CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Returns Forecasting in Pakistan: A Study by Using ARIMA Modelling

by

Hanna Azhar

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Management & Social Sciences
Department of Management Sciences

2022

I wish to dedicate this effort to The Holy Prophet Hazrat Muhammad (P.B.U.H). Also, dedicated to My Parents for their prayers, understanding supports as an essential pillars who always guide me, unconditional love. My Teachers and Mentors who are always enlighten my paths with their support and guided my unfocused words into Clear ideas.

CERTIFICATE OF APPROVAL

Returns Forecasting in Pakistan: A Study by Using ARIMA Modelling

by

Hanna Azhar

(MMS181011)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Nadeem Iqbal	AU, Islamabad
(b)	Internal Examiner	Dr. Jaleel Ahmed Malik	CUST, Islamabad
(c)	Supervisor	Dr. Arshad Hassan	CUST, Islamabad

Dr. Arshad Hassan
Thesis Supervisor
April, 2022

Dr. Lakhi Muhammad
Head
Dept. of Management Sciences
April, 2022

Dr. Arshad Hassan
Dean
Faculty of Management & Social Sci.
April, 2022

Author's Declaration

I, **Hanna Azhar** hereby state that my MS thesis titled "**Returns Forecasting in Pakistan: A Study by Using ARIMA Modelling**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Hanna Azhar)

Registration No: MMS181011

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Returns Forecasting in Pakistan: A Study by Using ARIMA Modelling**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Hanna Azhar)

Registration No: MMS181011

Acknowledgement

Then which of the Blessings of your Lord will you deny. In the Name of ALLAH ALMIGHTY The creator of universe, the most benevolent, and the only to be praised, whose blessing and exaltations flourished my thoughts and enabled me to improve my knowledge up to this stage. A special praise and Thousands of salutations and benedictions on the last prophet HAZRAT MOHAMMAD (P.B.U.H.), who enables us to recognize our creator and chosen-through by whom grace the sacred Quran was descended from the Most High.

It is a pleasure to thank those who made this project possible. Firstly, I would like to take this opportunity to pay my gratitude to a great teacher, mentor and my honorable supervisor, Prof. Dr. Arshad Hassan for his supervision, sage advice, guidance and support during my course and research work. I am lucky to have such supervisor who cares and responded my every irritating query, truly an inspiration for me and has no words to express my wishes. Also like to thanks to Farhan who helped me to coordinate my thesis work and especially in achieving the results and supports me.

I am extremely grateful and express my deepest and earnest thanks to my parents who are my strength, my support system, taking part in useful decisions, giving necessary advices, encouraging, support and motivation for completion of this degree and all. I am nothing without my Parents. I choose this moment to acknowledge your contributions appreciatively.

In the end I would like to take this opportunity to say thanks and express my cordial appreciation to my friends, fellows and best wishers for their prayers and support especially for the completion of this report. This was really a tough journey with endless hopes would never be possible without support of family, teachers, mentors and friends.

(HannaAzhar)

Abstract

Financial markets are complex due to a variety of elements including financial, economic, and other factors such as market expectations, political, and government policies. As a result of these considerations, financial markets are complicated when it comes to time series forecasting. The goal of this research is to look at using the ARIMA model to forecast stock returns in Pakistan. It is crucial in forecasting and assisting investors in making investment decisions. This research makes use of four Pakistani stock market indices (KSE-100, KSE-30, KMI-30 and KSE all share). Different test such as descriptive statistics, variance ratio, Augmented Dickey-Fuller, Phillips-Perron, Kwiatkowski Phillips Schmidt and Shin, and Autoregressive Integrated Moving Average (ARIMA) is used in this work. The ARIMA model was used to analyze daily stock returns for all indices in a time series, and the findings show that all indices' mean returns are positive, but close to zero. In the long run, this suggests a retrograde trend. KSE-100, KSE-30, KMI-30, and KSE all have anticipated values that are nearly identical to their actual values, with just minor differences. As a result, utilizing the historical values of all the indexes, the ARIMA model is capable of determining stock returns. In order to determine the seasonality, SARIMA is used. It's an ARIMA add-on that lets model the seasonal component of series directly. This study will aid investors and investment institutions in improving their stock market index forecasting outcomes, as well as policymakers in selecting the best forecasting model for other aspects of the economy.

Keywords: Financial markets, Time series forecasting, Pakistan stock market, Autoregressive Integrated Moving Average (ARIMA), Seasonal-Autoregressive Integrated Moving Average (SARIMA), Efficient market hypothesis, Stock market returns.

Contents

List of Figures

List of Tables

Abbreviations

ACF	Autocorrelation Function
ADF	Augmented Dickey Filler
AIC	Akaike Information Criteria
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
ASE	Amman Stock Exchange
BIC	Bayesian Information Criteria
BRIC	Brazil, Russia, India, China
BSE	Bombay Stock Exchange
CPI	Consumer Price Index
EMH	Efficient Market Hypothesis
ERW	Exponential Random walk
ES	Exponential Smoothing
FARIMA	Non-Integrated Values
FTSC	Financial Time Stock Exchange (100 index)
GDP	Gross Domestic Product
GNP	Gross National Product
KMI-30	KSE-MEEZAN index
KPSS	Kwiatkowski–Phillips–Schmidt–Shin
KSE all	Karachi stock exchange all share
KSE-100	Karachi stock exchange 100 index
KSE-30	Karachi stock exchange 30 index
LSTM	Long Short-Term Memory

MA	Moving averages
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MASE	Mean Absolute Squared Error
MSCI	American finance company Inc.
NIFTY 50	benchmark Indian stock market index
NSE	National Stock Exchange of India
NYSE	New York Stock Exchange
OGDCL	Oil and Gas Development Company Limited,
PACF	Partial Autocorrelation Function
PP	Phillips-Perron
PSO	Particle Swarm Optimization
RMSE	Root Mean Squared Error
ROI	Return on Investment
RW	Random walk
RWH	Exponential Random walk
RWI	Random walk innovation
S&P	Standard and Poor
SARIMA	Seasonal ARIMA
SIC	Schwarz information criterion
SVM	Support Vector Machine
TSF	Time Series Forecasting
URT	Unit Root Test
VARIMA	Vector ARIMA
VR	Variance Ratio Test
WMA	Weighted moving averages

Chapter 1

Introduction

One of the primary difficulties of investor or any other party is stock market return predictions. It is challenging and troublesome due to complex nature of the stock market and its manifestation in modern financial literature. It's an interesting area that will continue to engage and motivate the researchers for improving existing methods and develop new predictive models. The evolution of the stock market, as well as its role in economic growth, is a key field of financial research. The goal of academia and technical research is to discover the most effective forecasting tools and methodologies for making quick and accurate stock market predictions (Javier and Rosario 2003). The stock market, according to theoretical studies, aims to foster long-term economic growth through consumption and investment channels.

Forecasting is a continuous and dynamic process (Golden et al., 1994). It makes educated guesses about the direction of future trends based on inputs. It uses historical and current data to estimate some key variables. According to Cao et al. (2019) "Reliability and Accuracy of forecasting model is difficult and under a big question mark". Investment risk and uncertainty can be reduced by using a forecasting model or technique to predict market direction. It can also help policymakers and regulators make better decisions and take corrective action by increasing investment flows into stock markets. There are three types of forecasting techniques; qualitative technique, projection models and time series analysis. From all these techniques, time series analysis is the common better.

Although there are several strategies for forecasting time series, no one strategy is employed to anticipate the returns of all stock markets. Researchers differ about how to anticipate stock market returns, according to the literature. In time series method, historical data is used for estimating future outcomes. The assumption is that demand in the past is a good predictor of demand in the future. Moving Averages (MA), Weighted Moving Averages (WMA), Exponential Smoothing (ES), Auto Regressive Moving Averages (ARMA), and Auto Regressive Integrated Moving Averages (ARIMA) are some of the methods used for forecasting.

The future movement of the stock price is determined from stock market forecasting. There are many techniques to forecast stock market price movement. Therefore, this is divided into two main categories; ; technical analysis and fundamental analysis. "Technical analysis is a method of analyzing historical market data, historical securities data, and forecasting future prices based on the assumptions that stock prices are determined by market forces and that history tends to repeat itself "(Levy et al. 1967). Fundamental analysis is used to measure stock intrinsic value. It is founded on financial analyses of businesses or industries.

These approaches have been used for decades to make investment decisions, but they were challenged in the 1960s by random walk theory, also known as the EMH Fama (1970), which proposes that changes in future stock prices cannot be predicted based on past price changes. Some empirical studies have found "random walk" in stock price (Tong, 2012; Konak and Şeker, 2014). Some empirical studies, however, have discovered that stock prices are predictable (Darrat and Zhong, 2000; Al-Tabbakh et al., 2018; Lo and MacKinlay, 1989; Owido et al., 2013; Radikoko, 2014).

Stock markets are known as moving averages, because they go up and down on daily basis. These markets are so difficult to predict due to number of independent factors associated with the market environment and uncertainty of future movement, this nature can help to reduce investor risks. As a result, forecasting techniques can aid in making better investment decisions. A stock market that is aligned with a growing economy inspires investor confidence in more buying activity, which can lead to higher prices. When the value of a stock rises, people invest more in equity markets in order to increase their wealth.

Stock market forecasters are constantly attempting to develop successful techniques for forecasting the index value of prices. The main point is to make a lot of money by using the best trading strategies and a successful stock market prediction to get the best results and to minimize inaccurate forecast stock prices. The goal of investors is to devise any forecasting strategy that will assist them in easily benefiting from and limiting investment risk in the stock market. According to Wang et al. (2018) state "prediction should be possible from two perspectives: statistical and artificial intelligence techniques".

By improving models, investors and institutions gain the ability to make better investment decisions, achieve greater success, and plan and develop effective strategies for their current and future endeavors. Investors want to get their hands on any forecasting method that can guarantee easy profit while lowering investment risk in the stock market. The function of stock market is channeling the flow of funds from savers to investors. The prime question is that how to foresee the future market price of a stock. Mostly, stock price forecasting based on factors that affect value and profitability. The return of stock market depends on the predictability of stock movements. Financial returns play a critical role in investment decision-making, channeling effective investment to support economic growth and development.

Risk management and portfolio diversification are best accomplished by forecasting stock market returns. According to one famous quote; "don't put all eggs in one basket". According to theoretical and empirical research, there is a positive relationship between financial markets and economic growth (Levine, 1997; Rajan and Zingales, 1998; Rousseau and Wachtel, 2000; Kim et al., 2011; Challa et al., 2020; Siva Kiran Guptha and Prabhakar Rao, 2018; Mallikarjuna and Rao, 2019). Stock market returns forecasting foresees the future movement of the stock value. The varieties of elements have impact on stock market like macro-economic factors, international events and human behaviors. The changes in premium are explained by macroeconomic factors such as interest rate, inflation, unemployment, market regulation, and supervision. Other factors include market size, liquidity, volatility, concentration, integration with the global market, and the regulatory environment are also consider. That's why, stock market is troubleshooter due to volatility, because no one can exact predict the market.

Stock markets are known for their high volatility, dynamism, and complexity (Johnson and Soenen, 2003; Cristelli, 2014). Stock market is volatile towards new direction each time because of shrinking as well as growth, emotional and over-reacts, and because they do not follow a pattern, the stock market is volatile and exhibits high variability. "Volatility is an easy and sensitive concept that refers to unexpected return due to return as a result of unexpected events, resulting in massive price movement with non-constant variance." It means that market ups and downs have an impact on share price and demand. Volatility is one type of risk. High volatility means that the security price can change in any direction within short time period. It is the main feature of time series. This happens for a variety of reasons, including: Factors influencing the market, the economic environment, political scenarios, crisis, and war. That's why; Forecasting is difficult as a result of these factors.

The high probability of a declining market equates to high volatility, while the high probability of a rising market equates to low volatility. The good news is as volatility increase, the potential to make more money quickly increases but on the other hand the bad news is that as volatility increase, the risk also increases. The reasons for high stock market volatility are still being debated in the financial literature. As a result, the financial markets take on an unexpected behavior that may cause investors to become confused. Stock market volatility, on the other hand, may be a stumbling block in this process, particularly in an emerging economy where high price volatility leads to capital evaporation from the market.

The general public perception of the stock market is that it is either extremely risky to invest in or unsuitable for trading. Despite the fact that the stock market involves risk, many people are interested in investing. For any investor, the most important factor is to maximize the yields on their investment and achieve this element, and investors are constantly attempting to predict or forecast stock prices. To forecast stock prices, many savvy investors employ techniques such as fundamental and technical analysis. Many investors, on the other hand, rely on the opinions and recommendations of various stock market analysts and fundamental analysts. These fundamental or technical analyses, as well as prediction algorithms and functions, are used by some financial analysts and investors to forecast future share prices and performance.

There are many holistic studies that provide the clue to investors in order to increase investment returns. The accuracy of return estimates over a variety of time periods is the key focus. It also delves further into their developing market portfolio architecture, risk and return, performance, and trading efficiency. In the economy, the stock market is extremely momentous. The research on the stock market returns in the emerging economies is necessary because large investor attracts for high volatility and high returns. The purpose is to help investor who wants to invest. So, forecasting is not an easy function and the selection of best method is a difficult task. The efficient market hypothesis (EMH) tells about the prices reflection on the overall information. Also claims that market follows the random walk. There are various methods to capture volatility in financial environment.

There are various methods used for forecasting market returns. All are used for developed and emerging country market. These methods are ARMA, ARCH, GARCH models but almost to none (not much) work done with ARIMA model, especially Pakistan as an emerging stock market because it is consider as weak form of efficiency because it's still unexplored. There is a need to explore emerging country like Pakistan stock market by applying ARIMA for forecasting to develop the investor interest and create eager for the investment.

1.1 Theoretical Background

The researchers focused a lot on the efficiency of the stock market in the recent era. The term efficient market was introduced by Fama that states "on average the competition will cause the full effects of new information on intrinsic values to be reflected suddenly in actual prices". So basically it creates a linkage between information and stock prices. According to this, the EMH relates with the above scenario. According to this theory, a market is efficient when prices fully reflect both public and private information. It simply means that it exist timely and rapid incorporation of information to the stock prices. So, the behavior assessment of the stock market is significant.

The three types of market efficiency that are commonly considered in EMH are weak-form, semi-strong-form, and strong-form. The forecasted values cannot be influenced by historical prices, as shown by the weak form. In this form information such as earnings, merger and acquisition details are available to everyone. The semi strong form subjects to open access of the data. In this, the type of information is statement and dividend announcement. The strong form indicates that stock price movements affect all open and inside data. In this information is past, public and private information.

All past data is included in asset values in markets with low form efficiency (current prices of stocks fully incorporate all available information of previous share prices). It means that trading techniques based on previous pricing trends might not always give investors with higher profits. Prices are unpredictably volatile because they have limited memory and alter only when new information is received. As a result, asset prices appear to follow the random walk, implying that there is no average or correlation between successive changes, and price variations are unexpected and random.

Asset prices include all publicly accessible information in a market with semi-strong form efficiency (present prices of shares reflect all available publically information along with past prices of shares). The price level should represent all relevant historical, current, and forecast-able information gathered from public sources in this manner. It also affects the completeness and promptness with which fresh information is received. This diagram emphasizes the stock and flow part of data processing capabilities. The important element to remember is that it just requires data from public sources such as newspapers, press releases, and statistics.

In a highly efficient market, asset prices reflect all information, whether public or private. It's an extreme type of market because it implies that important company-specific information, such as pending takeover bids and dividend increases, is fully incorporated in asset prices with the very first trade after the information is generated, particularly after the board of director's votes in favor of dividend increases before it's announced publicly. For investment purposes, the EMH is regarded as a necessary tool. In most of the literature, this theory is used on weak-forms or semi strong-forms. Strong-forms efficiency is measured indirectly by looking at fund manager performance and determining whether they are able to earn profits net of risk premia after deducting the cost of obtaining private information.

1.1.1 Background of the Study

For the validity of the theory in developed and emerging economies, researches are conducted on stock markets. Because of their high volatility and high returns, emerging economies attract large investors (Akhtar and Khan, 2016). The research goal is more than just to add to the body of knowledge about emerging markets. It is also beneficial to investors who plan their investment strategies and decisions based on available information and market efficiency.

It is worthwhile to investigate the efficiency of the Pakistani stock market in order to develop a useful tool for forecasting stock prices and making investment decisions. As EMH theory claims that market has randomness, follows the random walk. According to the literature, the Pakistani market has a low level of efficiency. Various econometric tools are used to test the weak form of efficiency.

1.1.2 Motivation of the Study

In today's economy, the stock market is extremely important and has a significant impact on the economy. According to EMH, prices reflect the overall information. There are multiple methods like ARCH, GARCH and EGARCH to capture volatility as financial environment is highly uncertain and volatile. According to Zhou (1996), these are better correlated methods but complex properties to compute, as non-parametric models are more complex to compute but better predictive power. ARIMA is also used for forecasting volatility. Box explains that while all models are incorrect, some are useful. There are various models of ARIMA such as SVM, PSO, LSTM, VARIMA, SARIMA, and FARIMA.

In past years, many analysis use ARMA, ARCH, GARCH models but almost to none work done with ARIMA model especially for country like Pakistan in emerging stock market return forecasting. ARIMA is the best statistical forecasting method for obtaining timely and accurate stock predictions. Furthermore, it demonstrated that the evidence of series is integrated steps for stationary and differencing steps for non-stationary (Merh et al., 2010).

1.2 Gap Analysis

Forecasting is not an easy process. It is difficult to pick the best method for forecasting, predicting, and comparing stock prices in developed and emerging markets. Furthermore, the assurance of good long-term forecasting returns through the use of innovative technologies is always a point of contention. It is also crucial to provide information about portfolio construction, risk and returns, trading performance, and efficiency. As discussed, there are various methods used for forecasting market returns. All are used for developed country market and emerging country market as well.

In recent years, there are several studies that use ARMA, ARCH, GARCH but few or almost to none of that uses ARIMA model especially for emerging country markets. The research present on forecasting return of emerging market through ARIMA model are India, Mexico, Russia, Saudi Arabia, China and Brazil but this work is related to Pakistan only.

Furthermore, there is a need to evaluate the dataset's seasonality. Seasonality is a time series feature in which the data undergoes regular and predictable changes that repeat each calendar year. SARIMA is created by combining the ARIMA models given with the seasonal term. It is an ARIMA extension that allows univariate time series data with a seasonal component directly. Based on the notion of seasonal patterns, it is one step distinct than ARIMA. In other words, Seasonal fluctuations or patterns are defined as any predictable fluctuation or pattern that recurs or repeats over a one-year period.

1.3 Problem Statement

As previously mentioned, several studies using ARIMA and other models to forecast stock market returns have been published in the literature, particularly in developed markets. Few, on the other hand, have concentrated on developing and emerging markets. It becomes a part of emerging market with the passage of time that there is an interest of people then the insight of this phenomenon can increased. It's also providing a gateway to future researchers. Pakistan is still unexplored. Among existing models in various markets, the ARIMA model has proven to be more efficient and accurate. In addition, the ARIMA model is better suited to estimating short-term returns rather than long-term returns. The ARIMA model is frequently used to forecast long-term returns in research. SARIMA is also used to examine data for seasonality.

1.4 Research Questions

This study provides answers to the following questions:

1. Does random walk or martingale exists in returns of Pakistan's market?
2. Can ARIMA model forecast returns?
3. Does ARIMA model with seasoning effect is better in forecasting return?

1.5 Objectives of the Study

The following are the study's objectives:

1. To provide insight about the random walk process in stock returns.
2. To propose the ARIMA model for forecasting the return.
3. To capture the seasonality in returns through ARIMA model in forecasting returns.

1.6 Significance of Study

Forecasting's major purpose is to help investors make better investment selections, enhance investor accuracy, and boost efficiency. Investors, governments, and financial institutions must make dynamic decisions in order to get the most out of their investments. The investment decision has a big impact on achieving the expected returns from stock market forecasting. The stock market may fluctuate or disturb its consistency amid general unsettled situations. Using appropriate stock market strategies and accurate forecasting tools, uncertain conditions could be overcome (Zhang et al., 2019a,b). The most difficult aspect is accurately and quickly forecasting the stock market.

The movement of stock prices can be accurately predicted using a prediction model. The risk and uncertainty connected with the investing process might be decreased with this method. Investors would benefit from making educated investment decisions, and governments would benefit from taking necessary efforts to improve stock market investment. In time series analysis, ARIMA is one of the greatest statistical forecasting approaches for investors to generate rapid and accurate stock recommendations.

As already know that Pakistan is an emerging market. It is again categorized as an emerging market from last few years. It is facing severe challenges, instability and huge gap especially in recent times. We can see a high volatility and interest rate in Pakistan market. The government must maintain vigilance and will be able to formulate appropriate financial policies as a result.

The investor's main goal is diversification. From an investment standpoint, investor interest in the Pakistan stock market is increasing. It boosts a huge amount of investment from many international investors because they foresee many activities in this market. So this situation demands that the information about the market dynamics must be provided to the potential investors for efficient decision making. The regulator, institution and individual investor understands the advanced forecasting tools and techniques. For regulators or government, this will help to forecast the economic indicators and set the direction in an efficient track. For the investment institution and individuals investors efficient rate is the important.

With the various techniques to evaluate, analyze and forecasting tools in hand, the objective of the study is that this model of forecasting will be beneficial. This study not only contributes the ARIMA model in Pakistan context. The returns of stock markets become a common channel for investors and academicians. This gives a tool which helps them to forecast their investment returns more efficiently and help them to reallocate their invested funds for the better outcomes.

1.7 Plan of Study

This research is divided into five major chapters. The first three chapters concentrate on the theoretical aspects of the relevant topic, while the last two chapters cover the empirical aspects and interpretations of the study.

Chapter 1 contains an introduction, theoretical background, gap analysis, problem statement, questions, objectives, and significance of the research. Chapter 2 includes a review of previous studies' literature about theory, forecasting and method for forecasting. Chapter 3 discusses the current research study's research methodology. Chapter 4 delves into data analysis and results. Chapter 5 discusses the current research study's conclusions, recommendations, and limitations.

Chapter 2

Literature Review

This chapter discusses the evidence for the theory of efficient market hypothesis ”EMH” in time series forecasting of stock market returns using the ARIMA model.

2.1 Literature

This part of the study provides the detail research evidences about theory called efficient market hypothesis under umbrella of random walk in equity returns, comparison of index and exchange rate along time series forecasting in financial markets, study about ARIMA model and forecasting along usage of this model through numerous researches worldwide.

2.1.1 Random walk in Equity Returns

In 1970, Fama (1970) formally presented “Efficient market hypothesis”. According to him, in the efficient market prices fully reflects the information. It is a widely accepted financial theory that states that a market is efficient when the prices fully reflect the information. It claims attractive simplicity that financial assets are fully reflects to all available relevant information, including private information but investment was done requires prices to reflect publicly available information.

Jensen defines market efficiency as follows in 1978: ”**a market is efficient with respect to information set if it is impossible to make economic profits by**

trading on the basis of information set". Malkiel provided a closely related definition of market efficiency in 1992 "**a capital market is said to be efficient if it fully and correctly reflects all relevant information in determining security prices. The market is said to be efficient with respect to some information set, if security prices would be unaffected by revealing that information to all participants. Moreover, efficiency with respect to an information set implies that it is impossible to make economic profits by trading on the basis**".

The relevance of the information set utilized in testing, the capacity to use this information in a trading strategy, and the yardstick for establishing whether the EMH holds in terms of economic gains, such as risk adjusted and net transaction costs, are all highlighted in these definitions.

The EMH is a stumbling block. Logically, this may appear to be the conclusion of the story. However, the strength of the argument does not appear to be convincing. Rather than using series for forecasting, researchers constantly experiment with new prediction methods on the returns from speculative assets such as stock market prices. A significant amount of research appears on forecasting stock returns, usually with little success. It is unique in that investors' current and future forecasted payoffs affect their current and future trades, which in turn affect returns. It gives rise to appropriate forecasting models, posing a unique challenge in establishing successful forecasting and evaluation.

It also states that any change in the prices of securities is fully reflected in the prices of other securities at any given point in time. But the point is still there that all information concerned with factors like firm specific, industry and market is fully uncovered by stock price of a firm. It also refers that the prices of assets reflects the floating available information in market.

The EMH is defined by Gibson in 1888 for the first time. According to him, once the shares are publicly known in the open market, the value that develops there can be considered the best intelligence judgment about them. Some other researcher such as Banerjee (2014); Mandelbrot (1989); Ball and Brown (1968); Fama and French (1988); Rousseau and Wachtel (2000); Campbell and Shiller

(1988); Jegadeesh and Titman (1993); Welch and Goyal (2008); Chordia et al. (2014); McLean and Pontiff (2016) also contribute in this theory.

According to Reilly and Brown (2011), "every investor gets the desired return from their investment". According to Malkiel (2005), "the allocation of resources is based on the decision on the fair price discovery that can only be done when markets are efficient and reflect the relevant information". According to Dyckman and Morse (1986), "a market is efficient if the price of the traded security fully shows all the available information or these prices react immediately and form unbiased to new information". There may be a chance of mislead price to investors that affect the decision making process in security selection.

The criticism also exists about this theory. According to Aumeboonsuke and Dryver (2014), the element of inefficient market prevails, in the market and rejects the EMH. According to Malkiel (2003), the stock prices can't be predicted. But Fama study states in response that the prices adjust with new information and spread speedily without any delay. In response Malkiel argues that if the information speedily reflected on the stock prices then there is no link between current and future price because they are totally independent. But the analysis shows that the past prices change the expected future prices, which help the investor.

The three types of market efficiency namely allocation, operation and information. The allocation efficiency is a measure for allocation of resources to achieve the maximum benefits. The operation efficiency shows how well the investors achieve the cost minimization goals, and the information efficiency tells how rapidly the market adjusts with new information and prices. In addition, Fama divided empirical market efficiency tests into three categories: return predictability tests, events studies, and private information tests.

It is further divided into three categories in the test for return predictability; the first category is to test simple trading rules based on very recent stock returns, the day of the week, and the month of the year. The second category is to look for positive correlation in short-term returns over a short time horizon, such as on a weekly or monthly basis. The third category is long-horizon return predictability, which is based on a negative correlation observed over a longer period of time.

It was widely believed in 1950s that the perceptive investor could easily identify patterns in stock market movements that could be exploited through the adoption of a trading strategy. Therefore, it is quite surprising when Kendall and Hill (1953), then Roberts (1959); Fama et al. (1969) all documented that subsequent stock price changes were essentially uncorrelated with each other. It means that price followed a random walk. Because price trends are so volatile, a strategy of buying recent winners and shorting recent losers would be unprofitable.

Fama and Blume (1966) test a number of more sophisticated 'Filter rules,' such as "buy a stock after it has increased by x percent and do not sell it until it has decreased by y percent," and find that none of these strategies generate significant profits, even after trading costs are taken into account. In mid 1960s it is accepted that trading rules based on such simple recent information are not worth pursuing.

Most of the published researches on the short horizon predictable returns tend to fall in two categories. Stock prices first take a random walk, and then the underlying causes of the documented derivations are investigated. A second group of studies mount a much more open and aggressive attack on the EMH by explicitly testing the ability of investors to profit from the perceived tendency of stock prices to 'overreact' to news or events and then to revert back to equilibrium price levels over several time periods. By utilizing the fact that the variance of uncorrelated stock returns should increase in direct proportion to the sampling intervals, Lo and MacKinlay (1989) develop a unique and innovative method for testing for serial correlation in stock returns. They find strong evidence for a positive correlation between weekly stock returns, rejecting the random walk hypothesis.

Finally, they come to the conclusion that the autocorrelation patterns do not match the stationary mean reverting models of Summers (1986); Poterba and Summers (1988); Fama and French (1988); Lo and MacKinlay (1989), shows that less than 50 percent of profit from a contrarian investment strategy is due to overreaction. Positive autocorrelation patterns in weekly and monthly stock market returns are investigated by Conrad and Kaul (1988, 1989) and then Conrad, Kaul, and Nimalendran (1991).

There are two features in long horizon returns predictability. Firstly, there is significant negative autocorrelation in return intervals of between two to five years.

Secondly, with the increase in returns, the predictable component of the total variation of returns increases.

De Bondt and Thaler (1985); Poterba and Summers (1988); Fama and French (1988) have all documented these patterns (1988 a). (Reichenstein and Rich, 1994) discuss how investors can use this predictability to their advantage in their own investment portfolios. Fama and French results are typical of this literature that tells negative autocorrelation in returns beginning. On the other hand, McQueen (1992) doubts on the significance on the existence of long horizon return predictability. Roberts (1967); Fama and Blume (1966) introduced the three types of market efficiency, which are widely debated in EMH literature (1970). Weak, semi-strong, and strong are the three varieties. Forecasted values cannot be influenced by historical prices, as shown by the weak form. In this form information such as earnings, merger and acquisition details are available to everyone. According to Sensoy et al. (2017), information impounded and the traders with the skills enough to exploit the market to get the maximum returns. The semi strong form subjects to open access of data. In this the type of information is statement and dividend announcement. The strong form indicates that stock price movements affect all open and inside data. In this information is past, public and private information.

Fama (1970, 1991) concentrated on determining information efficacy. Fama (1992) finds more evidence of return predictability based on lagged values and publicly available data, but Fama (1970) concludes that the empirical evidence favors weak and semi-strong form efficiency. Bachelier (1990) claimed a century ago that the random walk theory basically asserts that series that would want to anticipate the returns from speculative assets are unforecastable in their most basic form. In the 1960s, Cootner (1964) validated this idea scientifically, and he has done so many times since. The EMH was suggested immediately after empirical data became available, based on the underlying reasoning that if a return can be foreseen, numerous investors will exploit it to create endless profits. Market players' activities generate EMH-compliant returns; otherwise, there would be an unlimited wealth "money machine," which is difficult to create in a stable economy.

Many studies mistakenly associate market efficiency with the random walk model for stock prices, but this is clearly not the case. It can be demonstrated that, under the so-called ”risk neutral” or equivalent martingale probability measures, stock prices plus cumulative dividends discounted at the risk free rate should follow a martingale process. Forecasting tests of the random walk model's validity on security prices, on the other hand, are insufficient to demonstrate market inefficiency because the EMH only implies market inefficiency in a limited set of circumstances.

Despite its simplicity, it is incredibly difficult to test and is frequently used in empirical studies. Forecasting experiments must include at least one of the following factors: the set of forecasting models available at any given time, as well as estimation methods. The technology used to find the best forecasting model or a combination of forecasting models.

The range of real-time data that is available, including public vs. private data and, ideally, the cost of obtaining it. A risk premium economic model that considers trade-offs and payoffs, transaction costs, existing trading technology, and any asset ownership restrictions. Even if the best forecasting model or set of models is disputed, the presence of a single successful prediction model is sufficient to demonstrate EMH violation. When model uncertainty is factored in, however, this is no longer the case until proof of a search technique that allows investors to find a successful model is presented.

Recognize that this is the type of forecasting method that may work even if the EMH is right. It appears self-evident that the usual constant-parameters models with a simple specification, such as the ARMA models presented in Box and Jenkins, are not suited to the task since they presuppose stationarity. Consideration should be given to rapidly changing models that can recognize and exploit any instance of transitory forecast-ability that may come and disappear as learning opportunities arrive and close.

Nonetheless, there is a significant difference. The loss function, essentially economic profits net of transaction costs, is normally straightforward to identify in EMH tests, but measuring economic profits is complicated by its reliance on the risk premium. Because evidence of predictability is only found to be poor, the joint hypothesis issue in assessing market efficiency in conjunction with a sustained specification for the economic risk premium is not regarded very essential by Fama 1970. As a result, Fama's previous survey assumed that the EMH

suggested that projected asset returns should be a "fair game," implying that risk premium fluctuations over time are not as essential.

Fama's perspective shifted dramatically in 1991, when he presented more substantial evidence of predictability and argued that an economic model for expected returns was necessary. He is particularly interested in the informational efficiency of security pricing. However, several recent research have discovered data that appears to support the theory of predictable patterns self-destruction in return, in that previously reported predictability vanished at a time when a consensus is forming that predictable patterns exist. So far, we have focused on market efficiencies that are concerned with the EMH's consequences for nonstationarities in returns.

Brock et al. (1992), for example, show that following technical trading principles may be profitable based on data from 1986. Despite the fact that their work was published in 1992, it was widely available prior to that. The same variables that alerted these writers to technical trading rules' apparent effectiveness may have attracted traders to abuse these tactics, resulting in their downfall.

According to Sullivan et al. (1999), "Technical trading rules' apparent historical capacity to create excess profits has broken down after 1986". According to Dimson and Marsh (1999), "The small cap premium vanished in the UK stock market when it became well recognized". Bossaerts and Hillion (1999) look at the predictability of monthly stock returns in a number of worldwide stock markets and find that the apparent in-sample predictability breaks down out of sample around 1990.

According to Aiolfi and Favero (2005), "Predictability in US equities, as established in prior research, appears to have vanished in the 1990s". Their findings support the idea that the best model chosen in 'real time' based on in-sample performance loses predictive value outside of the sample.

Individual forecasting models or stable combinations of them are unlikely to be beneficial in time series forecasting approaches. It is conceivable to revert to very basic approaches that can adapt or learn fast and that can be used in large numbers to a variety of return series in order to find potential 'hot spots' where forecasting capability is present.

According to Granger and Pesaran (2000), "thin modelling" is used when judgments are made based on just one "best" model, and "thick modelling" is used when decisions are made based on a mixture of outputs from models with statistically comparable outputs. The optimal technology for doing so is still being researched, and crucial quantities for the final output must be derived through bootstrap simulations. In macroeconomics, this is okay, but in high-speed finance, it is possibly not yet.

As a result, the EMH does not imply that all density variations are random. It does, however, need the absence of predictability in certain probability distribution functions. There is currently a lot of evidence that asset return volatility fluctuates over time in a predictable fashion. As a result, enhanced volatility forecasting models in the context of option pricing have sparked great attention Engle et al. (1992). Is this market efficient in its volatility? The answer is a resounding no unless a trading method can be devised that uses this data to detect under and overpriced options in the options markets. Option pricing should reflect the best volatility estimates at all times if options markets are efficient.

2.1.2 Forecasting in Financial Markets

The creation of investment policy relies heavily on stock return forecasting methods. The stock market may be predicted using a variety of methods. The basic goal is to be able to better precisely forecast stock market price movements. Information asymmetry, insider trading, and other abnormalities, on the other hand, may cause market inconsistency or affect the market's direction.

Forecasting is a fascinating field of study for academian and technical agents, and it will continue to intrigue them as they strive to improve current prediction models. Predictors are always attempting to create viable approaches for anticipating index value prices. The predictability of stock returns has been used in several research (Rapach et al., 2016; Zhu and Zhu, 2013; Pettenuzzo et al., 2014). Fama (1970, 1991); Lo and MacKinlay (1989) are only a few of the scholars that have looked into the effectiveness of stock market forecasting (1988). The reason for this is that everyone may invest based on their judgments and ability to plan and implement effective strategies for their current and future financial goals.

The main purpose of stock forecasting is to develop a trading strategy that identifies the actions to take in order to purchase or sell stocks at the best possible time. Examine future investment prospects, as well as the value of establishing and upgrading time series forecasting models, as well as their efficiency and success. The main goal is to make a lot of money by using the finest trading methods and forecasts to get the greatest results and to reduce the number of wrong stock price forecasts.

Investors' goal is to develop any forecasting approach that would enable them to profit from the stock market with ease while minimizing their investment risk. Investors try to predict stock values all of the time. Overconfidence and illusion of control, the narrative fallacy, anchoring bias, loss aversion, and herding mentality are all examples of investor biases that can lead to incorrect stock market price predictions. As a result, the underlying issues are estimating more accurate and quick stock price projections.

According to Neely et al. (2014); Wang et al. (2018); Challa et al. (2020), there are a variety of reasons for unexpected losses in invested capital, including investor errors in estimating their investments or portfolios. Evidence-based stock return forecasting is discussed by (Phan et al., 2015). Technical indicators are economically and statistically significant, according to Neely et al. (2014), who used them to estimate stock returns. Fundamental analysis and technical analysis, according to Chi (1999), are two analytical approaches for forecasting the stock market. Economic, industrial, and firm analyses are all topics covered by fundamental analysis. Technical analysis allows us to forecast future prices using market data from the past.Accurate stock price forecasting would not only minimize stock price uncertainty, but it would also provide investors a means to develop expectations and maybe avoid significant market swings. The projected stock market environment, as well as stock price forecasts, is critical for making decisions about stock investment timing and relative investment attractiveness among various market sectors (Fischer and Jordan, 2005; Pesaran and Timmermann, 1995).

According to Pai and Lin (2005), because of the complicated structure of the stock market, stock price forecasting is one of the most difficult tasks in financial forecasting. Suits (1962); Zotteri et al. (2005); Wen et al. (2019) explain how various econometric models have been used to predict stock market movements and consequently future stock prices and returns. Forecasting is a method of forecasting the direction of future trends by using previous data as inputs. It is based on historical and current data, and certain factors of importance may be estimated. Investment risk and uncertainty can be reduced by using a forecasting model or approach that predicts market direction.

Stock market efficiency grasps a lot of attention by the researchers. It creates a link between stock price and information set, and proposes their timely and rapidly incorporation. If stock prices completely represent relevant or accessible information at any one time, it is termed efficient. The assessment of stock market behavior is important. According to Boehmer and Kelley (2009), security price results to precede information because the information about prices helps in investing decisions. There are various forecasting techniques but from all these techniques, time series analysis is the best tool.

Time series analysis is thought to be a useful technique for forecasting trends. In business, understanding that ”time is money” is a crucial tool that applies everywhere money and time are linked. A time series is a collection of ordered observations made at evenly spaced time intervals on a quantitative feature of phenomena. One of the major goals of time series analysis is to avoid forecasting future values. However, one of the key drawbacks of the trends chart is that it may not always represent, anticipate, or maintain the market's continuous flow. According to Faisal (2012); Hawaldar and Mallikarjunappa (2009), ”Past patterns or flows, seasonal growth, or changes in stock prices all pique the curiosity of investors”. This broad view of the stock market must be comprehensive.

TSF is a dynamic research field which pulls the analysts group over last couple of decades. Tong (2012), “A suitable model fitting is required for an efficient TSF”. Analysts have put in a lot of time and effort over the years to develop capable models that can improve predicting accuracy. Over time, several major time series forecasting models have been developed in accordance.According to Raicharoen et al. (2003), "TSF is an example of foreseeing the future through comprehending the past". He stated that the essential element of time arrangement demonstrating is to meticulously collect and fully consider prior perceptions of a period arrangement in order to construct a good model that displays the arrangement's inalienable structure. This model is used to generate future values for the arrangement so that estimations may be made.

According to Zhang (2007), "TSF determines that the lines can be called as a proof of foreseeing the future via understanding the past". Because time series models are so important in so many disciplines, such as business, finance, economics, science and engineering, and so on, great attention should be made to match a suitable time series model to the needed setup.In the field of time series forecasting, there are several types of research projects. TSF is a prediction method that concentrates on the dependent variable's historical behavior. TSF models are a different way of analyzing and forecasting future developments based on the objective's historical behavior. Generally, the TSF model assumes that no causation influences the variable being forecasted. A basic deterministic model, such as linear extrapolation, or a complicated stochastic model, such as adaptive forecasting, may be used to generate a forecast.

According to Hipel and McLeod (1994), "stationarity is a type of statistical equilibrium". A stationary process' mean and variance characteristics are not time dependent. If there are more historical observations data, there is a higher chance that the time series will not be stationary, whereas we normally employ differencing or transformation in our model to keep it stationary for short time spans. According to Fang and Xu (2003), "because statistical-based approaches such as linear AR models are flexible enough to describe a wide range of stationary processes, time-series forecasting is commonly done using them". The ARMA model is one of these. First order non-stationarity is sometimes removed or reduced using methods like the linear ARIMA, which is based on the development of increments.

According to Bagnall (2004), "the research community is interested in stock price forecasting". A wide range of forecasting techniques are covered by time series analysis. Researchers have developed a number of variations on the fundamental ARIMA model and discovered that these strategies offer excellent results. Clustering time series from ARMA models with clipped data has been incorporated into the revisions.

2.1.3 ARIMA Model and Forecasting

The ARIMA model was first introduced in 1970 by George Box and Gwilym Jenkins. The Box-Jenkins approach is another name for it. This strategy, which was first described by Box and Jenkins (1976) in 1976, was further developed in the 1990s for the prediction of time series forecasting. Shumway and Stoffer (2011) suggested that in 1970's paper of Box and Jenkins "ARIMA is used with less volatility data that develops a systematic class of models to handle time correlated modeling and forecasting". It is a model that uses observed values to define a time series and may be used to anticipate future values. This approach is made up of a sequence of steps for detecting, estimating, and diagnosing ARIMA models using time series data. When these models are applied to any time series, they reveal non-seasonal patterns that are free of random white noise.

The goal is to develop an accurate stock forecasting model by comparing the accuracy of an auto ARIMA model with two customized ARIMA models. When attempting to make a secure investment, it is critical to have a thorough grasp of the current as well as projecting the future. It also aids in comprehending the importance of the ARIMA model for time series forecasting and the precision of its procedures.

Given the importance of developing and improving time series forecasting models, as well as studying their effectiveness and success, this study aims to get an accurate stocks forecasting model by comparing the results of accuracy of auto ARIMA and customize ARIMA models. Depending on the application, the current econometrics model has been updated (Zotteri et al., 2005). ARIMA, which is used to forecast financial time series data for the near term, is a more efficient and resilient econometrics model than order approaches like Artificial Neural Network (ANN) (Merh et al., 2010; Sterba and Hilovska, 2010). Short-term financial time series data may be forecasted well by using ARIMA models (Schmitz and Watts, 1970; Rajan and Zingales, 1998; Kyungjoo et al., 2007; Merh et al., 2010; Sterba and Hilovska, 2010).

In financial time series forecasting, ARIMA models are recognized to be more resilient and efficient than even the most popular ANN approaches, particularly for short-term prediction. When it comes to projecting short-term stock returns, the ARIMA technique is more accurate than when it comes to long-term market returns (Sabur and Haque, 1992). They are not the greatest models for understanding the past and predicting the future in most circumstances when it comes to financial time series because financial markets are primarily anticipated by other models. ARIMA may have some validity and yield decent findings, particularly when data is not volatile. Using ARIMA models to anticipate stock market returns has been used in a few research (Al-Shiab, 2006; Ojo and Olatayo, 2009; Adebayo et al., 2014).

Several research have found that ARIMA models for financial time series data give poorer forecasts (Zhang, 2003; Adebayo et al., 2014; Khandewal et al., 2015). ARIMA uses the dependent variables' historical values as well as the error term to estimate the most likely value of our variable of interest. In terms of predicting future stock values, it performs admirably. It is also used to figure out whether a model is better at short-term forecasting or in anticipating stock values over time.

When it came to producing short-term projections, ARIMA beat complicated structural models. The lengthy process of developing these models for predicting short-term stock prices is described. According to this model, the results gathered from real-world data showed that potential strength gives investors with short-term predictions that can help them make investment decisions. Model identification, parameter estimation, and diagnostic testing are the processes in developing an ARIMA predictive model. Despite its extensive use, numerous issues about ARIMA's accuracy remain unanswered. The questions are: (1) what is the optimal window length for generating regression parameters and forecast results; (2) should the window be fixed width sliding or increase; and (3) how successful are ARIMA predictions.

For decades, stock market return forecasting has been a hot topic. Forecasting financial markets is a topic that researchers are likely to be most interested in. Successful price projections provide a lot of advantages. ARIMA is one of the most common models. It is used to forecast future equity returns using data from the assets in question over time. This is one of the most used financial forecasting models. This has demonstrated a high level of efficiency in producing short-term projections. When it comes to short-term prediction, it consistently outperforms complicated structural models. A variable's future value is a linear mixture of its previous values and mistakes in this model.

There are just a few studies in the domain of stock price forecasting that use alternative models such as GARCH and ARIMA, and even fewer in emerging stock markets. Furthermore, the majority of the research focused solely on predicting stock price movement. It also failed to compare estimated and actual numbers to ensure that estimations were accurate (Zhang et al., 2019a,b).

The primary idea is that it forms a predicted item as a random series over a set amount of time in order to create a data series. To characterize the time series, a specific statistical approach based on autocorrelation analysis might be constructed. Following the establishment of the approach construct, future values of time series may be forecasted using the past and current values. Stock price prediction has been the subject of various theories and methodologies in recent years. From the standpoint of statistical models, ARIMA models are used. Prediction may be done using two approaches: statistical and artificial intelligence techniques, according to the literature. In the fields of economics and finance, it has been widely employed.

Regression, exponential smoothing and generalized autoregressive conditional heteroskedasticity are examples of other statistical models (GARCH). ARIMA models are widely utilized in the fields of economics and finance, as well as stock forecasting, from the standpoint of statistical models. However, due of its volatility and noise, stock market prediction in time series is regarded as one of the most difficult problems. Because stock price changes are non-linear and non-stationary, it is difficult to generate precise and dependable predictions.

To anticipate future returns, several researchers used ARIMA forecasting models (Khashei et al., 2009; Lee and Ko, 2011; Khashei et al., 2012). It is critical to choose a model that can indicate patterns and provide enough information for an investor to make an informed decision. According to Davi (2013) ARIMA is an algorithmic way to alter the series that is superior to predicting directly since it produces more accurate findings. Al Wadia and Ismail (2011) further indicate that the ARIMA model has a stable structure, was design exclusively for time series data. Because it is a univariate model, their forecasts are more accurate and dependable. As a result, leading indicators and explanatory factors cannot be used in this case.

Nochai and Nochai (2006) use ARIMA forecasting to time series data of oil palm prices and discover that the estimated ARIMA term is the most efficient for future returns. The three main factors in this model are stationarity, invariability, and parsimony, which are utilized for identification, estimation, and diagnostic testing, respectively (Asteriou and Hall, 2015). Ali et al (2011) employ ARIMA models to anticipate the stock market performance of Pakistan's oil and gas companies. Mondal et al. (2014) use the ARIMA model to anticipate the future returns of 56 Indian companies from various industries. According to their findings, 85 percent of their predictions were correct. Banerjee (2014) forecasts the Indian stock market index with the ARIMA model. The study finds that the model's short run prediction power is the greatest for achieving the best results.

Rapach et al. (2016) demonstrated the source of predictive power by decomposing a cash flow channel with vector auto regression. In addition, evidence of a link between short-sellers and traders has been discovered. Wang et al. (2018) finds a dynamic link between returns and volume on the basis of US stock returns. They discovered that following the volume curve does not yield significant benefit.Zhang et al. (2018) examine 18 macroeconomic and technical parameters to investigate oil price predictions. For a mean-variance investor, the findings indicated accurate forecasts and confidence equal return improvements. Zhang et al. (2019a,b) explain not only intraday stock movement trading behavior, but also U-shaped investment curve evidence. They discovered that morning returns may be used to forecast afternoon stock prices.

Using the most modern Auto-Regressive Integrated Moving Average Model (ARIMA), this predicts worsening circumstances in Iran, Europe as a whole, and particularly in Italy, Spain, and France. During mid-April 2020, the United States will surprise everyone and become the hotspot for new instances. For time series analysis, utilize four models to study the aggregate data set. This contains the ARIMA model, which combines the AR (Auto Regression) and MA (Moving Average) models into a single model. The ARIMA model was tested on publicly available Netflix stock data. Netflix daily stock price data is included in the dataset, which spans five years from April 7, 2015, to April 7, 2020. Autocorrelation Functions (ACFs), Partial Autocorrelation Functions (PACFs), and Mean Absolute Percentage Error (MAPE) are used to compare model accuracy and composition across studies. Durham (2002) believes that the stock market expansion in low-income nations such as Pakistan is overblown. It provides the stock market's influence on growth is exaggerated because high-income nations are included in cross-country regression analyses in various researches.

A stock market index is a figure that represents the stock market. It is calculated using the prices of a number of stocks (typically a weighted average). It is a tool that investors and financial managers use to define the market and compare the returns on various investments (Assaf, 2006). According to Husain and Mahmood (2001); Husain (2006), "Pakistan's stock market is still not matured enough to have a significant influence in shaping the economy's real sector". Furthermore, Pakistan's stock market cannot be considered a leading indication of economic activity. The economy and everyday lives are both affected by stock market fluctuations. The current economic catastrophe is more likely to be triggered by a drop in the price of a share. According to Zafar et al (2008), "The word volatility is defined as the variation in stock prices over time".

According to Kalu (2009), "stock price forecasting can help investors anticipate and prevent price risk". He forecasts the Nigerian Stock Exchange on a monthly basis using the ARIMA (p, d, q) model, with the fit sample spanning from January 1985 to December 2008, and the out-of-sample prediction period spanning from January 2009 to December 2009. The chosen model forecasts indices and growth rates that differ from the actual indices and rates. During the predicted period, the projections did not match market performance. As a consequence, the model's adequacy was tested by creating one-period predictions for the next 12 periods, and statistics revealed that the model forecast beat the naïve Model. As a result of the aberrations, the global economic catastrophe obliterated the correlation link between the NSE All-Share Index and its history.

Sultana et al. (2013) discuss that Forecasting time series is an essential issue in macroeconomics. We employ two strategies for analyzing time series data. Decomposition is one of the simplest and most fundamental methods for predicting time series. Breaking down a time series into its four components, trend, cycle, seasonality, and irregularity, is known as decomposing it. The ARIMA model is the foundation of the second technique. Based on decomposition of actual series of these variables and ARIMA model for monthly series from July 2008 to June 2013, they forecast macroeconomic variables CPI and LSM for the period July 2013 to September 2013 in this study. Compare the out-of-sample forecasts of two methodologies using the mean absolute deviation and the sum of squared errors to see which method gives the highest forecasting accuracy for policymakers to rely on when anticipating inflation (CPI) and economic growth (LSM).

According to Adebayo et al. (2014), "the forecasting time series stock market has piqued the interest of applied researchers due to the critical role it plays in the economy". They use common model selection criteria such as AIC, BIC, HQC, RMSE, and MAE to choose the ARIMA model for stock market forecasting in Botswana and Nigeria. The ARIMA (3, 1, 1) and ARIMA (1, 1, 4) models were found to be the best forecast models for the Botswana and Nigeria stock markets, respectively, according to the empirical investigation. Adebayo et al. (2014) discuss that the forecasting of stock price is important for the researchers of finance and economics field to improve prediction models throughout time ARIMA has been investigated in the literature for time series forecasting as a result of this motivated study. They go over the entire process of utilizing ARIMA to develop a predictive model. This study makes use of data from the New York Stock Exchange (NYSE) and the Nigeria Stock Exchange (NSE). The ARIMA model has a lot of potential for short-term prediction and can compete well with conventional stock price prediction approaches, according to the findings.

Wahyudi (2016) reports the forecasting of stock price volatility is an important. This study is about Indonesia stock price by using ARIMA because to its ease of use and widespread acceptance From January 4, 2010 to December 5, 2014, a daily CSPI was used to achieve this goal. ARIMA may be used to forecast Indonesia stock prices, according to the study. The findings showed that the model has a lot of promise for short-term prediction and that it can compete with other stock price prediction strategies. Using AIC criterion, the best ARIMA model for forecasting the Indonesia Composite Stock Price Index was identified. ARIMA (0, 0, 1) was shown to be the best model.

Gay et al. (2016) discusses the link between stock prices and macroeconomic indicators are well established (emerging) for the United States and other large countries. The purpose of this study is to use the Box-Jenkins ARIMA model to analyze the time series link between stock market index prices and macroeconomic factors such as exchange rate and oil price for the BRIC countries of Brazil, Russia, India, and China. Although there was no significant relationship between the respective exchange rate and oil price on the stock market index prices of either BRIC country, this could be due to the impact of other domestic and international macroeconomic factors on stock market returns, necessitating further research. Furthermore, no substantial association was seen between current and previous stock market results, suggesting that BRIC markets are inefficient.

Ashik and Kannan (2019) examines that the National Stock Exchange is India's largest and most fully automated trading system. The Nifty 50 is a stock market investor, with 50 businesses involved in the traders. One of the most well-known and important aspects of the Box-Jenkins approach to time series modelling is ARIMA. Using Box-Jenkins approach, the Nifty 50 stock market prices were examined and the pattern of the future trading day's stock market swings was forecasted. The R-Square value is 94 percent high, and the Mean Absolute Percentage Error is relatively tiny, according to the data. As a result, the accuracy of the forecast is better for the Nifty 50 closing stock price. As a result, it can be stated that the current study's closing stock price of Nifty 50 exhibits a moderate falling fluctuation tendency for the following trading days.

Dikshita and Singh (2017) explain that Volatility is used to forecast risk associated with an asset in an indirect manner. The study compares the accuracy of several volatility estimators against the accuracy of forecasting techniques. Close, Garman-Klass, Parkinson, Roger-Satchell, and Yang-Zhang techniques are used to estimate volatility, while the ARIMA approach is used to forecast volatility. Various volatility estimators were tested for efficiency and bias in this study. The accuracy of predicting using the best volatility estimator was determined by comparison analyses using several error measuring metrics such as ME, RMSE, MAE, MPE, MAPE, MASE, and ACF. The research determined that the Parkinson estimator was the most efficient volatility estimator out of five volatility estimators studied over a 10-year period and severely tested for forecasting volatility. Based on the results of MAE and RMSE, the analysis reveals that the anticipated values were correct. This study was carried out in order to address the demand from traders, option practitioners, and other stock market participants for an efficient volatility estimator that can estimate volatility with high accuracy.

Afeef et al. (2018) talks about Stock price predicting has always been and will continue to be one of the most important financial conjectures faced by investors. The forecasting strategy used in this study was based on the variable's prior values. The ARIMA approach was used on the stock prices of OGDCL, one of Pakistan's major firms. The company's daily adjusted closing stock prices were collected for over 15 years, from 2004 to 2018. Some of the ARIMA archetypes utilized in the study have a high potential for prediction in the short term, according to the findings. As a result, the ARIMA model is quite effective in predicting the future in the near term. The outcomes of the study might help stock investors improve their predicting skills.

Latha et al. (2018) discuss that the forecast S&P Bombay Stock Exchange (BSE) Sensex index return values are forecasted. The BSE Sensex is made up of 30 blue-chip corporations, or the top 30 companies listed on the stock exchange. Future returns are forecasted using the financial econometric methodology Auto ARIMA algorithm. ARIMA is used to fit ten years of historical data from April 2007 to March 2017 and to anticipate future return values from April 2017 to March 2019. The AIC value was used to assess various sorts of models. Validation was carried out by comparing anticipated and actual data values over a two-year period between April 2015 and March 2017. Accuracy is measured using both the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). Different investors might use the approach to select firms based on their expected returns.

Wadi et al. (2018), reports that closed price forecasting is a common practice in finance and economics, prompting academics to develop a fit model to improve forecasting accuracy. Many applications have been created and deployed using ARIMA. The model used in this study predicts closed time series data obtained from the Amman Stock Exchange (ASE) between January 2010 and January 2018. The ARIMA model provides strong short-term prediction results and will be useful for investments, according to the findings. Mihir (2019) suggest that RWH is a consequence of two GBM and EMH. The study examines the RWH for 20 major stocks for Indian Banking Sector. The RWH for 20 significant equities in the Indian banking sector is examined in this study. This information was gathered from NSE from April 2017 to March 2018 using ARIMA model. The study also runs ADF, URT and ARIMA. For which ADF supports RWH but auto ARIMA provide evidence against RWH.

Pandey and Bajpai (2019) argue that the ARIMA (p, d, q) model and artificial neural networks (ANN) have both been widely used to forecast time series data, with indications that ANN is superior. For comparing predicted accuracy, several combinations of ARIMA (p, d, q) and ANN were utilized. Using daily data over a ten-year period, this research seeks to develop improved ARIMA and ANN combinations for forecasting the Indian Stock Market, namely the NSE Nifty fifty. ARIMA (p, d, q) and ANN model prediction accuracy was examined using AAE, RMSE, MAPE, and MSPE statistical methods at the same time. The findings show that ARIMA (2, 1 and 2) and ANN (4-10-1) with both train functions GDX and BFG are the most accurate predictors, with ANN outperforming ARIMA in the end.

Dinardi (2020) suggests that anticipating stock returns is critical in finance, econometrics, and academic disciplines including time series analysis. Based on such data, he evaluates the firms listed on the S. Paulo Stock Exchange and forecasts future stock return behavior using a variety of forecasting approaches. ARIMA models are commonly employed in time series analysis and produce satisfactory results in most cases. Other approaches, particularly those belonging to the ARCH family, must be examined in light of the high volatility data. These methods are mostly used to forecast data from stock markets throughout the world. Accurate projections may help both the organizations that generate them and the stakeholders directly since they give enough knowledge to make better future judgments.

Siami-Namini et al. (2019) opine that time series prediction challenges are being addressed with machine and deep learning-based techniques. These methods have been found to yield more accurate findings than traditional regression-based modelling. Artificial RNNs with memory, such as the LSTM, have been shown to outperform ARIMA by a significant margin.

According to Deka and Resatoglu (2019), "the significant amount of uncertainty in the global foreign currency market is causing alarm among market players, dealers, and policymakers". It is necessary to develop trustworthy and complex forecasting models for foreign exchange rates and their drivers in order to anticipate future values and thereby avoid risks. The ARIMA model is used in this study to forecast Turkey's foreign exchange rate, with inflation being a prominent influence. The best ARIMA model for forecasting Turkey's foreign exchange rate is ARIMA (3, 1, 3), whereas the best ARIMA model for forecasting inflation is ARIMA (1, 1, 4). However, the recommended methodology for estimating Turkey's foreign exchange rate and inflation should be updated with new facts as time goes on. In the process of selecting the optimal model, ACF, PACF, AIC, and BIC, as well as predicting performance indicators like as MAE, MAPE, Bias percentage, RMSE, and Theil U statistics, are quite valuable.

Dong et al. (2020) declares that there is no obvious criterion for selecting a time series window to match an ARIMA model. Furthermore, no definite conclusions have been reached on whether the sample's older data should be discarded. As a result, drawing a firm judgment on the ARIMA model's predictive ability is impossible. The goal of this work is to fill in this knowledge gap. It compiles about two million ARIMA estimates of future daily returns, based on data from January 3, 1996 to May 12, 2017. Different model parameter settings are used in the forecasts. With a yearly over-optimistic error of 2.6561 percent, they conclude that the five-year sliding fixed-width window best reflects US equity market asset values. When positive and negative return situations are separated, however, the ARIMA models produce forecasting errors of 0.0009% and 0.011 percent, respectively, and both underestimate gain and loss. For low volatility equities, these inaccuracies are lower. They come to the conclusion that the ARIMA model's lack of nonlinearity is not a big worry, and that the model's validity is not compromised provided the data windows are carefully chosen. As a result, the finding is consistent with the weak form market efficiency theory and holds up in a transaction-cost setting.

Boye et al. (2020) develop that the Box-Jenkins technique was used to create a short-term stock exchange forecasting model. They used the ARIMA model to analyze monthly data from the Ghana Stock Exchange market from March 2013 to February 2018. The Bayesian Information Criterion (BIC) data is fitted to the ARIMA (0, 2, 1) model (BIC). The residuals of the fitted model were uncorrelated, according to diagnostic testing. This model is used to forecast over a six-month period. The Ghana Stock Exchange's performance is expected to improve significantly during the next six months, according to the anticipated numbers.

The research by Challa et al. (2020) anticipates the return and volatility characteristics of the Bombay Stock Exchange's S&P BSE Sensex and S&P BSE IT indexes. They analyze daily stock returns time series using descriptive statistical tests such as variance ratio, Augmented Dickey-Fuller, Phillips-Perron, and Kwiatkowski Phillips Schmidt and Shin; and ARIMA. The results show that both indexes' mean returns are positive but close to zero. In the long run, this indicates a regressive trend. With just a few variances, the anticipated numbers are nearly identical to the actual ones. As a result, the ARIMA model can forecast medium and long-term horizons using historical S&P BSE Sensex and S&P BSE IT values.

Khan and Alghulaiakh (2020) states that the growing availability of historical data and the need to anticipate, which involves making investment decisions, building plans and strategies for future initiatives, and the challenge of predicting the stock market due to its complex elements , Auto ARIMA is used in this investigation. Using Netflix stock history data for five years, ARIMA (p, d, q) is utilized to obtain an accurate stock forecasting model. ARIMA (1, 1, 3) outperformed the other two models in terms of MAPE calculation and holdout testing, demonstrating the ARIMA model's utility in stock forecasting.

Dhyani et al. (2020) discusses that shares, bonds, securities, and currencies are all exchanged on a regular basis in the financial market, making the datasets time series data. ARIMA is a time series analysis tool that aids in stock price forecasting by extracting significant data. It also aids in comprehending what has occurred in the past, as well as predicting future data behavior. Time series is a unique attribute that requires a unique set of prediction algorithms. Basic, Trend-Based, and Wavelet-Based ARIMAs are the three types of ARIMA. The component of time series data has been discussed and implemented using ARIMA model for gathered NIFTY daily data of Nifty 50 index, with the goal of predicting future stock value.

Meher et al. (2021) examines that many investors utilize numerous strategies such as fundamental research and technical analysis to anticipate stock prices. They also rely on the comments supplied by various stock market analysts at times. ARIMA is a time series analysis technique used in prediction algorithms, and the purpose of this study is to use the ARIMA model to forecast the share prices of chosen pharmaceutical firms listed on the NIFTY 100 in India. For each company, the sample size begins on January 1, 2017 and ends on December 31, 2019. The ADF test is performed to see if the data is stationary. Significant spikes in the correlogram of ACF and PACF have been noticed for ARIMA model estimation, and several models have been built using distinct AR and MA terms for each selected business.

Then, using the Volatility Adjusted R-squared and Akaike Information Criterion, the top 5 models were chosen, and the required inculcation of various AR and MA terms was made to adjust the models and determine the best adjusted ARIMA model for each business. In future research endeavors, the findings might be utilized to study stock prices and forecasts in greater depth.The stock exchange indices fluctuate every day. Stock indexes are frequently reported to decrease or rise owing to politically and economically unstable situations. As a result, a significant amount of finance research has been done to investigate the essential components of stock indices. It is difficult to quantify political events or changes, thus much past research has concentrated on financial and economic factors for which secondary data is readily available.

Chapter 3

Research Methodology

In this Chapter of the study, the data description, population, variable, empirical test of forecasting models and econometric model are all discussed briefly.

3.1 Methodology

With the rising availability of historical data and the necessity for forecasting, time series forecasting (TSF) has gained popularity. TSF provides a sequence of predicting future values, which overcomes traditional forecasting's constraints such as complexity and time. TSF uses current and historical data to forecast system behavior in the future. TSF's significance in real-world problems like network traffic, petroleum, weather forecasting, and financial markets is so important.

Investing and developing plans or strategies for future undertakings is decided by investment institutions and people. As a result, current forecasting areas in the researchers' domain have been working to improve models over time. Particularly when the decision-making process is viewed as inaccessible in general owing to the requirement to absorb and extract information from vast amounts of data.

Forecasting stock prices has become an appealing endeavor for investors looking for the finest stock market results. As a result, in recent years, a number of models and strategies for predicting stock prices have been created. Data in time series are listed as points in time order, which are discrete time sequences that are uniformly spaced in time, and forecasting is done by evaluating observed points in the series to anticipate the future.

Forecasting of financial markets probably attracts most attentions from researchers. The benefits of forecasts are obvious. In this context, many researches are based on GARCH and related models, but by adding ARIMA models, an attempt is made to test and forecast stock prices. The ARIMA model must be used with validation and testing, which has not been done in the majority of past research. This model is suitable for reliably forecasting stock returns using market techniques. It is used to forecast future equity returns using data from the assets in question over time.

There are various forms of ARIMA. ARIMA-SVM model is used to predict the stock index returns. Further, it is modify through PSO. SVM and LSTM, both are hybrid models that are used to predict stock market returns. VARIMA model is used to detect multiple time series vector. SARIMA model is used to identify seasonal effect and Fourier effects. It is based on error corrective approach; also enhance the order of AR and MA model part. FARIMA is used to predict non-integer values.

It is critical to have a good awareness of both the current and future predictions, particularly while looking for a secure investment. Additionally, it aids in understand the role of the time series forecasting model ARIMA and accuracy of technique. ARIMA is most commonly employed with data that has a low level of volatility. To deal with time-correlated modelling and forecasting, Box and Jenkins (1970) established a systematic class (Shumway and Stoffer 2011). Despite its widespread use, there are still issues with ARIMA's accuracy. For instance, what window length should be used to create regression parameters and, as a result, anticipate outcomes, and if the window should be fixed width sliding or begin rising.

The primary idea of the ARIMA model is that a predicted object is generated as a random series to take a data series over a specific length of time. Based on autocorrelation analysis of the time series, a specific statistical approach for describing the series might be developed. Future values may be predicted using the past and current values of time series once the approach was established. An ARIMA (p, d, q) model is an I (d) process with a stationary ARMA (p, q) process as its integer difference.

3.2 Data Description

Emerging markets are not strongly synchronized with developed markets yet but attractive to foreign investor due to high ROI. Pakistan's stock market is considered as other an emerging market. It was categorized in May 2017; MSCI classified it as an emerging market, whereas FTSE classified it as a secondary emerging market. The Pakistan stock market is a tiny market with limited liquidity. This limits the role in boosting economy. It seems to be quite volatile, because of the noise trade and speculators. Also seems to yield gain to compensate for higher market volatility.

Because of its sensitivity to shocks and news, the Pakistan stock market is extremely risky and volatile. However, Pakistani stock market is resilient that bounce back. The behavioral quirks observed in individual investors and pricing anomalies and unexplained stock price swings resulted as a result of this.

Pakistan stock market is emerging with high returns; significant volatility, high market concentration, and a difficulty to mobilize fresh investment are all factors to consider. Moreover, the other aspect is that the Pakistan the stock market is divided and less susceptible to foreign shocks, allowing for worldwide diversification. The market also offers investors large gains, which compensate for market volatility. News articles, analyst forecast, change in earing of listed companies, manipulation by big investors, herd behavior, government policies and political scenarios are the factors that contribute to market volatility in Pakistan. A political scenario causes more volatility.

3.3 Population and Sample of the Study

The research focuses on Pakistan's stock exchange to find out that ARIMA model is fit, accurate and reliable for forecasting. For this purpose, the study examines through 4 indices of Pakistan stock exchange. These indices are the KSE-100 index, KSE-30 index, KMI-30 index and KSE all share. The nature of data is secondary. For this, the data has been obtained from "investing.com" and "business recorder". The duration of this secondary data is for the past 12 years that starts from 1 January 2009 to 30 June 2021(the start date is use to incorporate all the 4 indices of the study). In this study, daily data is used to ensure the larger size of the sample.

3.3.1 KSE-100

The KSE is the oldest and most popular stock exchanges in emerging market. In November 1991, the KSE-100 index was launched. The largest businesses from each of the 34 sectors are included in this market capitalization weighted index of equities, while the remaining 66 companies are chosen based on market capitalization, independent of industry sector. Ordinary stock, preferred stock, redeemable certificates, and term financing certificates are among the securities traded (corporate bonds). The lEC suggestion was adopted by the governing Board on April 24, 2012. Since June 11, 2012, it has been computed at full-capacity. Later, beginning October 15, 2012, it recomposes on a free-float technique.

The main goal is to establish a baseline against which the stock price performance over time may be measured. To put it another way, the KSE 100 is meant to provide investors an idea of how the Pakistan equities market is doing. As a result, the KSE100 is comparable to a number of economic indicators in Pakistan, including GNP and CPI.

3.3.2 KSE-30

In June 2005, The KSE-30 index is launched with a 10000-point base value. On September 2006, Index is based only on the free-float of share and Indices represented total return of the market. This provides the sense to investors that how well Pakistan's big corporations are performing on the stock exchange. As a result, the primary goal is to establish a baseline against which stock price performance may be measured throughout time. As a result, it keeps track of numerous aspects of the economy, such as GNP, CPI, and actual liquidity.

3.3.3 KMI-30

The KMI-30 was established in September 2008 (the index's base period is June 30, 2008, although it was established in 2009) with the primary objective of assessing the performance of shariah-compliant stock investments based on Islamic standards. The KSE and Al-Meezan Investment Bank have collaborated to produce it. It also serves as a research tool for the allocation of strategic assets. Its creation will promote investor trust and engagement, in addition to tracking the performance of Shariah compliant shares.

3.3.4 KSE all share

The KSE all share is based on the Full Cap approach, this index includes all businesses listed on the PSX.

3.4 Descriptions of Variables

As this is about forecasting of returns so there is only variable is return. The following formula is used to compute it:

$$R_{it} = \ln\left(\frac{p_t}{p_{t-1}}\right) \tag{3.1}$$

R_{it} is the is the index's return

ln is the logarithm of returns in natural logarithms.

P_t is the index's closing price at time t

P_{t1} is the index's closing price at time t-1

3.4.1 Variance Ratio Test

Lo and MacKinlay (1989) developed the VR test, it is used to anticipate asset values and to investigate time series data forecasting by comparing return variations at different intervals. According to Miranda Tabak (2003), "If the data follow a random walk, period variance must be equal to the times variance of a single period difference". So, whether the data follows a random walk or not, the VR test is predicated on that assumption.

To establish statistical significance, this study uses Lo, MacKinlay, and Kim's rank, rank score, and sign based forms. The letter S denotes the 1–7 series. The VR test developed by Lo and MacKinlay (1988, 1989) may be used in both homoscedastic and heteroscedastic random walks using asymptotic normal or wild bootstrap probability (Kim, 2006). For Lo and MacKinlay, S1 and S2 are the tests (1988). S1 denotes homoskedasticity, no bias correction, and random walk series, whereas S2 denotes heteroskedasticity, martingale series, and no bias correction. Kim's exams will be S6 and S7 (2006). S6 infers heteroskedasticity and random walk series from 1000 iteration, whereas S7 infers homoskedasticity and random walk series from 1000 iteration.

For statistical significance, the Wright (2000) tests were tested using the bootstrap. Rank and random walk series are defined in Wright (2000)'s S3. Wald and multiple comparison VR tests at many intervals are performed by (Richardson and Smith, 1991; Chow and Denning, 1993). The data in this study will be tested using a random walk series. The VR test for the rank score and random walk series is shown in S4; the VR test for the sign-based test and martingale series is shown in S5. The Curly brackets denote VR values, the Square brackets P values, and the Parenthesis Wald (Chi-Square) values.

3.5 Econometric Model

This study is based on forecasting tool known as ARIMA. This is a type of linear model that can describe both stationary and non-stationary time series and does not need independent variables in the data series to forecast. It is used when the data is stationary, but it may also be extended to non-stationary series by allowing the data series to be differentiated. It relies on autocorrelation patterns in the data. This model is different since no specific pattern in the past data of the series to be projected is assumed.

According to Pankratz (2009), this approach combines Autoregressive (AR) and

Moving Average (MA) models to anticipate future returns. The AR term denotes the current value of a time series calculated from prior values of the same series, whereas the MA term denotes the present value of a series calculated from a linear combination of past mistakes. p is the number of autoregressive terms, d is the number of differences, and q is the number of moving average terms in the generic non-seasonal ARIMA (p, d, q) model. Following is the model's mathematical formula:

$$R_{it} = \alpha_\circ + \sum \beta_i R_{t-1} + \sum \gamma_i e_{t-1} \tag{3.2}$$

$_t$ is the today value
$_{(t-i)}$ is the previous day lag value
$\gamma_i e_{t-1}$ lag value of error term- (MA) term
To check the seasonality of ARIMA;

$$R_t = \alpha_\circ + \sum \beta_i R_{t-1} + \sum \gamma_i e_{t-1} + \delta SAR + \delta SMA \tag{3.3}$$

$_t$ is the today value
$_{(t-i)}$ is the previous day lag value
$\gamma_i e_{t-1}$ lag value of error term- (MA) term
δSAR is the seasonal AR term
δSMA is the seasonal MA term

If the first degree difference is not used, the Box-Jenkins technique assumes that the time series is stationary. The ARIMA (p, d, q) model is used in this case, with d representing the degree of difference selection. ARIMA (p, d, q) becomes an ARMA (p, q) model if the time series already has stationarity.

According to Miswan et al (2014); Pahlavani and Roshan (2015), "Researchers feel that the GARCH and EGARCH models, when compared to ARIMA models, cannot produce the best results since ARIMA is the best model for forecasting and modelling stock prices". Furthermore, by assuming symmetric or asymmetric effects, several mixed models such as ARIMA-GARCH, TGARCH, EGARCH, or GJR may be used to determine stock return volatility.

According to Thushara (2018), “the ARIMA and ARIMA-GARCH models provide the same results throughout time, and volatility does not change”. As a result, the ARIMA model is used to forecast future returns, together with the mean and variance formulae. In a real-time context, the suitable model is based on four steps.

1. The first step is to *identify the suitable values of p, d, and q* using the correlogram and partial correlogram tools. The ADF test is also used to determine whether the data is stationary.
2. The *parameters of the chosen model are estimated* using the least squares approach in the second stage, which is estimation.
3. The third step involves doing a *diagnostic check* to see if the residuals from the fitted model include white noise. Accept the specified model if it already exists; if not, start over. As a result, this approach is iterative in nature.
4. The successful ARIMA model from step three is utilized both inside and outside the sample period to anticipate future stock price returns in the fourth stage, which is *forecasting performance.*

3.5.1 ARIMA Method Application

The ARIMA procedure is divided into two stages: the first is the creation of the ARIMA model, and the second is the comparison of anticipated outcomes to actual results. It is clear from the literature that a holdback time is necessary in order to confirm correct forecasts. The diagnostic and parameter significance tests are also used to determine if residuals are white sounds.

3.5.2 Correlogram to Determine the Appropriate Values of p, d, and q

The AC and PAC correlation coefficients for correlogram are two different forms of correlation coefficients. The autocorrelation function (ACF) shows how current initial differencing correlates with lags. The partial autocorrelation function

(PACF) shows the relationship between all of the study's data and their intermediate delays. The ACF and PACF methods are used to establish the kind of ARMA model and the required p and q variables. The following formula is used to compute the ACF:

$$\hat{P}_k = \frac{\gamma_t}{\gamma_\circ} \tag{3.4}$$

$\hat{P}_k$ Is the sample's ACF

γ_t At lag k, what is the covariance?

$\gamma_\circ$ The variance of the sample

3.5.3 Unit Root Tests

To check for series stationarity, the unit root test is utilized. The presence of unit roots is checked using three tests: ADF, PP, and KPSS in this investigation. The stock returns series null hypothesis, which asserts a unit root for ADF, PP, and KPSS, was rejected since the p-values were less than 5%. As a result, the stationary series was found to be devoid of unit roots in all three tests.

With an intercept, test the equation
With a trend and an intercept, test the equation.
Without the intercept, test the equation.

3.5.4 ARIMA Model Estimations

ARIMA is the result of combining the acronyms AR and MA. In order to calculate the best-fit values, a linear regression model is used. The ARMA might be used to predict future returns once it has been installed. This may be accomplished using either static or dynamic forecasting approaches. Static forecasting utilized current and delayed values, whereas dynamic forecasting used previously anticipated values. AIC and SIC (both) comparisons were used to identify the best fit of the time series data for future forecasting in the Auto ARIMA model estimate.

The validation stage is crucial for determining the anticipated values' correctness. A static forecasting instrument might be used in the ARIMA process to accomplish this. The authors sought to anticipate future returns after completing the estimating step by comparing forecasted returns to actual returns. The descriptive statistics reveals the mean returns, types of tendency, probability in returns and mean value of returns.

For index returns, the standard VR test, non-parametric VR test, multiple VR test, and modified version of multiple VR test statistics may reject the null hypothesis of a random walk or martingale. Also, it may conclude that there was no evidence in favor of the EMH in the study's findings in the long run and suggests that the past information of stocks the indicators show that EMH is semi-strong.

Chapter 4

Results and Discussions

The empirical study of econometric models in this chapter represents the models estimate and forecasting offered in the previous chapter. The talk includes several tests such as descriptive statistics for all of the indices in the research, variance ration test, unit root analysis, and check correlogram for model selection, ARIMA and SARIMA applied to examine the phenomena under discussion, and analyses the results obtained.

4.1 Data Analysis and Result Discussion

This chapter exhibits the empirical analysis of econometric models provided. This chapter contains four sections. First check the stationarity of data and generate graphs and then discusses about the descriptive statistics of all the four indices in the study. In the second section, the variance ratio test is used to determine whether indices have a random walk. Apply three alternative specifications, including random walk, exponential random walk, and random walk innovations with graphs, to accomplish this.

The third section is about unit root analysis to investigate the stationarity of data. There are the different tests in Augmented Dickey-Fuller (ADF), Phillip-Perron (PP), and Kwiatkowski Phillip Schmidt and Shin (KPSS) tests are types of unit root testing. ADF with the holding period or users specified lags 2 4 8 16 is tested at level and 1st difference with combinations of none, intercept and trend intercept along with AIC and SIC with lag 5. PP is tested at level and 1st difference with combinations of none, intercept and trend intercept along with AIC and SIC with lag 5. With lag 5, KPSS is examined at the level and 1st difference using a mix of intercept and trend intercept along AIC and SIC. Then apply ARIMA model for forecasting returns of four indices. For this, first check the correlogram of the selected data at level and 1st difference with lags 36.

Once the model is selected, run the equation with the selected model. The fourth section examines the market efficiency of Pakistan's stock market indexes, such as the KSE-100, KSE-30, KMI-30, and KSE all-share indices, using the auto regressive integrated moving averages (ARIMA) model.

4.1.1 Graphical Representations

The basic step of analysis is to see the behavior of the data. In other words, need to check the stationarity of 4 indices data. It must be stationary for further forecasting analyses. The KSE 100 index has a growing trend in first graph, while the returns of the KSE 100 index are stationarized in the second graph.

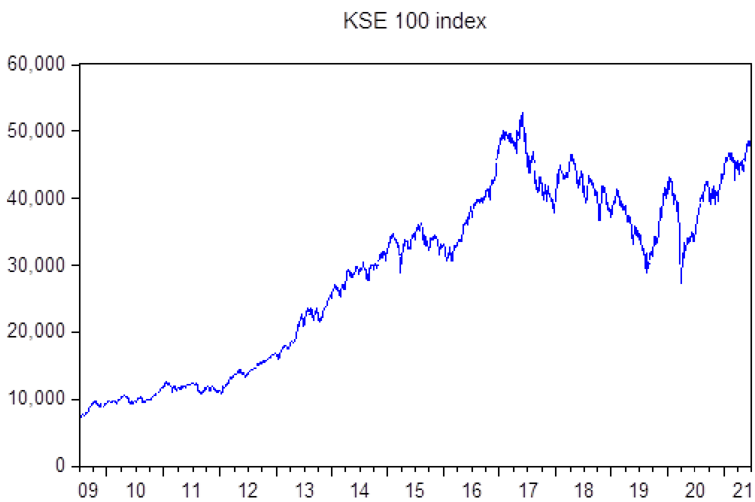


FIGURE 4.1: KSE100 Index

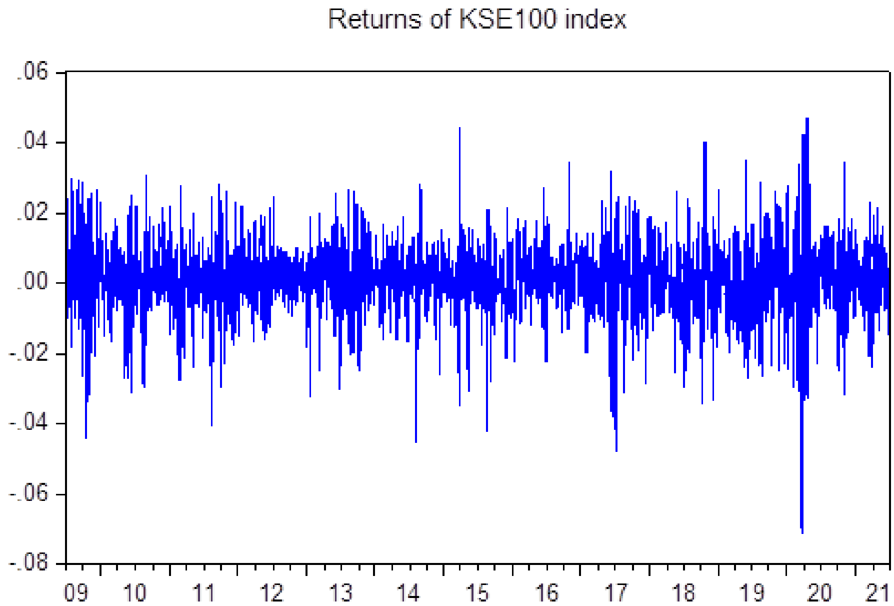


FIGURE 4.2: Returns of KSE100 Index

The KSE 30 index has a growing trend in first graph, while the returns of the KSE 30 index are stationarized in the second graph.

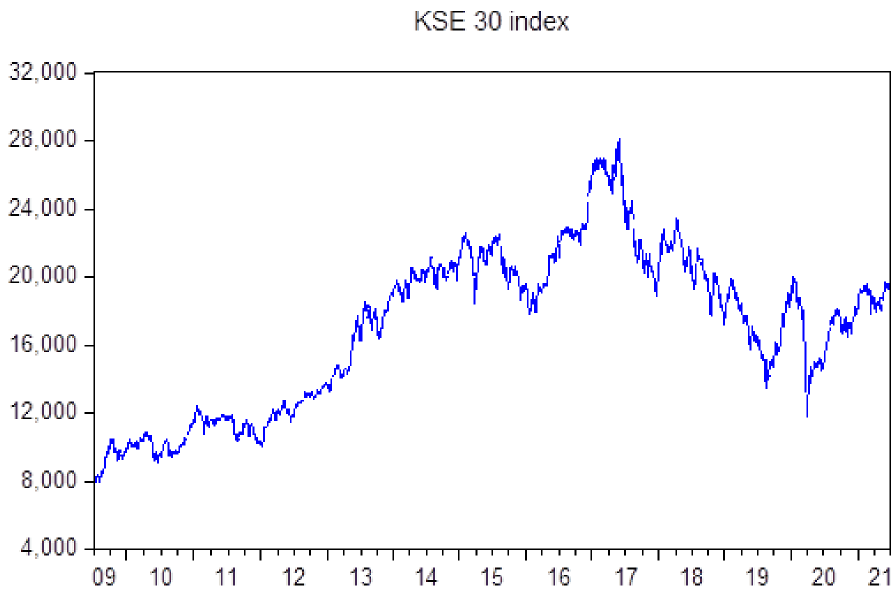


FIGURE 4.3: KSE30

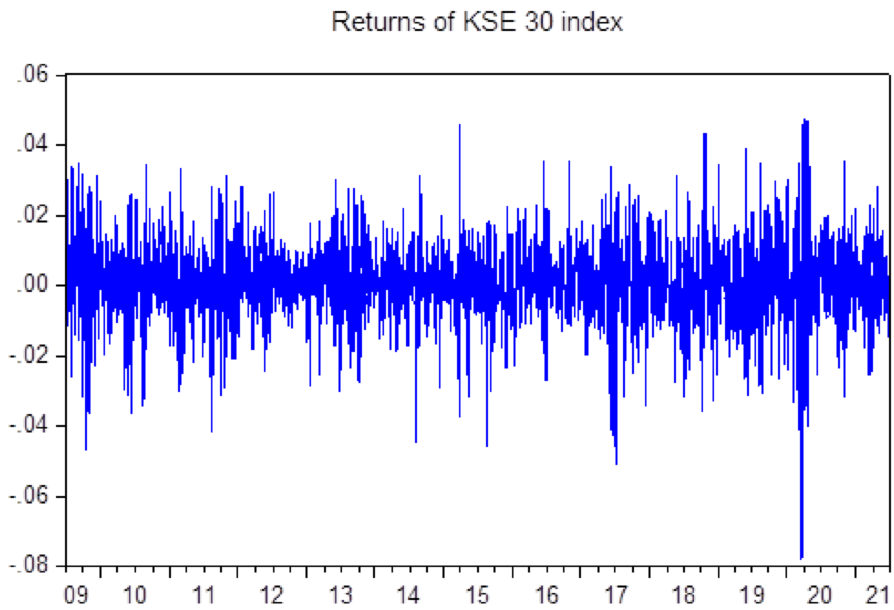


FIGURE 4.4: Returns of KSE30 Index

The KMI 30 index has a growing trend in first graph, while the returns of the KMI 30 index are stationarized in the second graph.

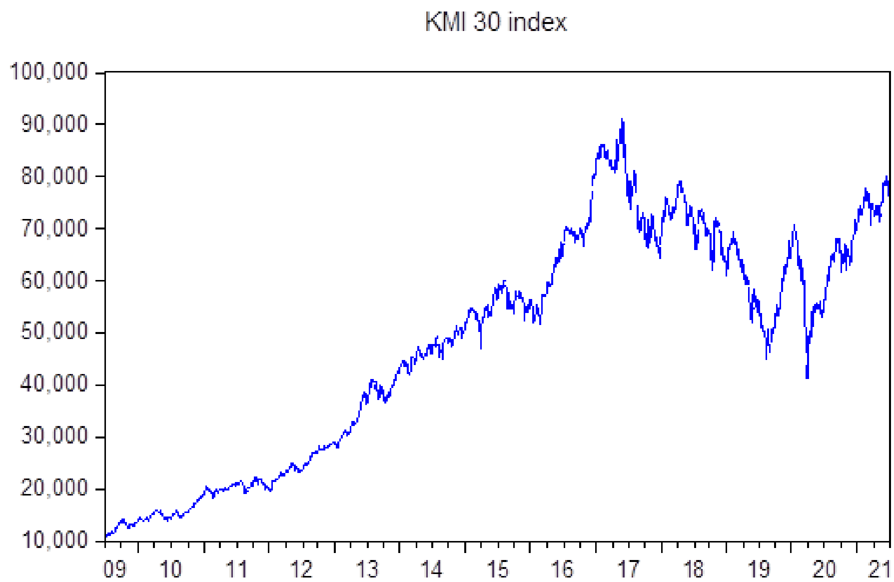


FIGURE 4.5: KMI30

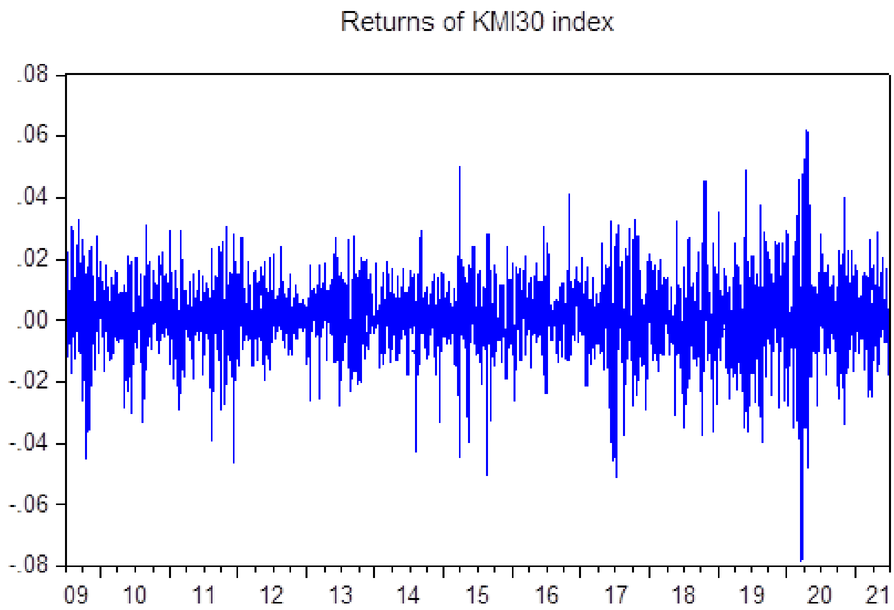


FIGURE 4.6: Returns of KMI30 Index

The KSE all-share index has a growing trend in first graph, while the returns of the KSE all-share index are stationarized in the second graph.

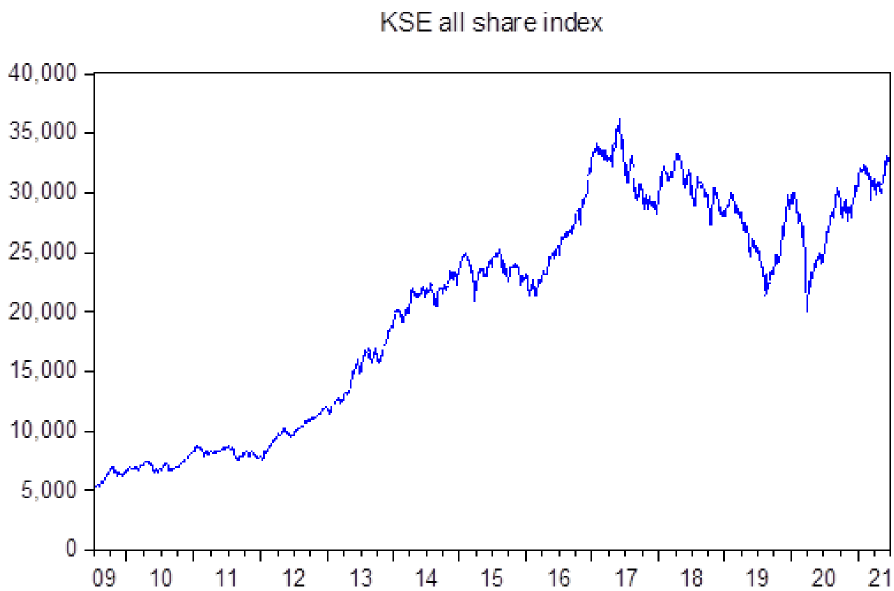


FIGURE 4.7: KSE all Shares

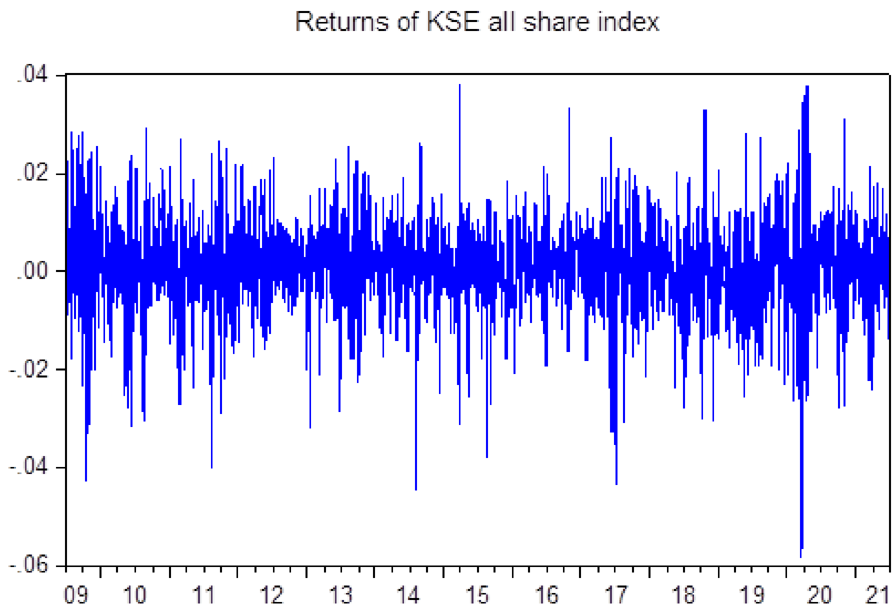


FIGURE 4.8: Returns of KSE all Shares Index

4.2 Descriptive Statistics

The next thing is to examine the behavior of data through descriptive statistics of the each index. The descriptive statistics is used to summarize a given set of data by using brief descriptive coefficients that can represent the entire sample. This is broken into two measures: measurements of central tendency and variability measures. The mean, median, and mode are measures of central tendency, whereas the standard deviation, variance, minimum and maximum variables, skewness and kurtosis, and Jarque-Bera are measures of variability. In other words, it is a summary of key moments or statistics such as mean, median, variance, standard deviation, skewness, kurtosis, and jarque-bera values that may be used to examine and guarantee that the data is normal and free of outliers.

Mean values shows the average indices points for the selected time period. The median value is the value that divides the sample data into two halves. Basically it's the middle value of the data. Similarly, standard deviations explain the data desperation from its mean. To measure the location of data we generally interpret the skewness and kurtosis figures. The asymmetry of a real-valued random variable's probability distribution around its mean is measured by skewness. It is defined as "in the set of data it is the degree of distortion from the symmetrical bell curve or normal distribution". The value of skewness shows the location of the data is negative or positive. The skewness value ranges from -1 (negative skewness), to +1 (positive skewness).

Kurtosis is a statistic that indicates whether a distribution is heavy-tailed or high-tailed in comparison to a normal distribution. It demonstrates the shape of the data (measure of peakedness/atness) and estimate (equal to) = 3 demonstrate that data is normally distributed, with a mesokurtic pattern. For values (more than) 3, a leptokurtic pattern is used, while for values larger than 3, a pattern with fat tails is used. When the Kurtosis value is (less than) < 3, the tails are less peaked and the Kurtosis value is called platykurtic. Jarque-Bera normality test indicates none of variable is normally distributed.

The table 4.1 covers the descriptive statistics for the all 4 indices with 2973 observations for this study. The sample period is taken of 12 years daily data starting from 1/7/2009 to 30/6/2021.

Table 4.1: Descriptive Statistics

Indices	KSE-100	KSE-30	KMI-30	KSE all share
Mean	0.00063	0.000303	0.000657	0.000617
Median	0.0008	0.000176	0.000415	0.000756
Maximum	0.04684	0.047279	0.061936	0.037986
Minimum	-0.071024	-0.077801	-0.078312	-0.057981
Standard Deviation	0.010421	0.011637	0.011842	0.009333
Skewness	-0.530186	-0.413651	-0.386516	-0.529612
Kurtosis	6.887575	6.616235	6.892224	6.176889
Jarque-Bera	2011.436	1704.716	1950.658	1389.206

The mean value of KSE-100 is 0.000630, KSE-30 is 0.000303, KMI-30 is 0.000657 and KSE all share is 0.000617. So, the all of the indices' mean values are positive, according to the research and the highest mean value is KMI-30 and the lowest value is KSE-30. It also shows the mean average returns of all the indices during this period are positive, but close to zero, indicating a long-term regressive trend. The median value of KSE-100 is 0.000800, KSE-30 is 0.000176, KMI-30 is 0.000415 and KSE all share is 0.000756. The highest median value is of KSE-100 and the lowest median value is of KSE-30.

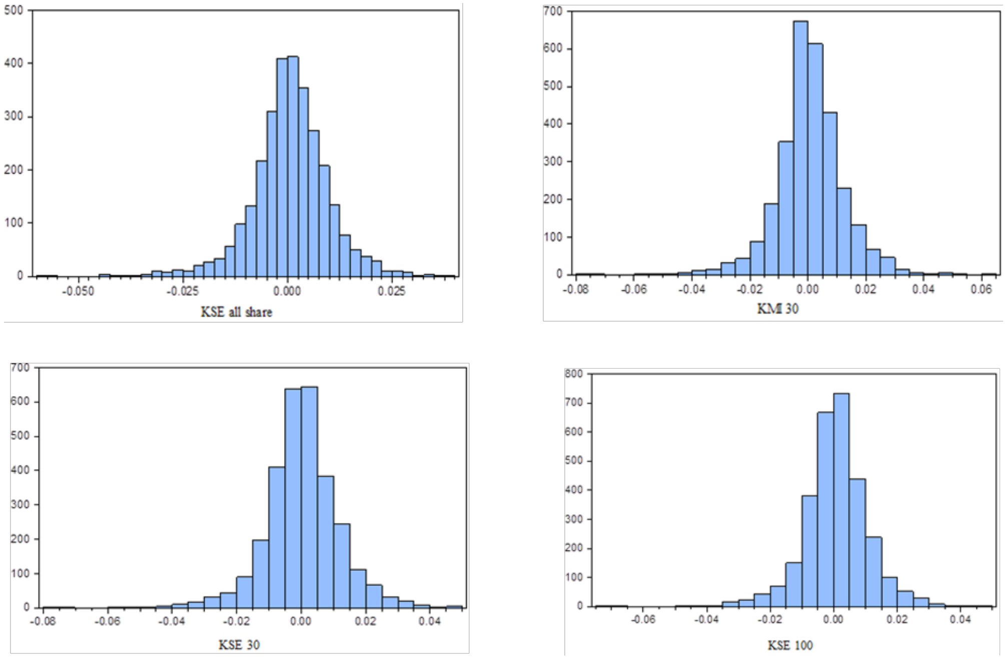


FIGURE 4.9: Descriptive Statistics

The maximum value of KSE-100 is 0.046840, KSE-30 is 0.047279, KMI-30 is 0.061936 and KSE all share is 0.037986. The highest maximum value is of KMI-30 and the lowest maximum value is of KSE all share. The minimum value of KSE-100 is -0.071024, KSE30 is -0.077801, and KMI30 is -0.078312, KSE all share is -0.057981. The highest minimum value is of KMI-30 and the lowest minimum value is of KSE all share. The standard deviation value of KSE-100 is 0.010421, KSE-30 is 0.011637, and KMI-30 is 0.011842, KSE all share is 0.009333. The standard deviation value is of KMI-30 and the lowest standard deviation value is of KSE all share. These values show the high volatility in the all 4 indices especially in the chosen period.

The skewness value of KSE-100 is -0.530186, KSE-30 is -0.413651, KMI-30 is -0.386516 and KSE all share is -0529612. So it's easy to predict that they are negatively skewed at left and represents as asymmetric tail. The kurtosis value of KSE-100 is 6.887575, KSE-30 is 6.616235, KMI-30 is 6.892224 and KSE all share is 6.176889. The kurtosis for all indices is (greater than) > 3 that indicates the presence of the fat tail distribution of stock returns. As a result, all of the indices that describe the sharp peak and fat tail distribution are larger than 3. This indicates that the time series is not distributed normally. The value of jarque-bera KSE-100 is 2011.436, KSE-30 is 1704.716, KMI-30 is 1950.658 and KSE all share is 1389.206. All the indices values are significantly greater than those found in a typical normal distribution. The graphical representation is shown in fig 4.9.

In precise, the descriptive statistics indicated that the mean returns were positive but close to zero, indicating a regressive trend in long-term values. Because the value of skewness is bigger than the mean value of returns, an asymmetric tail suggests a high possibility of earning in returns with a high risk. The data is not normally distributed as indicated by Jarque-Bera test which is based on skewness and kurtosis.

4.2.1 Variance Ratio Test (VR Test)

A common way to look at return fluctuations over time is to use a regression model. Furthermore, if the data follows a random walk, period variance must be measured in times variances of a single period difference. The variance ratio test assumes the data follows a random walk or not. In each of the three tests, the short holding durations 2, 4, 8, 16 are employed such as random walk, exponential random walk and random walk innovations are considered for the VR tests of all 4 indices. The table 4.2 shows results and graphical representations are shown in fig 4.10, 4.11, and 4.12 for all the 3 tests:

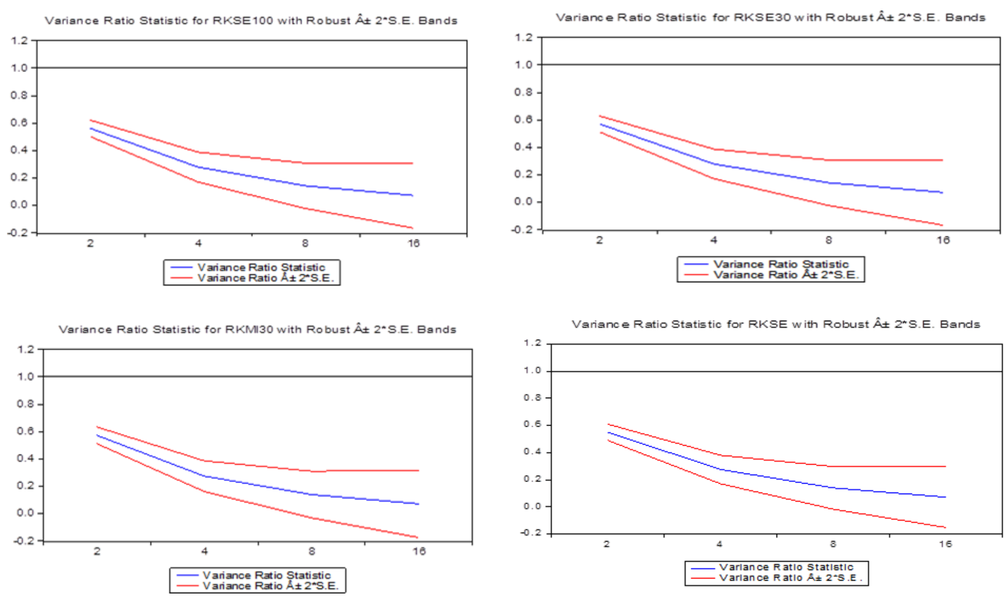


FIGURE 4.10: Random Walk

Table 4.2: Variance Ratio Test

			2	4	8	16
Returns of KSE 100	RW	VR	0.5616	0.2795	0.1431	0.0729
		Z-stats	-14.5724	-14.5724	-10.3849	-7.8241
		Prob	4.2064	4.5476	2.9034	5.1141
	ERW	VR	0.5530	0.2648	0.1568	0.0812
		Z-stats	-14.7923	-15.0580	-12.8163	-10.2275
		Prob	1.6420	3.0572	1.3293	1.4933
	RWI	VR	1.1239	1.2114	1.3268	1.4177
		Z-stats	4.7312	4.1102	3.9534	3.4639
		Prob	2.2322	3.9537	7.7049	0.0005
Returns of KSE 30	RW	VR	0.5707	0.2791	0.1428	0.0727
		Z-stats	-14.6207	-13.3536	-10.3611	-7.7548
		Prob	2.0731	1.1293	3.7263	8.8501
	ERW	VR	0.5637	0.2776	0.1533	0.0734
		Z-stats	-14.8995	-14.6369	-12.2607	-10.0664
		Prob	3.3186	1.6345	1.4724	7.7741
	RWI	VR	1.1198	1.1800	1.2716	1.3316
		Z-stats	4.5021	3.4636	3.2573	2.7201
		Prob	6.7277	0.0005	0.0011	0.0065
Returns of KMI 30	RW	VR	0.5732	0.2745	0.1400	0.0716
		Z-stats	-13.8799	-12.9598	-10.0731	-7.5360
		Prob	8.3842	2.0667	7.2649	4.8443
	ERW	VR	0.4974	0.2546	0.1249	0.0719
		Z-stats	-15.1622	-13.7788	-12.0208	-9.8329
		Prob	6.2928	3.4168	2.7615	8.1274
	RWI	VR	1.1036	1.1285	1.1924	1.2439
		Z-stats	3.8641	2.4442	2.2797	1.9703
		Prob	0.0001	0.0145	0.0226	0.0488
Returns of KSE all share	RW	VR	0.5503	0.2758	0.1403	0.0721
		Z-stats	-15.1387	-13.7797	-10.9398	-8.3038
		Prob	8.9952	3.3788	7.4386	1.0080
	ERW	VR	0.4850	0.2401	0.1245	0.0693
		Z-stats	-10.8654	-10.2873	-9.3294	-7.9944
		Prob	1.6841	8.0412	1.0652	1.3026
	RWI	VR	1.1166	1.2150	1.3483	1.4841
		Z-stats	4.7840	4.5032	4.5746	4.3664
		Prob	1.7185	6.6937	4.7704	0.0000

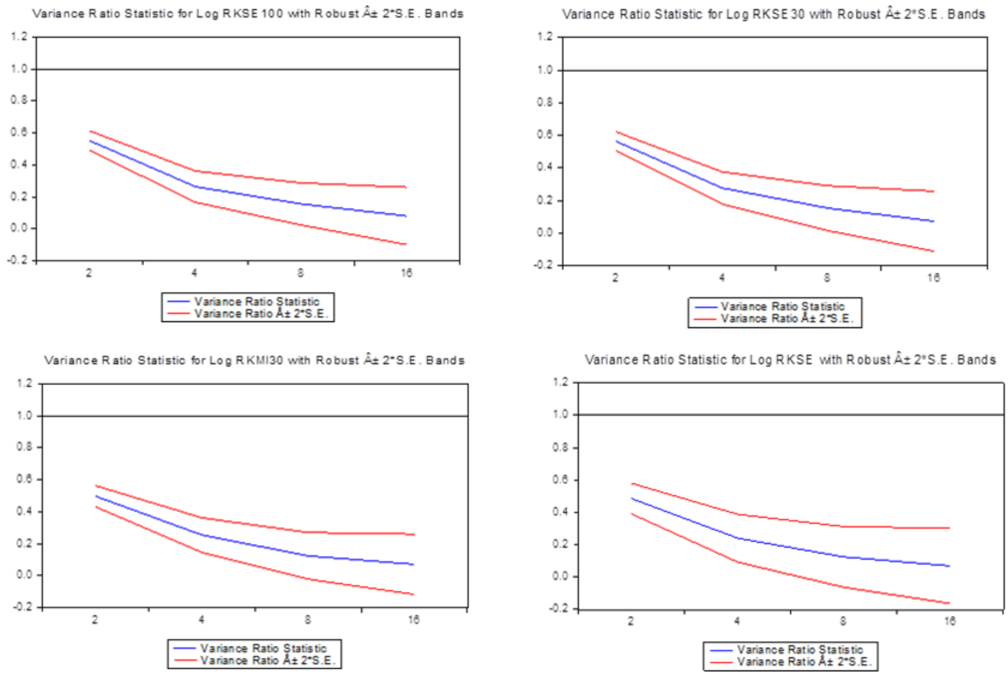


FIGURE 4.11: Exponential Random Walk

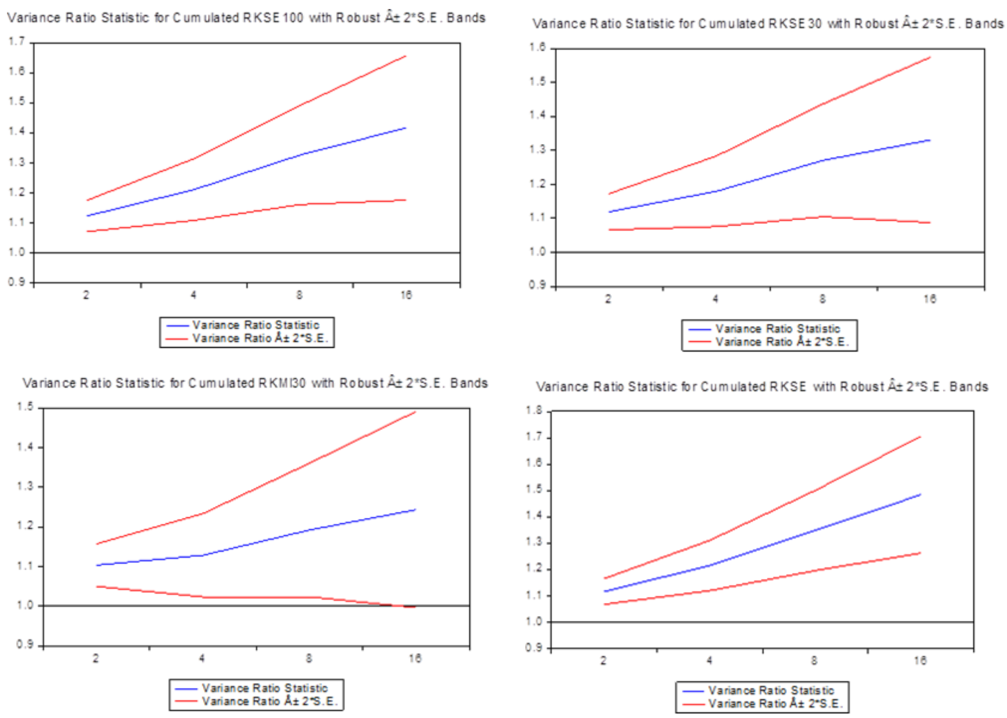


FIGURE 4.12: Random Walk Innovations

The Z-Statistic, Variance Ratio, and their accompanying p-values (probability) for this test are shown in the table above for 2, 4, 8, and 16 holding periods. The null hypothesis was rejected at the 1% level, and the returns could be strongly predicted using past prices. As a result, these indexes might be determined to be inefficient. Rapach et al. (2013) verified that the weak form was rejected using the same methodology.

4.2.2 Unit Root Test (UR Test)

Time series data is not normally distributed mostly, that's why stationarity may not be there. The stationarity of the series is checked using unit root tests because many previous studies report that it is essential to check the stationarity of data before applying the statistical method. This is also required for the trend in data and estimation of seasonality. In current analysis of the 4 indices, there are 3 techniques used to check the stationarity in all indices individually through unit root test such as ADF, PP, and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests have been enhanced. The endogenous and exogenous variables in the investigation are determined using these tests.

To examine for stationarity and seasonality in our data, we used the Augmented Dickey Fuller (ADF) test. It is employed with the assumptions of none, constant, and trend at the level and initial difference along with selection of AIC and SIC with maximum lags 5. AIC and SIC, both are the preferred to choose the model. The Phillip-Peron (PP) test rejects the null hypothesis. It is employed at the level and first difference with the assumptions of none, constant, and trend, as well as the default Bartlett Kernel with Newey-West bandwidth selection.

The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test is used to check if a time series is stationary around a deterministic trend. It is utilized at the level and first difference with the assumption of constant and trend, as well as the default Bartlett Kernel with automated Newey-West bandwidth selection. The null hypothesis that ADF, PP, and KPSS have a unit root was rejected since less than 5% of the p-values were positive. All of the strategies used in the desired indices were subjected to a unit root test, which is shown in the table 4.3.

Table 4.3: Unit Root Test

			KSE100		KSE30		KMI30		KSE all share	
			T-stats	**Prob**	**T-stats**	**Prob**	**T-stats**	**Prob**	**T-stats**	**Prob**
ADF	**Level**	None	-48.0073	0.0001	-48.3329	0.0001	-37.7203	3.149	-48.3261	0.0001
		Intercept	-48.1496	0.0001	-48.3513	0.0001	-37.8658	9.198	-48.5002	0.0001
		T&I	-48.187	3.3887	-48.3778	6.3887	-37.9232	6.3887	-48.5546	6.3887
	1st diff	None	-37.4704	6.79	-37.8551	7.1853	-38.0945	3.0769	-37.3363	2.9676
		Intercept	-37.4641	5.0272	-37.8487	8.2848	-38.0881	4.742	-37.33	1.8953
		T&I	-37.4577	6.3887	-37.8423	6.3887	-38.0817	6.3887	-37.3237	6.3887
PP	**Level**	None	-48.4972	0.0001	-48.4944	0.0001	-49.1167	0.0001	-49.2097	0.0001
		Intercept	-48.4882	0.0001	-48.499	0.0001	-49.1532	0.0001	-49.1244	0.0001
		T&I	-48.4894	6.3887	-48.5086	6.3887	-49.1776	6.3887	-49.1219	6.3887
	1st diff	None	-906.2505	0.0001	-801.2825	0.0001	-1331.0365	1	-890.627	0.0001
		Intercept	-906.1607	0.0001	-801.2846	0.0001	-1332.4217	1	-890.4581	0.0001
		T&I	-908.4651	0.0001	-802.8146	0.0001	-1340.2392	1	-892.9518	0.0001
KPSS	**Level**	Intercept	0.3395	0	0.2658	0	0.4074	0	0.4175	0
		T&I	0.052	0	0.0363	0	0.0441	0	0.057	0
	1st diff	Intercept	0.1073	0	0.1104	0	0.1539	0	0.1049	0
		T&I	0.0518	0	0.0548	0	0.0803	0	0.0453	0

For the best results, employ both the AIC and the SIC since both are desirable models to use. In the KMI 30 index, the results are almost same, however SIC outcomes are greater than AIC. This comparison of AIC and SIC reveals that for predicting, it is preferable to use the AIC which leads the lower order model. The smaller the AIC, the better the model fits. It is a formula for calculating the proxy function's estimate. A lesser degree of information is indicated by a negative AIC than by a positive AIC.

4.2.3 ARIMA Model

ARIMA is an acronym that combines two terms: AR and MA. ARIMA is the name for the generic model (p, d, q). The number of autoregressive terms (p) is the number of terms in the model. q is the number of moving average terms, and d is the number of differences.

The AR phrase refers to a present value of a time series that may be calculated using prior values from the same series. The AR order is represented by p, which is the lag value after which the PACF plot crosses the upper confidence interval for the first time, and the MA term is a present value of series, which is defined as a linear combination of past errors; assume the errors are distributed independently with a normal distribution. The MA order is represented by the number q, which can be found on the ACF plot; this is the time lag after which the ACF passes the upper confidence interval for the first time. The correlogram was measured to identify the best-fit values.

4.2.3.1 Correlogram

The correlogram is a tool that used to checking the randomness in a data set. Auto-correlation Function (ACF) and Partial Auto-correlation Function (PACF) are the two forms of correlation coefficients used in correlogram (PACF). An ACF measures the average correlation in a time series data and previous values of the series measured for different lag lengths, whereas an ACF measures the average correlation in a time series data and previous values of the series measured for different lag lengths. PACF indicates the correlation between the total observations of the study and their intermediate lags or we can say that it is used to correlation controls for any correlation between observations of a shorter lag length.

4.2.3.2 ACF and PACF Through Correlogram

In this analysis, the correlogram is an image that describes the correlation statistics of the data set at level and first-differencing with 10 lags. This method is used to calculate the proper p and q parameters for the ARMA model. The tables of correlogram of all the 4 indices are shown below; first the KSE 100 index is as under Table 4.4:

Table 4.4: ACF and PACF (KSE-100)

ACF and PACF at Level						
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
|*******	|*******	1	0.9988	0.9988	2969.73	0
|*******	| |	2	0.9976	-0.0006	5933.19	0
|*******	| |	3	0.9963	-0.0048	8890.34	0
|*******	| |	4	0.9951	-0.0022	11841.1	0
|*******	| |	5	0.9938	-0.0051	14785.5	0
|*******	| |	6	0.9926	-0.0104	17723.4	0
|*******	| |	7	0.9913	-0.0093	20654.6	0
|*******	| |	8	0.99	0.0032	23579.2	0
|*******	| |	9	0.9887	0.0107	26497.4	0
|*******	| |	10	0.9875	-0.0017	29409.1	0
ACF and PACF at 1ST DIFFERENCE						
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
|* |	|* |	1	0.1246	0.1246	46.2032	1.00E-11
| |	| |	2	0.0176	0.0021	47.1297	6.00E-11
| |	| |	3	0.0134	0.0111	47.6607	3.00E-10
| |	| |	4	0.0227	0.02	49.199	5.00E-10
| |	| |	5	0.03	0.0249	51.8821	6.00E-10
| |	| |	6	0.0303	0.0235	54.6253	6.00E-10
| |	| |	7	0.0064	-0.0011	54.747	2.00E-09
| |	| |	8	0.0008	-0.0012	54.7489	5.00E-09
| |	| |	9	-0.0181	-0.0199	55.7257	9.00E-09
| |	| |	10	0.0167	0.02	56.5619	2.00E-08

The above correlogram findings for identifying the model using ACF and PACF for the KSE-100 index show that lag 1 is adequate for running the ARIMA model.

The correlogram for the KSE-30 index are as under Table 4.5:

TABLE 4.5: ACF and PACF (KSE-30)

ACF and PACF at Level						
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
|*******	|*******	1	0.9981	0.9981	2965.68	0
|*******	| |	2	0.9961	-0.0219	5920.59	0
|*******	| |	3	0.9941	-0.0001	8864.75	0
|*******	| |	4	0.9922	0.005	11798.3	0
|*******	| |	5	0.9902	-0.0099	14721.1	0
|*******	| |	6	0.9881	-0.0202	17632.7	0
|*******	| |	7	0.986	-0.0188	20532.7	0
|*******	| |	8	0.9839	0.0033	23421.2	0
|*******	| |	9	0.9818	0.0212	26298.8	0
|*******	| |	10	0.9798	0.0023	29165.4	0
ACF and PACF at 1ST DIFFERENCE						
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
|* |	|* |	1	0.1207	0.1207	43.386	4.00E-11
| |	| |	2	-0.0028	-0.0176	43.4087	4.00E-10
| |	| |	3	0.0048	0.0074	43.4772	2.00E-09
| |	| |	4	0.02	0.0188	44.6724	5.00E-09
| |	| |	5	0.0269	0.0226	46.8246	6.00E-09
| |	| |	6	0.0288	0.0235	49.2944	7.00E-09
| |	| |	7	0.0074	0.0015	49.4589	2.00E-08
| |	| |	8	-0.0016	-0.0027	49.4667	5.00E-08
| |	| |	9	-0.0214	-0.0222	50.8392	7.00E-08
| |	| |	10	0.0082	0.0121	51.0407	2.00E-07

The above correlogram findings for identifying the model using ACF and PACF for the KSE-100 index show that lag 1 is adequate for running the ARIMA model.

The correlogram for the KMI-30 index are as under Table 4.6:

TABLE 4.6: ACF and PACF (KMI-30)

ACF and PACF at Level						
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
|*******	|*******	1	0.9987	0.9987	2969.06	0
|*******	| |	2	0.9973	-0.0009	5931.18	0
|*******	| |	3	0.996	0.0007	8886.38	0
|*******	| |	4	0.9947	-0.0014	11834.7	0
|*******	| |	5	0.9933	-0.0015	14776	0
|*******	| |	6	0.992	-0.0116	17710.3	0
|*******	| |	7	0.9906	-0.0072	20637.4	0
|*******	| |	8	0.9892	0.0041	23557.4	0
|*******	| |	9	0.9879	0.009	26470.4	0
|*******	| |	10	0.9865	-0.0018	29376.5	0

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
ACF and PACF at 1ST DIFFERENCE						
|* |	|* |	1	0.104	0.104	32.1765	1.00E-08
| |	| |	2	-0.0263	-0.0375	34.237	4.00E-08
| |	| |	3	-0.0033	0.0035	34.269	2.00E-07
| |	| |	4	0.0176	0.0168	35.1932	4.00E-07
| |	| |	5	0.0291	0.0257	37.7166	4.00E-07
| |	| |	6	0.023	0.0185	39.2961	6.00E-07
| |	| |	7	0.0019	-0.0008	39.3066	2.00E-06
| |	| |	8	-0.0005	0.0005	39.3072	4.00E-06
| |	| |	9	-0.0079	-0.0088	39.4919	9.00E-06
| |	| |	10	0.0067	0.0072	39.6277	2.00E-05

The above correlogram findings for identifying the model using ACF and PACF for the KMI-30 index show that lag 1 is adequate for running the ARIMA model.

The correlogram for the KSE all-share index are as under Table 4.7:

Table 4.7: ACF and PACF (KSE All-Share Index)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
ACF and PACF at Level						
|*******	|*******	1	0.999	0.999	2971.03	0
|*******	| |	2	0.9979	-0.026	5936.78	0
|*******	| |	3	0.9969	-0.0036	8897.21	0
|*******	| |	4	0.9958	-0.0145	11852.2	0
|*******	| |	5	0.9947	-0.006	14801.6	0
|*******	| |	6	0.9935	-0.0151	17745.2	0
|*******	| |	7	0.9924	-0.0107	20683	0
|*******	| |	8	0.9912	-0.0026	23615	0
|*******	| |	9	0.9901	-0.0016	26541	0
|*******	| |	10	0.9889	0.0013	29461.2	0
ACF and PACF at 1ST DIFFERENCE						
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
|* |	|* |	1	0.1413	0.1413	59.3885	1.00E-14
| |	| |	2	0.0166	-0.0034	60.2115	8.00E-14
| |	| |	3	0.0172	0.0156	61.0929	3.00E-13
| |	| |	4	0.0152	0.0109	61.781	1.00E-12
| |	| |	5	0.0541	0.0513	70.513	8.00E-14
| |	| |	6	0.0398	0.0253	75.2356	3.00E-14
| |	| |	7	0.0128	0.003	75.7249	1.00E-13
| |	| |	8	0.0054	0.0014	75.8121	3.00E-13
| |	| |	9	-0.0026	-0.0056	75.8322	1.00E-12
| |	| |	10	0.0246	0.0229	77.6317	1.00E-12

The above correlogram findings for identifying the model using ACF and PACF for the KSE all-share index show that lag 1 is adequate for running the ARIMA model.

4.2.4 ARIMA Model (p, d, q) as (1, 1, 1)

After deciding on a model, the parameters must be calculated. This model is built using the least squares approach. There is, however, no easy method for estimating the parameters for MA components models. The estimates must be found iterative in LS method. An iterative approach is utilized instead. This entails beginning with a rough estimate and iteratively fine-tuning it until the sum of squared errors is as small as possible. Before forecasting, it is important to double-check the model's accuracy. The residuals' ACF and PACF patterns may indicate a problem for best fitted model and also tells about how the model can be improved.

As per the correlogram findings, using the best fitted ARIMA model (p, d, q) as (1, 1, 1) for the four indexes KSE-100, KSE-30, KMI-30, and KSE all shares. In the table 4.8 below, the equation is run and the results are presented.

TABLE 4.8: ARIMA (1,1,1)

	KSE100	**KSE30**	**KMI30**	**KSE all share**
Significance	0	0	2	2
Sigma Q	0	0	0	8.58
Adj R Sq	0.01	0.01	0.01	0.01
AIC	-6.3	-6.08	-6.04	-6.52
SBC	-6.29	-6.07	-6.03	-6.51

According to this, the values of all the important parameters are based on the decision rules. The significance value depends on the number of variable lesser than 5% confidence intervals (0.05). The result shows that the KSE-100 and KSE-30 has zero significant number of variables whereas KMI-30 and KSE all share both have the two significant numbers of variables. As we have only one equation (p, d, q) as (1, 1, 1). The KSE-100 index is 0.000106, the KSE-30 index is 0.000133, the KMI-30 index is 0.000138, and the KSE all index is 8.583720. The Sigma Q value for the KSE-100 is the lowest.

R-squared adjusted for the number of predictors in the model is known as adjusted R-squared. According to the rules, the Adjusted R-squared value must be highest. The KSE-100 index is 0.014580, the KSE-30 index is 0.013960, the KMI-30 index is 0.011109, and the KSE all index is 0.01320 and this value is high in KSE-100. The

AIC and SBC, both values must be minimum to fit in the model for forecasting. The value of AIC for KSE-100 index is -6.303329, the KSE-30 index is -6.081915, the KMI-30 index is -6.044194, and the KSE all index is -6.522485, while The value of SBC for KSE-100 index is -6.295260, the KSE-30 index is -6.073846, the KMI-30 index is -6.036125, and the KSE all index is -6.514416 and both the values are lesser in KMI-30.

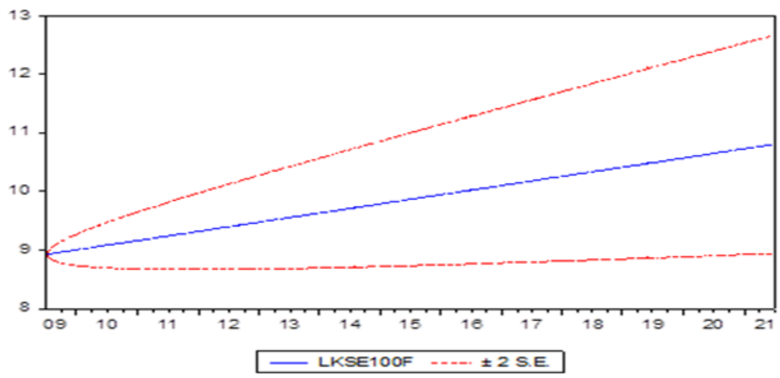
For estimating future returns, the ARIMA model is now employed. This may be accomplished using both the technique (dynamic and static) of forecasting, which employs both current and lagged variables.

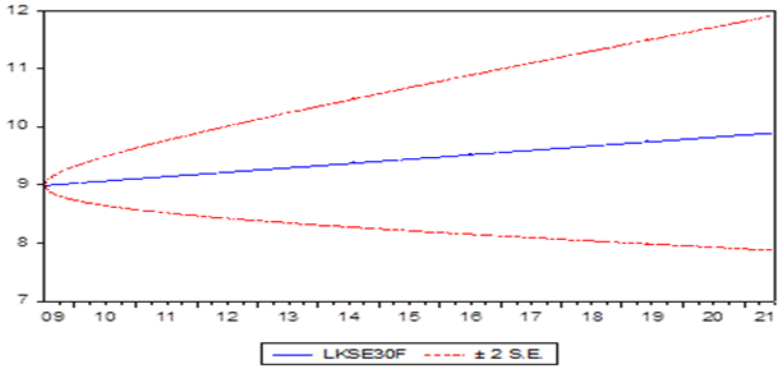
(a) ARIMA Through Dynamic Model

The table 4.9 of ARIMA model through dynamic method is attached in fig 4.13 with the graphs of all 4 indices.

TABLE 4.9: ARIMA Through Dynamic Model

	KSE100	KSE30	KMI30	KSE all share
RSME	0.3649	0.3427	0.4297	0.3743
MAE	0.3017	0.2912	0.3667	0.3065
MAPE	2.9361	2.9686	3.411	3.0809
Theil	0.0182	0.0179	0.0205	0.0194





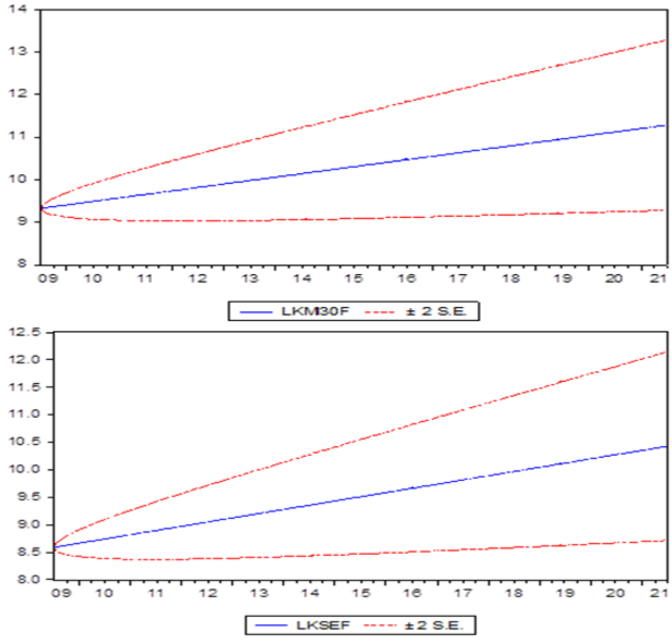


FIGURE 4.13: ARIMA through Dynamic Model

The measures utilized to isolate the forecasting model more accurately were root mean square error (RMSE) and mean absolute error (MAE). The MAE and RMSE were determined based on the differences in predicted and actual data. For the chosen period of 12 years, the selected ARMA models produce more accurate findings. Using a dynamic forecasting approach in the ARIMA process, this phase is critical for determining the accuracy of anticipated values.

The blue line represents the forecasted values whereas the red dotted line refers to first degree values. The difference between the actual and anticipated lines has widened. Because both values are not crossing at the same time, the projected and actual values are not identical.

(b) ARIMA Through Static Model

The table 4.10 of ARIMA model through static method is attached below in fig 4.14 with the graphs of all 4 indices.

TABLE 4.10: ARIMA (1,1,1)

	KSE100	**KSE30**	**KMI30**	**KSE all share**
RSME	0.0103	0.0115	0.0118	0.0093
MAE	0.0074	0.0083	0.0084	0.0067
MAPE	0.0732	0.0857	0.0786	0.0691
Theil	0.0005	0.0006	0.0006	0.0005

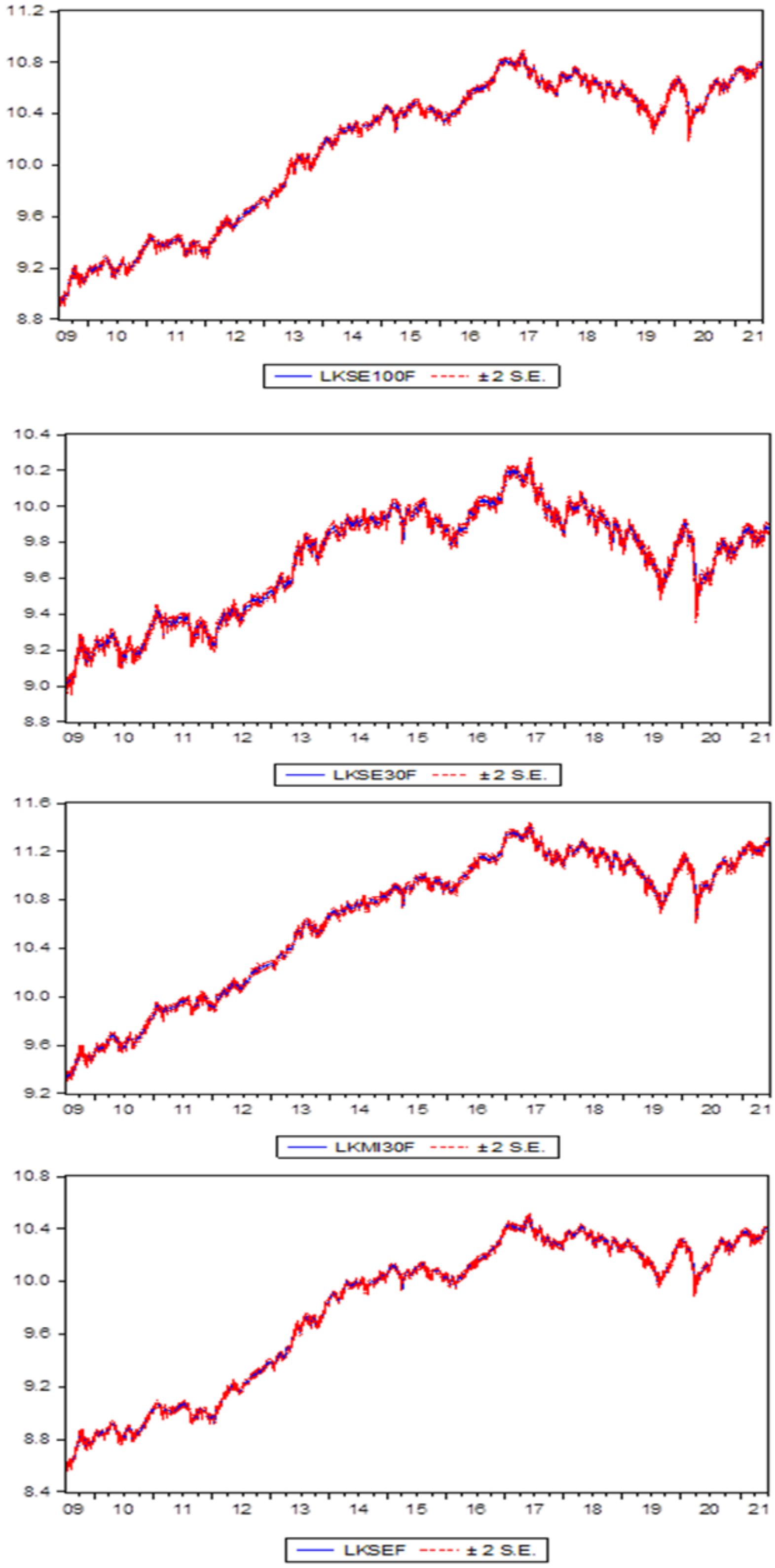


FIGURE 4.14: ARIMA through Static model

The measures utilized to isolate the forecasting model more accurately were root mean square error (RMSE) and mean absolute error (MAE). The MAE and RMSE were determined based on the differences in predicted and actual data. For the chosen period of 12 years, the selected ARMA models produce more accurate findings. Using a static forecasting approach in the ARIMA process, this phase is critical for determining the accuracy of anticipated values.

The blue line represents the forecasted values whereas the red dotted line refers to first degree values. Because both numbers are crossing at the same time, the projected and actual values are almost identical.

4.2.4.1 Seasonal ARIMA Model (p, d, q) as (1, 1, 1)

As per the correlogram findings, using the best fitted Seasonal ARIMA model (p, d, q) as (1, 1, 1) for the four indexes KSE-100, KSE-30, KMI-30, and KSE all shares. In the table 4.11 below, the equation is run and the results are presented.

TABLE 4.11: Seasonal ARIMA (1,1,1)

	KSE100	**KSE30**	**KMI30**	**KSE all share**
Significance	0	0	0	0
Sigma Q	0.0001	0.0001	0.0001	8.5689
Adj R Sq	0.0149	0.0133	0.0105	8.5689
AIC	-6.303	-6.0806	-6.043	-6.5229
SBC	-6.2909	-6.0685	-6.0308	-6.5108

According to this, the values of all the important parameters are based on the decision rules. The significance value depends on the number of variable lesser than 5% confidence intervals (0.05). The result shows that all the indices have zero significant number of variables. As we have only one equation (p, d, q) as (1, 1, 1). The KSE-100 index is 0.000107, the KSE-30 index is 0.000133, the KMI-30 index is 0.000138, and the KSE all index is 8.568940. The Sigma Q value for the KSE-100 is the lowest.

R-squared adjusted for the number of predictors in the model is known as adjusted R-squared. According to the rules, the Adjusted R-squared value must be highest. The KSE-100 index is 0.014878, the KSE-30 index is 0.013296, the KMI-30 index is 0.010545, and the KSE all index is 8.568940 and this value is high in KSE all share.

The AIC and SBC, both values must be minimum to fit in the model for forecasting. The value of AIC for KSE-100 index is -6.302958, the KSE-30 index is

-6.080571, the KMI-30 index is -6.042952, and the KSE all index is -6.522860, while The value of SBC for KSE-100 index is -6.290854, the KSE-30 index is -6.068467, the KMI-30 index is -6.030849, and the KSE all index is -6.510756 and both the values are lesser in KMI-30.

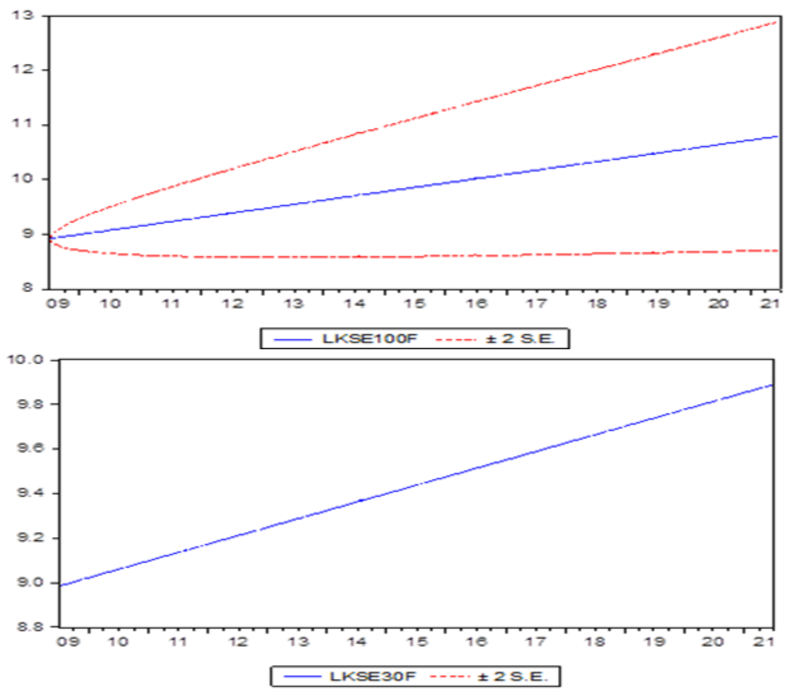
For estimating future returns, the Seasonal ARIMA model is now employed. This may be accomplished using both the technique (dynamic and static) of forecasting, which employs both current and lagged variables.

(a) Seasonal ARIMA through Dynamic model

The table 4.12 of Seasonal ARIMA model through dynamic method is attached below fig 4.15 with the graphs of all 4 indices.

TABLE 4.12: Seasonal ARIMA through Dynamic model

	KSE100	KSE30	KMI30	KSE all share
RSME	0.3688	0.3476	0.4363	0.0092
MAE	0.3056	0.2957	0.3728	0.0067
MAPE	2.9749	3.0155	3.4691	0.069
Theil	0.0184	0.0182	0.0208	0.0005



The measures utilized to isolate the forecasting model more accurately were root mean square error (RMSE) and mean absolute error (MAE). The MAE and RMSE were determined based on the differences in predicted and actual data. For the

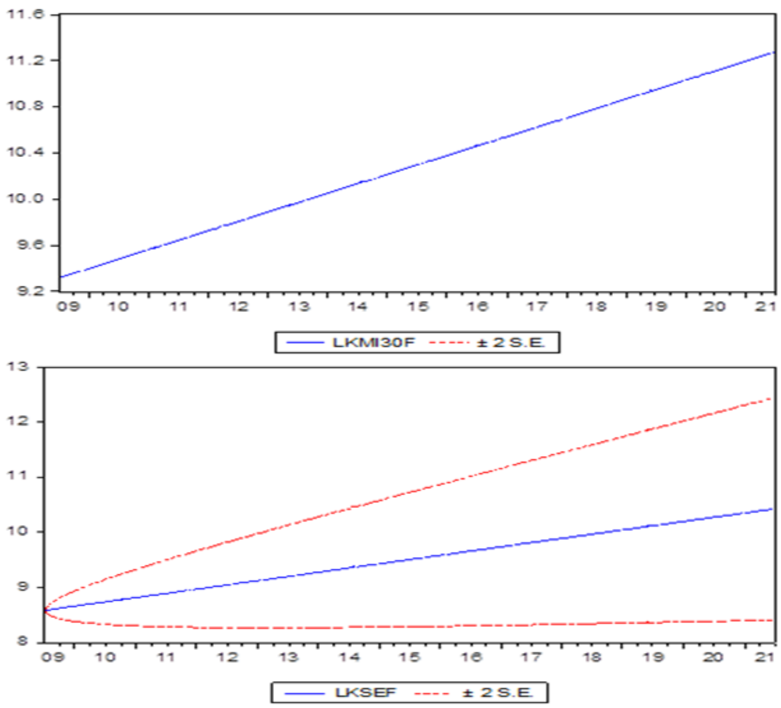


Figure 4.15: Seasonal ARIMA through Dynamic Model

chosen period of 12 years, the selected ARMA models produce more accurate findings. Using a dynamic forecasting approach in the seasonal ARIMA process, this phase is critical for determining the accuracy of anticipated values.

The blue line represents the forecasted values whereas the red dotted line refers to first degree values. The difference between the actual and anticipated lines has widened. Because both values are not crossing at the same time, the projected and actual values are not identical.

(b) Seasonal ARIMA through Static model

The table 4.13 of Seasonal ARIMA model through static method is attached below fig 4.16 with the graphs of all 4 indices.

Table 4.13: Seasonal ARIMA through Static model

	KSE100	**KSE30**	**KMI30**	**KSE all share**
RSME	0.0103	0.0115	0.0118	0.3755
MAE	0.0074	0.0083	0.0084	0.3081
MAPE	0.0732	0.0857	0.0786	3.0983
Theil	0.0005	0	0.0006	0.0194

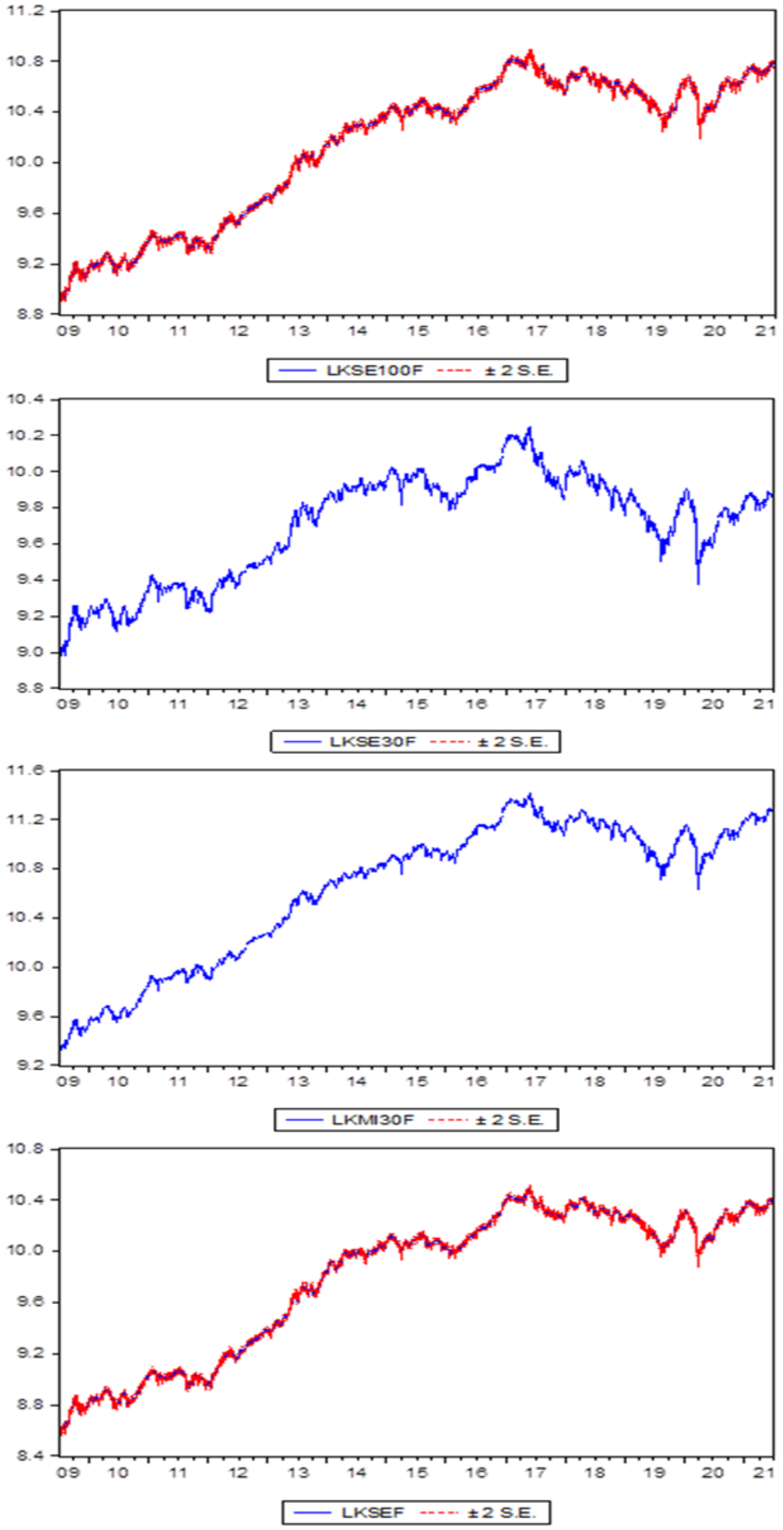


Figure 4.16: Seasonal ARIMA Through Static Model

The measures utilized to isolate the forecasting model more accurately were root mean square error (RMSE) and mean absolute error (MAE). The MAE and RMSE were determined based on the differences in predicted and actual data. For the chosen period of 12 years, the selected ARMA models produce more accurate findings. Using a static forecasting approach in the ARIMA process, this phase is critical for determining the accuracy of anticipated values. The blue line represents the forecasted values whereas the red dotted line refers to first degree values. Because both numbers are crossing at the same time, the projected and actual values are almost identical.

Chapter 5

Conclusion and Recommendations

5.1 Conclusion

Pakistan is a developing country with a rapidly growing economy. In recent years, it has faced tremendous problems, instability, and a widening divide. In Pakistan's market, there is a lot of volatility and interest rate. Pakistan is a country that has yet to be discovered with ARIMA model. It is worthwhile to look at the Pakistani stock market's efficiency in order to create a helpful tool for predicting stock prices and making investment decisions. The random walk is not followed because the EMH hypothesis states that the market is random. The Pakistani market is said to be inefficient, according to the literature. To examine the weak form of efficiency, many econometric approaches are applied.

The ARIMA technique is used for forecasting stock market returns. This paradigm is known as the Box-Jenkins (BJ) approach. It may be used to look at historical data and moving averages of random error factors. SARIMA is an ARIMA modification that enables direct seasonal component modelling of series, particularly for univariate models like return. ARIMA (p, q) as (1, 1) was used for all indices (KSE-100, KSE-30, KMI-30, and KSE all share), and this analysis found uncertainty, especially over extended periods of time. The study shows how well the method predicts complex and volatile stock data series. The ARIMA model's accuracy and speed were demonstrated using time series data.

The findings show that all of the indices have positive mean returns, but they are near to zero. This indicates a long-term downward tendency. The predicted values for the KSE-100, KSE-30, KMI-30, and KSE are almost equal to the actual values with reduced volatility. These findings have far-reaching implications. The expected returns studied in this study can be used to make investment decisions. Furthermore, by investing in profitable stocks, investors may construct a healthy portfolio.

Dynamic and static models are used to evaluate the ARIMA model. The distance between the actual and forecasted lines is widening in the dynamic model, whilst the actual and forecasted lines are overlapping in the static model. The outcomes are significant. The Seasonal ARIMA is also put to the test using dynamic and static models. The difference between the actual and forecasted lines is widening in the dynamic model, but the actual and forecasted lines are synced in the static model. The outcome is inconsequential or insignificant.

The market in Pakistan is said to be inefficient since it is predictable and does not follow a random walk. In comparison to the dynamic model, the static model performs better in terms of prediction. Seasonality does not occur, as evidenced by the negligible outcomes.

Despite this, the study has a few drawbacks. The corporate sector is represented by the KSE-100, KSE-30, KMI-30, and KSE all share. Under these indexes, there are several sectorial indices that use a more holistic approach and give investors with tips on how to get higher returns on their investments. Furthermore, the research may have focused on comparing the accuracy of return estimation over different time periods.

5.2 Recommendations

Researchers, corporations, investors, and governments may benefit from this study's findings in making educated stock market judgments. To investigate time series prediction, researchers can employ genetic models, nanotechnology models, and nonlinear regression models. Companies can design efficient methods to guarantee that their investments pay off. Individual investors may create their ideal portfolios, and policymakers can make well-informed decisions to maintain the stock market in good shape.

5.3 Limitations

The Future study can include stock price forecasting and comparisons in developed and emerging stock markets including Argentina, Brazil, China, Egypt, Indonesia, Iran, Mexico, Poland, Russia, Saudi Arabia, South Asia, South Korea, Taiwan, Thailand, and Turkey. Furthermore, long-term forecasting employing cutting-edge technologies will yield substantial profits. The focus will be on comparing a variety of sectorial indicators between Pakistan and other nations in order to obtain insight. Additionally, the focus will be on comparing a range of sectorial indicators inside the Pakistan stock exchange index as well as the various sectors to better comprehend portfolio structure, risk, return, performance, and trading efficiency.

Bibliography

Adebayo, F. A., Sivasamy, R., Shangodoyin, D. K., et al. (2014). Forecasting stock market series with arima model. *Journal of Statistical and Econometric Methods*, 3(3):65–77.

Afeef, M., Ihsan, A., and Zada, H. (2018). Forecasting stock prices through univariate arima modeling. *NUML International Journal of Business & Management*, 13(2):130–143.

Aiolfi, M. and Favero, C. A. (2005). Model uncertainty, thick modelling and the predictability of stock returns. *Journal of Forecasting*, 24(4):233–254.

Akhtar, S. and Khan, N. U. (2016). Modeling volatility on the karachi stock exchange, pakistan. *Journal of Asia Business Studies*, 10(3):253–275.

Al-Shiab, M. (2006). The predictability of the amman stock exchange using the univariate autoregressive integrated moving average (arima) model. *Journal of Economic and Administrative Sciences*.

Al-Tabbakh, S. M., El-Zahed, H., et al. (2018). Machine learning techniques for analysis of egyptian flight delay. *Journal of Scientific Research in Science*, 35(part 1):390–399.

Al Wadia, M. and Ismail, M. T. (2011). Selecting wavelet transforms model in forecasting financial time series data based on arima model. *Applied Mathematical Sciences*, 5(7):315–326.

Ashik, A. M. and Kannan, K. S. (2019). Time series model for stock price forecasting in india. In *Logistics, Supply Chain and Financial Predictive Analytics*, pages 221–231. Springer.

Assaf, A. (2006). Dependence and mean reversion in stock prices: The case of the mena region. *Research in International Business and Finance*, 20(3):286–304.

Asteriou, D. and Hall, S. (2015). Applied econometrics third edition.

Aumeboonsuke, V. and Dryver, A. L. (2014). The importance of using a test of weak-form market efficiency that does not require investigating the data first. *International Review of Economics & Finance*, 33:350–357.

Ball, R. and Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of accounting research*, pages 159–178.

Banerjee, D. (2014). Forecasting of indian stock market using time-series arima model. In *2014 2nd international conference on business and information management (ICBIM)*, pages 131–135. IEEE.

Boehmer, E. and Kelley, E. K. (2009). Institutional investors and the informational efficiency of prices. *The Review of Financial Studies*, 22(9):3563–3594.

Bossaerts, P. and Hillion, P. (1999). Implementing statistical criteria to select return forecasting models: what do we learn? *The Review of Financial Studies*, 12(2):405–428.

Box, G. and Jenkins, G. (1976). Time series aoalysis: Fgrecastlng bnd cootrol. *San Franciscol Holden-DaY*.

Boye, P., Ziggah, Y. Y., et al. (2020). A short-term stock exchange prediction model using box-jenkins approach. *Journal of Applied Mathematics and Physics*, 8(05):766.

Brock, W., Lakonishok, J., and LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of finance*, 47(5):1731–1764.

Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3):195–228.

Cao, J., Li, Z., and Li, J. (2019). Financial time series forecasting model based on ceemdan and lstm. *Physica A: Statistical Mechanics and its Applications*, 519:127–139.

Challa, M. L., Malepati, V., and Kolusu, S. N. R. (2020). S&p bse sensex and s&p bse it return forecasting using arima. *Financial Innovation*, 6(1):1–19.

Chordia, T., Subrahmanyam, A., and Tong, Q. (2014). Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Accounting and Economics*, 58(1):41–58.

Chow, K. V. and Denning, K. C. (1993). A simple multiple variance ratio test. *Journal of Econometrics*, 58(3):385–401.

Cootner, P. (1964). The random character of stock market prices cambridge.

Cristelli, M. (2014). Web queries can predict stock market volumes. In *Complexity in Financial Markets*, pages 151–161. Springer.

Darrat, A. F. and Zhong, M. (2000). On testing the random-walk hypothesis: a model-comparison approach. *Financial Review*, 35(3):105–124.

De Bondt, W. F. and Thaler, R. (1985). Does the stock market overreact? *The Journal of finance*, 40(3):793–805.

Deka, A. and Resatoglu, N. G. (2019). forecasting foreign exchange rate and consumer price index with arima model: The case of turkey. *International Journal of Scientific Research and Management*, 7(8):1254–1275.

Dhyani, B., Kumar, M., Verma, P., and Jain, A. (2020). Stock market forecasting technique using arima model. *International Journal of Recent Technology and Engineering*, 8(6).

Dimson, E. and Marsh, P. (1999). Murphy's law and market anomalies. *The Journal of Portfolio Management*, 25(2):53–69.

Dinardi, F. B. (2020). *Forecasting the stock market using ARIMA and ARCH/-GARCH approaches*. PhD thesis.

Dong, H., Guo, X., Reichgelt, H., and Hu, R. (2020). Predictive power of arima models in forecasting equity returns: a sliding window method. *Journal of Asset Management*, 21(6):549–566.

Engle, R. F., Hong, C.-H. T., Kane, A., and Noh, J. (1992). Arbitrage valuation of variance forecasts with simulated options. *University of California at San Diego, Economics Working Paper Series*.

Faisal, F. (2012). Forecasting bangladesh?s inflation using time series arima models. *World Review of Business Research*, 2(3):100–117.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417.

Fama, E. F. (1991). Efficient capital markets: Ii. *The journal of finance*, 46(5):1575–1617.

Fama, E. F. and Blume, M. E. (1966). Filter rules and stock-market trading. *The Journal of Business*, 39(1):226–241.

Fama, E. F., Fisher, L., Jensen, M. C., and Roll, R. (1969). Review väno96? *International Economic Review*, 10(1):1–21.

Fama, E. F. and French, K. R. (1988). Dividend yields and expected stock returns. *Journal of financial economics*, 22(1):3–25.

Fang, Y. and Xu, D. (2003). The predictability of asset returns: an approach combining technical analysis and time series forecasts. *International Journal of Forecasting*, 19(3):369–385.

Fischer, D. E. and Jordan, R. J. (2005). Security analysis and portfolio management.

Gay, R. D. et al. (2016). Effect of macroeconomic variables on stock market returns for four emerging economies: Brazil, russia, india, and china. *International Business & Economics Research Journal (IBER)*, 15(3):119–126.

Golden, J., Milewicz, J., and Herbig, P. (1994). Forecasting: Trials and tribulations. *Management Decision*.

Granger, C. W. and Pesaran, M. H. (2000). Economic and statistical measures of forecast accuracy. *Journal of Forecasting*, 19(7):537–560.

Hawaldar, I. T. and Mallikarjunappa, T. (2009). Indian stock market reaction to the quarterly earnings information. *Iqbal and T. Mallikarjunappa (2009).?Indian Stock Market Reaction to the Quarterly Earnings Information?, Indian Journal of Finance*, 3(7):43–50.

Hipel, K. W. and McLeod, A. I. (1994). *Time series modelling of water resources and environmental systems*. Elsevier.

Husain, F. (2006). Stock prices, real sector and the causal analysis: The case of pakistan. *Journal of Management and Social Sciences*, 2(2).

Husain, F. and Mahmood, T. (2001). The stock market and the economy in pakistan. *The Pakistan Development Review*, pages 107–114.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91.

Johnson, R. and Soenen, L. (2003). Economic integration and stock market co-movement in the americas. *Journal of Multinational Financial Management*, 13(1):85–100.

Kendall, M. G. and Hill, A. B. (1953). The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A (General)*, 116(1):11–34.

Khan, S. and Alghulaiakh, H. (2020). Arima model for accurate time series stocks forecasting. *(IJACSA) International Journal of Advanced Computer Science and Applications*, 11(7).

Khandewal, D., Ballal, N. V., and Saraswathi, M. V. (2015). Comparative evaluation of accuracy of 2 electronic apex locators with conventional radiography: an ex vivo study. *Journal of Endodontics*, 41(2):201–204.

Khashei, M., Bijari, M., and Ardali, G. A. R. (2009). Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (anns). *Neurocomputing*, 72(4-6):956–967.

Khashei, M., Bijari, M., and Ardali, G. A. R. (2012). Hybridization of autoregressive integrated moving average (arima) with probabilistic neural networks (pnns). *Computers & Industrial Engineering*, 63(1):37–45.

Kim, J. H. (2006). Wild bootstrapping variance ratio tests. *Economics letters*, 92(1):38–43.

Kim, J. H., Shamsuddin, A., and Lim, K.-P. (2011). Stock return predictability and the adaptive markets hypothesis: Evidence from century-long us data. *Journal of Empirical Finance*, 18(5):868–879.

Konak, F. and Şeker, Y. (2014). The efficiency of developed markets: Empirical evidence from ftse 100. *Journal of Advanced Management Science Vol*, 2(1).

Kyungjoo, L., Sehwan, Y., and John, J. (2007). Neural network model vs. sarima model in forecasting korean stock price index. *Issues in Information Systems*, 8(2):372–378.

Latha, C. M., Nageswararao, K. S., and Venkataramanaiah, M. (2018). Forecasting time series stock returns using arima: Evidence from s & p bse sensex. *International Journal of Pure and Applied Mathematics*, 118(24):1–21.

Lee, C.-M. and Ko, C.-N. (2011). Short-term load forecasting using lifting scheme and arima models. *Expert Systems with Applications*, 38(5):5902–5911.

Levine, R. (1997). Financial development and economic growth: views and agenda. *Journal of economic literature*, 35(2):688–726.

Lo, A. W. and MacKinlay, A. C. (1989). The size and power of the variance ratio test in finite samples: A monte carlo investigation. *Journal of econometrics*, 40(2):203–238.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of economic perspectives*, 17(1):59–82.

Malkiel, B. G. (2005). Reflections on the efficient market hypothesis: 30 years later. *Financial review*, 40(1):1–9.

Mallikarjuna, M. and Rao, R. P. (2019). Evaluation of forecasting methods from selected stock market returns. *Financial Innovation*, 5(1):1–16.

Mandelbrot, B. B. (1989). Louis bachelier. In *Finance*, pages 86–88. Springer.

McLean, R. D. and Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, 71(1):5–32.

Meher, B. K., Hawaldar, I. T., Spulbar, C. M., and Birau, F. R. (2021). Forecasting stock market prices using mixed arima model: A case study of indian pharmaceutical companies. *Investment Management and Financial Innovations*, 18(1):42–54.

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.

Miranda Tabak, B. (2003). The random walk hypothesis and the behaviour of foreign capital portfolio flows: the brazilian stock market case. *Applied Financial Economics*, 13(5):369–378.

Mondal, P., Shit, L., and Goswami, S. (2014). Study of effectiveness of time series modeling (arima) in forecasting stock prices. *International Journal of Computer Science, Engineering and Applications*, 4(2):13.

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management science*, 60(7):1772–1791.

Nochai, R. and Nochai, T. (2006). Arima model for forecasting oil palm price. In *Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and applications*, pages 13–15.

Ojo, J. and Olatayo, T. (2009). On the estimation and performance of subset autoregressive integrated moving average models. *European Journal of Scientific Research*, 28(2):287–293.

Owido, P. K., Onyuma, S., and Owuor, G. (2013). A garch approach to measuring efficiency: a case study of nairobi securities exchange. *Research Journal of Finance and Accounting*, 4(4):1–16.

Pai, P.-F. and Lin, C.-S. (2005). A hybrid arima and support vector machines model in stock price forecasting. *Omega*, 33(6):497–505.

Pandey, V. S. and Bajpai, A. (2019). Predictive efficiency of arima and ann models: a case analysis of nifty fifty in indian stock market. *International Journal of Applied Engineering Research. ISSN*, pages 0973–4562.

Pankratz, A. (2009). *Forecasting with univariate Box-Jenkins models: Concepts and cases*. John Wiley & Sons.

Pesaran, M. H. and Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *The Journal of Finance*, 50(4):1201–1228.

Pettenuzzo, D., Timmermann, A., and Valkanov, R. (2014). Forecasting stock returns under economic constraints. *Journal of Financial Economics*, 114(3):517–553.

Phan, D. H. B., Sharma, S. S., and Narayan, P. K. (2015). Stock return forecasting: Some new evidence. *International Review of Financial Analysis*, 40:38–51.

Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of financial economics*, 22(1):27–59.

Radikoko, I. (2014). Testing the random walk (rw) behaviour of botswana?s equity returns. *Journal of Business Theory and Practice*, 2(1):84–99.

Raicharoen, T., Lursinsap, C., and Sanguanbhokai, P. (2003). Application of critical support vector machine to time series prediction. In *Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS'03.*, volume 5, pages V–V. IEEE.

Rajan, R. G. and Zingales, L. (1998). Power in a theory of the firm. *The Quarterly Journal of Economics*, 113(2):387–432.

Rapach, D. E., Ringgenberg, M. C., and Zhou, G. (2016). Short interest and aggregate stock returns. *Journal of Financial Economics*, 121(1):46–65.

Reichenstein, W. and Rich, S. P. (1994). Predicting long-horizon stock returns: Evidence and implications. *Financial Analysts Journal*, 50(1):73–76.

Reilly, F. K. and Brown, K. C. (2011). *Investment Analysis and Portfolio Management (Text Only)*. Cengage Learning.

Richardson, M. and Smith, T. (1991). Tests of financial models in the presence of overlapping observations. *The Review of Financial Studies*, 4(2):227–254.

Roberts, H. (1967). Statistical versus clinical prediction of the stock market. *Unpublished manuscript*.

Roberts, H. V. (1959). Stock-market" patterns" and financial analysis: methodological suggestions. *The Journal of Finance*, 14(1):1–10.

Rousseau, P. L. and Wachtel, P. (2000). Equity markets and growth: Cross-country evidence on timing and outcomes, 1980–1995. *Journal of Banking & Finance*, 24(12):1933–1957.

Sabur, S. and Haque, M. Z. (1992). Resource use efficiency and returns from some selected winter crops in bangladesh. *Economic Affairs (Calcutta)*, 37(3):158.

Schmitz, A. and Watts, D. G. (1970). Forecasting wheat yields: an application of parametric time series modeling. *American Journal of Agricultural Economics*, 52(2):247–254.

Sensoy, A., Fabozzi, F. J., and Eraslan, V. (2017). Predictability dynamics of emerging sovereign cds markets. *Economics Letters*, 161:5–9.

Shumway, R. H. and Stoffer, D. S. (2011). Time series and its applications: With r examples. *Springer Science*, pages 1–202.

Siami-Namini, S., Tavakoli, N., and Namin, A. S. (2019). The performance of lstm and bilstm in forecasting time series. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 3285–3292. IEEE.

Siva Kiran Guptha, K. and Prabhakar Rao, R. (2018). The causal relationship between financial development and economic growth: an experience with brics economies. *Journal of Social and Economic Development*, 20(2):308–326.

Sterba, J. and Hilovska, K. (2010). The implementation of hybrid arima neural network prediction model for aggregate water consumption prediction. *Aplimat?Journal of Applied Mathematics*, 3(3):123–131.

Suits, D. B. (1962). Forecasting and analysis with an econometric model. *The American Economic Review*, 52(1):104–132.

Sullivan, R., Timmermann, A., and White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *The journal of Finance*, 54(5):1647–1691.

Sultana, K., Rahim, A., Moin, N., Aman, S., and Ghauri, S. P. (2013). Forecasting inflation and economic growth of pakistan by using two time series methods. *International Journal of Business and Economics Research*, 2(6):174–178.

Summers, L. H. (1986). Does the stock market rationally reflect fundamental values? *The Journal of Finance*, 41(3):591–601.

Tong, H. (2012). *Threshold models in non-linear time series analysis*, volume 21. Springer Science & Business Media.

Wadi, S., Almasarweh, M., Alsaraireh, A. A., and Aqaba, J. (2018). Predicting closed price time series data using arima model. *Modern Applied Science*, 12(11):181–5.

Wahyudi, S. T. (2016). The arima model for the indonesia stock price. *International Journal of Economics & Management*, 11:74–88.

Wang, Q., Li, S., Li, R., and Ma, M. (2018). Forecasting us shale gas monthly production using a hybrid arima and metabolic nonlinear grey model. *Energy*, 160:378–387.

Welch, I. and Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508.

Wen, F., Xu, L., Ouyang, G., and Kou, G. (2019). Retail investor attention and stock price crash risk: evidence from china. *International Review of Financial Analysis*, 65:101376.

Zhang, G. P. (2003). Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50:159–175.

Zhang, G. P. (2007). A neural network ensemble method with jittered training data for time series forecasting. *Information Sciences*, 177(23):5329–5346.

Zhang, Y., Ma, F., Shi, B., and Huang, D. (2018). Forecasting the prices of crude oil: An iterated combination approach. *Energy Economics*, 70:472–483.

Zhang, Y., Ma, F., and Zhu, B. (2019a). Intraday momentum and stock return predictability: Evidence from china. *Economic Modelling*, 76:319–329.

Zhang, Y., Zeng, Q., Ma, F., and Shi, B. (2019b). Forecasting stock returns: Do less powerful predictors help? *Economic Modelling*, 78:32–39.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics*, 14(1):45–52.

Zhu, X. and Zhu, J. (2013). Predicting stock returns: A regime-switching combination approach and economic links. *Journal of Banking & Finance*, 37(11):4120–4133.

Zotteri, G., Kalchschmidt, M., and Caniato, F. (2005). The impact of aggregation level on forecasting performance. *International Journal of Production Economics*, 93:479–491.