



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم

بسم الله الرحمن الرحيم



HANAA ALY



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم



شبكة المعلومات الجامعية التوثيق الإلكتروني والميكروفيلم



HANAA ALY



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها
علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات



يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار



HANAA ALY



Faculty of Business
Statistics, Mathematics
& Insurance Department

Statistical Arbitrage in the Egyptian Stock Market (An Applied Study)

A thesis submitted in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in Applied Statistics

Submitted by

Hisham Mohamed Abdelaziz Saad

Assistant Lecturer- Statistics, Mathematics & Insurance Department
Faculty of Business, Ain Shams University

Under supervision of

Prof. Dr. Mostafa Galal Mostafa

Professor of Applied Statistics and Head of the
Department of Statistics, Mathematics, and Insurance
Faculty of Business, Ain Shams University

Prof. Dr. Mamdouh Abdelalim Saad Mowafy

Associate Professor of Applied Statistics and Vice Dean for
Community Service and Environmental Development
Faculty of Business, Ain Shams University

2021



Faculty of Business
Statistics, Mathematics
& Insurance Department

Ph.D. Thesis

Name of Student: Hisham Mohamed Abdelaziz Saad
Title of Thesis: Statistical Arbitrage in the Egyptian Stock Market
(An Applied Study)
Academic Degree: Ph.D. in Applied Statistics

This thesis is submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Applied Statistics has been approved by:

Examination Committee

Prof. Dr. Mustafa Galal Mustafa

Professor of Applied Statistics and Head of the
Department of Statistics, Mathematics, and Insurance
Faculty of Business
Ain Shams University

**Supervisor &
Chairman**

Prof. Dr. Mamdouh Abdelalim Saad Mowafy

Associate Professor of Applied Statistics and Vice Dean for
Community Service and Environmental Development
Faculty of Business
Ain Shams University

Co-supervisor

Prof. Dr. Amr Ibrahim Abdelrahman Elatraby

Professor of Applied Statistics, Vice President of Badr
University in Cairo, and Former Dean
Faculty of Business
Ain Shams University

Member

Mr. Mohamed Mahmoud Eletreby

Chairman of Banque Misr and Chairman of Federation of
Egyptian banks (FEB)

Member

Date of dissertation defense 18/12/2021 Approval date / /

Acknowledgment

First of all, all the praises and acclamations for Almighty **ALLAH**, The Most Merciful and Most Kind, Whose blessings enabled me to pursue and complete this study.

I sincerely acknowledge and express my deepest thankfulness to **Prof. Dr. Mustafa Galal Mustafa** for his patience, valuable time, and unlimited knowledge. I am very thankful that he made me learn statistics in all stages of my career form being an undergraduate student to a PHD candidate.

I owe my truthful thanks and gratitude to my respected supervisor **Prof. Dr. Mamdouh Abdelalim Mowafy** for his constant support, valuable advice, and encouragements throughout my study. Without his direction and support, I could never have been able to complete my thesis successfully.

My gratitude and sincere thanks are of course extended to **Prof. Dr. Amr Elatraby** for all what I learnt from him during all my academic years and for his never-ending support. I would like to thank him for agreeing to be a member of the examination committee and for his valuable time and effort in reviewing this research work.

I am also grateful to **Mr. Mohamed Eletreby** for his precious time and effort in evaluating this research work and for accepting to be a member of the examination committee. His advice added great value to this thesis.

Finally, I must mention that it was mainly due to my family's moral support and continuous encouragements during my entire academic career which enabled me to complete this academic work.

To the soul of my mother and father...

ABSTRACT

Hisham Mohamed Abdelaziz Saad

Statistical Arbitrage in the Egyptian Stock Market (An Applied Study)

**PHD Applied Statistics
Ain Shams University- Faculty of Business
Statistics, Mathematics & Insurance Department**

Statistical arbitrage, which is the subject of this study, covers a variety of short-term trading and investment strategies that employs numerous parametric and nonparametric statistical techniques to identify possible relative mispricing between pairs of stocks thereby generating positive and market neutral returns. This study aims to uncover potential arbitrage opportunities present in the Egyptian stock market (EGX) by constructing trading algorithms based on combining different statistical techniques and investigating whether the risk and returns generated using these techniques can be improved by creating and implementing a dynamic trading threshold. It is worth mentioning that this is the first study to investigate the performance of such strategies in the Egyptian stock market.

The study covers the period from January 2019 to December 2020 using the daily adjusted closing prices of stocks that are eligible by EGX for short selling. The in-sample formation period in which the arbitrage portfolios are constructed is set to 12 months and the out-of-sample trading period (testing period) in which the portfolios are traded is set to 6 months. The formation and trading periods are rolled over every month, thereby creating 7 trading cycles. The approaches applied are the nonparametric distance approach pioneered by Gatev et al. 1999, Johansen cointegration approach, and Principal component analysis combined with hierarchical clustering analysis using Ward's method (PCA/HCA).

In all approaches, fixed and dynamic trading thresholds for generating trading signals are constructed and implemented. The fixed threshold is based on a standard deviation metric. The dynamic threshold is based on the conditional standard deviation that is estimated recursively using a GARCH model. In addition, sensitivity analysis is performed to investigate the effect of the trading threshold value on trading statistics and returns generated by the different approaches. All algorithms, analysis, and model building are performed using the **R** programming language.

The results show that all approaches generate positive excess returns when applied to the Egyptian stock market. The cointegration approach outperforms other approaches in terms of excess returns. In terms of risk and market neutrality of the returns, the PCA/HCA approach is superior to the other approaches. The excess returns generated show a low level of risk with market neutral returns, as evidenced by the insignificant low market beta value. The study also finds that generating trading signals based on the dynamic threshold does not improve returns but achieves a risk reduction in most of the approaches. Sensitivity analysis shows that higher trading threshold values generate higher excess returns but on the other hand, substantially reduces the percentage of pairs traded and increases the strategy's risk. Thus, more consideration should be given to the choice of the trading threshold value.

The trading algorithms created, and approaches applied can be exploited by traders, investors, and fund managers to generate positive short-term, market neutral returns while hedging risks.

Keywords: Statistical arbitrage, Pairs trading, Mean reversion, Algorithmic trading, Stocks, Distance approach, Short-selling, Market neutral returns, Egyptian Exchange, Johansen cointegration, Principal component analysis, Wards hierarchical clustering, ARCH, GARCH.

SUMMARY

Statistical arbitrage is a short-term equity trading strategy that exploits parametric and nonparametric statistical techniques to identify the relative mispricing between various securities, thereby generating positive returns while hedging risks. The idea evolved from pairs trading, which is dependent on identifying a portfolio of two stocks that moves together in some historical manner. The strategy is to attain positive returns from the temporary deviations that may occur in their equilibrium prices by implementing a long-short (buy-sell) position for the two stocks in a pair. Trading decisions are automatically generated using trading algorithms. This in turn eliminates human subjectivity in taking decisions and enables the consideration of numerous portfolios and trades.

With the advancement of statistical techniques and computing power, not only a pair of stocks may be observed, but hundreds or even more can be monitored and combined in an arbitrage portfolio. These pairs can be chosen by sectors, industries, or advanced clustering techniques to effectively hedge numerous risk factors while generating positive returns.

The general objective of this study is to uncover potential arbitrage opportunities present in the Egyptian stock market (EGX) by constructing trading algorithms based on combining different statistical techniques and investigating whether the risk and return of these techniques can be improved by implementing a dynamic trading threshold.

The study covers the period from January 2019 to December 2020 using the daily adjusted closing prices of stocks that are eligible by EGX for short selling. The in-sample formation period in which the arbitrage portfolios

are constructed is set to 12 months and the out-of-sample trading period (testing period) in which the portfolios are traded is set to 6 months. The formation and trading periods are rolled over every month, thereby creating 7 trading cycles. The approaches applied are the nonparametric distance approach pioneered by Gatev et al. 1999, Johansen cointegration approach, and Principal component analysis combined with hierarchical clustering analysis using Ward's method (PCA/HCA).

In all approaches, fixed and dynamic trading thresholds for generating trading signals are constructed and implemented. The fixed threshold is based on a standard deviation metric. The dynamic threshold is based on the conditional standard deviation that is estimated recursively using a GARCH model. In addition, sensitivity analysis is performed to investigate the effect of the trading threshold parameter on trading statistics and returns generated by the different approaches. All algorithms, analysis and model estimations are performed using the **R** programming language.

The results show that all approaches generate positive excess returns when applied to the Egyptian stock market. The cointegration approach outperforms other approaches in terms of excess returns. In terms of risk and market neutrality of the returns, the PCA/HCA approach is superior to the other approaches. The excess returns generated show a low level of risk with market neutral returns, as evidenced by the insignificant low market beta value. The study also finds that generating trading signals based on the dynamic trading threshold does not improve returns but achieves a risk reduction in most of the approaches. Sensitivity analysis shows that higher trading threshold values generate higher excess returns but on the other hand, substantially reduces the percentage of pairs traded and increases the

strategy's risk. Thus, more consideration should be given to the choice of the trading threshold value.

The trading algorithms created, and approaches applied can be exploited by traders, investors, and fund managers to generate positive short-term, market neutral returns while hedging risks.

The study is organized into six chapters as follows:

Chapter One:

This chapter provides a general overview of the study. It includes an introduction, problem statement, importance of study, objectives of study, and limitations of study. It also includes a brief literature review regarding statistical arbitrage in different equity markets.

Chapter Two:

This chapter introduces some of the theories and mathematics that are related to statistical arbitrage. It also includes a brief history of the Egyptian Exchange highlighting the implementation stages of short selling in Egypt's capital market.

Chapter Three:

This chapter outlines the theoretical framework of the statistical models used. It includes a theoretical explanation for the Engle and Granger cointegration approach, Johansen Cointegration approach, GARCH model, Principal component analysis, and Hierarchical clustering analysis. The nonparametric distance approach is explained in the methodology.

Chapter Four:

This chapter provides a detailed description of the methodology applied and the algorithms created. It focuses on the empirical application of the techniques and stages that are followed.

Chapter Five:

This chapter presents the results that are obtained throughout the analysis. It contains the results of the distance approach, cointegration approach, and the PCA/HCA approach. In all approaches, fixed and dynamic thresholds are used in generating trading signals. The results comparison, sensitivity of returns and trading statistics to different trading threshold values are also presented in this chapter.

Chapter Six:

This chapter summarizes and concludes the study. Evaluation of the results, suggested recommendations, and future research are provided in this chapter.

Table of Contents

Chapter One: Introduction	1
1.1 Background	1
1.2 Nature of the problem	4
1.3 Importance of study.....	5
1.4 Objectives of study.....	6
1.5 Limitations of study	7
1.6 Literature review	7
1.7 Data and software used	21
1.8 Organization of study	22
Chapter Two: Theories and approaches of statistical arbitrage and a brief history of the Egyptian stock market	24
2.1 Introduction	24
2.2 Market neutral trading strategies.....	25
2.3 Mean-reversion.....	26
2.4 Statistical arbitrage and pairs trading	28
2.5 The Egyptian stock market.....	33
2.6 Short selling in Egypt.....	39
Chapter Three: Theory of statistical models.....	45
3.1 Stationarity and unit root testing	45
3.2 Cointegration.....	48
3.3 ARCH and GARCH Models	53

3.4 Principal component analysis.....	59
3.5 Hierarchical clustering analysis	64
Chapter Four: Methodology and Algorithms.....	67
4.1 Data adjustments and stock filtrations	67
4.2 Timing of formation and trading periods	68
4.3 Formation period	69
4.4 Trading period	74
4.5 Return calculation and evaluation	78
4.6 Benchmark comparison metrics	80
4.7 Algorithm’s overview	81
4.8 Sensitivity analysis.....	85
Chapter Five: Statistical analysis and results.....	86
5.1 Stock filtrations and selection	86
5.2 Results of the distance approach	90
5.3 Results of the cointegration approach	99
5.4 Results of the PCA/HCA approach.....	107
5.5 Results comparison	114
5.6 Threshold sensitivity analysis	118
Chapter Six: Conclusions and Recommendations	124
6.1 Study overview.....	124
6.2 Conclusions of the study	125
6.3 Recommendations	128