis is much wider than a fundamental analysis. According to Mark Taylor, and Hellen Allen (1992: 304–314), 90% of the investors polled used it for trading. It was the first time they empirically presented that a technical analysis is one of the most important tools when it comes to making good decisions not only about stock but also about the foreign exchange market.

We are going to focus on 5 major types of indicators that are used worldwide by traders to gain an upper hand in trading on these markets.

1. Bollinger Bands were developed by John Bollinger. This indicator is based on volatility in the form of bands placed above and below a moving average. Volatility is based on the standard deviation that responds to changes as volatility increases or decreases; the expansion in the bands takes place due to higher volatility and it shrinks to show lower volatility. There are 3 major lines in this indicator, the middle one is a simple moving average accompanied by upper and lower bands, which are typically two standard deviations from the middle line (Bollinger 1992: 47–51).

However, it can be modified based on the user's preference. The phenomenon of a squeeze shows bands coming close to each other, which signals low volatility in the present and traders consider it to be a sign of high volatility in the future, offering possible trading opportunities. In the opposite scenario, the widening up of the bands shows higher current volatility and greater chances of exiting the trade, almost 90–92% of the price action happens between the bands and anything outside them is considered an outlier. These outliers are mostly caused by the change in interest rates, release of earning reports by a company and major geopolitical events.

However, John Bollinger bands come with certain limitations: they only consider volatility as their primary focus point. Bollinger suggested that these must be used with at least two of their non-correlated indicators to provide more precision. These bands are primarily reactive, meaning they cannot be used as a predictive tool for the price action making them a lagging indicator (Lento, Gradojevic, and Wright 2007: 263–267).

2. The relative strength indicator (RSI) was developed by J. Welles Wilder and was published in the book titled „New concepts in technical trading systems" (Wilder 1978). It is a momentum-based indicator, which measures the velocity and magnitude of the price movement. These directions are used to measure the current and historical weakness or strength of the stock market based on the closing price of a recent trading period (Taran-Morosan 2011: 5855–5862).

This indicator is mostly used on the 14-day time frame, which ranges from the scale of 0 to 100. Typically, when the indicator line is above 70, it signals the instrument being overbought, and when it is below 30, it is considered being oversold. When it is above or below 80 and 20 respectively, it is considered that the prices are in a stronger momentum (Hamid, Akash and Asghar 2011: 6342–6349).

An overbought instrument indicates a potential selling opportunity and an oversold one signals a potential buying opportunity for traders. The RSI is directly proportional to the velocity of a change in the trend and the magnitude of the move. The divergence between the price action and relative strength index means that the market turning around is highly possible. However, the RSI can also remain overbought for extended time periods if the stock chart is in an uptrend and below 30 if the stock is in a downtrend. This can be confusing for a new trader: when in a downtrend, the RSI is more likely to peak up near the 50% level rather than 70%. This signals traders the long-term downtrend confirmation and vice versa. It is also helpful in improving the chances of getting successful trades. However, it is strongly advised not to solely rely on the RSI but to combine it with other trend-following indicators (Nitin 2020). The appearance of higher highs and higher lows signals the confirmation of a bullish trend on the chart, which is helpful in determining a stable long-term trend, and the same applies to the downtrend identification when the RSI is forming lower highs and lower lows. The concept of swing rejection shows an indicator line reaching the overbought territory above the level of 70%, and then crossing back below it.

In the next step, the RSI forms another high without entering the overbought territory and breaking its most recent low in a bearish swing rejection (Murphy 2009). In a bullish swing rejection, it falls into the oversold area, and then crosses back above 30%, forming another bump without entering an oversold territory. This ultimately breaks its most recent high, indicating a strong uptrend offering a buying opportunity.

The $RSI = 100 - 100/(1 + RS)$, where RS represents the average gain of up periods/the average loss of down periods over a session of 14 trading days (Sahin and Ozbayoglu 2014: 240–245). The RSI can be a very useful indicator for setting up entry and exit points, but it also comes with a set of disadvantages. In the real market conditions, it may not always line up and agree with other technical indicators, which is perfectly normal, as the markets are dynamic and more often irrational; therefore the RSI can only be used as an indicator and not as a predictor. This is also a lagging indicator; it shows the stock is overbought, which does not necessarily mean that it is the time to enter the shorts. The stock can stay oversold for extended periods of time until a reversal occurs; it has a success rate of 49–50%, which is like tossing a coin.

The RSI with given parameters and daily optimisation was compared with other strategies by Blanka Šedivá, and Patrice Marek (2017). It showed that the RSI was still able to produce significantly positive results but when it came to longer time periods, the simple buy and hold strategy also worked.

3. The next one on the list is volume-weighted average price indicator. Granger (1968) found a positive correlation between the daily volume and stock daily price height difference. It provides the average price of the traded security throughout the day based on the price and volume of the trading session. It looks like moving averages appearing as a single line on intraday charts (includes all time frames). The mechanisms to use this indicator are rather easy. If the VWAP is rising and the prices of the security are above it, this indicates an uptrend. Similarly, prices below the line and a declining VWAP are considered an indication of a downtrend.

Investors usually assess the price of this security using this indicator, if the price is higher than the VWAP line at the end of the day, they have overpaid and vice versa.

Osborne (1959: 145–173) studied the relationship between the price and volume, showing the price movement is a diffusion process through establishing a model. This model states that the variance depends on the number of transactions, which suggested positive correlation between an absolute value of changes in the price and volume.

Large institutions and big players on the market use the VWAP ratio to enter and exit markets without creating much impact. They buy below the VWAP level and prefer to sell when the prices are above it; this pushes the prices back to the average, whereas retail traders consider VWAP to confirm the overall trend by entering long positions when prices are above the line and taking shorts when they are below it. Investors asking for the VWAP execution agree to postpone or sequence their trades to reduce their trading cost and market impact while buying and selling large quantities of shares.

The VWAP is calculated as the summation of the dollar value traded for every transaction divided by total shares. The VWAP not only shows traders the trend confirmations but also the value of the security that is being traded.

Another advantage of this indicator is that it cannot be manipulated; hence, it improves both market transparency and efficiency (Cushing, and Madhavan 2001: 12–19). However, it also comes with certain limitations. This is a single day indicator and resets itself as the new day begins, this could indicate that the average price can diverge from the actual VWAP reading. The VWAP is based on the historical values and volume; therefore, it is incapable of predicting the price of the security, and high seasonality hampers its proper functioning. One way to prevent this problem is to focus on the market or the transaction market scale instead of the calendar timings. Andrew Lo and Jang Wang (2000: 257–300) were the first to propel the CAPM model to the volume. In a paper published by edrzej Bialkowski, Serge Darolles, and Gaëlle Le Fol (2008: 1709–1722), it was found that by separating the market part from the observed volume, two additional goals were obtained. Firstly, the liquidity measure for a firm is a more accurate indicator of the arbitrage activity than the observed volume. Secondly, decomposition helps to measure changes in seasonality more accurately in recent years. Many scholars from China started to use the analytical tools and theory backing them away from western scholars. They made a significant analysis on the volume price relation in the stock market of China. Chen and Song (2000: 62–68) did a multi-level empirical research into the relationship between price change and trading volume in Chinese stock market by using a random sample of 31 stocks. Peiyuan, and Donghui (2002: 64–70) have empirically tested the linear and non-linear Granger causality relationship between the price and volume for the Chinese stock market. They presented the results showing that there is a linear Granger causality relationship between volume and bi-directional non-linear Granger causality and between these time series. Zhou Weixing (2010) empirically analysed the volume price variation and the trading volume relation on a deep level by using high-frequency data and found that the price variation and trading volume are correlated along with a non-linear convex function shown by the volume price curve.

4. The simple moving average is a calculation examining the data points by the creation of a series of means of various subsets of full datasets.

The SMA is the unweighted mean of previous *n* data in the series also known as a rolling or moving mean. The time that is selected for trading opportunities can be short, medium, and long. In financial terms, moving-average levels can be explained as support when the market is falling and resistance when the market is rising. When the given data are not circled around the average, a simple moving average lags the single datum point by half-width of the sample. A simple moving average is highly susceptible to the disproportionate effect of old datum points, which are going out or if the new data are coming in (Chande 1992).

If the data have a periodic fluctuation, the application of a simple moving average will discard the variation while the mean will always contain one complete cycle. However, a perfectly regular cycle is rarely encountered (Ellis and Parbery 2005: 399–411).

With a wide range of applications, it is crucial to mitigate the shifting caused by using only historical data; therefore, a central moving average should likely be calculated for better predictions. These are equally spaced on either side of the series point where the average is computed, which requires using an odd number of datum points in the sample window.

The downside of the simple moving average is overlooking a significant number of signals shorter than the window length making the situation worse by inverting it. This creates peaks and troughs, which further leads to the result being less smooth for the analysis. This happens due to the unorganised frequencies appearing during the session (Johnston et al. 1999: 1267–1271).

5. The exponential moving average is a type of moving average, which focuses on most recent data points, also known as exponentially weighted moving average (Klinker 2011: 97–107).

Also, unlike the simple moving average, it reacts more significantly to recent price changes while the SMA applies equal weight to all observations in that period. Just like every other indicator we discussed above, it also provides us with signals for buying and selling based on the divergence and crossovers from the historical average price points.

The EMA can be used on the chart with different time periods like 20, 30, 90 and 200 days, depending on the time perspective for trading (Li, and Zhu 2014: 436–439). If a trader is scalping, it is best to use shorter time frames and medium time frames for day trading.

On the other hand, when it comes to swing trading or investing, larger time frames provide a better idea for the buying and selling signals. The 12-and 26-day exponential moving average is most popular among traders who look for short-term opportunities. When the indicator line crosses a 200-day moving average, it signals that a reversal has occurred for longer period opportunities. It is also used to create indicators like a price percentage oscillator and moving average convergence divergence (MACD). By nature, all the moving averages that are used in technical analysis are lagging indicators by default. Predicting the price solely based on these can create more havoc if they are misinterpreted even by a slight difference. The EMA is having an edge over other moving averages, as it grabs the recent price movement more firmly. This makes it a very insightful indicator for entry points in trading (Kolková 2017).

The EMA indicators are better for trading in trending markets. A strong bullish market is shown by an upward trending EMA line and vice versa. A trader should not only look at the direction of the line but also the rate of change of the line from one bar to another. In a strong uptrend, the rate of the EMA heading upwards is strong; when the trend starts to fade, it starts to grow at a diminishing rate until the rate of change is zero followed by the reversal of the trend.

Metastock is primarily considered the most complete and complex tool to trade on the markets; it can show the graphics along with the stock prices in real time (Rosillo, Fuente and Brugos 2013: 1541–1550).

Traders use it primarily on minute and hourly time frames to adjust the operating schemes to analyse a real time scenario. It requires expert-level knowledge to operate due to more than 150 indicators suited for different trading styles. In addition to providing alerts via mobile phone and email, it is widely used by analysts to figure out a reversal scenario.

Visual Chart focuses on providing a vast variety of products and services at its consumer's disposal showing exchange inversion, meeting the consumer's needs. The main charts are Visual chart V, Visual chart Java edition, Visual chart pocket station and Visual chart Direct Access; all these tools require fast and reliable internet connection to function properly. The third tool is called a personal broker, and it is extensively used in the analysis of financial markets and calculation of investor's profitability portfolio (Rosillo, Fuente and Brugos 2013: 1541–1550). However, this comes with a price of 90 euros for quotes with a fixed annual cost of 45 euros for updating and renovation.

Wing-Keung Wong, Meher Manzur, and Boon-Kiat Chew (2003: 543–551) described the role of a technical analysis and how rewarding it is in the real market scenario. It is used to test the performance of the most popular and established trend-following moving averages and the counter-trend indicator RSI – relative strength index.

The results obtained by the researchers show that the indicators for trading on the markets can produce significantly positive returns. The Singapore stock exchange (SES) members, stockholders and traders enjoy substantial profits through these technical indicators. Many firms have their own teams of traders; these teams significantly rely on a technical analysis of real market conditions. However, despite the enormous literature on a technical analysis, it is still far from being clearly understood.

A technical analysis is a norm for financial markets; therefore, the entire topic of financial management remains fascinating as concerns market efficiency in real-time.

Jeffrey Frankel and Kenneth Froot (1990: 181–185) documented that practitioners, i.e. professionals who forecast the big moves, depend heavily on a technical analysis. They found out a shift away from a fundamental analysis to a technical one in the 1980s.

In the real market conditions, the existence of a technical analysis is stated by the fact that most real-time financial services like Telerate and Reuters provide elaborate, comprehensive, and latest technical information.

The ARIMA stands for the autoregressive integrated moving average, and it has numerous applications in varied fields which is used as a tool to predict the future value of a series based entirely on its own inertia. Kalid Yunus, Torbjörn Thiringer, and Peiyuan Chen (2016: 2546–2556) used the ARIMA models to encapsulate time correlation and possibility distribution of determined wind-pace time collection records (Vaccaro et al. 2015).

Several researchers have been developing models such as the regression model, exponential method and GARCH approaches. However, there are a few works that implemented the ARIMA model in predicting stock market data (Kenny, Meyler, and Quinn 1998: 23–43).

The forecasting accuracy of the ARIMA model gradually decreases at the stage of the growth process, depending on the assumed period. This method is applicable to cases of the high-technology market, especially for the financial institutions and banks since it gives a reliable indication for the future (Almasarweh, and Alwadi 2018).

Generally, it is reported in literature that prediction can be done from two perspectives: statistical and artificial intelligence techniques (Wang 2012: 33–38).

They are also considered robust and efficient in forecasting time series for financial data especially for the short-term prediction. The stock prices can be predicted with larger accuracy compared to other models than even the most popular ANNs techniques (Merh, Saxena, and Pardasani 2010: 23–43).

2. Data and Methodology

This paper uses IXIC (NASDAQ Composite Index) and SPY (SPDR S&P 500 ETF Trust) data collected in the period from April 2000 till early May 2020. The IXIC index is a large market-capitalisation weighted index of over 2500 stocks. Since it encapsulates a large percentage of traded stocks on the US market, it gives a good indicator of movement of all traded stocks. SPY is the largest Exchange-Traded Fund in the world, which is referred to as S&P 500 index[1] in financial terms. The paper first does tests for stationarity, corrects non-stationarity, finds the appropriate ARIMA model and runs various tests to find the most efficient ARIMA (p, d, q) model. The software R has been used for visualisation and an analysis.

Thereafter, major indicators used by traders (Bollinger Bands, RSI, VWAP, SMA, EMA) are used for computation of expected prices. A detailed graphical analysis and accuracy tests are done to ascertain the most efficient indicator. The forecasts from the ARIMA model are also compared with these results to gain a better understanding of the market movements.

2.1. Empirical Analysis

Tests for stationarity

1) Plotting the series: the data of IXIC index prices for the entire range are plotted. The graph is observed to check if there are any trends or patterns of seasonality. If these are observed, then non-stationarity exists in the data.
2) Augmented Dickey-Fuller test (ADF): ADF tests for the presence of a unit root in a data set. The presence of a unit root indicates non-stationarity in a data set. It allows for higher order auto-regressive processes.

$$\Delta y_t = \alpha + \beta t + \gamma y_t - 1 + \delta \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \ldots$$
$$H_0: \gamma = \delta = \delta_2 = 0$$

The null hypothesis is the presence of a unit root and therefore non-stationarity in the data set, while the alternative hypothesis is a stationary time series.

[1] S&P 500 market index is a good indicator of the US stock prices since the market capitalisation of its stocks constitutes 80% of the free float market capitaliszaion of all listed stocks on the US stock markets.

3) Kwiatkowski-Phillips-Schmidt-Shin test (KPSS): KPSS tests check the null hypothesis that the series is stationary against the alternative that it is not.

$$y_t = \alpha + \beta t + \mu_t + u_{\mathrm{t}}$$
$$\mu_t = \mu_{t-1} + \varepsilon_t,\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

The KPSS test statistic is the Lagrange multiplier and used to test:

H_0: $\sigma_\varepsilon^2 = 0$
H_1: $\sigma_\varepsilon^2 < 1$

Based on the results of these tests, we can conclude whether the time series under consideration is stationary or non-stationary.

Transforming from non-stationary to stationary

If the tests indicate the presence of non-stationarity, the data series is differenced, and the tests are run again to check if the results indicate the presence of stationarity. In cases where the variance is volatile, logarithmic transformations can stabilise the time series.

Identifying the appropriate ARIMA model

After a series has been made stationary, the next step is to identify the correct model.

1) Auto correlation function plots (ACF): Autocorrelation function plots give us the correlation of any series with its lagged values.
2) Partial autocorrelation function (PACF) plots: A partial autocorrelation function plots the autocorrelation of the residuals of the series and its lagged values along, meaning it gives the correlation of a series and its lagged values that are not explained by all lower order lags.
3) If the ACF plot ends geometrically, the PACF plot ends sharply and is insignificant after p lags; then we can say that it follows AR(p).
4) If the PACF plot ends geometrically, the ACF plot ends sharply and is insignificant after p lags; then we can say that it follows MA(p).
5) If both plots are geometrically declining, then it likely follows an ARMA process.

To identify the correct ARIMA model, R also uses maximum likelihood estimation (MLE). It tries to maximise the log-likelihood for given values of p, d and q while finding a parameter; it estimates to maximise the probability of obtaining the data that we have observed.

Testing the ARIMA model

1) AIC test:
   The test statistic is given by:

   $$AIC = (1/n - 1) \sum_{i=1}^{n} (x_i - \bar{x})^2 + 2k$$

   where n is the number of sample points of data x, and k is the number of parameters to be estimated. A lesser value is generally considered substantial support for the model.

2) BIC test:
   The test statistic is given by:

   $$BIC = (1/n - 1) \sum_{i=1}^{n} (x_i - \bar{x})^2 + k \ln(\ )$$

   where n is the number of sample points of data x, and k is the number of parameters to be estimated. From the two given models, the lower value of *BIC* is preferred. It tells us that the model is a better fit.

3) Ljung-Box test: Ljung-Box tests for autocorrelation between the different lag terms. The test statistic Q is given by:

   $$Q = T(T + 2) \sum_{k=1}^{s} [r_k^2/(T - K)]$$

   where T is the number of observations, s is the number of lags to test autocorrelation, and r is the autocorrelation coefficient. The hypothesis tested is that the residual is white noise against the alternative that it is not.

Since we are dealing with financial data here, it is very common to find heteroskedasticity or non-constant variance. It calls for more sophisticated models that can capture this volatility. The ARCH (Auto Regressive Conditional Heteroskedasticity) model captures this real-world volatility. There are other extensions like GARCH (Generalized Autoregressive Conditional Heteroskedasticity), EGARCH (Exponential GARCH), IGARCH *et cetera*. The GARCH model is often used in the market to model stock volatility, returns and price.

Once the fitted ARIMA model fails diagnostic checks, the following step is to check if there are any ARCH effects in the time series. This is tested by the LM ARCH Test.

LM ARCH Test

To test the presence of the ARCH effects, a Lagrange Multiplier Test is conducted, where a mean equation is estimated. The mean equation can be a regression of the variable on other variables or even a constant. The residuals from this regression equation are squared and regressed on their lagged terms. The number of lags it is regressed on determines the order of ARCH.

The following equation is referred to as the mean equation. AR and MA terms are also included in some models in this equation:

$$R_t = \mu + \varepsilon_t$$

The squares of ε_t are regressed on its lag values. If the coefficient of the LAD (Least absolute deviation) term is significant to zero, there is absence of the ARCH effects (Null Hypothesis).

The LM test Statistic:

$$(T–q)\ R\hat{\ }2$$

where T is the sample size, q is the number of squared error terms in the regression equation, and R^2 is distributed as chi squared distribution with q (order of lag) degrees of freedom. The null hypothesis is rejected when the test statistic value is greater than the tabulated value.

Estimating an appropriate model

The ARCH models are estimated using the Maximum Likelihood method. The GARCH models can be represented by the mean equation and the following equations. σ_t is the variance for the residuals from the mean equation. Since z follows standard normal distribution, the variance of ε_t is σ_t^2. This is also the conditional variance of stock returns, which is clear from the mean equation. Thus, equation (*) gives us the desired property of variance, and large (small) variances are followed by large (small) volatility changes.

$$\varepsilon_t = \sigma_t * z_t$$
$$\sigma_t^2 = \omega + \alpha_1\, \sigma_{t-1}^2 + \beta_1\, \varepsilon_{t-1}^2 \ (*)$$
$$z_t \sim N\ (0,1)$$

The objective is to capture the changing variance of the returns of the stock, which is accomplished in all orders and forms of GARCH. Often, the ARMA terms are also included in the mean equation.

In both time series considered here, the ARFIMA or Autoregressive Fractionally Integrated Moving Average Terms will be included. This is a generalisation of ARIMA and uses non-integral values for a differencing component. This is popular in series that have a long memory property.

Residual diagnostics

Ljung-Box test is to check if the residuals behave as white noise. The null hypothesis is that there is no autocorrelation in the residuals. This is rejected only if the p value is less than 0.05 (level of significance).

Another test to check the presence of any remaining ARCH effects is the LM ARCH test. Furthermore, plots of empirical distribution of error terms and QQ norm plots are observed. These help to determine whether the errors follow a standard normal distribution. The QQ plot is often a straight line if there is presence of the said theoretical distribution.

2.2. Analytical and Graphical Approach

Stock market indicators

A detailed analysis of the predicted values from indicators like Bollinger Bands, MACD, RSI, SMA, EMA and VWAP is presented in the result section along with accuracy measures to assess the relatively efficient indicator. The concept and the steps involved in the computation of these indicators for the corresponding 20-year period are presented below.

1. Bollinger Bands

Bollinger Bands indicator was developed in the 1980s by technical analyst John Bollinger. It comprises three bands, namely, upper, middle and lower, which indicate the pricing channels and incorporates volatility in the series. The idea of plotting moving averages was taken a step further by using the concept of standard deviations to define upper and lower rate boundaries. Standard deviation is a measure of dispersion. It helps to assess volatility in the series. 68% of the observations lie in one standard deviation from the mean on the bell curve. 95% lie in two standard deviations. About 99.7% observations lie in three standard deviations from the mean.

2. $\boldsymbol{\sigma = \sqrt{\sum(x_i - \bar{x})/n}}$

Middle Band calculates a simple moving average of the 20-day price. A longer period may also be considered, but this will result in increasing the number of standard deviations employed too.

$\bar{x} = (\sum x_i)/n$

Upper band is the summation of the moving average and twice the value of standard deviation for the corresponding time period.

Upper band $= \bar{x} + 2\sigma$

Lower band is the difference between the moving average and twice the value of standard deviation for the corresponding time period.

Lower Band $= \bar{x} - 2\sigma$

The actual prices are plotted along with the three bands to assess the accuracy of Bollinger Bands in predicting price movements.

3. SMA (Simple Moving Average)

This indicator helps traders to ascertain the trend for short-, medium- and long-term investments. It is calculated like a simple arithmetic mean. However, one limitation of this indicator is that it gives equivalent weight to all the data points for which the average is calculated.

Short-term trends and trading are analysed best using 5–20 period averages. Where medium term traders prefer the 20–60 period moving average, and long-term investors may consider 100 or even more time periods for the computation.

Here, a time period of 10 is taken to compute the moving average for the entire time period under study.

4. EMA (Exponential Moving Average)

As opposed to SMA, EMA eliminates the lag in the indicator and hence, reflects the price movements faster than SMA. It gives higher weightage to the recent observations. The typical time period taken by short-term traders is 12 or 26 days, whereas long-term investors use 50-day or 200-day EMA.

The paper uses the 12-day look back period and computes EMA using the following formula:

$$EMA = [Closing\ price * multiplier] + [EMA(previous\ period) * (1 - multiplier)].$$

where, $multiplier = 2/(n + 1)$
Here, we take n time period as 12.

5. VWAP (Volume Weighted Average Price)

The inputs involved in computation of this index are the high, low and closing price along with the volume traded for the intraday period. The typical price is calculated by taking an average of the intraday high, low and closing price.

This typical price is multiplied by the volume. Thereafter, a cumulative summation of this product (Volume* Typical Price) is taken. Similarly, a cumulative sum or a running total of the volume is calculated. Finally, a ratio of the running total of the price-volume to the running total of the volume provides us with the indicator values.

6. MACD (Moving Average Convergence Divergence)

MACD or the hybrid indicator inculcates both trends along with momentum concepts. It can generate buy and sell signals indicating the zones above the zero line as bullish and below it as bearish.

The MACD line is formed by taking a difference of the values of a 12-day EMA and a 26-day EMA. Further, signals are generated by taking a 9-day EMA of theMACD line/values. A histogram can be plotted by taking the difference between the MACD values and the signal values.

7. RSI (Relative Strength Indicator)

The RSI value oscillates between 0 and 100 indicating overvalued (or overbought) and undervalued (or oversold) stocks or other assets. A value over 70 indicates overvaluation and a value under 30 indicates undervaluation conditions. An overvalued stock/asset may be in for a pullback in prices reversing the trend. This kind of analytical conclusions from RSI helps traders make buy-sell decisions.

First, day over day percentage changes in price are calculated; these changes in price help to distinguish between gains and losses for each time point. Thereby, a moving average of 12-day gains and losses is calculated separately. Relative strength or ratio of average gains to average losses is then computed for each time point. RSI is then calculated using the following formula:

$$RSI = 100 - [100/(1 + RS)]$$

However, if the average loss is zero at any time point, then *RSI* is equal to 100.

A graphical analysis against actual price observations is carried out for all 6 indicators. Since the period under study (2000–2020) is quiet long, to track and compare the actual prices against the projected values, we need to consider a slimmer time frame. Hence, the graphs for the financial crisis period, i.e., 2007–2008, and the recent period (2019–2020), are studied in detail. The period of financial crisis is chosen because it was highly volatile for all securities.

The recent period is chosen to gain a better understanding of the market in the current scenario.

The accuracy measures such as RMSE and MSE further assist in ascertaining the most efficient indicator of the price movements on the market.

3. Identification and Results

3.1. Results for SPY ETF

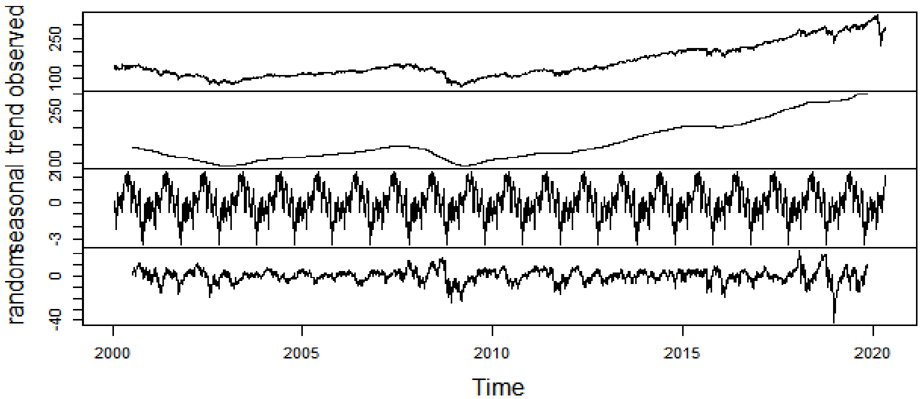
The results for SPY ETF have been presented. The econometric analysis is followed by the graphical analysis of different indicators. The forecasts from the fitted model are thereby analysed with the predicted values from the indicators.

Tests for non-stationarity

1. Plot of adjusted closing price of IXIC indicates that there is a trend in the data. This means that the data exhibits non-stationarity. Furthermore, from the decomposition, we can see there is a seasonal component and the variance is not constant throughout. This gives a further indication that the data is non-stationary.

Figure 1.1

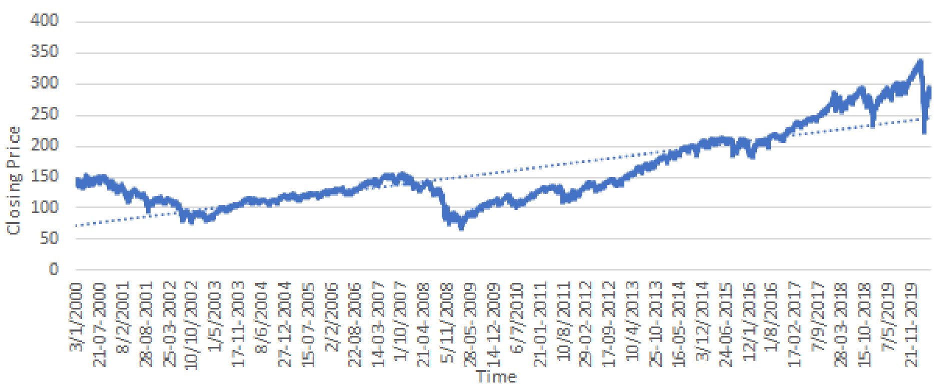
Decomposition of additive time series from 2000–2020



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Figure 1.2

Closing price of SPY ETF



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

2. Augmented Dickey-Fuller test

Dickey-Fuller	Lag order	p-value
-2.300	17	0.4513

Source: author's own elaboration using R software.

3. KPSS test

KPSS trend	Truncation lag parameter	p-value
8.8067	10	0.01

Source: author's own elaboration using R software.

A high p-value is observed in the ADF test and the null hypothesis is accepted. However, a low p-value for the KPSS test leads to a rejection of the null hypothesis. Thus, the results of the KPSS and ADF tests indicate that there is presence of non-stationarity.

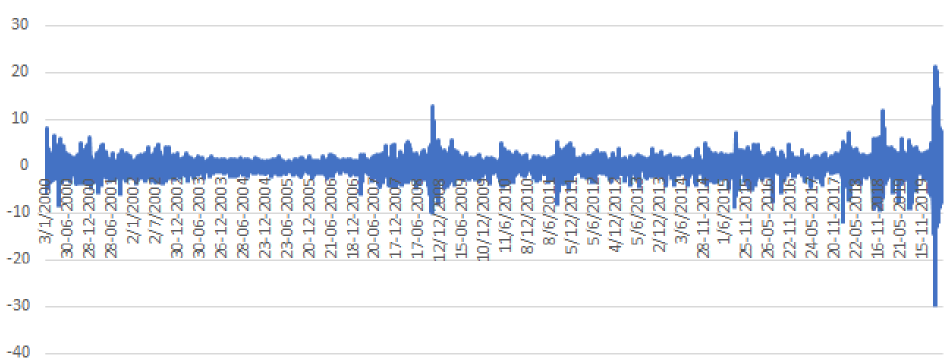
Transforming non-stationary series to stationary series

Since non-stationarity has been identified, first differencing is done to see if it is converted into a stationary series. Tests for non-stationarity are run again and the following results are obtained.

1. Change in stock prices after first differencing: The plot appears to be randomly distributed around zero indicating the possibility that the log transformed first differenced series is stationary.

Figure 1.3

Change in prices of SPY ETF



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

2. ADF test after differencing

Dickey-Fuller	Lag order	p-value
-82.685	0	0.01

Source: author's own elaboration using R software.

3. KPSS test after differencing

KPSS trend	Truncation lag parameter	p-value
0.025028	10	0.1

Source: author's own elaboration using R software.

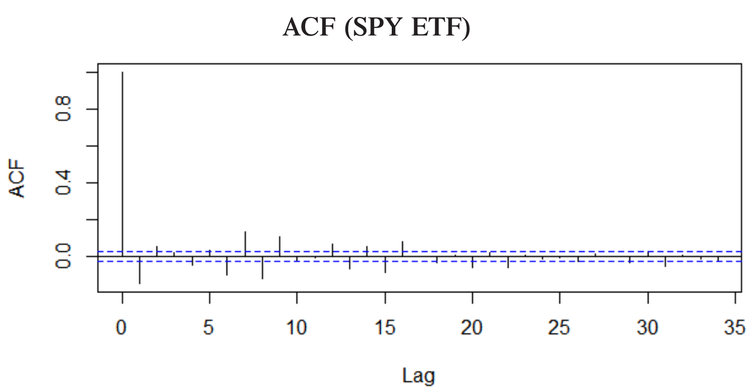
In comparison to the results obtained before first differencing, we obtain the opposite results, indicating that the first differenced series is stationary. Subsequently, the paper proceeds to use the ACF and PACF plots of the differenced series to identify the correct model.

Identifying the model

To identify the p, d, q values of the ARIMA model, the ACF and PACF are observed.

1) ACF plot of transformed first differenced series

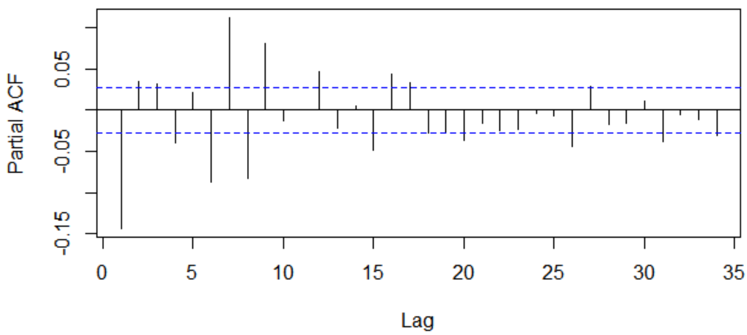
Figure 1.4



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Figure 1.5

PACF (SPY ETF)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

The PACF plot is insignificant after the seven-time lag indicating a value for p (AR terms). However, the significant lines in the ACF plot are not enough to determine the value for q (MA terms).

Testing to find the most efficient model

The results for ARIMA (2,1,0), ARIMA (4,1,1), ARIMA (3,1,1), ARIMA (5,1,0) & ARIMA (5,2,0) are computed[2]. According to the results, ARIMA (5,2,0) has the best fit since it has the lowest AIC and BIC values. Now, a check needs to be done so as to conclude it is the best fit model.

Best fit model

ARIMA (5,1,4) with drift

	VALUES
Sigma Squared	3.703
Log Likelihood	-10608.18
AIC	21236.36
AICc	21236.4
BIC	21301.77

Source: author's own elaboration using R software.

---

[2] See Appendix I.

	AR1	AR2	AR3	AR4	AR5
Coefficients	-0.290043	0.759003	-0.262654***	-0.861737	-0.041298**
	MA1	**MA2**	**MA3**	**MA4**	
Coefficients	0.182671	-0.752859	0.358009***	0.753016	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1

Source: author's own elaboration using R software.

A lower BIC & AIC value implies that ARIMA (5,1,4) is the best fit model for our analysis.

Before interpreting this result, first diagnostic checks are carried out to check the robustness of the model. Since most of the coefficients are not significant even at 5% level and the financial time series is known to have increasing variance, these checks are important before finalising the model.

Diagnostic checking

The next step is to check the significance of autocorrelation coefficients. The output for the Ljung-Box test is as Table.

Q*	DF	p-value
37.211	3	4.152e-08

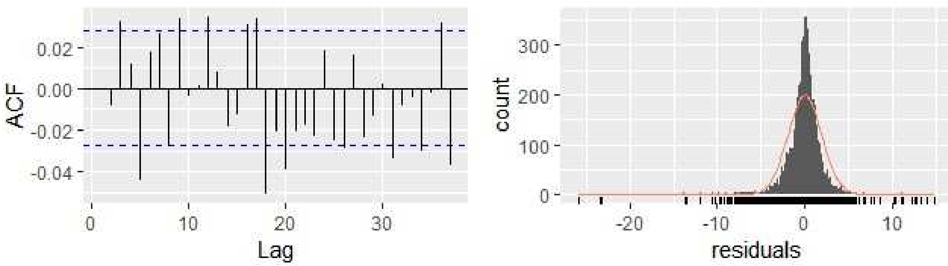
Source: author's own elaboration using R software.

The null hypothesis is rejected at 1% level of significance implying that the residuals do not follow a white noise process; hence, it is not completely random.

Additionally, an introspection of the histogram of residuals indicates that the residuals follow normal with mean zero and non-constant variance.

Figure 1.6

ACF and Lag residuals (SPY ETF)



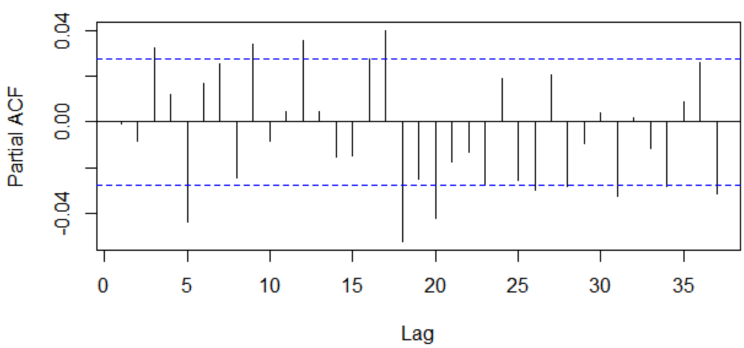
Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

ACF and PACF of residuals

If there are spikes outside the insignificant zone for both ACF and PACF plots, we can conclude that residuals are non-random with information in them. From the plots, we can see that the mean of the residuals is very close to zero; however, the correlation between residuals is non-zero. The time plot of the residuals shows that the variation in the residuals will differ over time.

Figure 1.7

Partial ACF vs. PACF (SPY ETF)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

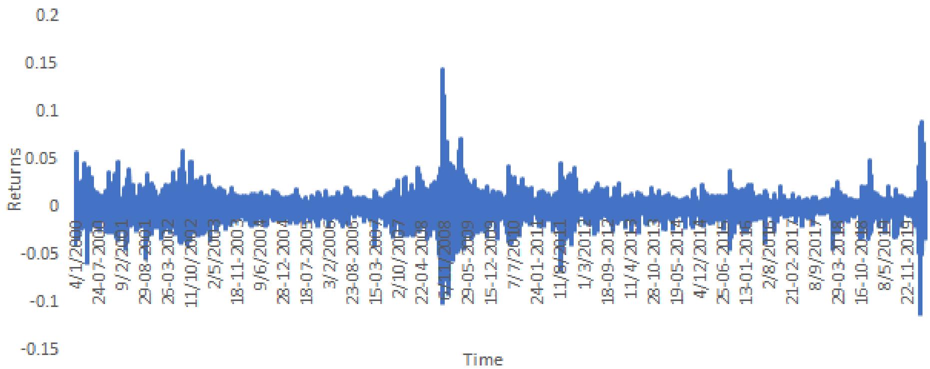
Since this confirms autocorrelations in the residuals, a model that captures the non-constant variance feature of the time series must be employed.

Volatility models

To examine the series again, a graph of daily returns of SPY is plotted over time.

Figure 1.8

Daily returns of SPY ETF (2000–2020)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Here, we observe a period of large changes followed by further large changes and those of smaller changes followed by further smaller changes. The values fluctuate unpredictably from period to period, hence indicating a volatile time series.

Testing the ARCH Effects

ARCH LM test

The mean equation is estimated, and the residuals obtained are squared and regressed on one lagged squared residual.

Chi-squared	df	p-value
336.88	1	< 2.2e-16

Source: author's own elaboration using R software.

A low p-value is observed in the ARCH LM test and the null hypothesis of no ARCH effect is rejected. Thus, the results of the test indicate that there is presence of the ARCH effects.

Estimating GARCH models

ARFIMA (1,0,1) GARCH (1,1)

	mu	omega	Alpha 1	Beta 1	AR 1	MA 1
Coeffi-cients	0.000582**	0.000000***	0.092664***	0.916989***	0.786767.	-0.830544*
SE	0.000089	0.000001	0.014780	0.012656	0.151314	0.137052

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1

Source: author's own elaboration using R software.

Information criterion statistics

	VALUES
HQIC	-6.4381
AIC	-6.4408
BIC	-6.4331
SIC	-6.4408

Source: author's own elaboration using R software.

$R_t = 0.000582 \quad 0.786767\, R_{t-1} - 0.830544\, \varepsilon_{t-1} + \varepsilon_t$
$(\varepsilon_t = \sigma_t * z_t)$
$\sigma_t^2 = 0.000000 + \ 0.092664\, \sigma_{t-1}^2 + 0.916989\, \varepsilon_{t-1}^2$-----*
$(z_t \sim N(0,1))$

A \$ 1 increment in the previous period in SPY returns with respect to its previous period. On average, it will lead to a \$ 0.79 increase in the returns value of SPY in the current period.

Also, from the third equation, small (or large) volatility changes are followed by similar smaller (or larger) changes.

Diagnostic checking (Standardised residuals tests)

The next step is to check the significance of coefficients. The output for the Ljung-Box tests, LM ARCH test and a few other tests are as Table.

Tests	Statistic	p-value
Weighted Ljung-Box Test [on standardised residuals]	0.03183	0.8584
Weighted Ljung-Box Test [on standardised squared residuals]	1.252	0.26311
LM ARCH Test [TR^2]	1.049	0.3057

Source: author's own elaboration using R software.

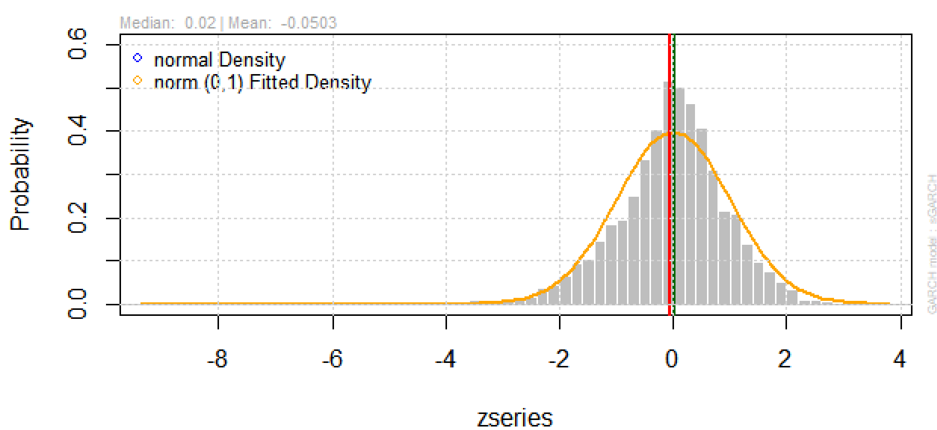
Ljung-Box test on standardised and standardised squared residuals and LM ARCH test accept the null hypothesis and are significant at 5% level of significance.

Thus, it is concluded that the errors store no additional relevant information and have no serial correlation. The errors behave like a white noise process. Also, there are no further arch effects.

Thus, the model ARFIMA (1,0,1) GARCH (1,1) passes the diagnostic checking and is the best fit model for the SPY ETF time series. Further, to confirm the robustness of the model over the time series, we plot the following graphs.

Figure 1.9

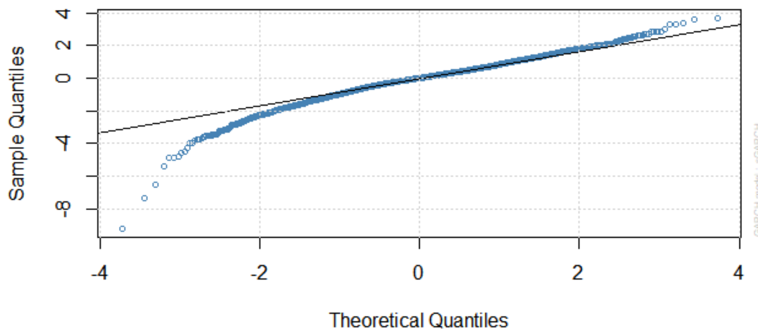
Empirical density of standardised residuals



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Figure 1.10

Norm QQ plot (SPY ETF)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

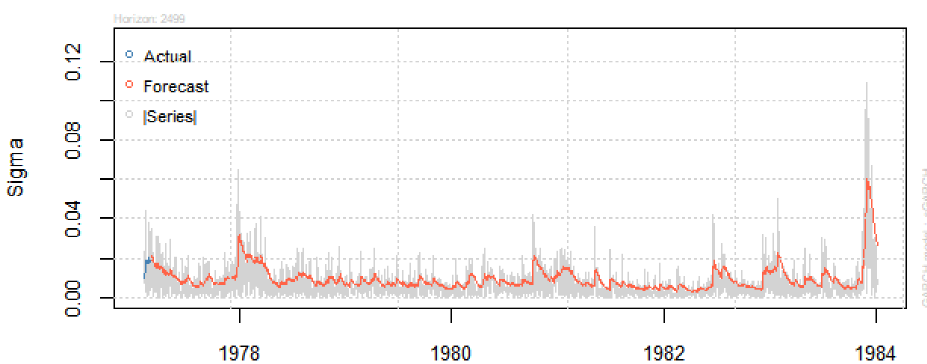
The two graphs shown above confirm the randomness of the residuals. The QQ plot of residuals, further suggests the normal distribution of errors. The QQ plot resembles a straight line (with a few exceptions of outliers) suggesting the standard normal distribution with the mean zero and standard deviation one. This plot arranges the sample data in an ascending order and plots them against quantiles from a theoretical distribution (here, standard normal distribution).

These plots and the results from Ljung-Box tests confirm no autocorrelation in the errors; therefore, the robustness of the model ARFIMA (1,0,1) GARCH (1,1).

Forecast

Figure 1.11

Forecast rolling sigma vs. [series] (SPY ETF)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Here, the predicted conditional volatility is represented as Forecast Rolling Sigma. This is a conditional mean of the series at time $t+h$. The model efficiently predicts the volatility paths such as spikes and dips, and validates the long-term mean of the series. However, the initial spike or dip is not predicted since it depends on the past values of the time series.

Results from various indicators for SPY

The detailed analysis of the values predicted from indicators such as Bollinger Bands, MACD, RSI, SMA, EMA and VWAP is hereby presented along with a few accuracy measures to assess the relatively efficient indicator.

Since the period under study is 2000–2020, this is to track and compare the actual prices against the projected values and we need to consider a slimmer time frame. Hence, the graphs for the financial crisis period, i.e. 2007–2008 and the recent period (2019–2020) are presented for each indicator.

The period of the financial crisis is chosen because it was highly volatile for all securities; however, the recent period is chosen to gain a better understanding of the market in the current scenario.

1. Bollinger Bands

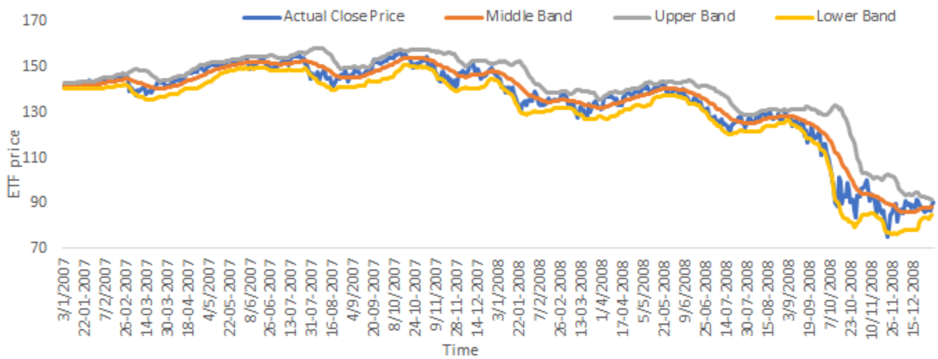
For the entire period, Bollinger Bands **studied under (2000–2020)**[voc/st] are not presented as a longer time scale compresses the observations at each time point. It is unable to give a clear picture of the movement of the bands. A magnified version for the years 2007–2008 and 2019–2020 is presented below (Figure 1.12.).

The blue band, indicating the actual prices is observed to be quite volatile, especially towards the end of 2008. All the bands initially, at the beginning of 2007, are quite close and spread out later in 2008 incorporating the higher volatile nature of the market. One useful observation here is that the actual prices oscillate between the upper and lower bands but never cross these boundary rates (Figure 1.13.).

One major similarity that is observed is the expansion of the Bollinger bands with the onset of financial stress on the market. However, this time, volatility is much higher as the lower and upper band widen up considerably in the first quarter of 2020. Increased volatility in the first quarter of 2020 is successfully incorporated in the wider ranges of the upper and lower bands. Once again, it is worth noting the actual price movements never cross the boundary bands, hence proving the efficiency of the volatility indicator.

Figure 1.12

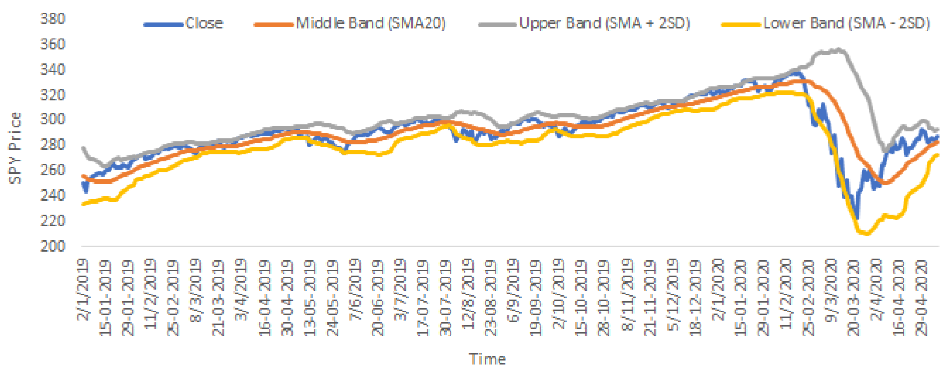
Actual prices of SPY v/s Bollinger Bands (SPY ETF)



Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

Figure 1.13

Actual prices of SPY v/s Bollinger Bands (SPY ETF)



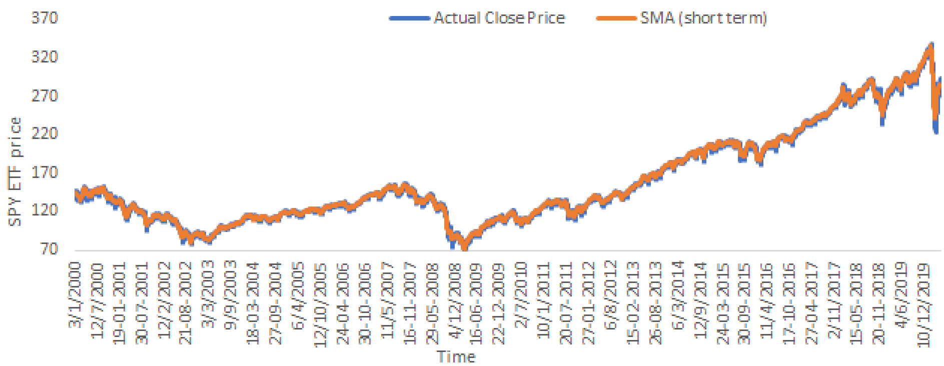
Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

2. SMA (Simple Moving Average)

The graph comparing actual prices and the SMA projections for SPY ETF is presented for the period under study, i.e. 2000–2020.

Figure 1.14

Actual prices of SPY v/s SMA projections (SPY ETF).

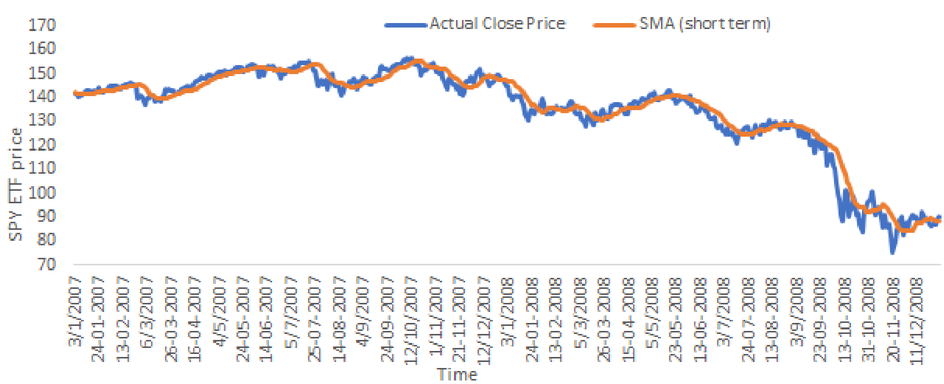


Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Since the time period under study is very long, it is unable to paint a clear picture when it comes to forecast errors. A detailed and microscopic analysis is necessary to visualise the effectiveness of the indicator. The following graphs are for the period of 2007–2008 and 2019–2020.

Figure 1.15

Actual prices of SPY v/s SMA projections (SPY ETF)



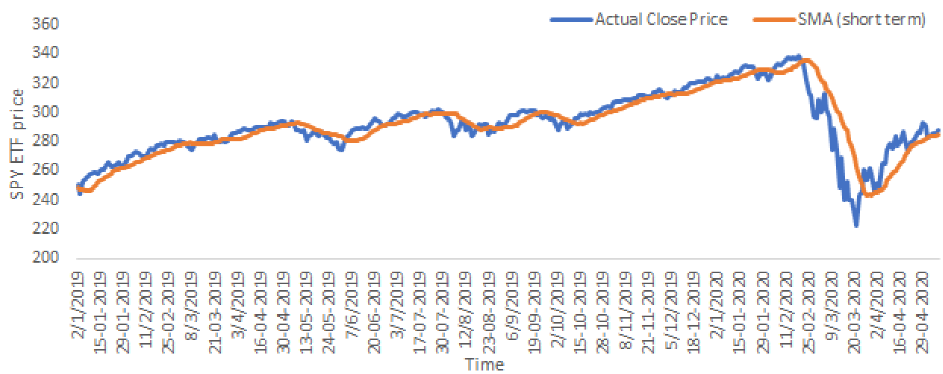
Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

It is observed that the actual closing price has a much more volatile motion than the projected values. It is also very useful to observe that any trend change is incorporated in the SMA line a few periods later. It is thus deductible that SMA is a little slower in realising the change in price movements.

This is also consistent with the theoretical fact that it assigns equal weightage to all observations and hence reflects any opposite movement with a lag.

Figure 1.16

Actual prices of SPY v/s SMA (Short term)



Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

A similar observation is made in the graph for the period of 2019–2020. A lag in the realisation of sudden and adverse price movements is noticed in the predicted values. A significant drop in prices witnessed from February 2020 onwards is attributed to the COVID-19 pandemic.

Accuracy Measures

	ME	RMSE	MAE	MPE	MAPE
Training set	0.1803078	3.944103	2.566009	0.0460829	1.704767

Source: author's own elaboration using R software.

Above, there are a few accuracy measures for the model. For this analysis, the following is observed.

Mean Percentage Error (MPE)

The Mean Percentage Error is the computed average of the percentage errors by which a forecast of a model differs from actual values of the quantity being forecasted. The MPE value obtained is 0.0460829, implying that the difference between the forecasted values and the actual values is not significant[3].

Root Mean Squared Error (RMSE)

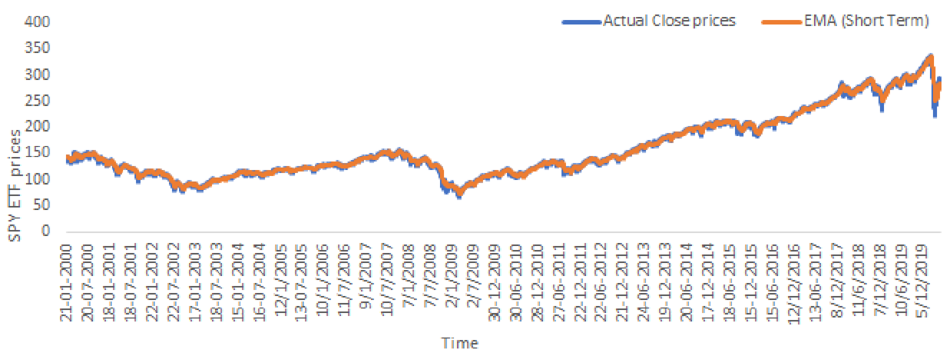
The Root Mean Squared Error is a measure of the concentration of the actual data points around the best fit. It is the standard deviation of the errors. A lower value depicts an overall good fit. The RMSE value obtained is 3.944103 implying a good fit.

3. EMA (Exponential Moving Average)

EMA for a 12-day period is computed and plotted. This is specifically useful for short-term traders. This look back period is increased when medium-or long-term investments are considered.

Figure 1.17

Actual close prices of SPY v/s EMA projections



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Once again, the long time period makes it difficult to visually track the actual and predicted movements. Thus, two different shorter time frames are studied separately for 2007–2008 and 2019–2020 (Figure 1.18).

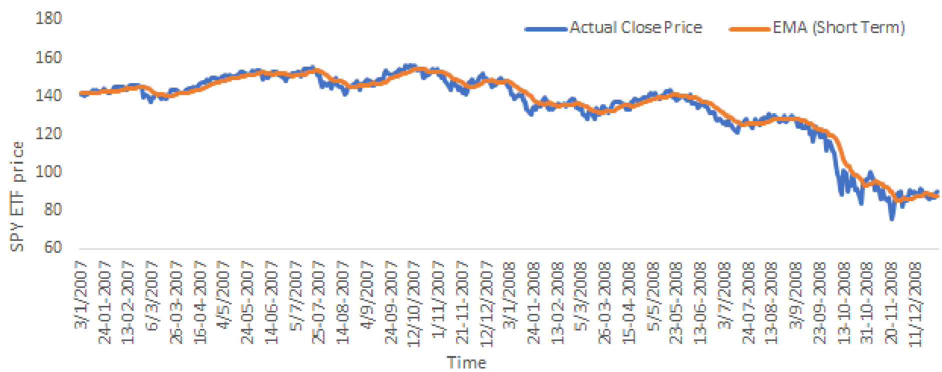
EMA is faster in incorporating price changes than SMA since it gives more weightage to recent observations (Figure 1.19).

---

[3] See Appendix II.

Figure 1.18

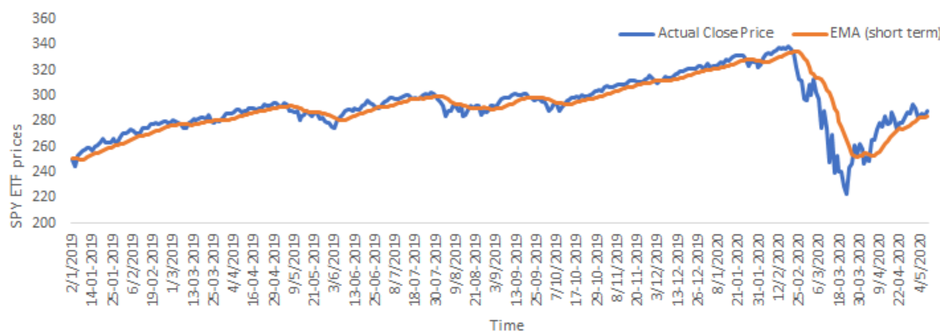
Actual close prices of SPY v/s EMA projections



Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

Figure 1.19

Actual close prices of SPY v/s EMA projections



Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

The shock from the black swan event is realised with a slightly small lag. The lag times in the realisation of adverse movements of the series are present but are lesser than in the SMA projections.

Accuracy measures

	ME	RMSE	MAE	MPE	MAPE
Training set	0.2591323	4.684283	2.55217	0.09049056	1.698478

Source: author's own elaboration using R software.

Above, there are a few accuracy measures for the model. For this analysis, the following is observed.

Mean Percentage Error (MPE)

The Mean Percentage Error is the computed average of the percentage errors by which a forecast of a model differs from actual values of the quantity being forecasted. The MPE value obtained is 0.09049056 implying that the difference between the forecasted values and the actual values is not significant[4].

Root Mean Squared Error (RMSE)

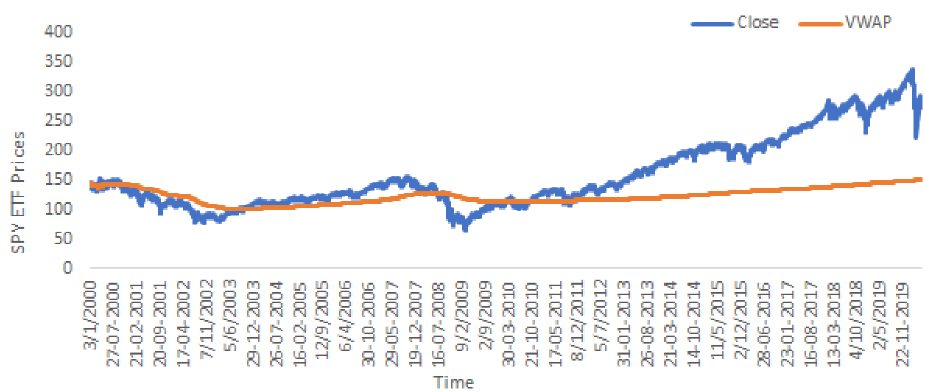
The Root Mean Squared Error is a measure of the concentration of the actual datum points around the best fit. It is the standard deviation of the errors. A lower value depicts an overall good fit. The RMSE value obtained is 4.684283 implying a good fit.

4. VWAP (Volume Weighted Average Price)

A ratio of running the cumulative of price-volume to that of the volume is plotted along with actual price movements for the entire period under study (2000–2020).

Figure 1.20

Actual close prices of SPY v/s VWAP projections



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

4 See Appendix II.

The VWAP indicator line does not reflect the trend or volatility of the actual price movements effectively. The indicator line is rather flat and is not at all conclusive. It reflects the overall trend but not at a similar scale to that of the price. The gap between actual and predicted values widens after 2013.

Accuracy measures

	ME	RMSE	MAE	MPE	MAPE
Training set	36.31693	63.11259	43.72535	14.71358	22.44715

Source: author's own elaboration using R software.

A few accuracy measures have been shown above for the model. For this analysis, the following MPE and RMSE are observed.

Mean Percentage Error (MPE)

The Mean Percentage Error is the computed average of the percentage errors by which forecast of a model differs from actual values of the quantity being forecasted. The MPE value obtained is 14.71358 implying that the difference between the forecasted values and the actual values is significant[5].

Root Mean Squared Error (RMSE)

The Root Mean Squared Error is a measure of the concentration of the actual datum points around the best fit. It is the standard deviation of the errors. A lower value depicts an overall good fit. The RMSE value obtained is 63.11259 implying not a good fit.

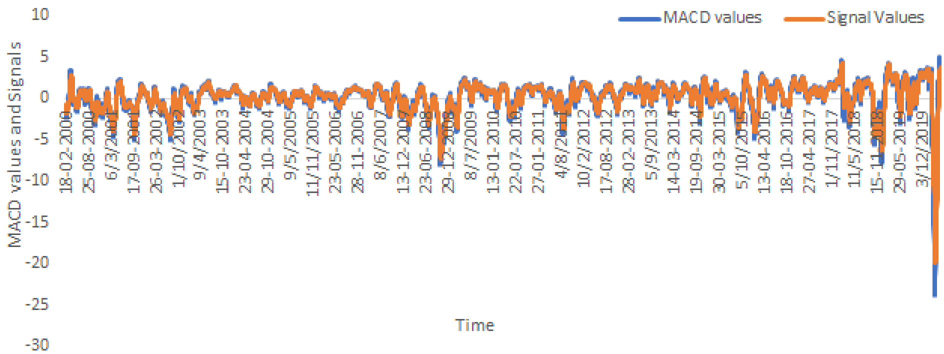
5. MACD (Moving Average Convergence Divergence)

The following plot depicts the MACD line, i.e. difference between slow and fast averages for the entire period under study (2000–2020). It also plots necessary signal lines for the corresponding period. A positive or upward momentum in the price movement of the stock is marked when the signal line crosses over MACD line and a negative movement in the price when MACD line crosses the signal line.

5 See Appendix II.

Figure 1.21

MACD and signal lines for SPY ETF

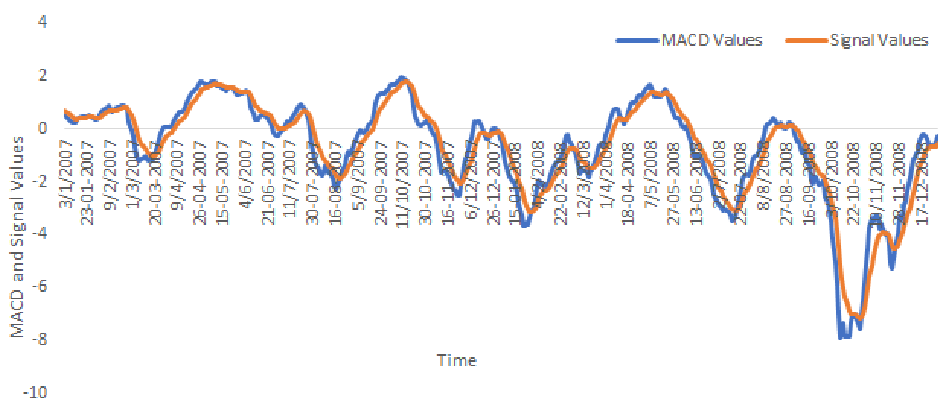


Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

The consolidated graph for 20 years does not clearly indicate the bullish or bearish nature due to its sheer scale. The following graphs consider 2007-2008 and 2019–2020 periods for a comprehensible analysis.

Figure 1.22

MACD and signal lines for SPY ETF



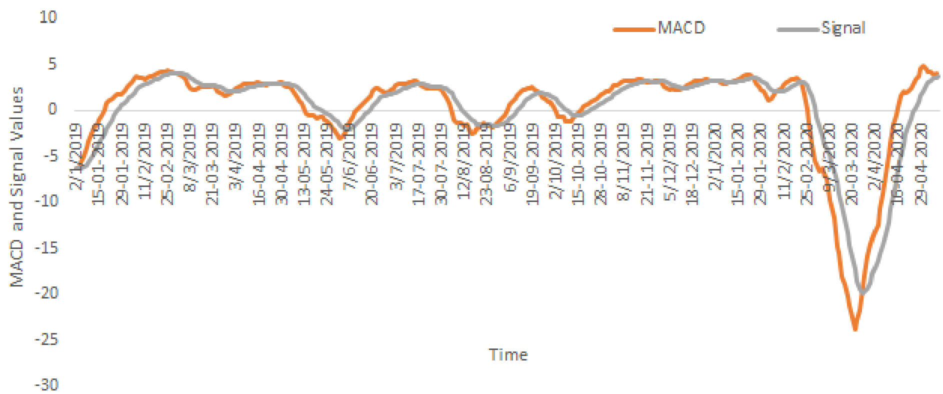
Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

Buying opportunities are presented in mid-2007 and April–May 2008. A buying opportunity is recognised when the MACD line crosses over the signal line in the first quadrant.

Selling opportunities are presented in the periods of July–August 2007, December 2007–January 2008 and August–October 2008. When the MACD line crosses under the signal line, it is taken as an accurate signal for selling.

Figure 1.23

MACD and signal lines for SPY ETF



Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

The plot for the 2019–2020 period indicates that there were selling opportunities towards the end of the first quarter of 2020. This may be since the market was able to realise the consequence of the pandemic that may have lasting effects throughout the year in the market, hence indicating a potentially capped loss if one sells as early as possible.

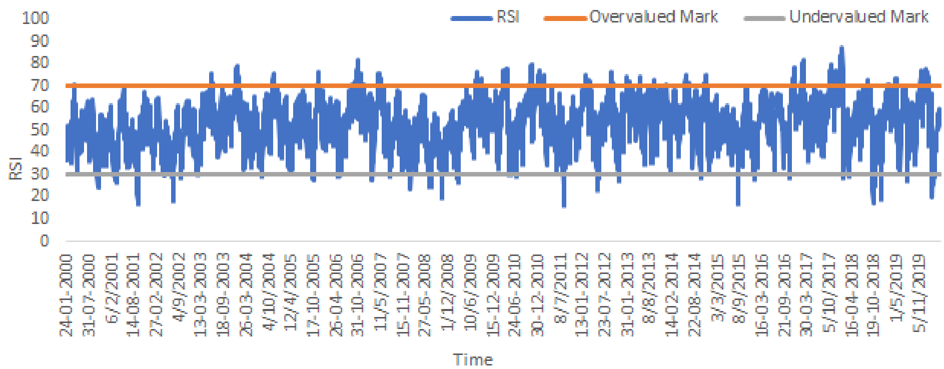
The market shows signs of the bullish nature throughout 2019 with a few exceptions. However, it clearly entered the bearish phase from early 2020 onwards.

6. RSI (Relative Strength Indicator)

RSI is on a scale of 0 to 100. It indicates whether a security or asset is overvalued or undervalued. Hence, it gives an idea whether it may reverse the trend and indicate a potentially correct time to buy or sell (Figure 1.24).

Figure 1.24

RSI of SPY ETF



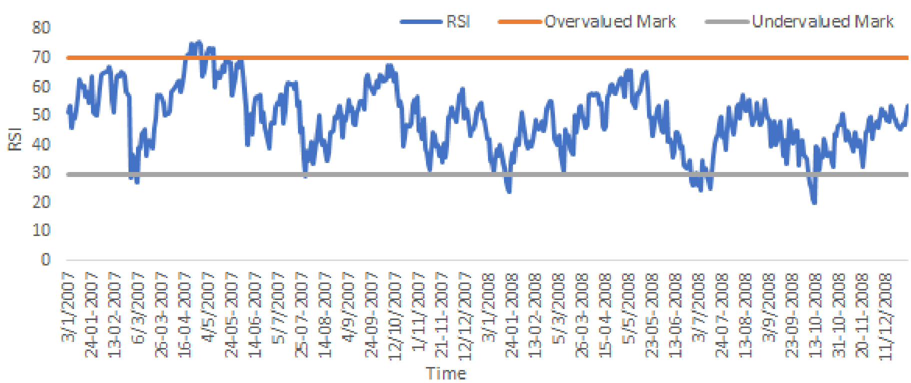
Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

A comprehensive graph over 20 years shows that there are relatively fewer times when the index price was undervalued compared to when it was overpriced.

To get clear signals to buy or sell, RSI for a relatively lower time scale is plotted below.

Figure 1.25

RSI of SPY ETF

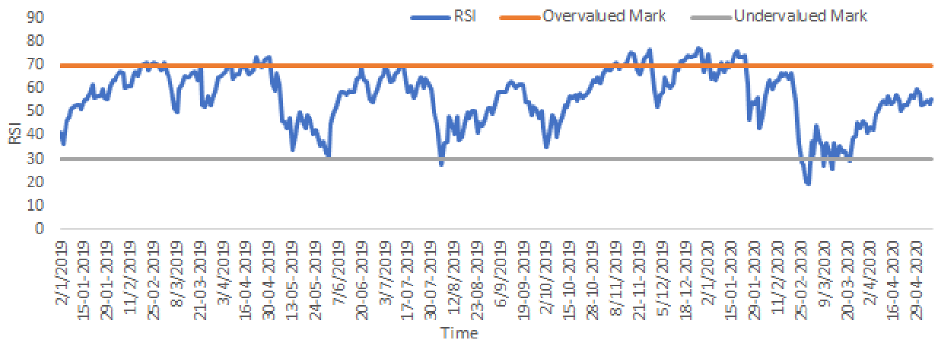


Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

For the period of 2007–2008, the SPY ETF price was neither overvalued nor undervalued with a few exceptions (April–May 2007, July and October 2008). RSI for SPY fluctuates between 30 and 70, indicating it was fairly valued during the financial crisis.

Figure 1.26

RSI of SPY ETF



Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

The SPY ETF value is observed being overpriced during November 2019, January 2020 and April 2019. The price shows signs of undervaluation in February–March 2020 followed by a rise in the RSI, which brings it back to the fair valuation bracket.

After a comprehensive and holistic analysis of all 6 indicators, it is concluded that Bollinger Bands are quite efficient in predicting the pricing channels. Since they incorporate the volatility in the data series, it is a robust indicator of the stock price movements. A comparative analysis of the actual prices indicates that the actual price in real time can never cross over the upper and lower band; hence, giving a robust boundary or range for the price movements.

The indicators like MACD and RSI can also give buy-sell cues by prompting the presence of the bearish-bullish market or the overvalued-undervalued stock.

Furthermore, the choice of an indicator for trading purposes also depends on the time period of investments (short/medium/long-term) and the risk appetite of the investor.

However, coupling 2 or 3 indicators together, for instance Bollinger Bands (measuring volatility) and MACD (incorporating trend and momentum) can give robust signals to make a trade.

3.2 Results for the IXIC Index

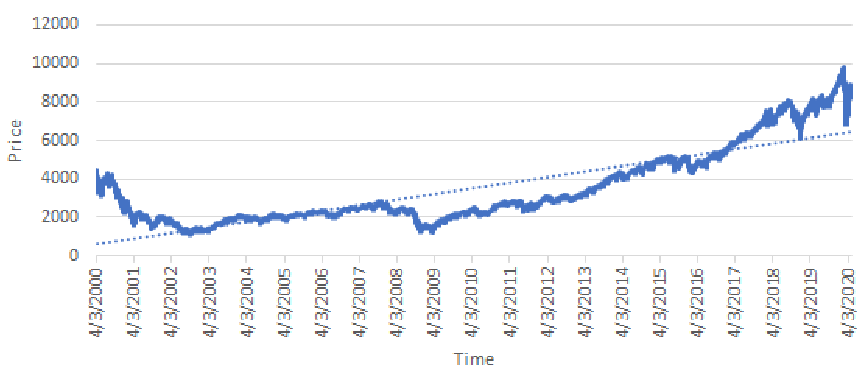
First, the results for the IXIC index have been presented. The econometric analysis is followed by the graphical analysis of different indicators. The forecasts from the fitted model are thereby analysed with the values predicted from the indicators.

Tests for non-stationarity

1. The plot of adjusted closing price of IXIC indicates that there is a trend in the data. This means that the data exhibits non-stationarity. Furthermore, from the decomposition, we can see there is a seasonal component and the variance is not constant throughout. This gives a further indication that the data is non-stationary.

Figure 2.1

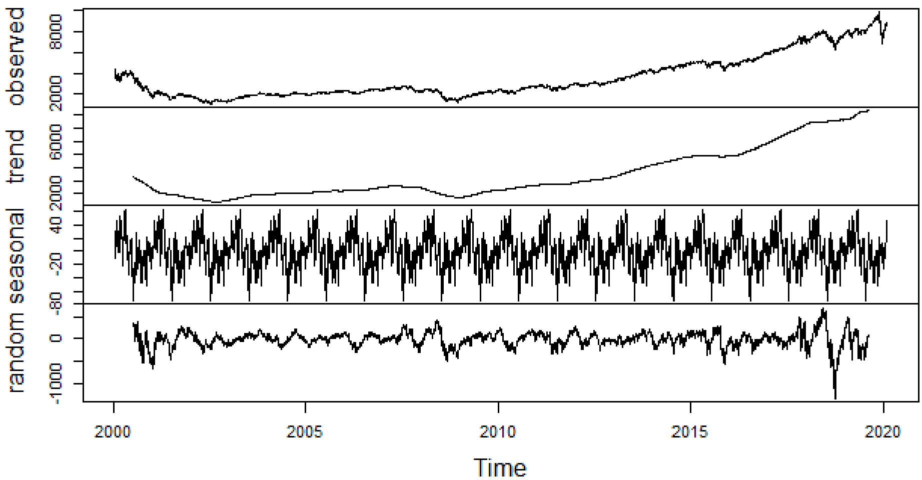
Adjusted closing price of IXIC



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Figure 2.2

Decomposition of additive time series (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

2. Augmented Dickey-Fuller test

Dickey-Fuller	Lag order	p-value
-2.254	17	0.4708

Source: author's own elaboration using R software.

3. KPSS test

KPSS trend	Truncation lag parameter	p-value
9.0788	10	0.01

Source: author's own elaboration using R software.

A high p-value is observed in the ADF test and the null hypothesis is accepted. However, a low p-value for the KPSS test leads to a rejection of the null hypothesis. Thus, the results of the KPSS and ADF tests indicate that there is presence of non-stationarity.

Transforming a non-stationary series to a stationary series

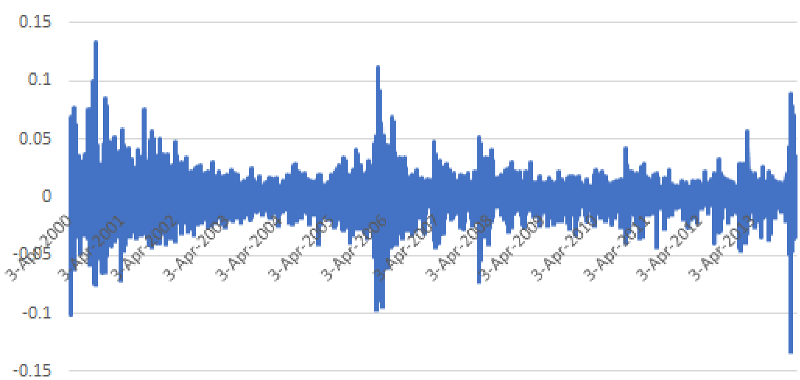
Since non-stationarity has been identified, log transformation and first differencing are done to see if it is converted into a stationary series. Tests for non-stationarity are run again and the following results are obtained.

1) Change in stock prices after logarithmic transformation and first differencing

The plot appears to be randomly distributed around zero indicating the possibility that the log transformed first differenced series is stationary.

Figure 2.3

Change in log prices (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

2) ADF test after differencing

Dickey-Fuller	Lag order	p-value
-9.6148	0	0.01

Source: author's own elaboration using R software.

3) KPSS test after differencing

KPSS trend	Truncation lag parameter	p-value
0.14245	10	0.0568

Source: author's own elaboration using R software.

In comparison to the results obtained before the first differencing, we obtain the opposite results, indicating that the first differenced series is stationary. Subsequently, the paper proceeds to use the ACF and PACF plots of the log differenced series to identify the correct model.

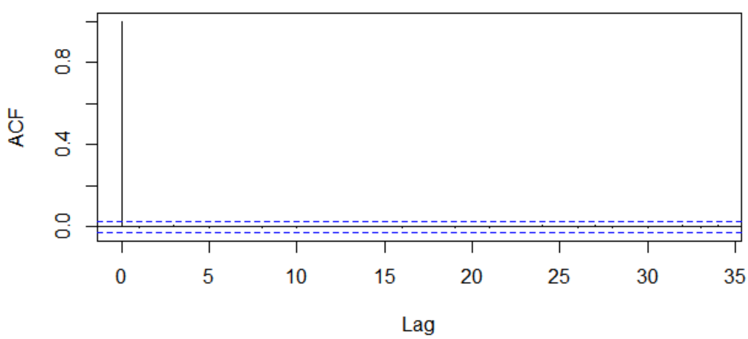
Identifying the model

To identify the p, d, q values of the ARIMA model, the ACF and PACF are observed.

1) ACF plot of log-transformed first differenced series

Figure 2.4

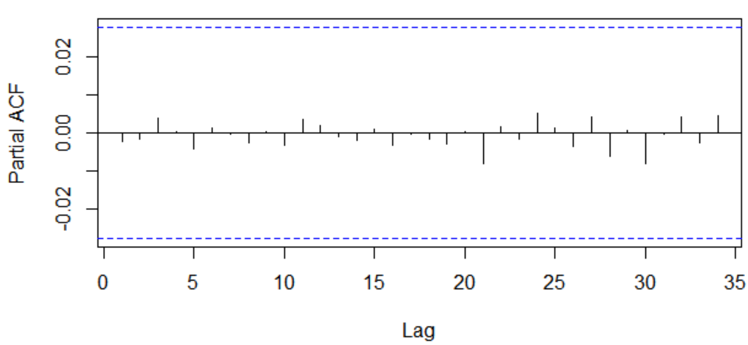
ACF (IXIC)



Source: author’s own elaboration based on SPY ETF (2000–2020) using R software.

Figure 2.5

PACF (IXIC)



Source: author’s own elaboration based on SPY ETF (2000–2020) using R software.

The ACF and PACF plot are insignificant at all time lags, indicating a value for p (AR terms) and q (MA terms). Since there is no significant line above the confidence interval of the ACF and PACF plot, this indicates the zero value for both p and q and hence, a random walk process.

Testing to find the most efficient model

The results for ARIMA (0,1,0), ARIMA (1,1,1), ARIMA (1,1,0), ARIMA (0,1,1) & ARIMA (5,2,0) are computed[6]. According to the results, ARIMA (5,2,0) has the best fit since it has the lowest AIC and BIC values. Now a check needs to be done to conclude it is the best fit model.

Best fit model

ARIMA (5,2,0) with drift

	VALUES
Sigma Squared	0.0002523
AIC	-27521.72
AICc	-27521.72
BIC	-27515.19

Source: author's own elaboration using R software.

	AR1	**AR2**	**AR3**	**AR4**	**AR5**
Coefficients	-0.8848***	-0.7235***	-0.5204***	-0.3296***	-0.1572***
SE	0.0139	0.0181	0.0194	0.0181	0.0139

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Source: author's own elaboration using R software.

A lower BIC and AIC value implies that ARIMA (5,2,0) is the best fit model for our analysis[7].

[6] See Appendix I.

[7] Even though the ACF and PACF plots suggest an ARIMA (0,1,0) model, R software function (auto. arima) gives the best fit model based on the lowest AIC, BIC, Maximum Likelihood Estimation and a few unit root tests combined.

Thus, the final model for our analysis is given by:

$$y_t = -0.8848y_{t-1} - 0.7235y_{t-2} - 0.5204y_{t-3} - 3296y_{t-4} - 0.1572y_{t-5}$$

where, y_i is log of prices.

A 1% change in previous period's index price, on average, will lead to a 0.88% decline in the value of IXIC in the current period. Similarly, the percentage decline in the current period's index price can be viewed with respect to changes in prices of the previous 5 periods.

A very important and interesting insight from these results is that the impact of past prices declines as the lag increases. This is consistent with the stochastic property of the asset prices.

Diagnostic checking

The next step is to check the significance of autocorrelation coefficients. The output for the Ljung-Box test is presented below:

Q*	DF	p-value
333.94	5	< 2.2e – 16

Source: author's own elaboration using R software.

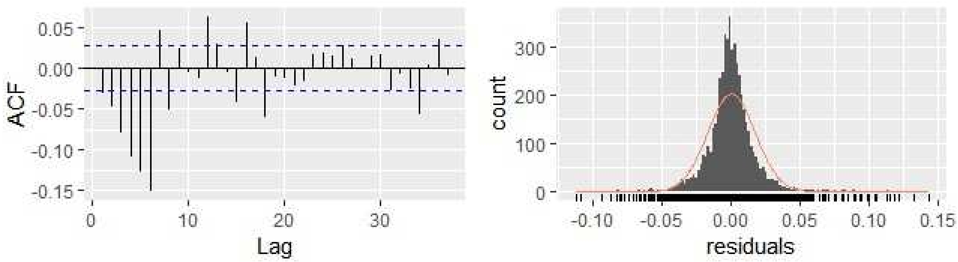
The null hypothesis is rejected at 1% level of significance implying that the residuals do not follow a white noise process and hence, are not completely random.

Additionally, an introspection of the histogram of residuals indicates that the residuals follow normal with the mean zero and a non-constant variance[8].

[8] Typically, financial data haves eteroskedasticity or non-constant variance. This is clearly observed in the first graph where we plotted prices over time. Here, logarithmic transformations were initially taken to capture this increasing variance. However, autocorrelations in variance calls for fitting ofre sophisticated models (like ARCH, GARCH, EGARCH, etc.) toonhe financial time series.

Figure 2.6

ACF v/s Lag & Count v/s Residuals (IXIC)



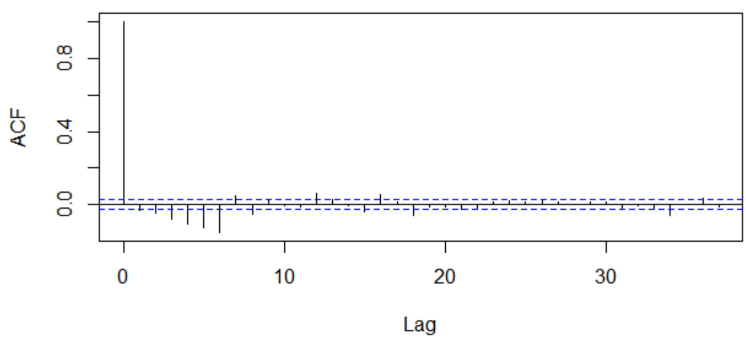
Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

ACF and PACF of residuals

If there are spikes outside the insignificant zone for both the ACF and PACF plots, we can conclude that residuals are non-random as concerns the information in them. From the plots, we can see that the mean of the residuals is very close to zero; however, the correlation between residuals is non-zero. The time plot of the residuals shows that the variation in the residuals varies over time.

Figure 2.7

ACF Residuals (IXIC)

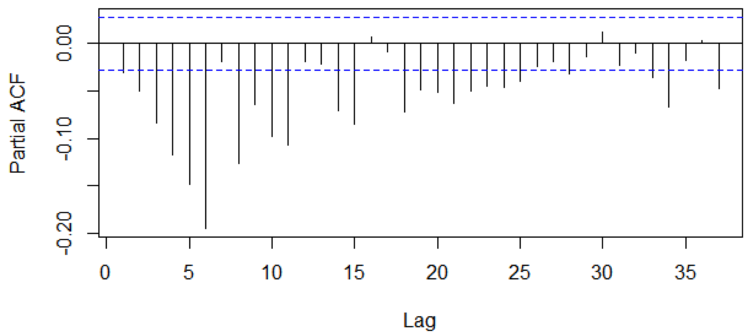


Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Since, the partial ACF confirms autocorrelations in the residuals, a model that captures the non-constant variance feature of the time series must be employed (Figure 2.8).

Figure 2.8

PACF residuals (IXIC)



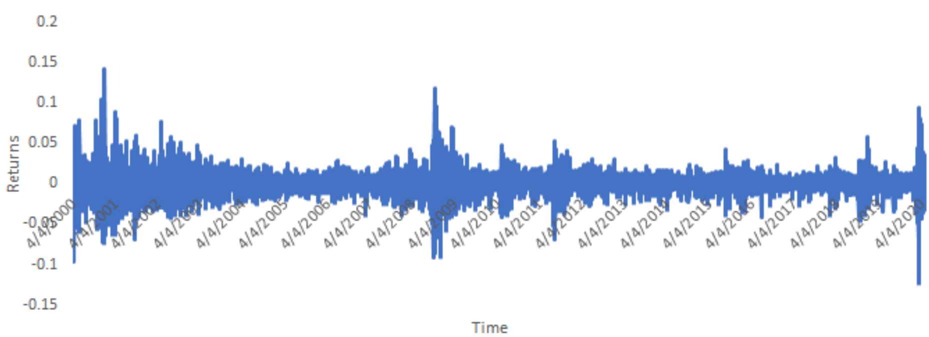
Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Volatility models

To examine the series again, a graph of daily returns of IXIC is plotted over time.

Figure 2.9

Daily returns (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Here, we observe a period of large changes followed by further large changes and those of smaller changes followed by further smaller changes. The values fluctuate unpredictably from period to period, hence indicating a volatile time series.

Testing ARCH Effects

ARCH LM test

The mean equation is estimated, and the residuals obtained are squared and regressed on one lagged squared residual.

Chi-squared	df	p-value
410.27	1	< 2.2e – 16

Source: author's own elaboration using R software.

A low p-value is observed in the ARCH LM test and the null hypothesis is rejected. Thus, the results of the test indicate that there is presence of the ARCH effects.

Estimating GARCH models

ARFIMA (1,0,1) GARCH (1,1)

	mu	omega	Alpha 1	Beta 1	AR 1	MA 1
Coefficients	0.000830***	0.000003***	0.109776***	0.876520***	0.944054***	-0.963788***
SE	0.000092	0.000001	0.009903	0.010643	0.004667	0.000704

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Source: author's own elaboration using R software.

Information criterion statistics

	VALUES
HQIC	-5.9839
AIC	-5.9866
BIC	-5.9789
SIC	-5.9866

Source: author's own elaboration using R software.

$R_t = 0.000830 + 0.944054R_{t-1} - 0.963788\varepsilon_{t-1} + \varepsilon_t$
$(\varepsilon_t = \sigma_t * z_t)$
$\sigma_t^2 = 0.000003 + 0.109776\sigma_{t-1}^2 + 0.876520\varepsilon_{t-1}^2$- ---- *
$(z_t \sim N(0,1))$

A \$1 increment in the previous period in IXIC returns, on average, will lead to a \$ 0.94 increase in the returns value of IXIC in the current period.

Also, from the third equation, small (or large) volatility changes are followed by similar smaller (or larger) changes.

Diagnostic checking (standardised residuals tests)

The next step is to check the significance of coefficients. The output for the Ljung-Box tests, the LM ARCH test and a few other tests are presented below.

Tests	Statistic	p-value
Weighted Ljung-Box Test [on standardised residuals]	0.08484	0.7708
Weighted Ljung-Box Test [on standardised squared residuals]	2.654	0.10331
LM Arch Test [TR^2]	0.08859	0.7660

Source: author's own elaboration using R software.

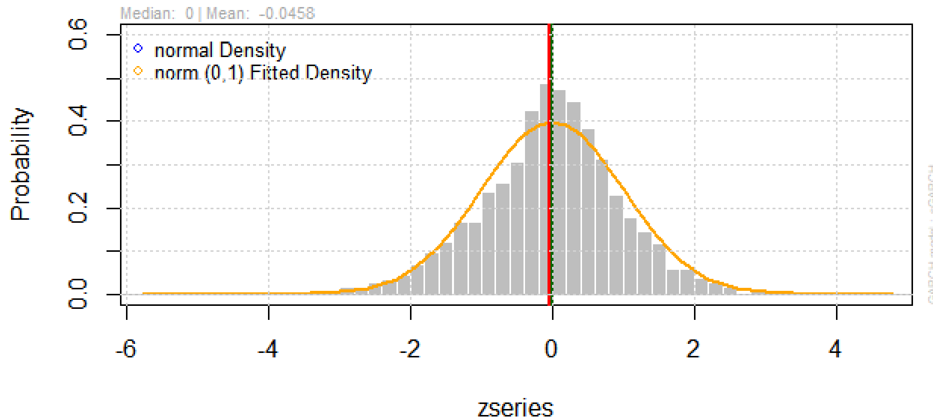
Ljung-Box test on standardised and standardised squared residuals and the LM ARCH test accept the null hypothesis and is significant at 5% level of significance.

Thus, it is concluded that the errors store no additional relevant information and have no correlation. The errors behave like a white noise process. Also, there are no further Arch effects.

Thus, the model ARFIMA (1,0,1) GARCH (1,1) passes the diagnostic checking and is the best fit model for the IXIC index time series. Furthermore, to confirm the robustness of the model over the time series, we plot the following graphs.

Figure 2.10

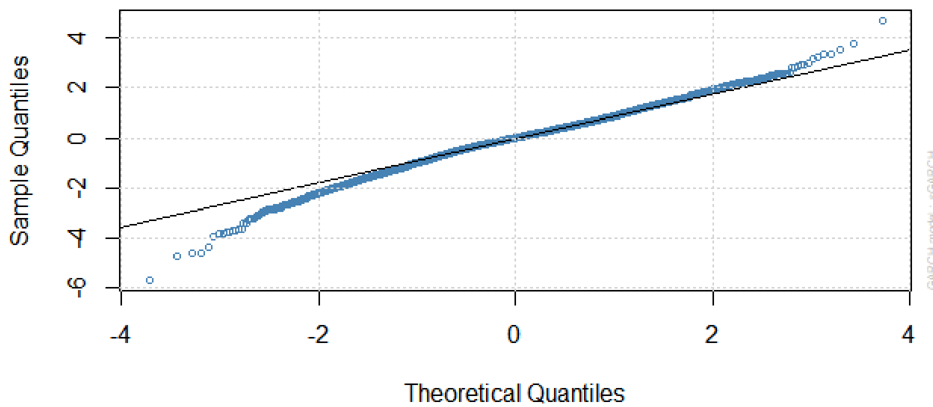
Empirical density of standardised residuals (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Figure 2.11

Norm QQ plot (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

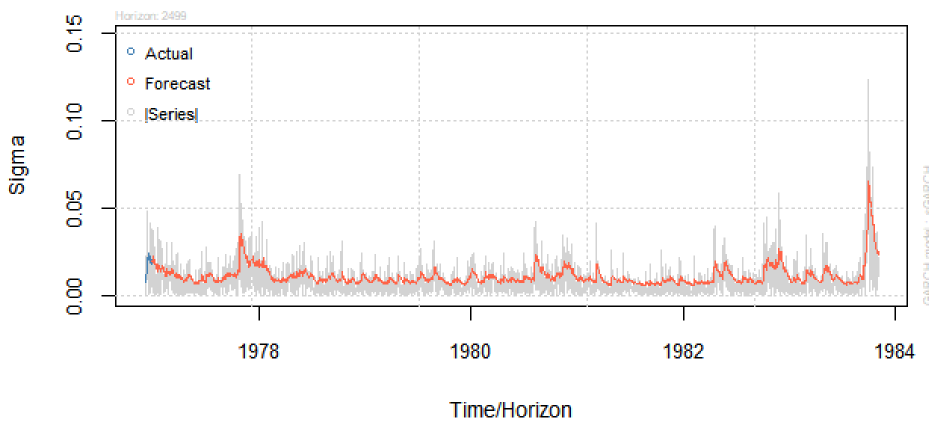
The two plots confirm the randomness of the residuals. The QQ plot of residuals further suggests the normal distribution of errors. The QQ plot resembles a straight line suggesting standard normal distribution with the mean zero and standard deviation one. This plot arranges the sample data in an ascending order and plots them against quantiles from a theoretical distribution (here, standard normal distribution).

These plots and the results from Ljung-Box tests confirm no autocorrelation in the errors and hence the robustness of the model ARFIMA (1,0,1) GARCH (1,1).

Forecast

Figure 2.12

Forecast Rolling sigma v/s [series] (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Here, the predicted conditional volatility is represented as Forecast Rolling Sigma. The conditional mean is the series at time t+h. The model predicts efficiently the volatility path, i.e. spikes and dips. However, the initial spike or dip is not predicted since it depends on the past values of the time series.

Results from various indicators for IXIC

Detailed analysis of the predicted values from the indicators like Bollinger Bands, MACD, RSI, SMA, EMA and VWAP is hereby presented along with a few accuracy measures to assess the relatively efficient indicator.

Since the period under study is 2000–2020, to track and compare the actual prices against the projected values, we need to consider a slimmer time frame. Hence, the graphs for the financial crisis period, i.e., 2007–2008, and the recent period 2019–2020, are presented for each indicator.

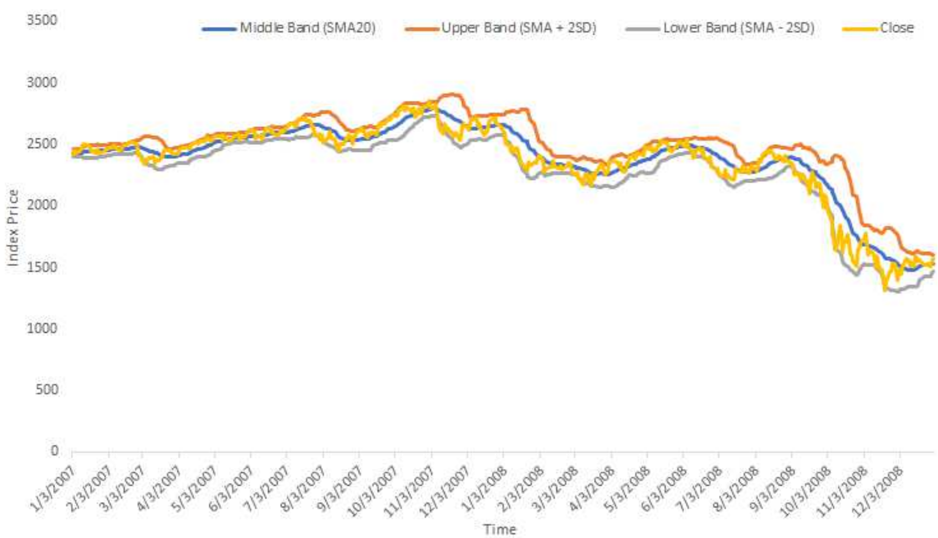
The period of the financial crisis is chosen because it was a highly volatile time period for all securities. The recent period is chosen to gain a better understanding of the market in the current scenario.

1. Bollinger Bands

The Bollinger Bands for the entire period under study (2000–2020) are not presented as a longer time scale compresses the observations at each time point and does not give a clear picture of the movement of the bands. A magnified version **for 2007–2008 and 2019–2020** is presented as follows.

Figure 2.13

Actual Index price v/s Bollinger bands (IXIC)



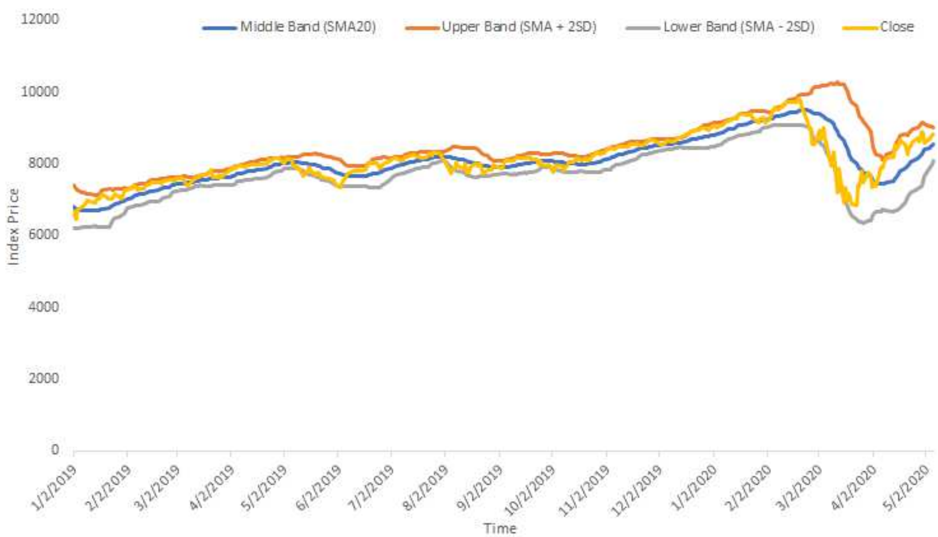
Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

The yellow band, indicating the actual prices is observed to be quite volatile throughout. All three bands are quite close at the beginning and spread out later in the year, next incorporating the higher volatile nature of the market. One useful observation here is that the actual prices oscillate between the upper and lower bands but never cross these boundary rates (Figure 2.14).

One major similarity that is observed is the expansion of the Bollinger Bands with the onset of financial stress on the market. As this was later observed in 2007 and 2008, a similar trend is observed from February 2020 onwards. Increased volatility in the first quarter of 2020 is successfully incorporated in the wider ranges of the upper and lower bands. Once again, it is worth noting that the actual price movements never cross the boundary bands, hence proving the efficiency of the volatility indicator.

Figure 2.14

Actual Index price v/s Bollinger bands (IXIC)



Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

2. SMA (Simple Moving Average)

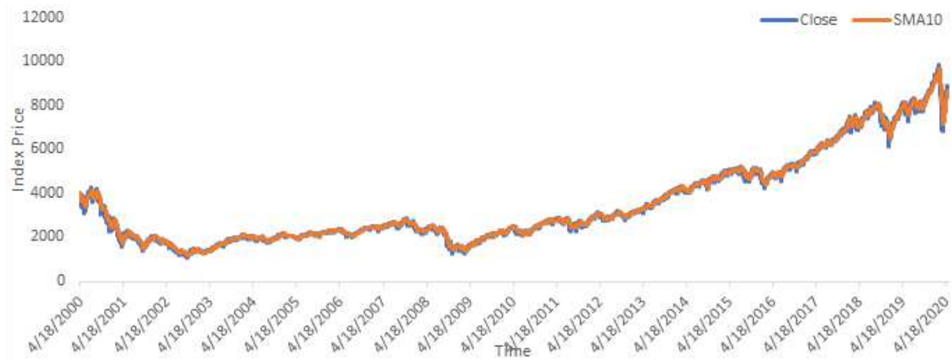
The graph comparing the actual prices and the SMA projections is presented for the period under study, i.e., 2000–2020 (Figure 2.15).

Since the time period under study is very long, it does not paint a clear picture when it comes to forecast errors. A detailed and microscopic analysis is necessary to visualise the effectiveness of the indicator. The following graphs are for the period of 2007–2008 and 2019–2020 (Figure 2.16).

It is observed that the actual closing price has a more volatile motion than the projected values. It is also very useful to observe that any trend change is incorporated in the SMA line a few periods later. It is deductible that SMA is a little slower in realising the change in price movements.

Figure 2.15

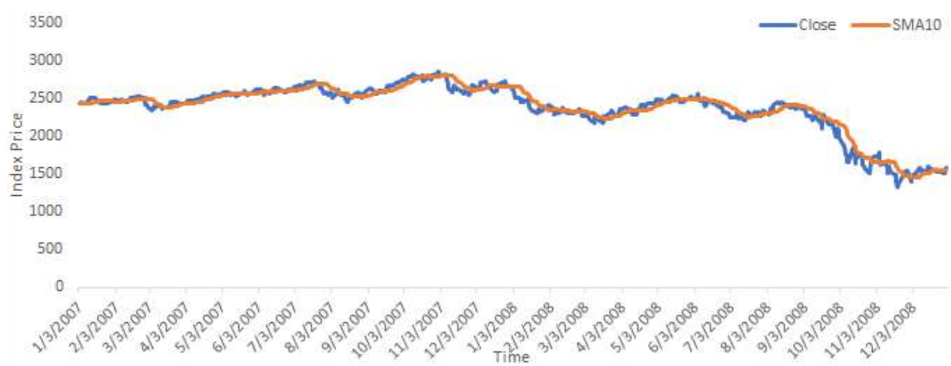
Actual Index price v/s SMA projections (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Figure 2.16

Actual Index price v/s SMA projections (IXIC)



Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

This is also consistent with the theoretical fact that it assigns equal weightage to all observations and hence reflects any opposite movement with a lag (Figure 2.17).

A similar observation is made in the graph for the period 2019–2020. A lag in the realisation of sudden and adverse price movements is noticed in the predicted values.

Figure 2.17

Actual Index price v/s SMA projections (IXIC)

Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

Accuracy measures

	ME	RMSE	MAE	MPE	MAPE
Training set	6.328891	115.653	73.8213	0.03834196	2.272291

Source: author's own elaboration using R software.

A few accuracy measures for the model are mentioned in the table above. For this analysis, the MPE and RMSE are observed.

Mean Percentage Error (MPE)

The Mean Percentage Error is the computed average of the percentage errors by which the forecast of a model differs from actual values of the quantity being forecasted. The MPE value obtained is 0.03834196 implying that the difference between the forecasted values and the actual values is not significant.

Root Mean Squared Error (RMSE)

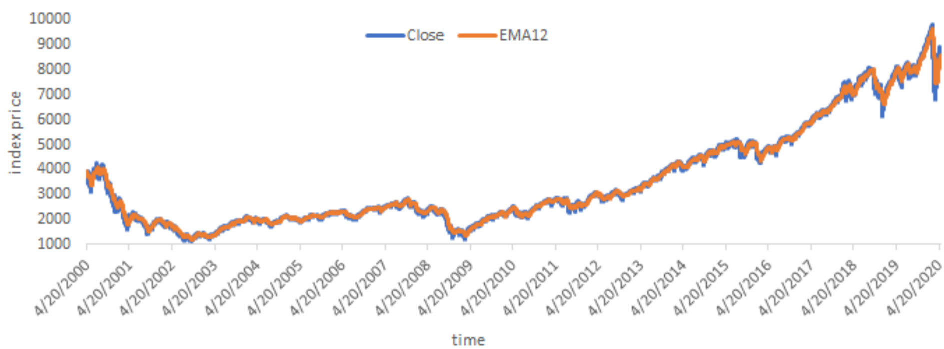
The Root Mean Squared Error is a measure of the concentration of the actual datum points around the best fit. It is the standard deviation of the errors. A lower value depicts an overall good fit. The RMSE value obtained is 115.653, implying a not good fit.

3. EMA (Exponential Moving Average)

EMA for a 12-day period is computed and plotted. This is specifically useful for short-term traders. This look back period is increased when medium-or long-term investments are considered.

Figure 2.18

Actual Index price v/s EMA projections (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

Once again, the long time period makes it difficult to visually track the actual and predicted movements. Thus, two different shorter time frames are studied separately for two periods (2007–2008 and 2019–2020).

Figure 2.19

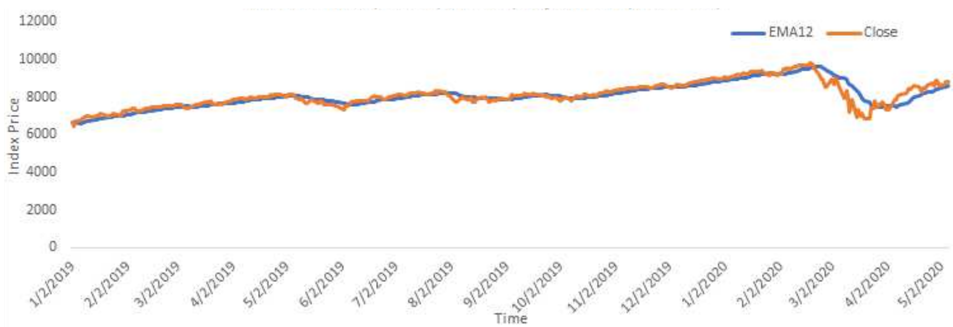
Actual Index price v/s EMA projections (IXIC)

Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

EMA is faster in incorporating price changes than SMA since it gives more weightage to recent observations. The effects of the financial crisis are clearly visible in the early 2008.

Figure 2.20

Actual Index price v/s EMA projections (IXIC)



Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

The shock from the black swan event is not effectively captured in the early 2020. However, the downward trend is realised after a small lag and so is the (temporary) recovery in April 2020. The volatile motion, however, is not mirrored on the indicator line.

Accuracy measures

	ME	RMSE	MAE	MPE	MAPE
Training set	8.690561	139.5241	73.72134	0.07838845	2.264921

Source: author's own elaboration using R software.

A few accuracy measures for the model are mentioned in the table above. For this analysis, the following is observed.

Mean Percentage Error (MPE)

The Mean Percentage Error is the computed average of the percentage errors by which forecast of a model differs from the actual values of the quantity being forecasted. The MPE value obtained is 0.07838845, implying that the difference between the forecasted values and the actual values is not significant.

Root Mean Squared Error (RMSE)

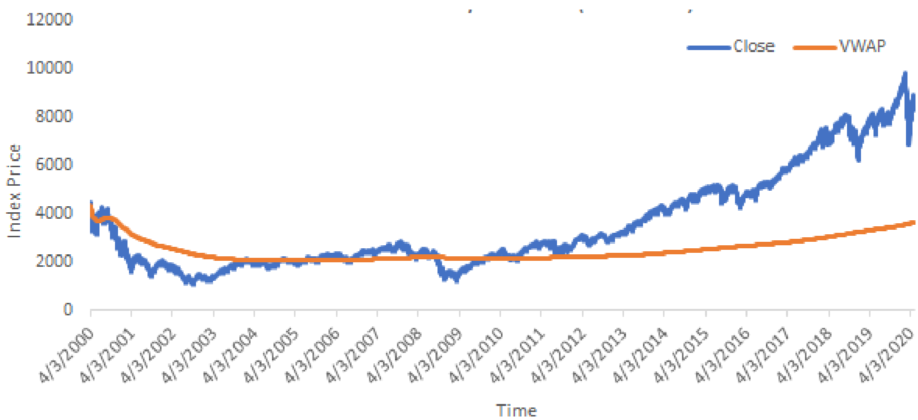
The Root Mean Squared Error is a measure of the concentration of the actual datum points around the best fit. It is the standard deviation of the errors. A lower value depicts an overall good fit. The RMSE value obtained is 139.5241, implying a not good fit.

4. VWAP (Volume Weighted Average Price)

A ratio of running- cumulative of price-volume to that volume is plotted along with actual price movements for the entire period under study (2000–2020).

Figure 2.21

Actual Index price v/s VWAP (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

The VWAP indicator line does not reflect the trend or volatility of the actual price movements effectively. The indicator line is rather flat and is not at all conclusive. It reflects the overall trend but not at a similar scale to that of the price. The gap between actual and predicted values widens after 2013–2014. The closing price upward trend is followed by the VWAP but is not at the same scale.

Accuracy measures

	ME	RMSE	MAE	MPE	MAPE
Training set	1039.626	2000.735	1358.959	12.73405	31.62165

Source: author's own elaboration using R software.

Above, there are a few accuracy measures for the model. For this analysis, the MPE is observed.

Mean Percentage Error (MPE)

The Mean Percentage Error is the computed average of the percentage errors by which forecast of a model differs from actual values of the quantity being forecasted. The MPE value obtained is 12.73405 implying that the difference between the forecasted values and the actual values is significant.

Root Mean Squared Error (RMSE)

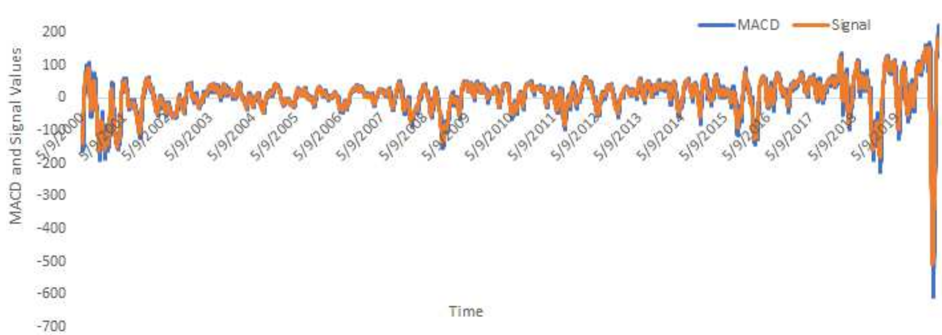
The Root Mean Squared Error is a measure of the concentration of the actual datum points around the best fit. It is the standard deviation of the errors. A lower value depicts an overall good fit. The RMSE value obtained is 2000.735, implying it is not a good fit.

MACD (Moving Average Convergence Divergence)

The following plot depicts the MACD line, i.e. the difference between slow and fast averages for the entire period under study (2000–2020). It also plots a signal line for the corresponding period. A positive or upward momentum in the price movement of the stock is marked when the signal line crosses over the MACD line, whereas a negative movement in the price when the MACD line crosses the signal line.

Figure 2.22

MACD and signal lines (IXIC)

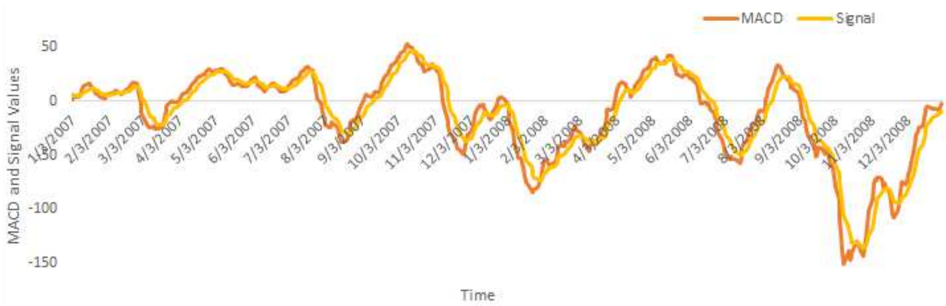


Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

The consolidated graph for 20 years does not clearly indicate the bullish or bearish nature due to its sheer scale. The other graphs present a period between **(2007–2008 and 2019–2020)** for a comprehensible analysis.

Figure 2.23

MACD and signal lines (IXIC)



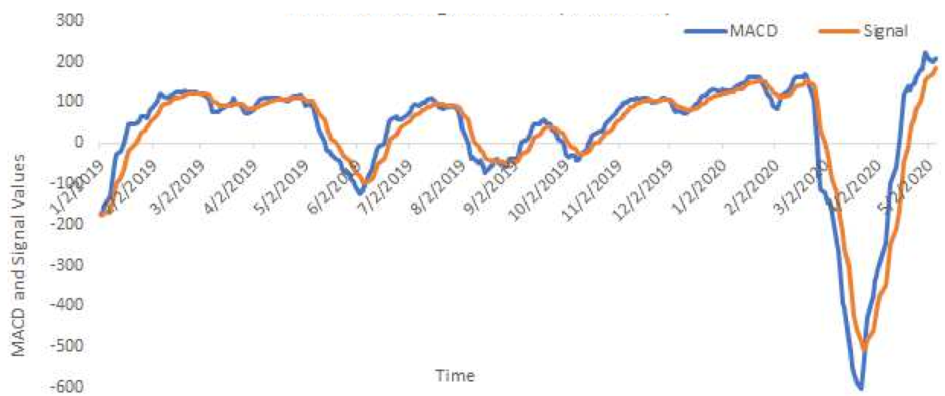
Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

Buying opportunities are presented **in March–April 2007, September–October 2007 and April–May 2008**. A buying opportunity is recognised when the MACD line crosses over the signal line in the first quadrant.

Selling opportunities are presented in **January–February 2008, October–November 2008.** When the MACD line crosses under the signal line, it is taken as an accurate signal for selling.

Figure 2.24

MACD and signal lines (IXIC)



Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

The plot for the 2019–2020 period indicates that there were selling opportunities towards the end of the first quarter of 2020. This may be since the market was able to realise the consequence of the pandemic, which may have lasting effects on the market throughout the year, hence, indicating a potentially capped loss if one sells as early as possible.

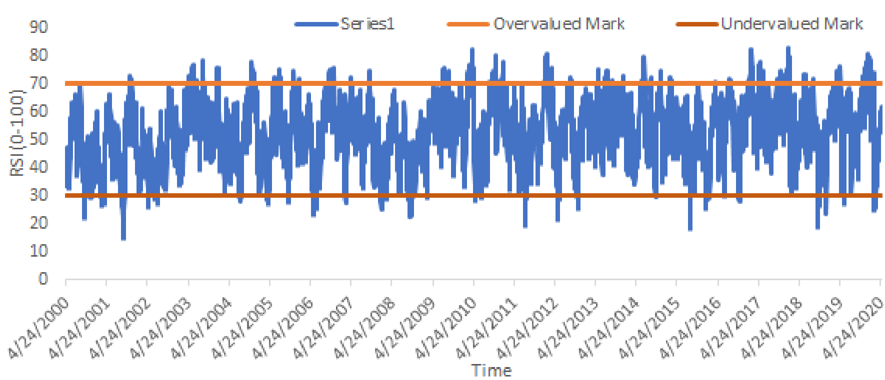
The market shows signs of the bullish nature throughout 2019 with a few exceptions. However, it clearly entered the bearish phase from the early 2020 onwards with a little sign of recovery from the very latest data (May 2020).

RSI (Relative Strength Indicator)

RSI is on a scale of 0 to 100. It indicates whether a security or asset is overvalued or undervalued. Hence, it gives an idea whether it may reverse the trend and indicate a potentially correct time to buy or sell.

Figure 2.25

RSI (IXIC)



Source: author's own elaboration based on SPY ETF (2000–2020) using R software.

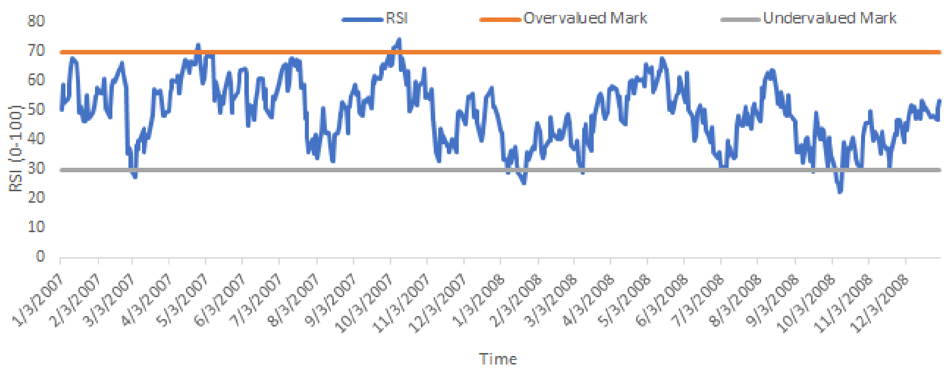
A comprehensive graph over 20 years shows that there are a relatively fewer times when the index price was undervalued than it was overpriced.

To get clear signals to buy or sell, RSI for a relatively lower time scale is plotted as follows.

For the period of 2007–2008, the IXIC index price was neither overvalued nor undervalued with a few exceptions (March 2008). RSI for the IXIC index fluctuates between 30 and 70, indicating it was fairly valued during the financial crisis (Figure 2.27).

Figure 2.26

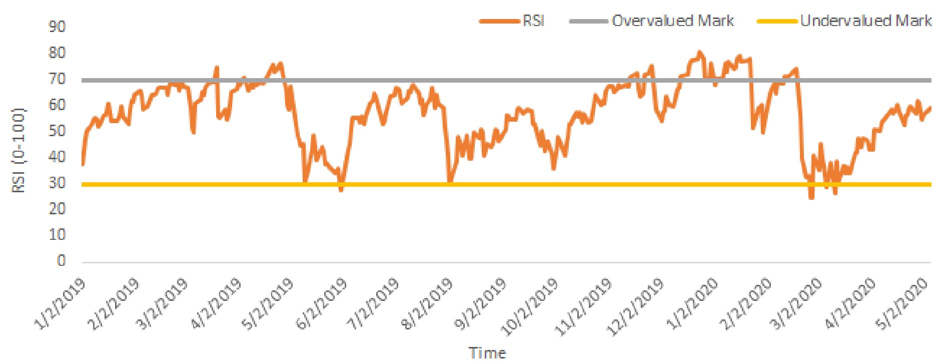
RSI (IXIC)



Source: author's own elaboration based on SPY ETF (2007–2008) using R software.

Figure 2.27

RSI (IXIC)



Source: author's own elaboration based on SPY ETF (2019–2020) using R software.

The IXIC index value is observed being overpriced during **December 2019 to February 2020 and in March and May 2019**. The index price shows signs of undervaluation in March 2020 followed by a rise in the RSI, which brings it back to the fair valuation bracket.

After a comprehensive and holistic analysis of all 6 indicators, it is concluded that Bollinger Bands are quite efficient in predicting the pricing channels. Since this incorporates the volatility in the data series, it is a robust indicator of the stock price movements. A comparative analysis with the actual prices indicates that the actual price in real time can never cross over the upper and lower band. Hence, giving a robust boundary or range for the price movements.

Indicators like MACD and RSI can also give buy-sell cues by prompting the presence of the bearish-bullish market or overvalued-undervalued stock.

Furthermore, the choice of an indicator for trading purposes also depends on the time of investments (short/medium/long- term) and the risk appetite of the investor.

However, coupling 2 or 3 indicators together, for instance, Bollinger Bands (measuring volatility) and MACD (incorporating trend and momentum), can give robust signals to make a trade.

Conclusion

The research aimed to filter out the most efficient indicator for the stock market. Two time series, IXIC (NASDAQ Composite Index) and SPY ETF, for the time from 2000 to 2020 were extensively analysed for this purpose. Advanced econometrics (time series) models were used to fit the data and to make further predictions about the prices. ARFIMA along with the GARCH model were the best fit models for both time series (with separate orders). The models passed the diagnostic checks and hence the forecasts from them are considered to be quite accurate.

Traditional indicators, like Bollinger Bands, SMA, EMA, MACD, VWAP and RSI, used by traders on the financial markets, were also analysed alongside the actual closing prices to ascertain the best practice to comprehend price movements. Bollinger Bands, which capture the volatility efficiently, sorted to be a distinguished indicator. Pricing decisions should be based on thorough analysis using a combination of 2–3 indicators, like Bollinger Bands (to capture volatility) and MACD (to have trend and momentum components). The RSI indicator can also give an explicit valuation of an asset or security.

However, it should be noted that the models and indicators do not absorb new qualitative information and are a part of technical analysis. The econometric models were proven to be an efficient predictor for pricing channels but are mathematically tedious for traders. Traditional indicators are quick and simpler replacements with marginally higher forecast errors.

List of References

Adebayo, F.A., Sivasamy, R., and Shangodoyin, D.K. (2014) 'Forecasting stock market series with ARIMA Model'. *Journal of Statistical and Econometric Methods* 3(3), 65–77.

Almasarweh, M., and Alwadi, S. (2018) 'ARIMA model in predicting banking stock market data'. *Modern Applied Science* 12(11).

Białkowski, J., Darolles, S., and Le Fol, G. (2008) 'Improving VWAP strategies: A dynamic volume approach'. *Journal of Banking & Finance* 32(9), 1709–1722.

Bollinger, J. (1992) 'Using bollinger bands'. *Stocks & Commodities* 10(2), 47–51.

Chande, T.S. (1992) 'Adapting moving averages to market volatility'. *Stock & Commodities* 10(3).

Chen, Y.L., and Song, F.M. (2000) 'An empirical study on the relationship between price changes and trading volume in China stock market'. *Journal of Management Sciences in China* 3(2), 62–68.

Chong, T.T.L. and Ng, W.K. (2008) 'Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30'. *Applied Economics Letters* 15(14), 1111–1114.

Cushing, D., and Madhavan, A.N. (2001) 'The hidden cost of trading at the close'. *Trading* 1, 12–19.

Edwards, R.D., Magee, J., and Bassetti, W.C. (2018) *Technical analysis of stock trends*. CRC Press.

Ellis, C.A., and Parbery, S.A. (2005) 'Is smarter better? A comparison of adaptive, and simple moving average trading strategies'. *Research in International Business and Finance* 19(3), 399–411.

Frankel, J.A., and Froot, K.A. (1990) 'Chartists, fundamentalists, and trading in the foreign exchange market'. *The American Economic Review* 80(2), 181–185.

Granger, C.W. (1968) 'Some aspects of the random walk model of stock market prices'. *International Economic Review* 9(2), 253–257.

Hamid, K., Akash, R.S.I., Asghar, M., and Ahmad, S. (2011) 'Corporate social performance, financial performance and market value behavior: An information asymmetry perspective' *African Journal of Business Management* 5(15), 6342–6349.

Johnston, F.R., Boyland, J.E., Meadows, M., and Shale, E. (1999) 'Some properties of a simple moving average when applied to forecasting a time series'. *Journal of the Operational Research Society* 50(12), 1267–1271.

Kenny, G., Meyler, A., and Quinn, T. (1990) *Forecasting Irish inflation using ARIMA models*. Central Bank of Ireland, Technical Papers 3/RT/98 (No. 3/RT/98).

Klinker, F. (2011) 'Exponential moving average versus moving exponential average'. *Mathematische Semesterberichte* 58(1), 97–107.

Kolková, A. (2017) 'Testing EMA indicator for the currency pair EUR/USD'. *International Journal of Entrepreneurial Knowledge* 5(1), 35–40.

Lento, C., Gradojevic, N., and Wright, C.S. (2007) 'Investment information content in Bollinger Bands'. *Applied Financial Economics Letters* 3(4).

Li, G., and Zhu, J. (2014) June, Research on the Effectiveness of Technical Indicators with the Volume, in *2014 International Conference on Education, Management and Computing Technology (ICEMCT-14)*, Atlantis Press, 436–439.

Lo, A.W., and Wang, J. (2000) 'Trading volume: definitions, data analysis, and implications of portfolio theory'. *The Review of Financial Studies* 13(2), 257–300.

Malkiel, B.G., and Fama, E.F. (1970) 'Efficient capital markets: A review of theory and empirical work'. *The Journal of Finance* 25(2).

Menkhoff, L., and Taylor, M.P. (2007) 'The Obstinate Passion of Foreign Exchange Professionals: Technical Analysis'. *Journal of Economic Literature* 45(4), 936–972.

Merh, N., Saxena, V.P., and Pardasani, K.R. (2010) 'A Comparison Between Hybrid Approaches of ANN and ARIMA for Indian Stock Trend Forecasting'. *Business Intelligence Journal* 3(2), 23–43.

Murphy, J.J. (2009) *The Visual Investor: How to Spot Market Trends* (Vol. 443). John Wiley & Sons.

Nitin, K. (2020) *Leveraging Fundamental and Technical Analysis for Predicting the Stock Market Movements*.

Osborne, M.F. (1959) 'Brownian motion in the stock market'. *Operations Research* 7(2), 145–173.

Peiyuan, S., and Donghui, S. (2002) 'CAPM Based Study of Herd Behavior: Evidence from Chinese Stock Market and Discussion with Song Jun and Wu Chongfent' [J]. *Economic Research Journal 2*, 64–70.

Rosillo, R., de la Fuente, D., and Brugos, J.A.L. (2013) 'Technical analysis and the Spanish stock exchange: testing the RSI, MACD, momentum and stochastic rules using Spanish market companies'. *Applied Economics* 45(12), 1541–1550.

Sahin, U., and Ozbayoglu, A.M. (2014) 'TN-RSI: Trend-normalized RSI indicator for stock trading systems with evolutionary computation'. *Procedia Computer Science* 36, 240–245.

Šedivá, B., and Marek, P. (2017) 'Stability analysis of optimal mean-variance portfolio due to covariance estimation', *Procedings of the 35th International Conference Mathematical Methods in Economics*, Univerzita Hradec Králové.

Ţăran-Moroşan, A. (2011) 'The relative strength index revisited'. *African Journal of Business Management* 5(14).

Taylor, M.P., and Allen, H. (1992) 'The use of technical analysis in the foreign exchange market'. *Journal of International Money and Finance* 11(3), 304–314.

Vaccaro, A., El-Fouly, T.H.M., Cañizares, C.A., and Bhattacharya, K. (2015) 'Local learning-ARIMA adaptive hybrid architecture for hourly electricity price forecasting'. *IEEE Eindhoven PowerTech* (1–6). IEEE.

Wang, C., and Wu, C. (2002) 'Linear and nonlinear Granger causality test of stock price-volume relation: Evidence from Chinese markets'. *Journal of Management Sciences 5*(4), 7–12.

Wang, Y.F. (2002) 'Predicting stock price using fuzzy grey prediction system'. *Expert Systems With Applications* 22(1), 33–38.

Weixing, G.L.Z. (2010) 'An Empirical Analysis of the Volume-Price Relation Using High-Frequency Data in the Chinese Stock Market'. *Chinese Journal of Management* 8.

Wilder, J.W. (1978) *New concepts in technical trading systems*. Trend Research.

Wong, W.K., Manzur, M., and Chew, B.K., (2003) 'How rewarding is technical analysis? Evidence from Singapore stock market'. *Applied Financial Economics* 13(7), 543–551.

Yunus, K., Thiringer, T., and Chen, P. (2015) 'ARIMA-based frequency-decomposed modeling of wind speed time series'. *IEEE Transactions on Power Systems* 31(4), 2546–2556.

Determining the Most Efficient Technical Indicator of Investing in Financial Markets based on Trends, Volume, Momentum and Volatility

Abstract

In this paper, we investigate the most efficient technical indicator of the top five selected indicators of investing in financial markets. The indicators are based on a trend on the entire market, volume in each time period, momentum, and volatility of the financial instruments, e.g. Bollinger Bands, SMA, EMA, VWAP, MACD, RSI respectively. Our primary focus is on financial markets of the United States; however, our research findings can be applied to any index throughout the world. This dissertation strongly supports the idea of utilisation of technical analysis to have an edge while trading on financial markets. This paper uses IXIC (NASDAQ Composite Index) and SPY (SPDR S&P 500 ETF Trust) data concerning the period from 2000 to 2020. To empirically test the predictability of these indicators, however, we used the time spans of 2007–2008 and 2019–2020. Due to the economic recession and the COVID-19 pandemic in the respective periods, the time is suitable for testing our indicators because of comparatively higher volatility on the markets.

Advanced econometrics (time series) models are used to fit the data and to make price predictions. ARFIMA along with GARCH are the best fitted models for both time series.

Key words: financial markets, indicators, volatility, Bollinger Bands, SMA, EMA, MACD, RSI, SPY, IXIC, ARIFMA, GARCH

Określenie najskuteczniejszego technicznego wskaźnika inwestowania na rynkach finansowych na podstawie trendu, wielkości, tempa i zmienności

Streszczenie

W niniejszym opracowaniu przeprowadzamy badanie najskuteczniejszego wskaźnika technicznego spośród pięciu najlepszych wybranych wskaźników inwestowania na rynkach finansowych. Są one oparte na trendzie na całym rynku, wielkości w każdym okresie, impecie oraz zmienności rynków finansowych, np. odpowiednio wstędze Bollingera, SMA, EMA, VWAP, MACD i RSI. Nasza uwaga skupia się głównie na rynkach finansowych w USA; jednakże, nasze badania mogą mieć zastosowanie do każdego indeksu na świecie. Niniejsza rozprawa stanowi silne poparcie idei wykorzystania analizy technicznej do zdobycia przewagi na rynkach finansowych. W opracowaniu wykorzystane są dane IXIS (NASDAQ Composite Index) oraz SPY (SPDR S&P 500 ETF Trust) dotyczące lat 2000–2020. Jednakże w celu empirycznego przetestowania przewidywalności tych wskaźników zastosowaliśmy przedziały czasowe 2007–2009 oraz 2019–2020. Ze względu na recesję i pandemię COVID-19 w odpowiednich okresach, są one szczególnie przydatne do przetestowania naszych wskaźników, gdyż rynki cechuje stosunkowo wysoka zmienność w tym czasie.

Modele (szeregi czasowe) zaawansowanej ekonometrii stosowane są by dopasować dane i przewidywać ceny. ARFIMA oraz GARCH są modelami najlepiej dopasowanymi do obu szeregów czasowych.

Słowa kluczowe: rynki finansowe, wskaźniki, zmienność, wstęga Bollingera, SMA, EMA, MACD, RSI, SPY, IXIC, ARIFMA, GARCH

Cite as:

Lakhwan, D., Dave, A. (2020) 'Determining the most efficient technical indicator of investing in financial markets based on trends, volume, momentum and volatility'. *Myśl Ekonomiczna i Polityczna* 3(70), 64–137. DOI: 10.26399/meip.3(70).2020.17/d.lakhwan/a.dave

Cytuj jako:

Lakhwan D., Dave A., *Determining the most efficient technical indicator of investing in financial markets based on trends, volume, momentum and volatility* [*Określenie najskuteczniejszego technicznego wskaźnika inwestowania na rynkach finansowych na podstawie trendu, wielkości, tempa i zmienności*]. „Myśl Ekonomiczna i Polityczna" 2020 nr 3(70), 64–137. DOI: 10.26399/meip.3(70).2020.17/d.lakhwan/a.dave

APPENDICES

Appendix I

The results for the other ARIMA models fitted over the SPY ETF time series, which were rejected due to the relatively higher AIC and BIC values, are as follows.

1. ARIMA (4,1,1)

	AR1	AR2	AR3	AR4	MA1
Coefficients	-0.892088	-0.089231	-0.021481	-0.010926	0.824186
SE	0.091413	0.019817	0.019045	0.014559	0.090344

	VALUES
Sigma Squared	0.000251
Log Likelihood	13777.89
AIC	-27543.78
AICc	-27543.77
BIC	-27504.62

2. ARIMA (2,1,0)

	AR1	AR2
Coefficients	-0.0690	-0.0325
SE	0.0141	0.0141

	VALUES
Sigma Squared	0.000251
Log Likelihood	13775.82
AIC	-27545.64
AICc	-27545.63
BIC	-27526.06

3. ARIMA (3,1,1)

	AR1	AR2	AR3	MA1
Coefficients	0.1523	-0.017	0.0108	-0.221

	VALUES
Sigma Squared	0.0002511
Log Likelihood	13775.86
AIC	-27541.71
AICc	-27541.7
BIC	-27509.07

4. ARIMA (5,1,0)

	AR1	AR2	AR3	AR4	AR5
Coefficients	-0.0688	-0.0321	0.0052	-0.0062	-0.0199
SE	0.0141	0.0141	0.0141	0.0141	0.0141

	VALUES
Sigma Squared	0.0002511
Log Likelihood	13776.96
AIC	-27541.93
AICc	-27541.91
BIC	-27502.76

The results for the other ARIMA models fitted over the IXIC Index time series, which were rejected due to relatively higher AIC and BIC values, are as follows.

1. ARIMA (1,1,1)

	AR1	MA1
Coefficients	0.2881	-0.3583
SE	0.1837	0.1797

	VALUES
Sigma Squared	0.0002511
Log Likelihood	13775.38
AIC	-27544.77
AICc	-27544.76
BIC	-27525.18

2. ARIMA (0,1,0)

	VALUES
Sigma Squared	0.0002523
Log Likelihood	13761.86
AIC	-27521.72
AICc	-27521.72
BIC	-27515.19

3. ARIMA (1,1,0)

	AR1
Coefficients	-0.0668

	VALUES
Sigma Squared	0.0002512
Log Likelihood	13773.15
AIC	-27542.29
AICc	-27542.29
BIC	-27529.24

4. ARIMA (0,1,1)

	MA1
Coefficients	-0.0709
SE	0.0141

	VALUES
Sigma Squared	0.0002512
Log Likelihood	13773.86
AIC	-27543.72
AICc	-27543.72
BIC	-27530.66

Appendix II

Training set error measures:

Apart from Ljung-Box tests and other criteria that measure autocorrelations and normality in errors, there are a few accuracy and precision measures for the forecasts, which are listed as follows. The values for these measures are tabulated and are also tabulated below.

MAPE (Mean Absolute Percentage Error) – it is the average of the percentage errors. The errors are divided by the respective demand and summed over and divided by the total number of observations.

MAE (Mean Absolute Error) – it is the average of the absolute errors. In order to understand the implication of the value, it should be divided by average demand. The only demerit is, hence, the scale of the measure, which can be resolved by dividing the average demand to attain MAE in percentage terms.

MPE (Mean Percentage Error) – The Mean Percentage Error is the computed average of the percentage errors by which a forecast of a model differs from actual values of the quantity being forecasted.

RMSE (Root Mean Squared Error) – the Root Mean Squared Error is a measure of the concentration of the actual datum points around the best fit. It is the standard deviation of the errors. A lower value depicts an overall good fit. The square root of the average of squared errors is computed for this measure.

ME (Margin of Errors) – this measure depicts the difference between the actual data and the forecasts values in percentage terms. A larger value indicates lesser accurate forecasts.

The values for these accuracy measures are given below for the SMA, EMA and VWAP projections for the IXIC and SPY data series.

1. SMA projections for SPY ETF time series

	ME	RMSE	MAE	MPE	MAPE
Training set	0.1803078	3.944103	2.566009	0.0460829	1.704767

2. EMA projections for SPY ETF time series

	ME	RMSE	MAE	MPE	MAPE
Training set	0.2591323	4.684283	2.55217	0.09049056	1.698478

3. VWAP projections for SPY ETF time series

	ME	RMSE	MAE	MPE	MAPE
Training set	36.31693	63.11259	43.72535	14.71358	22.44715

4. SMA Ppojections for IXIC index series

	ME	RMSE	MAE	MPE	MAPE
Training set	6.328891	115.653	73.8213	0.03834196	2.272291

5. EMA projections for IXIC index series

	ME	RMSE	MAE	MPE	MAPE
Training set	8.690561	139.5241	73.72134	0.07838845	2.264921

6. VWAP projections for IXIC index series

	ME	RMSE	MAE	MPE	MAPE
Training set	1039.626	2000.735	1358.959	12.73405	31.62165