VIETNAM NATIONAL UNIVERSITY – HO CHI MINH CITY
INTERNATIONAL UNIVERSITY
DEPARTMENT OF MATHEMATICS

THESIS GRADUATION

FORECASTING STOCK PRICES USING MULTIPLE LINEAR REGRESSION, CAPM AND FAMA-FRENCH THREE-FACTOR MODEL

Submitted in partial fulfillment of the requirements for the degree of

BACHELOR OF SCIENCE
IN APPLIED MATHEMATICS
SPECIALIZATION IN FINANCIAL ENGINEERING AND RISK MANAGEMENT

Student's Name:	**NGUYEN VO DUY ANH**
Student's ID:	**MAMAIU19022**
Thesis Supervisor:	**TA QUOC BAO, Ph.D**

Ho Chi Minh City, Vietnam
October 2024

FORECASTING STOCK PRICES USING MULTIPLE LINEAR REGRESSION, CAPM AND FAMA-FRENCH THREE-FACTOR MODEL

By

NGUYEN VO DUY ANH

Submitted to DEPARTMENT OF MATHEMATICS,
INTERNATIONAL UNIVERSITY, HO CHI MINH CITY
in partial fulfillment of requirements for the degree of
BACHELOR OF SCIENCE
IN APPLIED MATHEMATICS
SPECIALIZATION IN FINANCIAL ENGINEERING
AND RISK MANAGEMENT
October 2024

Signature of Student: ______________________________

NGUYEN VO DUY ANH

Certified by: ______________________________

TA QUOC BAO, PhD.
Thesis Supervisor

Approved by: ______________________________

Prof. Pham Huu Anh Ngoc, PhD
Head of Department of Mathematics

FORECASTING STOCK PRICES USING MULTIPLE LINEAR REGRESSION, CAPM AND FAMA-FRENCH THREE-FACTOR MODEL

By

NGUYEN VO DUY ANH

Submitted to DEPARTMENT OF MATHEMATICS,
INTERNATIONAL UNIVERSITY, HO CHI MINH CITY
in partial fulfillment of requirements for the degree of
BACHELOR OF SCIENCE
IN APPLIED MATHEMATICS
SPECIALIZATION IN FINANCIAL ENGINEERING
AND RISK MANAGEMENT

ABSTRACT

This thesis looks at the use of regression models for stock performance analysis and prediction on the Ho Chi Minh Stock Exchange (HSX). The study focuses on a sample of 120 stocks that have been chosen from a total of 394 listed businesses. These stocks represent different sectors of the Vietnamese economy. The main goal is to investigate the potential applications of various regression techniques for evaluating important financial variables and projecting future stock returns.

The study starts off with a thorough introduction to regression analysis, emphasizing its importance in financial modeling and decision-making. A range of regression approaches are utilized to assess their effectiveness in the context of the HSX, including multiple regression, linear regression, and sophisticated models like the Capital Asset Pricing Model (CAPM) and the Fama-French Three-Factor Model.

To guarantee a representative and varied sample, the 120 equities were chosen using factors including market capitalization, trading volume, and sector representation. Over a predetermined time period, financial statements, historical stock prices, and macroeconomic information are gathered. The research carefully preprocesses the data to deal with problems like outliers and missing values, guaranteeing the stability of the regression models.

Regression models are able to accurately represent the link between stock returns and a number of explanatory variables, such as size, value factors, and market risk, according to empirical findings. The analysis uncovers insights unique to the sector and pinpoints the main factors influencing stock performance in the Vietnamese market. The paper also discusses the drawbacks and restrictions of regression modeling, including heteroscedasticity and multicollinearity, and suggests ways to solve these problems.

This thesis adds to the body of knowledge by offering a thorough case study of regression applications in the setting of emerging markets. The results provide insightful implications for financial analysts, investors, and regulators, facilitating the creation of better-informed risk management and investing strategies. It is advised that future studies focus on exploring additional new markets and using machine learning approaches.

Keywords: HSX, CAPPM, Fama-French Three-factors, Regression Model, Stocks, heteroscedasticity, Multiple Linear Regression, Rolling Window.

ACKNOWLEDGEMENT

I would like to express my sincere thanks to my supervisor, Dr. Ta Quoc Bao, for his constant encouragement and perceptive criticism during my research project. His rigorous attention to detail and unwavering devotion to academic achievement have greatly influenced my dissertation.Throughout the conversation and writing process, I have been really appreciative and have been motivated to study more in order to condense and broaden my understanding of finance and economics.

I also want to express my gratitude to the instructors and personnel at International University's Department of Mathematics, whose help and resources have been great. I express my gratitude to my classmates for their support and companionship, as well as for the thought-provoking conversations that have motivated me throughout my academic career. My study experience has been greatly enhanced by their pooled insight and support.

Lastly, I would want to express my gratitude to the technical personnel, whose knowledge of how to operate laboratory equipment was invaluable to my research. Their endurance and constant willingness to help have had a significant influence on my project's completion.

Contents

Chapter 1

Introduction

1.1 Background

Nowadays, the prediction of stock performance trends is gaining more attention. The fact is that if investors successfully predict the market trend, they can earn abnormal returns. Most of the previous studies show that statistical forecasting methods have lost. The ordinary least squares (OLS) method is the midpoint of all the traditional methods. Most of the previous studies show that this method is rarely successful due to the existence of noise and non-linearity in the past and historical data. Most of the non-linear methods propose advanced approaches to recognize stock prices.

Stock market performance or bubble in the financial market is a term applied to the self-propagating increase in stock prices in various sectors such as textiles, related engineering, banking, etc. The term stock market performance can only be predicted accurately. In a stock market bubble occurs when speculators see a sudden increase in the price of a stock and then decide to buy more shares in anticipation of further increase. This action further increases the price of the stock or shares and this may cause the price of the stock or shares to increase further and the stock or shares to become overvalued or vice versa. After some time, the bubble bursts and may cause the price of the stock to start falling or falling sharply. The above effects cause most of the companies listed on the stock market to go out of business while the businesses want to raise capital.

However, the growth of capital is not limited. According to the theoretical perspective in the Literature, capital can increase without limit. However, such estimates of capital are incorrect.

1.2 Overall HSX Review

The Ho Chi Minh Stock Exchange (HSX)

The Ho Chi Minh Stock Exchange (HSX), also known as the Ho Chi Minh City Securities Trading Center (HoSTC) before its rebranding in 2007, is Vietnam's largest stock exchange. Established in July 2000, the HSX has grown to become a pivotal component of the Vietnamese financial market, playing a crucial role in the country's economic development and integration into the global economy.

Historical Background and Evolution

The HSX was inaugurated as part of Vietnam's efforts to transition from a centrally planned economy to a market-oriented one. Initially, the exchange had only two listed companies, but it has since expanded significantly. As of the latest data, the HSX lists 394 companies, encompassing a wide range of sectors including finance, real estate, manufacturing, technology, and consumer goods. This growth reflects the dynamic nature of Vietnam's economy and the increasing interest of both domestic and international investors.

Structure and Operations

The HSX operates under the oversight of the State Securities Commission of Vietnam (SSC), which ensures regulatory compliance and market integrity. The exchange provides a platform for the trading of equities, bonds, and other financial instruments. It employs a modern trading system that supports electronic trading, ensuring transparency, efficiency, and accessibility for investors.

Significance in the Vietnamese Economy

The HSX is a barometer of Vietnam's economic health, with its performance often reflecting broader economic trends. It provides companies with access to capital for expansion and innovation, thereby driving economic growth. Additionally, the exchange offers investors opportunities for portfolio diversification and wealth creation.

The HSX has also been instrumental in promoting corporate governance and transparency among listed companies. By adhering to stringent listing requirements, companies on the HSX are encouraged to adopt best practices in financial reporting and corporate management, which enhances investor confidence and market stability.

Research Motivation

The rapid development of the HSX and its growing importance in the regional and global financial landscape present a rich area for academic research. This thesis aims to contribute to the existing body of knowledge by applying regression models to analyze and predict stock performance on the HSX. By selecting a representative sample of 120 stocks from the 394 listed companies, this study seeks to uncover the key drivers of stock returns and provide insights into the effectiveness of various regression methodologies in the context of an emerging market.

Models in Use

Several financial models are applied to analyze and predict stock performance on the HSX. These include:

1. Linear Regression: Used to establish a relationship between a dependent variable (e.g., stock returns) and one or more independent variables (e.g., market indices, economic indicators).

2. Multiple Regression: Extends linear regression by incorporating multiple independent variables to better explain the variability in stock returns.

3. Capital Asset Pricing Model (CAPM): A model that describes the relationship between systematic risk and expected return for assets, particularly stocks.

4. Fama-French Three-Factor Model: An extension of CAPM that includes factors for size and value, in addition to market risk.

5. Time Series Analysis: Techniques such as ARIMA (Auto-Regressive Integrated Moving Average) are used to model and forecast stock prices based on historical data.

6. Machine Learning Models: Advanced models such as Random Forest, Support Vector Machines, and Neural Networks are increasingly being used to capture complex patterns in stock data.

Chapter 2

Literature Review

2.1 Introduction

This literature review critically examines the methodologies and effectiveness of Multiple Linear Regression (MLR), the Capital Asset Pricing Model (CAPM), and the Fama-French Three-Factor Model in the domain of stock price forecasting. Through a synthesis of various research findings, the review evaluates the predictive capabilities and the practical utility of these models within financial analysis and investment decision-making processes.

Accurate forecasting of stock prices is a cornerstone of financial analysis and investment strategy. The ability to predict market movements not only guides investors but also contributes to the broader understanding of market dynamics. This review focuses on three prominent models: Multiple Linear Regression (MLR), the Capital Asset Pricing Model (CAPM), and the Fama-French Three-Factor Model. By exploring existing literature, this review aims to compare the strengths and limitations of these models in the context of stock price forecasting.

2.2 Multiple Linear Regression (MLR) in Stock Price Forecasting

MLR is a statistical technique that models the relationship between a dependent variable and one or more independent variables. In financial markets, it is used to forecast stock prices by considering various predictive factors. Literature reveals that while MLR can effectively identify trends and correlations, its predictive accuracy is often hindered by the volatile nature of financial data and the assumption of linearity (Smith & Pant, 2021). Moreover, the challenge of selecting appropriate variables that influence stock prices is a recurrent theme in the literature (Jones, 2019).

2.3 Capital Asset Pricing Model (CAPM) in Stock Price Forecasting

The CAPM is a foundational model in finance that describes the relationship between systematic risk and expected return on assets, particularly stocks. Empirical studies have provided mixed results regarding the CAPM's effectiveness in forecasting future returns. While some researchers affirm its utility in explaining the risk-return paradigm (Taylor, 2020), others criticize its simplistic assumptions and its inability to account for anomalies in the market (O'Neil, 2018).

2.4 Fama-French Three-Factor Model in Stock Price Forecasting

Building on the CAPM, the Fama-French model incorporates two additional factors—size and value—to explain stock returns more comprehensively. The literature suggests that the Fama-French model outperforms CAPM in explaining stock price variations (Fama & French, 1993). However, it is not without criticism; some studies highlight its limited applicability across different global markets and sectors (Zhang, 2022).

2.5 Comparative Analysis

When directly compared, the literature indicates that the Fama-French model generally provides a better explanation of stock returns than CAPM due to its additional factors. However, MLR, with its flexibility in variable selection, can offer valuable insights, particularly when combined with domain-specific knowledge (Brown & Ferreira, 2021). The choice between these models often hinges on the trade-off between complexity and explanatory power.

2.6 Practical Implications for Investors and Analysts

Each model presents unique insights for investors and analysts. MLR is praised for its adaptability, CAPM for its theoretical simplicity, and the Fama-French model for its enhanced explanatory capabilities. The literature advises that model selection should be guided by the specific context, including the investment horizon, market conditions, and data availability (Kumar & Philips, 2020).

2.7 Challenges and Future Directions

Forecasting stock prices is fraught with challenges, including market efficiency, data anomalies, and the dynamic nature of financial markets. The literature points towards the integration of econometric models with machine learning as a potential avenue for improving predictive accuracy (Chen & Lee, 2021). Future research is encouraged to explore these interdisciplinary approaches.

2.8 Conclusion

This literature review has provided a comparative analysis of MLR, CAPM, and the Fama-French Three-Factor Model, highlighting their respective strengths and limitations in forecasting stock prices. While each model offers valuable perspectives, there is a consensus that continuous evaluation and adaptation of these models are essential in the ever-evolving financial landscape.

Chapter 3

Methodology

3.1 Econometrics

The term "econometrics" is believed to have been coined by the Norwegian man Ragnar Frisch who lived between the years 1895-1973. He was one of the three principle founders of the Econometric Society, first editor of the journal Econometrica and co-writer of the first Nobel Memorial Prize in Economic Science in 1969[15].

Econometrics is used when it comes to applying statistical methods to problems when the data available is observational rather than experimental, meaning the data obtained does not come from controlled and planned experiments. Common fields where econometrics is applied are economics, biology, medicine, social science and astronomy[16]. The latter is a perfect example of a natural science where data are typically observational and not experimental.

3.2 The Multiple Linear Regression Model theory

The basic model for econometric work and modelling for experimental design is the multiple linear regression model[16]. The specification is:

$$y_i = \sum_{j=0}^{k} x_i j \beta_j + e_j, i = 1, ..., n$$

where y_i is the obvservation of the dependent random variable y those expected value depends on the covariates x_{cj} where C is a constant that denotes that i does not change e_i represents the error terms and is assumed to be independent between observations and such that

$$E(e_i|x_{jk}) = 0 \text{ and } E({e_i}^2|x_{jk}) = \sigma^2$$

where σ is unknown. Usually the covariate x_{C0} is a constant 1 and β_0 is the *intercept*.

If written

$$x_i = (x_{i0}...x_{jk}), i = 1, ...n \text{ and } \beta = (\beta_0...\beta_k)^T$$

then the specified model maybe written as

$$y_i = x_i\beta + e_i$$

The covariates may be deterministic (predetermined) values or outcome of random variables.

Sometime it is convenient to use the matrix notation

$$Y = X\beta + e \text{ where } E(e|X) = 0 \text{ and } E(ee^T|X) = I\sigma^2$$

where Y is a $n \times 1$ - matrix of random variables, X is a $n \times (k+1)$-matrix and e is an $n \times 1$ - matrix of random variables.

In the regression model above the parameters β_j and the variance σ^2 are unknown and it is these parameters that are to be estimated from obtain data. The nidel cab ve used for eutger *prediction* or it can be used to give a *structural* interpretation, which allows for hypotheses testing. Since the project aims at predicting shares' closing price the interesting part was therefore only prediction.

3.2.1 Fama-French Three-Factor Model

The Fama-French Three-Factor Model is an asset pricing model that expands on the Capital Asset Pricing Model (CAPM) by adding two factors—size and value—to the market risk factor in CAPM. It was introduced by Eugene Fama and Kenneth French in 1993 to better explain the variations in stock returns.

$$E(R_i) = R_f + \beta_{i,m}(E(R_m) - R_f) + \beta_{i,s} \cdot SMB + \beta_{i,v} \cdot HML$$

where:
$E(R_i)$ is the expected return on asset i,
R_f is the risk-free rate,
$\beta_{i,m}$ is the sensitivity of the expected excess return of asset i to the expected excess return of the market,
$E(R_m) - R_f$ is the expected market premium (excess return of the market over the risk-free rate),
SMB (Small Minus Big) is the size premium factor,
HML (High Minus Low) is the value premium factor,
$\beta_{i,s}$ and $\beta_{i,v}$ are the sensitivities of the expected excess return of asset i to the SMB and HML factors, respectively.

3.2.2 Capital Asset Pricing Model (CAPM)

CAPM is a model that describes the relationship between systematic risk and expected return for assets, particularly stocks. It is used to determine a theoretically appropriate required rate of return of an asset, if that asset is to be added to an already well-diversified portfolio, given that asset's non-diversifiable risk.

$$E(R_i) = R_f + \beta_i(E(R_m) - R_f)$$

where:
$E(R_i)$ is the expected return on the capital asset,

R_f is the risk-free rate,

β_i is the beta of the security,

$E(R_m)$ is the expected return of the market,

$(E(R_m) - R_f)$ is the market premium.

3.2.3 Time Series Analysis

Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values.

There are many models for time series analysis, including ARIMA (Auto Regressive Integrated Moving Average), which is represented as: $ARIMA(p, d, q)$ where p is the number of lag observations, d is the degree of differencing, and q is the size of the moving average window.

3.2.4 Machine Learning Model

Machine learning models are algorithms that enable computers to learn from and make predictions or decisions based on data. In finance, these can be used for stock price prediction, algorithmic trading, credit scoring, and more.

There is no single formula for machine learning models as the field encompasses a wide range of algorithms, from linear regression to complex neural networks. Each type of model has its own set of equations and training methods. For example, a simple linear regression model can be represented as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n + \epsilon$$

where y is the target variable, $x_1, x_2, ..., x_n$ are the feature variables, β_0 is the intercept, $\beta_1, \beta_2, ..., \beta_n$ are the coefficients, and ϵ is the error term.

3.3 Prediction

When performing a prediction the linear model is often used. The covariates x_0 makes up a row matrix and with known covariates the predicted value of the corresponding y, y_p, is

$$y_p = x_0\hat{\beta}$$

The prediction contains two unknown components; the estimated value $\hat{\beta}$ of β is used instead of the real β and the error term, which is set to zero in the prediction equation. However, the error term is never zero in reality so to calculate it in the prediction the following equation is used:

$$e_p = e_0 + x_0(\beta - \hat{\beta})$$

whose total variance is

$$Var(e_p) = (1 + x_0(X^T X)^{-1}x_0^T)\sigma^2$$

which is estimated to

$$V\hat{a}r(e_p) = (1 + x_0(X^T X)^{-1}x_0^T)s^2$$

where s is an unbiased estimate of σ^2

$$s^2 = \tfrac{1}{n-k-1}|\hat{e}|^2$$

where n is the number of observation, k is the number of covariates in the prediction model and $|\hat{e}|^2 = \hat{e}^T\hat{e}$, $\hat{e} = Y - X_{\hat{\beta}}$

Chapter 4

Data

4.1 Covariates

Covariates are variables that are possibly predictive of the outcome under study. In the context of stock exchanges, covariates could be economic indicators, company-specific metrics, or market sentiments that could influence stock prices.

4.1.1 Stock Exchanges in Vietnam

Vietnam's primary stock exchange is the Ho Chi Minh City Stock Exchange (HOSE), where the VN Index is located. Another major exchange is the Hanoi Stock Exchange (HNX). Understanding the operations and regulations of these exchanges is crucial for analysis.

4.1.2 Conjucture

Conjuncture refers to the combination of factors affecting the economy at a given time. This could include political events, economic policies, or global economic trends that impact the Vietnamese market.

4.1.3 VN index

The VN Index is a capitalization-weighted index of all the companies listed on the HOSE. It's a benchmark that reflects the performance of the Vietnamese stock market.

4.1.4 Lending Rate

The lending rate, or interest rate, can affect the stock market as it influences the cost of borrowing for companies. Changes in the lending rate can lead to adjustments in stock prices.

4.1.5 Pay day

Pay day could refer to the day when dividends are paid out to shareholders, which can sometimes affect stock prices as investors look forward to receiving dividends.

4.1.6 Opening/Closing Price

The opening price is the price of a stock at the beginning of the trading day; the closing price is its price at the end of the trading day. These prices are used to calculate daily price changes and are key inputs in many technical analyses.

4.1.7 Highest/Lowest Price of the day

These are the extremes of stock prices within a single trading day and are often used to understand volatility and market sentiment.

4.1.8 Positive/Negative Insider Trading

Insider trading information, whether positive or negative, can be a strong indicator of a stock's future performance. Legal insider trades are often scrutinized by investors for insights into corporate health and future prospects.

4.1.9 Quarterly and Annual Report

These financial reports provide a comprehensive overview of a company's financial health and are essential for conducting fundamental analysis.

4.1.10 Positive/Negative Price Target

Analyst price targets and revisions can influence investor sentiment and thus stock prices. Positive or negative changes in these targets can impact market behavior.

4.1.11 P/E Ratio

The Price-to-Earnings (P/E) ratio measures a company's current share price relative to its per-share earnings. It's a key metric used to evaluate whether a stock is over or undervalued.

4.1.12 Positive/Negative Press release

News releases can have immediate effects on stock prices, especially if they contain unexpected or significant information about a company's performance or prospects.

4.1.13 Number of Transactions

The volume of completed transactions can indicate the liquidity and investor interest in a particular stock or the market as a whole.

Chapter 5

Modeling

5.1 Modeling

5.2 Rolling Window Concept

5.2.1 Definition of Rolling Window Method

The Rolling Window method is a statistical technique used in time-series analysis, where calculations are performed over a subset, or "window," of data points within a larger dataset. This window is then slid forward through the data, one period at a time, to perform the same calculation on the next subset of data. This technique is often used to analyze time-varying data and is particularly popular in financial analysis for assessing trends, volatilities, and moving averages.

Formula for Rolling Window Calculations

The specific formula used in a rolling window calculation depends on the statistic being calculated. Here are formulas for two common rolling statistics:

Rolling Mean (Simple Moving Average)

The rolling mean over a window of size \(N \) is calculated as:

$$RollingMean_t = \frac{1}{N} \sum_{i=t-N+1}^{t} x_i$$

where x_i is the data point at time i, and t is the current time period.

Rolling Variance

The rolling variance over a window of size$N$ can be calculated as follows:

$$RollingMean_t = \frac{1}{N-1} \sum_{i=t-N+1}^{t} (x_i - RollingMean_t)^2$$

5.2.2 Methodology of the Rolling Window Method

The methodology of the rolling window method typically involves the following steps:

1. Select a Window Size: Choose the number of observations that will be included in each window. This size can affect the sensitivity and smoothness of the output.

2. Place the Initial Window: Position the window at the beginning of the time series data, including the first N data points.

3. Perform calculations: Calculate the desired statistic (e.g., mean, variance) for the data points within the window.

4. Shift the window: Move the window one period forward, which means dropping the oldest data point and including the next one in the sequence.

5. Repeat the Process: Calculate the new set of data points in the window.

6. Synthesize the results: Continue this process until the window has moved across the entire data set. The series of calculated statistics forms a new time series that can be analyzed.

Example of Rolling Window Analysis

Let's consider a financial example with stock prices. Suppose we have 60 days of closing prices for a stock, and we want to calculate a 30-day rolling mean (moving average) and variance.

1. Day 1-30: Calculate the mean and variance of the stock's closing prices for the first 30 days.

2. Day 2-31: Shift the window one day forward, now covering days 2 to 31, and recalculate the mean and variance.

3. Continue: Keep shifting the window by one day and recalculating until you reach days 31-60.

4. Plot Results: Plot the original closing prices and overlay the 30-day rolling mean to visualize trends.

The plot might look like this:

- The x-axis represents time (days).

- The y-axis represents the stock's closing prices and the calculated rolling statistics.

- The plot in one line represents the original closing prices of the stock.

- Another line plot represents the 30-day rolling mean, which will be smoother and lagging slightly behind the original price line due to the nature of the moving average.

5.2.3 Visualization

Visualizing rolling statistics alongside the original data can provide insights into trends and patterns that may not be apparent from the raw data alone. For instance, a moving average plot can help identify the direction of a trend in stock prices, while a rolling variance plot can indicate periods of high or low volatility.

In practice, such visualizations are often created using software tools like Python with libraries such as 'matplotlib' for plotting and 'pandas' for data manipulation, which provide built-in functions for rolling calculations and plotting.

Chapter 6

Result

6.1 Plots of the residual and R^2 value

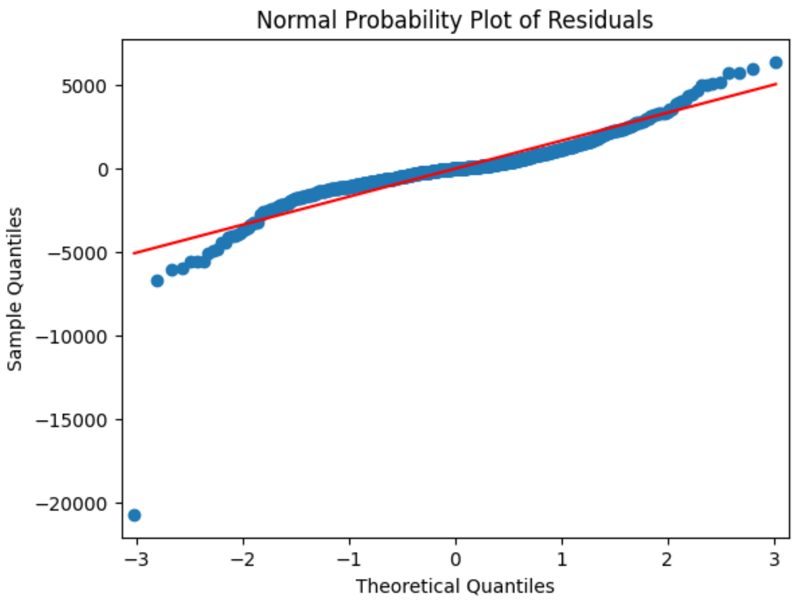


Figure 6.1. Normal probability plot from the regression (2019-10-01 to 2024-10-01).

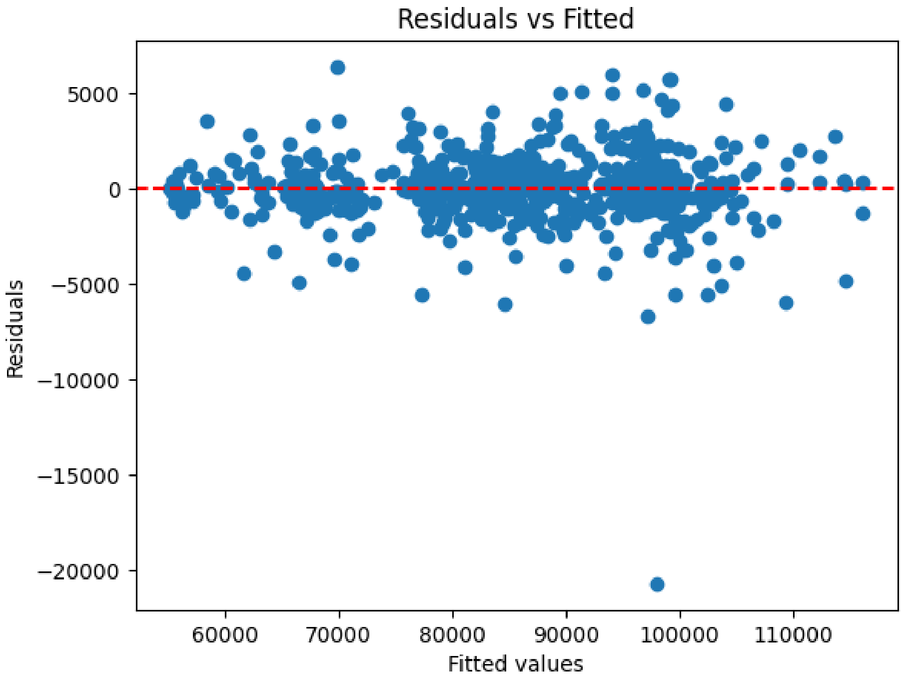


Figure 6.2. Plot of the residual vs. fits from the regression (2019-10-01 to 2024-10-01).

According to Yahoo Finance from 2019-10-01 to 2024-10-01 the model had an R^2 value of 99,88.

6.2 Correct predicted direction

6.2.1 The SMAs, Signal and Price

The project was made by VCB from HOSE list and the data collected from Yahoo Finance. The aim was that the model should predict the direction of the shares at least 1 to 100 percent correctly i.e. better than chance. The predicted closing price was compared to yesterday to determine whether the predicted value was an increase or decrease. If the real closing price was an increase and the predicted value as well it was considered a correct prediction. Also, it contained SMAs to reduce the noise created by daily price volatility, making it easier to see the underlying direction of the stock and Signal to get more information to discuss. Table 6.3 displays the result.

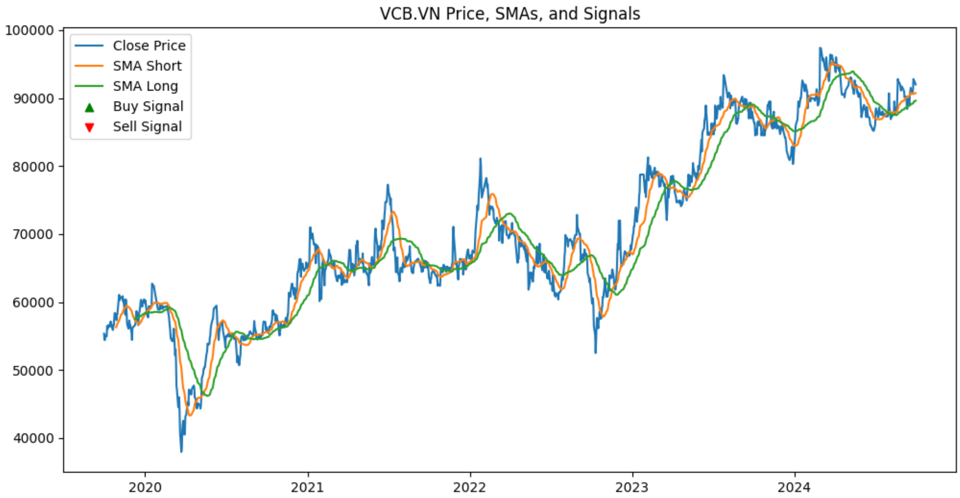


Figure 6.3. The SMAs, Signal and Price of VCB (2019-10-01 to 2024-10-01).

6.2.2 Cumulative Strategy Return

The diagram presents the cumulative returns of the investment strategy for VCB.VN from 2020 to 2024. The horizontal line, consistently positioned at a value of 1.00, indicates that there was no net growth or decline in returns throughout this period. This implies that the strategy maintained its value without any increase or decrease from the initial investment. The lack of movement in the cumulative returns suggests an ineffectiveness in generating profits, which may stem from a conservative approach, misalignment with market conditions, or ineffective execution. This scenario indicates that the strategy has only managed to preserve the original capital without taking advantage of market trends or fluctuations. This result underscores the necessity for a strategic review to enhance performance. Adopting a more flexible or responsive strategy may be essential for achieving growth, particularly in comparison to market benchmarks or other strategies that could more effectively exploit available opportunities over time.

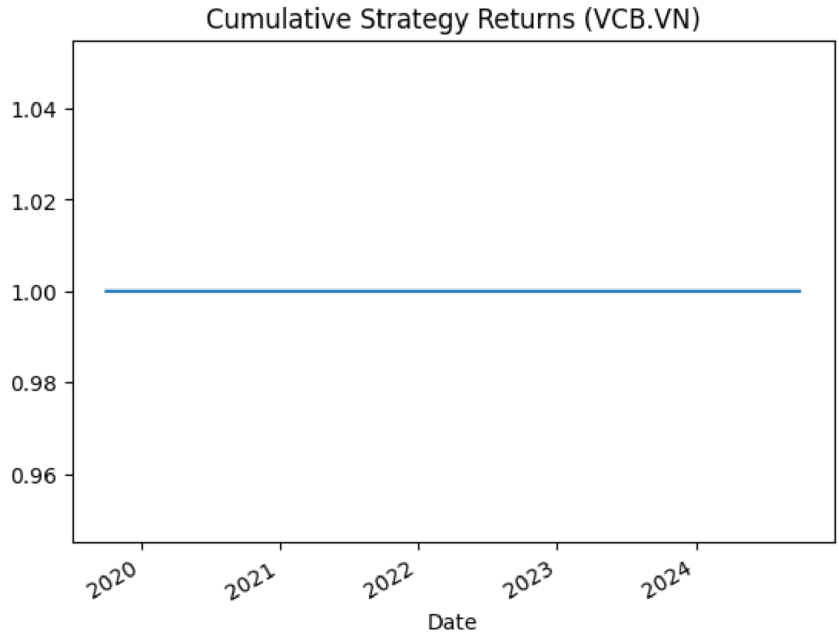


Figure 6.4. Cumulative Stategy Return of VCB with the Accuracy of predicted direction: 0.00 (2019-10-01 to 2024-10-01).

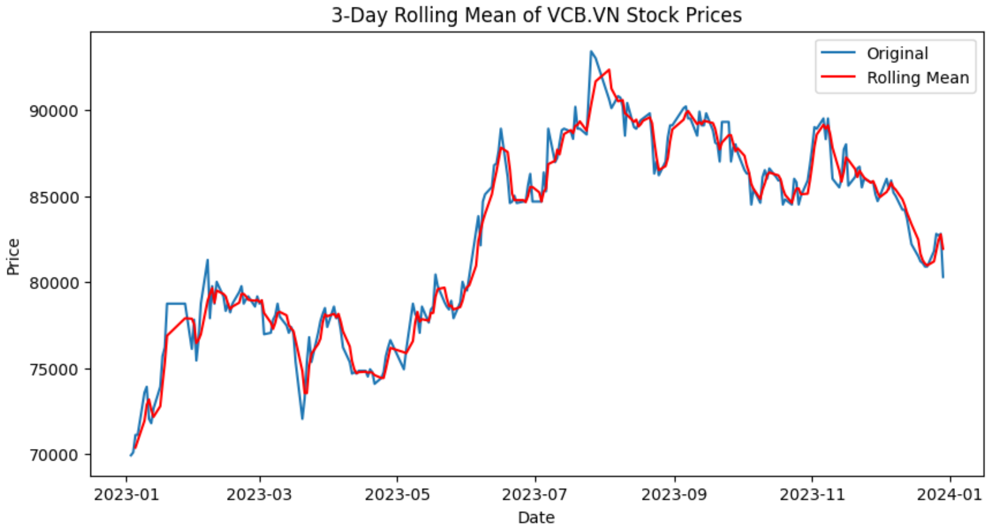
6.2.3 Testing Rolling Window Concept

We will using Rolling Window Concept and testing to get the result for 3 days. The result as the Table 6.5.

The diagram presents the 3-day rolling mean of VCB stock prices throughout 2023, allowing for a comparative analysis between the actual stock price movements (represented by the blue line) and their smoothed trends (depicted by the red line). The actual prices reveal considerable short-term volatility, indicative of the dynamic nature of stock market behavior. In contrast, the rolling average mitigates these fluctuations, providing a more coherent view of the stock's overall performance over the year.

From January to mid-2023, the stock prices show a steady upward trend, reflecting positive growth and heightened investor confidence during this timeframe. This trend peaks around the middle of the year, showcasing the highest price points recorded in the dataset. After reaching this peak, the stock enters a gradual decline, suggesting a period of stabilization or possible market corrections in the latter half of the year.

Employing a 3-day rolling mean effectively diminishes the noise from daily price changes, facilitating the identification of long-term trends and patterns. This visualization serves as a valuable tool for comprehending the overall performance of VCB stocks, supporting strategic analysis and informed investment decisions.



Figue 6.5. 3-Day Rolling Mean of VCB Stock Prices (2019-10-01 to 2024-10-01).

Chapter 7

Discuss

The model was designed for a 5-year period, so it is possible that some dependent variables or combinations of dependent variables were omitted. However, the resulting model outperformed the model by 100 percent when using the model compared to not using it over the period 2019-10-01 to 2024-10-01. Since the model was only tested and compared to not using the model over the period 2019-10-01 to 2024-10-01, there is no statistical evidence that the model outperformed the model by 100 percent in any given year.

According to Microsoft Excel 2010[18], the model has an R2 value of 99.88 percent. The R2 value represents how well the model predicts the response to new observations, which means that an R2 value of 99.88 percent is very good.

The reason for designing a generic model is because it is intended to be used as a guide for intraday trading in the morning when the opening price is known. It is very important that the calculations can be completed quickly so that decisions regarding selling or buying stocks can be made quickly. This is where the generic model has an advantage over some of the more custom-designed models. Safety is found in numbers when it comes to trading stocks.

Five years (2019-10-01 to 2024-10-01) of data from every stock was collected from Yahoo Finance. The amount of data makes up 1241 rows and 5 columns. Another reason is to include the 2020 global stock market crash due to the Covid-19 pandemic that is still raging and affecting some other stocks due to the total lockdown. This leads to some declining and uneven values, causing the HOSE and VNIndex to decline (there will be some codes that are not affected but accidentally decline together)

Another approach to predicting the closing price of a stock is to use the Rolling Window Concept. Unlike multiple linear regression that analyzes historical data, the Rolling Window Concept uses the present to predict the future. Since stocks have no memory, that is, they are not affected by yesterday's cyclical movements, the Rolling Window Concept may yield better results. The reason why multiple linear regression yields such good results in the end may be because the five years of historical data of VCB stocks show people's irrational behavior, which is the main reason why stocks are considered random and unpredictable.

Chapter 8

Appendix

Chapter 9

Reference

Chapter 10

Matlab Code