

**MAXIMIZING PROFIT PER UNIT TIME IN  
COINTEGRATION BASED PAIRS TRADING**

A THESIS  
SUBMITTED TO THE DEPARTMENT OF INDUSTRIAL  
ENGINEERING  
AND THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE OF  
BILKENT UNIVERSITY  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
MASTER OF SCIENCE

by  
Duygu Tatal  
January, 2014

I certify that I have read this thesis and that in my opinion it is full adequate, in scope and in quality, as a dissertation for the degree of Master of Science.

---

Assoc. Prof. Savaş Dayanık (Advisor)

I certify that I have read this thesis and that in my opinion it is full adequate, in scope and in quality, as a dissertation for the degree of Master of Science.

---

Assoc. Prof. Sinan Gezici

I certify that I have read this thesis and that in my opinion it is full adequate, in scope and in quality, as a dissertation for the degree of Master of Science.

---

Asst. Prof. Emre Nadar

Approved for the Graduate School of Engineering and Science

---

Prof. Dr. Levent Onural

Director of the Graduate School of Engineering and Science

# ABSTRACT

## MAXIMIZING PROFIT PER UNIT TIME IN COINTEGRATION BASED PAIRS TRADING

Duygu Tatal

M.S. in Industrial Engineering

Advisor: Assoc. Prof Savaş Dayanık

January, 2014

Pairs Trading is a modest but persistent statistical arbitrage strategy. It is based on identifying a pair of stocks whose prices are driven by the same economic forces, and trade according to the spread between their prices. In this thesis, we study on a pairs trading method which maximizes the profit per unit time. We identify the pairs using cointegration analysis and build the method using Markov Chains. After constructing the method, we examine its performance on both simulated and real data. We use banking stocks from Istanbul Stock Exchange as real data.

**Keywords:** Pairs Trading, Cointegration, Time Series Analysis, Markov Chain.

# ÖZET

## BİRİM ZAMANDAKİ KARI MAKSİMUMA ULAŞTIRAN EŞBÜTÜNLEŞME BAZLI EŞLİ ALIM-SATIM METODU

Duygu Tural

Endüstri Mühendisliği, Yüksek Lisans

Tez Yöneticisi: Doç. Dr. Savaş Dayanık

Ocak, 2014

Eşli Alım-Satım tekniği az miktarda ama devamlı kar sağlayan bir istatistiksel arbitraj yöntemidir. Bu metod, fiyatları aynı ekonomik sebeplerden etkilenen bir çift hisse senedi belirleyip, bu iki hisse senedinin fiyatları arasındaki farka göre alım-satım yapmaya dayanır. Bu tezde birim zamandaki karı maksimize etmeye dayanan bir eşli alım-satım metodu üzerinde çalışılmıştır. Hisse senedi çiftleri eşbütünleşme yöntemi kullanılarak belirlenmiş ve metod Markov Zinciri'nin özelliklerinden faydalanılarak oluşturulmuştur. Çalışmanın sonunda metod, hem simüle edilmiş, hem de gerçek veriler üzerinde denenerek metodun performansı ölçülmüştür. Gerçek veri olarak Borsa İstanbul'un bankacılık hisse senetleri kullanılmıştır.

**Anahtar Sözcükler:** Eşli Alım-Satım, Eşbütünleşme, Zaman Serileri Analizi, Markov Zinciri.

To my family...

## **ACKNOWLEDGEMENT**

First and foremost, I would like to express my gratitude to my advisor, Assoc. Prof. Savaş Dayanık.

I would like to thank the members of my thesis committee Assoc. Prof. Sinan Gezici and Asst. Prof. Emre Nadar for critical reading of this thesis and for their valuable comments.

I am grateful to my precious friend Ezel Ezgi Budak Ayar for being such a long time in my life. I feel she is always there for me with her great heart and endless support. This thesis would not be possible without her help, patience and support. I also wish to thank Müge Güçlü, Burak Ayar and Mustafa Karakul for their academic assistances. I am also indeed grateful to Fatih Cebeci for his encouragement, patience and support. I am also thankful to Nur Budak and Özcan Budak for their great support. I am also grateful to all other friends that I failed to mention here.

Last but not the least; I also would like to express my deepest gratitude to my mother Yurdagül Tural, father Mustafa Tural, and sister Beste Tural for their eternal love, support and trust at all stages of my life and especially during my graduate study.

# Contents

<b>1. Introduction</b> .....	<b>1</b>
<b>2. Literature Review</b> .....	<b>5</b>
2.1 Cointegration-based Strategies .....	5
2.2 Pairs Trading and Statistical Arbitrage Studies .....	6
<b>3. Problem Definition</b> .....	<b>10</b>
3.1 Maximizing Profit per Unit Time in Pairs Trading.....	10
3.2 Cointegration.....	12
<b>4. Proposed Pairs Trading Method</b> .....	<b>14</b>
<b>5. Case Study</b> .....	<b>20</b>
5.1 Stock Price Data from BIST .....	20
5.2 Profit per Unit Time for $X_0 = 0$ .....	32
5.3 Profit per Unit Time for a Random Starting Point.....	34
<b>6. Numerical Results</b> .....	<b>41</b>
6.1 Performance Measurement Using Simulated Data.....	41
6.2 Performance Comparison between Two Methods .....	44

6.3 Performance Measurement Using BIST Banking Stock Data .....	50
<b>7. Conclusion and Future Research.....</b>	<b>54</b>
<b>Bibliography .....</b>	<b>57</b>
<b>Appendix A: Time Series Plots of the Banking Stock Pairs.....</b>	<b>60</b>
<b>Appendix B: Model Outputs of the Cointegration Equations.....</b>	<b>68</b>
<b>Appendix C: Profit per Unit Time Graphs for <math>X_0 = 0</math>.....</b>	<b>72</b>
<b>Appendix D: Density, Empirical CDF and Q-Q Plots of Starting Points...75</b>	
<b>Appendix E: Expected Profit per Unit Time for Random Starting Point..88</b>	
<b>Appendix F: Profit per Unit Time Values for Simulated Samples.....91</b>	
<b>Appendix G: Estimated Parameter Values of Simulated Data.....97</b>	
<b>Appendix H: Comparison of Two Methods.....102</b>	



# List of Tables

Table 4.1: Probability matrix representation for process of the strategy .....	17
Table 5.1: Cointegration test output for stocks Akbank and Fortis.....	23
Table 5.2: Normalized cointegration coefficients for Akbank and Fortis.....	23
Table 5.3: Cointegration test output for stocks Fortis and Vakıfbank.....	24
Table 5.4: Normalized cointegration coefficients for Fortis and Vakıfbank.....	24
Table 5.5: Cointegration test output for stocks Fortis and Şekerbank.....	25
Table 5.6: Normalized cointegration coefficients for Fortis and Şekerbank.....	26
Table 5.7: Cointegration test output for stocks Vakıfbank and Şekerbank.....	27
Table 5.8: Normalized cointegration coefficients for Vakıfbank and Şekerbank.....	27
Table 5.9: Cointegration test output for stocks Akbank and Vakıfbank.....	28
Table 5.10: Normalized cointegration coefficients for Akbank and Vakıfbank.....	28
Table 5.11: Cointegration test output for stocks Akbank and Şekerbank.....	29
Table 5.12: Normalized cointegration coefficients for Akbank and Şekerbank..	30

Table 5.13: Parameter estimations of the AR(1) models.....	31
Table 5.14: Optimal trade-in points for $\emptyset$ values between 0.95 and 0.99 for a random starting point process.....	39
Table 6.1: Optimal trade-in points for $\emptyset$ values between 0.95 and 0.99 for simulated data.....	43
Table 6.2: Estimated parameter values for $\emptyset=0.95$ .....	48
Table B.1: Model outputs for the pair Fortis-Akbank.....	68
Table B.2: Model outputs for the pair Fortis-Şekerbank.....	69
Table B.3: Model outputs for the pair Vakıfbank-Şekerbank.....	69
Table B.4: Model outputs for the pair Akbank-Vakıfbank.....	70
Table B.5: Model outputs for the pair Akbank-Şekerbank.....	70
Table B.6: Model outputs for the pair Fortis-Vakıfbank.....	71
Table F.1: Profit per unit time values for 20 samples for $\emptyset=0.95$ series.....	92
Table F.2: Profit per unit time values for 20 samples for $\emptyset=0.96$ series.....	93
Table F.3: Profit per unit time values for 20 samples for $\emptyset=0.97$ series.....	94
Table F.4: Profit per unit time values for 20 samples for $\emptyset=0.98$ series.....	95
Table F.5: Profit per unit time values for 20 samples for $\emptyset=0.99$ series.....	96
Table G.1: Estimated parameter values of simulated data for $\emptyset=0.96$ .....	98
Table G.2: Estimated parameter values of simulated data for $\emptyset=0.97$ .....	99
Table G.3: Estimated parameter values of simulated data for $\emptyset=0.98$ .....	100
Table G.4: Estimated parameter values of simulated data for $\emptyset=0.99$ .....	101

Table H.1: Profit per unit time values of 20 samples for  $\emptyset=0.95$ ,  $a_2=0.28$  and  $a_2=0.54$ .....103

Table H.2: Profit per unit time values of 20 samples for  $\emptyset=0.96$ ,  $a_1=0.60$  and  $a_2=0.58$ .....104

Table H.3: Profit per unit time values of 20 samples for  $\emptyset=0.97$ ,  $a_1=0.65$  and  $a_2=0.66$ .....105

Table H.4: Profit per unit time values of 20 samples for  $\emptyset=0.98$ ,  $a_1=0.67$  and  $a_2=0.76$ .....106

Table H.5: Profit per unit time values of 20 samples for  $\emptyset = 0.99$ ,  $a_1 = 0.99$  and  $a_2=0.98$ .....107

# List of Figures

Figure 4.1: Illustration of trading process.....	15
Figure 5.1: Time series graph of Fortis and Akbank stock prices.....	21
Figure 5.2: Time series graph of Yapı Kredi Bank and Vakıfbank stock prices.....	22
Figure 5.3: Time series graph of Yapı Kredi Bank and Garanti Bank stock prices..	30
Figure 5.4: Profit per unit time graph for $X_0 = 0$ and $\sigma = 0.05$ .....	34
Figure 5.5: Density graph for $\phi = 0.95$ , $a = 0.135$ and $\sigma = 0.05$ .....	37
Figure 5.6: Q-Q plot for $\phi = 0.95$ , $a = 0.135$ and $\sigma = 0.05$ .....	37
Figure 5.7: Profit per unit time graph for a random starting point and $\sigma = 0.05$ .....	40
Figure 6.1: Profit per unit time graph using simulated data.....	44
Figure 6.2: The timeline for the investor's process.....	46
Figure A.1: Time series graphs of Fortis and Akbank .....	60
Figure A.2: Time series graphs of Fortis and Şekerbank.....	61
Figure A.3: Time series graphs of Vakıfbank and Şekerbank.....	61
Figure A.4: Time series graphs of Akbank and Vakıfbank.....	62
Figure A.5: Time series graphs of Akbank and Şekerbank.....	62

Figure A.6: Time series graphs of Fortis and Vakıfbank.....	63
Figure A.7: Time series graphs of Fortis and Garanti Bank.....	63
Figure A.8: Time series graphs of Garanti Bank and Vakıfbank .....	64
Figure A.9: Time series graphs of Garanti Bank and Şekerbank.....	64
Figure A.10: Time series graphs of Yapı Kredi Bank and Fortis.....	65
Figure A.11: Time series graphs of Yapı Kredi Bank and Şekerbank.....	65
Figure A.12: Time series graphs of Yapı Kredi Bank and Vakıfbank .....	66
Figure A.13: Time series graphs of Yapı Kredi Bank and Akbank.....	66
Figure A.14: Time series graphs of Yapı Kredi Bank and Garanti Bank.....	67
Figure A.15: Time series graphs of Akbank and Garanti Bank.....	67
Figure C.1: Profit per unit time for $\emptyset$ in $[0.95, 0.98]$ , $\sigma = 0.1$ and $\Delta b = 0.005$ ....	72
Figure C.2: Profit per unit time for $\emptyset = 0.99$ , $\sigma = 0.1$ and $\Delta b = 0.01$ .....	73
Figure C.3: Profit per unit time for $\emptyset$ in $[0.95, 0.98]$ , $\sigma = 0.2$ and $\Delta b = 0.01$ .....	73
Figure C.4: Profit per unit time for $\emptyset = 0.99$ , $\sigma = 0.2$ and $\Delta b = 0.02$ .....	74
Figure D.1: Density graph for $\emptyset = 0.95$ , $a = 0.135$ and $s = 0.05$ .....	76
Figure D.2: Empirical CDF graph for $\emptyset = 0.95$ , $a = 0.135$ and $s = 0.05$ .....	77
Figure D.3: Q-Q plot for $\emptyset = 0.95$ , $a = 0.135$ and $s = 0.05$ .....	77
Figure D.4: Density graph for $\emptyset = 0.99$ , $a = 0.245$ and $s = 0.05$ .....	78
Figure D.5: Empirical CDF graph for $\emptyset = 0.99$ , $a = 0.245$ and $s = 0.05$ .....	79
Figure D.6: Q-Q plot for $\emptyset = 0.99$ , $a = 0.245$ and $s = 0.05$ .....	79
Figure D.7: Density graph for $\emptyset = 0.95$ , $a = 0.275$ and $s = 0.1$ .....	80

Figure D.8: Empirical CDF graph for $\emptyset = 0.95$ , $a = 0.275$ and $s = 0.1$ .....	81
Figure D.9: Q-Q plot for $\emptyset = 0.95$ , $a = 0.275$ and $s = 0.1$ .....	81
Figure D.10: Density graph for $\emptyset = 0.99$ , $a = 0.49$ and $s = 0.1$ .....	82
Figure D.11: Empirical CDF graph for $\emptyset = 0.99$ , $a = 0.49$ and $s = 0.1$ .....	83
Figure D.12: Q-Q plot for $\emptyset = 0.99$ , $a = 0.49$ and $s = 0.1$ .....	83
Figure D.13: Density graph for $\emptyset = 0.95$ , $a = 0.55$ and $s = 0.2$ .....	84
Figure D.14: Empirical CDF graph for $\emptyset = 0.95$ , $a = 0.55$ and $s = 0.2$ .....	85
Figure D.15: Q-Q plot for $\emptyset = 0.95$ , $a = 0.55$ and $s = 0.2$ .....	85
Figure D.16: Density graph for $\emptyset = 0.99$ , $a = 0.98$ and $s = 0.2$ .....	86
Figure D.17: Empirical CDF graph for $\emptyset = 0.99$ , $a = 0.98$ and $s = 0.2$ .....	87
Figure D.18: Q-Q plot for $\emptyset = 0.99$ , $a = 0.98$ and $s = 0.2$ .....	87
Figure E.1: Expected profit per unit time for $\emptyset$ in $[0.95,0.98]$ , $\sigma = 0.1$ and $\Delta b = 0.005$ .....	88
Figure E.2: Expected profit per unit time for $\emptyset = 0.99$ , $\sigma = 0.1$ and $\Delta b = 0.01$ ..	89
Figure E.3: Expected profit per unit time for $\emptyset$ in $[0.95,0.98]$ , $\sigma = 0.2$ and $\Delta b = 0.01$ .....	89
Figure E.4: Expected profit per unit time for $\emptyset = 0.99$ , $\sigma = 0.2$ and $\Delta b = 0.02$ .....	90

# Chapter 1

## Introduction

Pairs trading is a statistical arbitrage strategy with a long history. In the 1980s, Nunzio Tartaglia and his colleagues, a group of scientists from mathematicians, physicists and computer scientists who did not have a solid financial background, worked on Wall Street data to reveal the potential arbitrage opportunities in the equities markets. They used sophisticated statistical methods to develop technical trading programs to search for the arbitrage opportunities and to execute trades automatically, which made them the pioneers in this field. Meanwhile, their programs identified pairs of securities, mostly substitutes of each other, whose prices tended to move together. They traded those pairs with a great success in 1987 and made a \$50 million profit. After their great success, this trading method, called *pairs trading*, has become very popular (Gatev et al. [12]).

Pairs trading is a speculative investment strategy which stems from mispricing one of two related stocks. It offers modest but persistent profits (Peskin and Bourdreau [25]; Gatev et al. [12]). This is a market-neutral strategy which depends on relative pricing of stocks. The strategy is based on identifying pairs of shares whose prices are driven by the same economic forces, i.e. two shares whose prices tend to move together, and then trading them when any temporary deviations occur from their long-run average

relationship. The intuition behind this method is that the deviation will be reversed eventually and they will return to their equilibrium relationship since the stocks' prices tend to move together (Gillespie and Ulph [13]). The deviation in the prices is often because one asset is overvalued to the other asset. When we notice a deviation, we invest in a two-asset portfolio, which we call a pairs trading. We sell the overvalued asset (take short position) and buy the undervalued asset (take long position) (Puspaningrum [26]). When the market returns to the equilibrium, in other words when the deviation of the stock prices from their equilibrium values is revealed, we take opposite positions for both assets and obtain profit from the trade. Although this strategy generally does not create big profits, it creates a hedge against the market so that you can gain profit with a low risk.

In order to be successful in pairs trading, it is crucial to be the first to observe the opportunity. Otherwise, the opportunity is exploited by others, and the spread between pair prices becomes tighter and the potential profit becomes smaller. Another important issue in pairs trading is to decide on the appropriate prices, or deviation, to buy and sell the shares.

Selecting correct pairs is crucial; however, since it is possible to detect related pairs with simple statistical techniques, this also becomes an advantage of pairs trading. Correlation analysis, regression analysis, cointegration or some non-parametric rules can be used to choose the related pairs of stocks. In this study, we identify the stock pairs using cointegration because it is based on mean reversion, which guarantees that the spread between the stocks is temporary and will be closed in the end. In other words, mean reversion assures that although anomalies arise among stock prices in the short-term, they will be corrected in the long-term (Ehrman [8]). Cointegration can be defined as in Definition 1.1.



*Definition 1.1.* Let  $X_t$  and  $Y_t$  be two integrated time series. If any linear combination of those series becomes a time series having a lower order of integration, then  $X_t$  and  $Y_t$  are called cointegrated.

*Definition 1.2.* The time series  $X_t$  is integrated of order  $d$ , which is represented as  $X_t \sim I(d)$ , if  $(1 - L)^d X_t$  is a stationary process, where  $L$  is the lag operator.

Many analysts rely on correlation, in order to select the pairs and seek a positively correlated pair. Often, those pairs appear in the same or related sectors (Lin et al. [18]). However, Banerjee et al (1993) and Vidyamurthy (2004) showed that the cointegration technique is more effective than correlation technique for revealing potential profit in temporary pricing anomalies between two stock prices driven by common underlying factors. Correlation and regression analysis yield imprecise long-run equilibrium relationships between the shares and do not necessarily imply mean-reversion; on the other hand, the pairs formed by cointegration technique assures that the spread between the pairs are temporary or reverting (Gatev et al. [12]) which is why we prefer to use cointegration technique in our study to choose the pairs.

In this study, we follow Lin et al (2006) and Puspaningrum et al (2009). In those papers, cointegration is used for choosing the pairs and the goal is to find the conditions under which a preset minimum profit per trade is reached. However, the studies do not consider how long it takes to obtain this preset level and it is very important to know this duration since human beings cannot live forever. In this thesis we also take time into account and propose a pairs trading method that maximizes the long-run profit per unit time.

The remainder of the thesis is organized as follows:

In Chapter 2, we review the literature on cointegration and pairs trading. We summarize the studies dealing with cointegration in order to illustrate the problems this technique is used in, other than pairs trading. In addition, we give examples of existing pairs trading

studies. In Chapter 3, we introduce our problem. Then, we elaborate the concept of cointegration. In Chapter 4, we propose our resolution method. Chapter 5 presents a case study on which we apply our method. We use six banks' stock price data from BIST (Istanbul Stock Exchange). In Chapter 6, we provide the results of the computational experiments using both simulated and historical data to evaluate our proposed method and discuss the success of this study. In Chapter 7, we summarize our findings and discuss possible future research topics.

# Chapter 2

## Literature Review

The literature on pairs trading problems is not very extensive. In addition, most of the studies are applications of specific case studies. Thus, the specific type of pairs trading we work on, which is pairs trading using cointegration method, is rare in literature.

In this chapter, firstly we will mention about cointegration based strategies in literature. We will present the studies using cointegration strategy from the fields other than pairs trading in this section. In the second part, we will mention about some basic studies about pairs trading in literature. We will present both type of studies, the studies using cointegration as a pair selection method and the studies using methods other than cointegration, in this section.

### **2.1 Cointegration-based Strategies**

Walls (1994) studied the linkages between natural gas spot prices at various production fields, pipeline hubs, and city markets in US Natural Gas Industry using cointegration rank test. Natural gas spot market data are used for likelihood based tests for cointegration. In that study, Johansen method is used for spatial market linkages. The results of that study shows that the natural gas spot markets are strongly correlated. 19

market pairs are examined and 13 of the pairs satisfied the rules for perfect market integration.

Chan, Benton and Min (1997) compare 18 nations' stock markets over a 32-year period. Johansen's cointegration test is used to analyze the cross country market efficiency. They show that a small number of stock markets are cointegrated.

Bala and Mukund (2001) conducted a study based on the linkage between the US and Indian stock markets. In this study, the theory of cointegration is used to show the interdependence level of Bombay stock exchange and BSE & NYSE & NASDAQ. They concluded that the markets are not interdependent during the selected time period.

Long-run equilibrium relationship and short-run dynamic relation between the Indian stock market and the developing countries' stock markets are studied by Wong, Agarwal and Du (2005). They examine the Granger causality relationship and the pairwise, multiple and fractional cointegrations between the Indian stock market and the other stock markets in US, UK, Japan. They concluded that the Indian market is closely cointegrated with US, UK and Japan stock markets. With a common fractional, nonstationary component Indian stock index and the other stock indices form fractionally cointegrated relationship in the long run. That study also uses the Johansen method to identify the cointegration relationships.

Narayan (2005) studied the saving investment correlation for China over periods 1952-1998 and 1952-1994. The study revealed that saving and investment are correlated for China for the both periods. They applied residual based tests for cointegration between China's saving and investment for the two periods.

## **2.2 Pairs Trading and Statistical Arbitrage Studies**

In 2003, Nath studied securities in the highly liquid secondary market for U.S. government debt over 1994-2000. While most of the studies in this field use daily data,

intraday data are used in this study. Four different opening and closing thresholds are used to discover the patterns of the security market. He concluded that 15<sup>th</sup> percentile as opening trigger and 5<sup>th</sup> percentile as closing trigger is profitable for U.S. security market.

Hong and Susmel (2003) conducted a study on pairs trading strategies in which they used cointegration as a pair detection strategy. They worked on 64 Asian shares listed in their local markets and listed in U.S. as American Depository Receipts. They found that if an investor waits for one year before closing his trade-in pairs trading, he can obtain an annualized profit of more than 33%. Also, they stated that market frictions, such as different trading hours, make the strategies risky and hard to implement.

Lin et al. (2006) study pairs trading using cointegrated series to reveal pairs. They develop a pairs trading method to satisfy a preset minimum profit per trade. They illustrate the success of their method on both simulated and real data. They use daily closing price data for two Australian Stock Exchange quoted bank shares, ANZ and ADB, over 18 months to test the results. We will elaborate on that paper because we will use it later to take a further step. Two share prices  $P_1(t)$  and  $P_2(t)$ , chosen using cointegration technique, are taken at time  $t$  and the number of shares in short position and in long position at time  $t$  are denoted by  $N_1(t)$  and  $N_2(t)$ , respectively. Let  $P_1(t) + \beta P_2(t) = \varepsilon_t$  denote the cointegration equation where  $\varepsilon_t \sim I(0)$ , i.e.  $\varepsilon_t$  is integrated of order 1, and  $t \geq 1$  for some  $\beta < 0$ . Then two points  $a$  and  $b$  are chosen such that  $\varepsilon_{t_o} > a > 0$  and  $\varepsilon_{t_c} < b$ , where  $t_o$  shows the time of opening a trade and  $t_c$  shows the time of closing a trade, and the integer  $n > \frac{K|\beta|}{(a-b)}$  is chosen, where  $K$  is the minimum preset profit per trade. The trading rule is that we open a trade at  $t_o$  when  $P_1(t_o) > P_2(t_o)$  and  $\varepsilon_{t_o} > a > 0$ , buy  $n$  shares of  $S_2$  and sell  $\frac{n}{|\beta|} + 1$  shares of  $S_1$  at time  $t_o$ . Then, we close out the trade at time  $t_c$  when  $\varepsilon_{t_c} < b$  which ensures that the gain is at least  $K$  monetary units.

Do et al. (2006) propose an approach for pairs trading in a continuous time setting. They assume asset prices follow mean reverted process. They present an EM algorithm to estimate model parameters.

Gatev et al. (2006) conduct an empirical study on pairs trading with daily stock data over 1962-2002. They use a minimum distance of two historical standard deviations between normalized historical prices as the trigger of a trade. When the spread between the stock prices becomes more than this pre-specified threshold, they trade-in. When the spread becomes tighter, which is less than the pre-specified threshold, they trade-out. They tested their method in the presence of transaction costs and their study proved robustness to estimates of transaction costs. They find that the profits have a very small correlation with the spreads between small and large stocks, and between value and growth stocks.

Khandani and Lo (2007) study statistical arbitrage on a number of quantitative long/short equity hedge funds. They discuss the performance of the Lo-MacKinlay contrarian strategies in the context of the liquidity crisis of 2007. In that paper, market-neutrality is enforced by ranking stock returns by quantiles and trading “winners-versus-losers”, in a dollar-neutral fashion.

Papadakis et al. (2007) conduct a study on the impact of accounting information events, such as earnings announcements, on the profitability of pairs trading strategy proposed by Gatev et al. (2006). They find that trades are mostly triggered around accounting information events for the portfolio of U.S. stock pairs between years 1981 and 2006. In addition, they find that the trades opening after those events are less profitable than the trades opening during non-event periods.

Avellaneda and Lee (2008) present a systematic approach to statistical arbitrage and for constructing market-neutral portfolio strategies based on mean-reversion. The approach presented in their study is based on decomposing stock returns into systematic

components. The paper compares the ETF and the PCA methods by reproducing the results of Khandani and Lo (2007).

Mudchanatongsuk et al. (2008) propose a stochastic control approach for pairs trading strategy. They work on log-relationship between pairs and modeled them as an Ornstein-Uhlenbeck process. They test the method with simulated data.

Engelberg et al. (2009) discuss the concept of liquidity in pairs trading. They found that the profits are higher when the spread between the prices of the pairs is because of news that temporarily reduces the liquidity of one or both of the stocks in the pair. Also, the profits are lower when the spread between the prices is relevant to the news about the value of a stock in the pair.

Puspaningrum et al. (2009) conducted a study on pairs trading which they used cointegration technique to select the pairs for the trade. They use the stationarity properties of errors of the cointegrated pairs which follow AR(1) processes. They explore how the preset boundaries to open a trade affect the minimum total profit over a specified trading horizon. They estimated the number of trades over a specified trading horizon by using the average trade duration and the average inter-trade-interval. They developed numerical algorithm to estimate the average trade duration, the average inter-trade-interval, and the average number of trades and to maximize the minimum total profit.

Do and Faff (2010) discuss the performance of pairs trading over years. They reveal that as opposed to the continuing downward trend in profitability of pairs trading over years, that strategy worked well during periods of recent global financial crises.

# Chapter 3

## Problem Definition

In this study, our objective is to find a cointegration based pairs trading strategy maximizing the profit per unit time. We are inspired by the study of Lin et al (2006), which is about developing a pairs trading method to satisfy a minimum profit per trade. Although they ensure gaining a preset profit per trade in this strategy, time horizon to obtain that amount of profit is not considered, which may lead the cycle of a trade to be very long. Thus, we take the topic from this aspect in this study.

### 3.1 Maximizing Profit per Unit Time in Pairs Trading

We intend to find two cointegrated stock pairs,  $P_1$  and  $P_2$ , whose cointegrating error is integrated of order 0, i.e. follows a stationary process as in

$$P_1(t) + \beta P_2(t) = X_t, \text{ where } X_t \sim I(0). \quad (1)$$

This relationship ensures us that the two stocks follow similar patterns, that is decrease and increase at similar points of time, due to the nature of cointegration, which implies mean-reversion.



The cointegration equation identifies our portfolio, which consists of 1 stock of  $P_1$  and  $\beta$  stocks of  $P_2$ . We intend to build a method for pairs trading such that a preset price level for our portfolio initiates each trade. That is, before we start to trade, we set a point  $a$  such that whenever  $X_t$  exceeds point  $a$  we trade-in.  $a$  can be point that the spread between  $P_1$  and  $P_2$  are wide enough to earn from the trade. We short sell the outperforming stock and buy the underperforming stock at this point; in other words, we trade-in. To trade-out, we wait for  $X_t$  to reach below point 0, i.e.  $X_t \leq 0$ , and we take opposite positions for  $P_1$  and  $P_2$  at this point and close the trade.

Although any point that the spread between the prices of  $P_1$  and  $P_2$  are sufficiently wide works for pairs trading, in order to do a profitable trade, we intend to find the point  $a$  earning the most. Finding optimal  $a$ , namely the most profitable point is the objective of our study.

Let us show the calculation of profit from this strategy. If we trade-in at point  $a$  and trade-out at point 0, we gain a profit of  $a$  from one trading cycle, which makes the profit per unit time equal  $a/C$  where  $C$  represents cycle length of one trade of the series  $X_t$ . However, note that the first time  $X_t$  reaches to point  $a$  may realize at a point bigger than point  $a$ . In addition, the first time  $X_t$  reaches to point 0 may also realize at a point smaller than point 0. Thus, the profit per unit time for one trading cycle is at least as large as  $a/C$ . Our goal in this study is to maximize the expected profit per unit time. We benefit from the properties of renewal reward processes to estimate the expected profit per unit time as in (2).

$$\lim_{t \rightarrow \infty} \frac{1}{t} E [a/C(t)] = \frac{a}{E[C(t)]} \quad (2)$$

Although the process of stock pricing is continuous, because we use daily stock prices, we can use discrete state space and simplify the problem, which is an advantage for us since we will use Markov Chains to develop the method.

### 3.2 Cointegration

In this section, we explain a key topic of our study, cointegration, which we use for pairs identification.

Many econometric studies are related to long-term equilibrium relation generated by market forces, which makes forming long-term equilibriums an important goal for econometricians. Before the papers published by Granger and Newbold (1974), and Nelson and Plosser (1982), econometricians tended to remove trends and drifts from the models in order to make the variables stationary since most of the techniques in econometrics are based on stationarity. However, with those papers it is indicated that removing trends from the model does not necessarily provide constant unconditional mean and variance over time, which describe stationary series. (Dolado, Gonzalo and Marmol, [7]).

Cointegration is the field of study to seek a linear combination of nonstationary variables which forms a stationary process. In 1981, Granger introduced the notion of cointegration and in 1987, Engle and Granger elaborated it. Then, Johansen introduced the Johansen Test, which allows for more than one cointegration relationship between 2 time series.

Before elaborating the concept let us begin with the definitions of integration and stationarity.

*Definition 3.1.* A series whose mean, variance and autocorrelation are constant over time is called stationary.

*Definition 3.2.* A series with no deterministic component which has a stationary, invertible, ARMA representation after differencing  $d$  times, is said to be integrated of order  $d$ . (Engle and Granger [11])

Let us consider a vector of economic variables  $x_t$  and let  $\beta$  be the vector of coefficients of those variables. Then, the system is in equilibrium if  $\beta x_t = 0$ . If this system does not equal zero, but if there is an equilibrium error  $\varepsilon_t$  such that  $\beta x_t = \varepsilon_t$ , then the series  $\varepsilon_t$  must be stationary for the equilibrium to be meaningful (Enders [8]).

Let us restate the concept according to Engle and Granger's definition:

*Definition 3.3.* The components of the vector  $x_t = (x_{1t}, x_{2t}, \dots, x_{nt})'$  are said to be cointegrated of order  $d \geq 0$ ,  $b > 0$ , denoted by  $x_t \sim CI(d, b)$  if

- i. All components of  $x_t$  are integrated of order  $d$ .
- ii. There exists a vector  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$  such that the linear combination  $\beta_{xt} = \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_n x_{nt}$  is integrated of order  $(d - b)$ .

The  $\beta$  vector is called the cointegrating vector.

As Definition 3.3 states, cointegration does not necessarily remove integration. It is a method to decrease the level of integration. However, in this study we need our portfolio to be a stationary series, thus; we require the pairs be cointegrated of order  $(d, b)$  where  $d = b$ .

Note that there may be more than one cointegrating vector to decrease the level of integration of  $x_t$ . In fact, if there are  $n$  components of  $x_t$ , then the maximum number of cointegrating vectors is  $n - 1$ . Because we work with pairs, the vector  $x_t$  has two components, which leads only one cointegrating vector for each pair of stock price series.

Although cointegration can also be formed among more than 2 variables, in this study we use cointegration between 2 variables, because we work with two stock portfolios.

# Chapter 4

## Proposed Pairs Trading Method

Before starting to develop our method let us remind that our objective is to maximize long-run expected profit per unit time in trading process. Firstly, let us give a representation of the series we will analyze. Since we work on stationary linear combinations of nonstationary pairs, we ignore the pairs now on and continue with the whole portfolio, i.e. the series  $X_t$  which is represented as (1). Because the stock prices fit to autoregressive processes in many studies in literature, we assume that  $X_t$  follows an autoregressive process without intercept of order 1, i.e.  $AR(1)$ , as in

$$X_t = \phi X_{t-1} + \varepsilon_t, \tag{3}$$

where  $-1 < \phi < 1$  and  $\varepsilon_t \sim Gsn(0, \sigma_\varepsilon^2)$ . Note that an  $AR(1)$  process is stationary as long as its regression coefficient  $\phi$  takes a value strictly between  $-1$  and  $1$ . We assume this holds for now, but we will examine if this assumption holds for our empirical work in the next chapter. Let us remind that means and variances of stationary processes do not change over time which is why we prefer to use in pairs trading. This property of stationarity ensures us that our portfolio returns to mean eventually. In other words, we are confident that our process tend to return its mean  $0$ , i.e. the trade-out point, which is the point to close the trade in our strategy.

Consider now the following pairs trading strategy. We trade-in at time  $t_i$  which is the first time  $t$  such that  $|X_t| \geq a$ , where  $a$  is a preset trade-in point, and trade-out at time  $t_0$  which is the first time  $t > t_0$  such that

$$\begin{cases} X_t \leq 0 \text{ if } X_{t_i} > 0 \\ X_t \geq 0 \text{ if } X_{t_i} < 0 \end{cases} \quad (4)$$

Note that  $\mathbb{E}X_t = 0$  because the series  $X_t$  is a stationary without intercept  $AR(1)$  process and has a Normal distribution. Then, the distance from the mean to  $a$  and the distance from the mean to  $-a$  be the same. Thus, as soon as we decide the trade-in point  $a$ , we can use its symmetric point  $-a$  as another trade-in point for the process. The critical threshold  $a$  is a decision variable and has to be chosen so as to maximize the long-run expected total profit per unit time. Let us display this trading process in Figure 4.1.

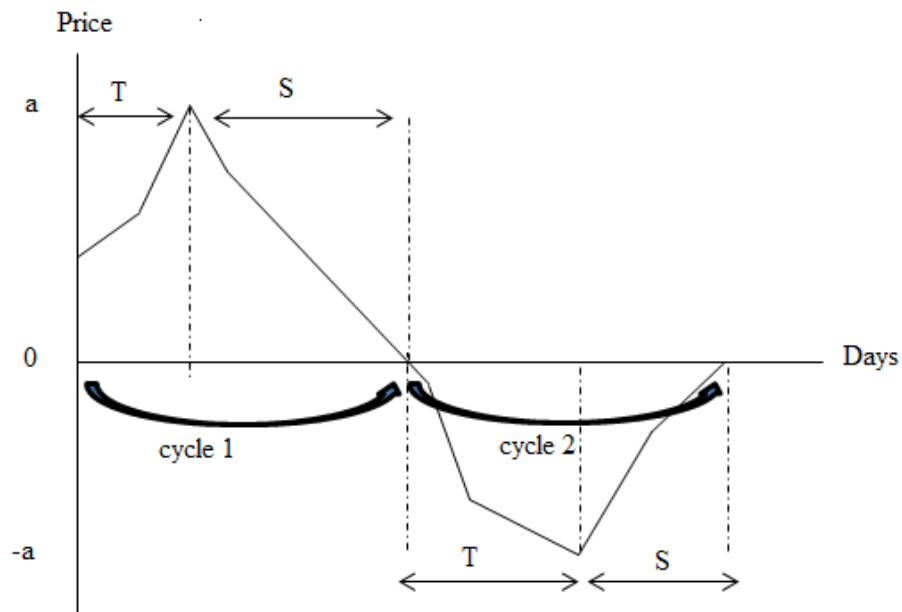


Figure 4.1: Illustration of trading process

Let  $T$  be the duration up to we trade-in and  $S$  be the duration between the points we trade-in and trade-out. Then we can define the expected cycle length of a trade as

$$\mathbb{E}C(x) = \mathbb{E}_x[T + \mathbb{E}_{X_T}S] \quad (5)$$

where  $X_0 = x$ ,  $-a < x < a$  and  $\mathbb{E}_x$  is the conditional expectation given  $X_0 = x$ . Then,  $a / \mathbb{E}C(x)$  becomes the long-run expected reward per unit time by renewal reward theorem. Notice the trade off in the choice of point  $a$ ; as  $a$  increases maximum profit increases; however,  $T$  and  $C(x)$  also increase. Thus, the relationship between  $a$  and the profit per unit time is not trivial.

To simplify the problem, we discretize both time and state space of  $X$  and assume that  $X$  is a Markov Chain. Let us tackle this problem in two parts. Firstly, let us consider the trade-in process, and then we will consider the trade-out process.

Let  $\hat{E}$  be the state space for trade-in process as (6).

$$\hat{E} = \{-a + 1, \dots, -1, 0, 1, \dots, a - 1, a, -a\} \quad (6)$$

Let  $\hat{P}$  be the one-step transition probability matrix for the process to trade-in. Note that  $a$  and  $-a$  be absorbing states for trade-in process. In other words, the trade-in process ends when  $X_t$  touches to points  $a$  or  $-a$ . Notice that  $a$  and  $-a$  are not absorbing states for the whole process; however we temporarily makes this assumption in order to calculate the length of trade-in process. We partition the  $\hat{P}$  matrix as in Table 4.1, so that the  $\hat{Q}$  component of the matrix has the entries  $ij$  where  $i$  and  $j$  are both transient states and the  $\hat{R}$  component has entries  $ij$  where  $i$  is transient but  $j$  is absorbing. In the matrix,  $\hat{O}$  shows the zero matrix and  $\hat{I}$  shows the identity matrix.

Table 4.1: Probability matrix representation for process of the strategy

$$\hat{P} = \begin{array}{|c|c|} \hline \hat{Q} & \hat{R} \\ \hline \hat{0} & \hat{I} \\ \hline \end{array}$$

Then the probability of  $X_t$  being at point  $a$  at time  $T$ , starting from point  $x$ , becomes

$$\hat{\mathcal{G}}(x, a) := P(X_T = a | X_0 = x) = \hat{W} \hat{R}_{(x,a)} \quad (7)$$

where

$$\hat{W} = \hat{I} + \hat{Q} + \hat{Q}^2 + \hat{Q}^3 + \dots = (\hat{I} - \hat{Q})^{-1} \quad (8)$$

as shown for example by Karlin and Taylor [16].

Now let us remove the assumption that  $a$  and  $-a$  are absorbing states and define our real state space as in (9).

$$E = \{-b, -b + 1, \dots, -a, -a + 1, \dots, -1, 0, 1, \dots, a, a + 1, \dots, b - 1, b\} \quad (9)$$

where  $b$  is defined by

$$b = 5\sigma / (1 - \phi) \quad (10)$$

We use  $b$  value as an approximate limit for the points  $X_t$  can take at most. The value of  $b$  ensures that the probability that  $X$  go outside  $[-b, b]$  is negligible, because,

$$X_1 | X_0 = x \sim \text{Gsn}(\phi x, \sigma^2) \quad (11)$$

We discretize  $[-b, b]$  in a regular grid with step size  $\Delta b$ . Then, the state  $E$  has a total of  $2b / \Delta b + 1$  states. Then, we can construct the  $P$  matrix by

$$\begin{aligned}
P_{ij} &= P \left\{ \frac{(-b + (j-1)\Delta b - b + j\Delta b)}{2} < X_1 < \frac{(-b + j\Delta b - b + (j+1)\Delta b)}{2} \mid X_0 \right. \\
&\quad \left. = -b + i\Delta b \right\} \\
&= \int_{\frac{(-b+(j-1)\Delta b - b + j\Delta b)/2}{\sigma\sqrt{2\pi}}}^{\frac{(-b+j\Delta b - b + (j+1)\Delta b)/2}{\sigma\sqrt{2\pi}}} \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left\{-\frac{(y - \phi(-b + i\Delta b))^2}{2\sigma^2}\right\} dy \\
&= \Phi\left(\frac{\frac{(-b+(j+1)\Delta b - b + j\Delta b)}{2} - \phi(-b + i\Delta b)}{\sigma}\right) - \Phi\left(\frac{\frac{(-b+(j-1)\Delta b - b + j\Delta b)}{2} - \phi(-b + i\Delta b)}{\sigma}\right) \tag{12}
\end{aligned}$$

where  $-b \leq j \leq b$ . Note that  $\sigma$  represents the standard error of regression of the process and the mean of the process becomes  $\phi(-b + i\Delta b)$  where  $\phi$  is the coefficient of the  $AR(1)$  process. If  $j = -b$ ,

$$\begin{aligned}
P_{ij} &= P\left\{X_1 < \frac{(-b-b+1)}{2} \mid X_0 = -b + i\Delta b\right\} \\
&= \Phi\left(\frac{\frac{(-b-b+1)}{2} - \phi(-b + i\Delta b)}{\sigma}\right) \tag{13}
\end{aligned}$$

If  $j = b$ ,

$$\begin{aligned}
P_{ij} &= P\left\{\frac{(b+b-1)}{2} < X_1 \mid X_0 = -b + ib\right\} \\
&= 1 - \Phi\left(\frac{\frac{(b+b-1)}{2} - \phi(-b + i\Delta b)}{\sigma}\right) \tag{14}
\end{aligned}$$

The expected cycle time of one trade is represented as shown in (2), where  $T$  is the time it takes until the process trades in and  $S$  shows the time it takes until the process trades out. Then,

$$EC(x) = \mathbb{E}_x T + \mathbb{E}_x [\mathbb{E}_{X_T} S] \tag{15}$$



Let us define  $V(x)$  and  $U(x)$  as (15) and (16) in order to restate  $\mathbb{E}C(x)$  in a closed form.

$$V(x) = \mathbb{E}_x T \quad (16)$$

$$U(x) = \mathbb{E}_x S \quad (17)$$

So,  $V(x)$  shows the expected time it takes until the process trades in and  $U$  function shows the expected time it takes until the process trades out when it starts from the trade-in point  $x$  at time  $T$ . Then, the expected cycle length of one trade becomes  $V(x) + \mathbb{E}_x[U(X)]$ . Then,

$$\mathbb{E}C(x) = V(x) + \sum_{-b \leq y \leq -a \text{ or } a \leq y \leq b} U(y) P\{X_T = y | X_0 = x\} \quad (18)$$

and

$$\mathbb{E}C(x) = V(x) + \sum_{-b \leq y \leq -a \text{ or } a \leq y \leq b} U(y) \mathfrak{C}(x, y) \quad (19)$$

Note that,  $V$  and  $U$  functions can also be shown as in (20) and (21) respectively, using the representation in (7). That is, (17) shows the expected number of times the process visits point  $j$  before being absorbed into  $\pm a$ , if the process starts from point  $i$ , where  $-a < i < a$  and  $-a < j < a$ . In addition, (18) shows the number of steps it takes the process to reach to point  $j$  if it starts from the point  $i$ , where  $i$  is between the points  $-b$  and  $-a$  or between the points  $a$  and  $b$ .

$$V(i) = \sum_{-a < j < a} \widehat{W}(i, j) \quad (20)$$

and

$$U(i) = \sum_{-b \leq j \leq -a \text{ or } a \leq j \leq b} W(i, j) \quad (21)$$

Note that  $\widehat{W}(i, j)$  in (20) comes from the calculations of trade-in process while  $W(i, j)$  in (21) comes from the calculations of trade-out process.

# Chapter 5

## Case Study

### **5.1 Stock Price Data from BIST**

In this study, we work on daily stock prices of six banks in BIST (Istanbul Stock Exchange), Akbank, Şekerbank, Vakıfbank, Garanti Bank, Fortis and Yapı Kredi Bank. We use these stocks for the implementation of the method we develop. We analyze the daily prices of the stocks between years 2008 and 2010.

We check whether there exists cointegration of order (1,1) between the banks' stock prices we fit to the stationary series, created by the cointegrating linear combinations of the cointegrated pairs and reveal the best-fit time-series models. We also check for the parameters of these time-series models to obtain the most common values those parameters take for the cointegrated stock series we concern. We decide which type of time series model the cointegrated series of Turkish banking stock prices fit the best. Then, we implement our pairs trading method according to that model and the most common estimated parameters. Finally, we measure the performance of our method on the cointegrated pairs of the banking stock prices. We use the statistical package program EViews 5 (Econometric Views) for the analyses of the data.

We first check the stocks for stationarity and we found that all stock price series are integrated of order 1. Then, we check for cointegration between the stock prices in order to detect possible pairs. The cointegration tests revealed, 6 cointegrated stock pairs. The cointegration test outputs and the cointegration equations can be found in Tables 5.1 to 5.12 for the pairs. The time series graphs of all the stock pairs are reported in Appendix A. Let us take a look at the patterns of two pairs, one of which has cointegration relationship and the other does not have, in order to illustrate the influence of cointegration on series. Figure 5.1 presents the time series graphs of Fortis and Akbank stock prices. Although there does not exist a perfect match between these two patterns, both stock prices decrease concurrently in the first 300 days, and then start increasing afterwards. The tests show that two stocks are cointegrated.

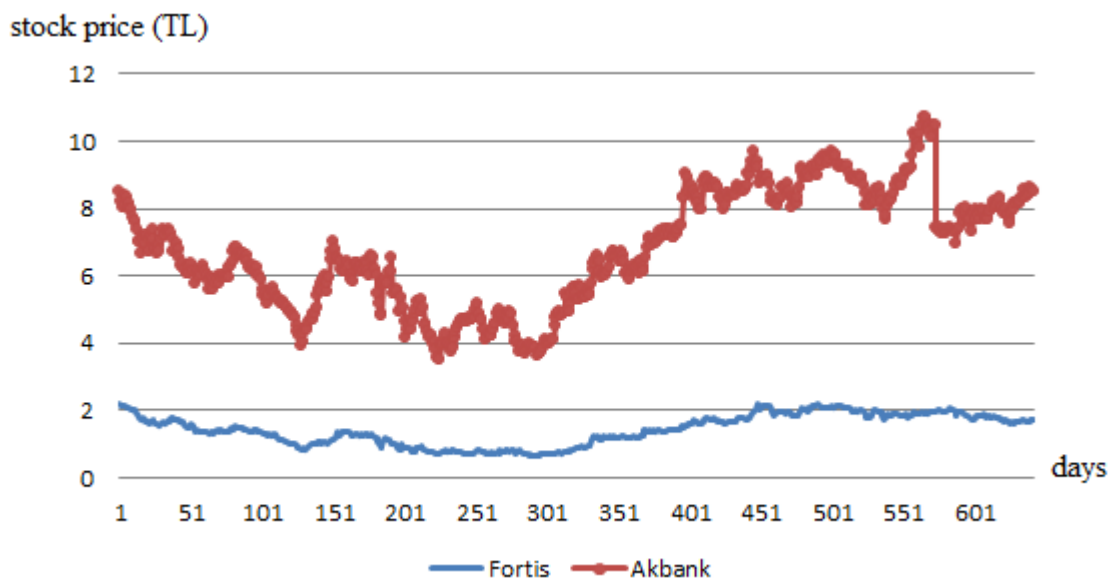


Figure 5.1: Time series graph of Fortis and Akbank stock prices

As a contrary example, let us inspect two stock prices which are not cointegrated. Figure 5.2 presents time series graphs of Yapi Kredi Bank and Vakıfbank stock prices. Although the stock prices fluctuate similarly, each follows a distinct pattern. For example, around the 350<sup>th</sup> day Yapi Kredi Bank stock price decreases while Vakıfbank

stock price increases abruptly. In addition, around 550<sup>th</sup> day Vakıfbank stock price suddenly decreases while Yapi Kredi Bank stock price has a continuous increase there. Those signs make a first warning for us for that these two stocks may not be cointegrated. After we conduct cointegration test, the results confirm that Yapi Kredi Bank and Vakıfbank stock prices are not cointegrated. Note that, although the time series graphs often give clear signs for the lack of cointegration between two series, we need to conduct cointegration tests to confirm or reject it.

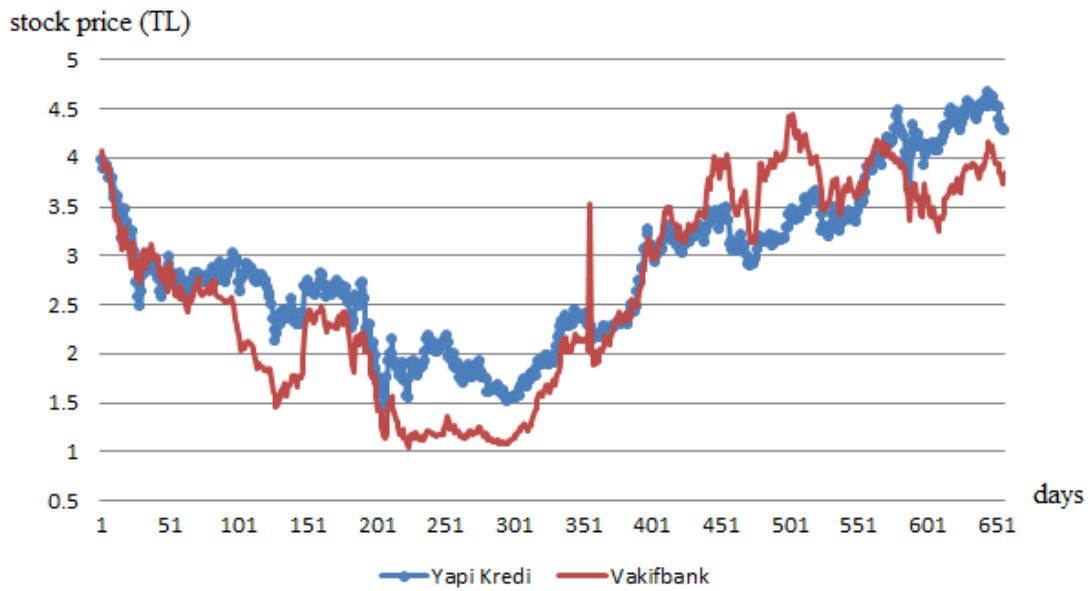


Figure 5.2: Time series graphs of Yapi Kredi Bank and Vakıfbank stock prices

Table 5.1 presents the Johansen Cointegration Test output for the pair Akbank-Fortis. In the first row, Johansen's Trace Test tests the null hypothesis that the series are not cointegrated, and rejects the null hypothesis with p-value=0.0002. In the second row, the test output for the null hypothesis that there is at most 1 cointegration vector is presented. We cannot reject the null hypothesis since the p-value is 0.7506. The cointegration coefficients are reported in Table 5.2. Thus,

$$Akbank - 3.734326 * Fortis - 1.499182 = X_t \quad (21)$$

is a stationary process.

Table 5.1: Cointegration test output for stocks Akbank and Fortis

Date: 08/28/10 Time: 12:11				
Sample (adjusted): 3,642				
Included observations: 640 after adjustments				
Trend assumption: No deterministic trend (restricted constant)				
Series: Akbank – Fortis				
Lags interval (in first differences): 1 to 1				
Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value	Prob. **
None *	0.051549	36.33882	20.26184	0.0002
At most 1	0.003847	2.466900	9.164546	0.7506
Trace test indicates 1 cointegrating egn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
** MacKinnon-Haug-Michelis (1999) p-values				

Table 5.2: Normalized cointegrating coefficients for Akbank and Fortis (standard errors in parentheses and c represents intercept coefficient)

Akbank	Fortis	c
1	-3.734326 (0.23982)	-1.499182 (0.36042)

We apply the cointegration test to Fortis-Vakıfbank pair in Table 5.3. The Trace Test rejects that null hypothesis that series are not cointegrated with  $p\text{-value}=0.0022$  at  $\alpha=0.05$ . On the other hand, the test output for the null hypothesis that there are at most 1

cointegration vector cannot be rejected because the p-value is 0.8413. Table 5.4 presents the cointegrating coefficients for that pair. Thus,

$$Fortis - 0.442035 * Vakıfbank - 0.243589 = X_t \quad (22)$$

is a stationary process.

Table 5.3: Cointegration test output for stocks Fortis and Vakıfbank

Date: 08/28/10 Time: 12:35				
Sample (adjusted): 3,657				
Included observations: 655 after adjustments				
Trend assumption: No deterministic trend (restricted constant)				
Series: Fortis - Vakıfbank				
Lags interval (in first differences): 1 to 1				
Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value	Prob. **
None *	0.041251	29.26624	20.26184	0.0022
At most 1	0.003847	1.673774	9.164546	0.8413
Trace test indicates 1 cointegrating egn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
** MacKinnon-Haug-Michelis (1999) p-values				

Table 5.4: Normalized cointegrating coefficients for Fortis and Vakıfbank (standard errors in parentheses and c represents intercept coefficient)

Fortis	Vakıfbank	c
1	-0.442035 (0.01913)	-0.243589 (0.05491)

Table 5.5 presents the Johansen's Test output for Fortis and Şekerbank pair. The Trace Test rejects that null hypothesis that there is at most 0 cointegration equations with p-value=0.0424 at  $\alpha= 0.05$ . Yet, the test output for the null hypothesis that there are at most 1 cointegration vector cannot be rejected since the p-value is 0.8573. Table 5.6, shows the cointegrating coefficients for Fortis-Şekerbank pair. So, the cointegrating equation becomes

$$Fortis - 1.308419 * Şekerbank + 1.035947 = X_t, \quad (23)$$

where  $X_t$  is a stationary process.

Table 5.5: Cointegration test output for stocks Fortis and Şekerbank

Date: 08/28/10 Time: 12:45				
Sample (adjusted): 3,657				
Included observations: 655 after adjustments				
Trend assumption: No deterministic trend (restricted constant)				
Series: Fortis - Şekerbank				
Lags interval (in first differences): 1 to 1				
Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value	Prob. **
None *	0.027050	19.54978	20.26184	0.0424
At most 1	0.002422	1.588192	9.164546	0.8573
Trace test indicates 1 cointegrating egn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
** MacKinnon-Haug-Michelis (1999) p-values				

Table 5.6: Normalized cointegrating coefficients for Fortis and Şekerbank  
(standard errors in parentheses and c represents intercept coefficient)

Fortis	Şekerbank	c
1	-1.308419 (0.25553)	1.035947 (0.58714)

Table 5.7 presents the test output for Vakıfbank-Şekerbank. Johansen's Trace Test tests the null hypothesis that there are at most 0 cointegration vectors in the first row, and rejects that null hypothesis with p-value=0.0498. However, we cannot reject the null hypothesis that there are at most 1 cointegration vector since the p-value is 0.8335. In Table 5.8, we present the coefficients of that cointegration vector, whose explicit equation is

$$Vakifbank - 3.257224 * Şekerbank + 3.375803 = X_t, \quad (24)$$

where  $X_t$  is a stationary process.



Table 5.7: Cointegration test output for stocks Vakıfbank and Şekerbank

Date: 08/28/10 Time: 13:06				
Sample (adjusted): 3,661				
Included observations: 659 after adjustments				
Trend assumption: No deterministic trend (restricted constant)				
Series: Vakıfbank - Şekerbank				
Lags interval (in first differences): 1 to 1				
Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value	Prob. **
None *	0.026153	19.17940	20.26184	0.0498
At most 1	0.002599	1.715075	9.164546	0.8335
Trace test indicates 1 cointegrating egn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
** MacKinnon-Haug-Michelis (1999) p-values				

Table 5.8: Normalized cointegrating coefficients for Vakıfbank and Şekerbank  
(standard errors in parentheses and c represents intercept coefficient)

Vakıfbank	Şekerbank	c
1	-3.257224 (0.69325)	3.375803 (1.59043)

In Table 5.9, we observe that the Trace Test rejects that null hypothesis that there is at most 0 cointegration equations with p-value=0.0002 at  $\alpha=0.05$  for Akbank and Vakıfbank stock prices. On the other hand, the test output for the null hypothesis that there are at most 1 cointegration vector cannot be rejected because the p-value is 0.7506. Table 5.10 presents the cointegrating coefficients and (25) shows the cointegrating equation for that pair.

$$Akbank - 1.658613 * Vakıfbank - 2.389100 = X_t, \quad (25)$$

where  $X_t$  is a stationary process.

Table 5.9: Cointegration test output for stocks Akbank and Vakıfbank

Date: 08/28/10 Time: 13:38				
Sample (adjusted): 3,642				
Included observations: 640 after adjustments				
Trend assumption: No deterministic trend (restricted constant)				
Series: Akbank - Vakıfbank				
Lags interval (in first differences): 1 to 1				
Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value	Prob. **
None *	0.050353	35.20108	20.26184	0.0002
At most 1	0.003331	2.135571	9.164546	0.7506
Trace test indicates 1 cointegrating egn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
** MacKinnon-Haug-Michelis (1999) p-values				

Table 5.10: Normalized cointegrating coefficients for Akbank and Vakıfbank (standard errors in parentheses and c represents intercept coefficient)

Akbank	Vakıfbank	c
1	-1.658613	-2.389100
	(0.09609)	(0.27273)

Finally, Table 5.11 presents the Johansen's Test outputs for Akbank and Şekerbank, in which we observe that the Trace Test rejects that null hypothesis that there is at most 0 cointegration equations with p-value=0.0108 at  $\alpha=0.05$  for Akbank and Şekerbank

stock prices. Yet, the test output for the null hypothesis that there are at most 1 cointegration vector cannot be rejected since the p-value is 0.4205. Table 5.12, shows the cointegrating coefficients for Akbank-Şekerbank pair. Additionally, the cointegrating equation is shown in (26).

$$Akbank - 3.404745 * Şekerbank - 0.416574 = X_t, \quad (26)$$

where  $X_t$  is a stationary process.

Table 5.11: Cointegration test output for stocks Akbank and Şekerbank

Date: 08/28/10 Time: 13:44				
Sample (adjusted): 3,642				
Included observations: 640 after adjustments				
Trend assumption: No deterministic trend (restricted constant)				
Series: Akbank - Şekerbank				
Lags interval (in first differences): 1 to 1				
Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value	Prob. **
None *	0.032165	24.86763	20.26184	0.0108
At most 1	0.006143	3.943856	9.164546	0.4205
Trace test indicates 1 cointegrating egn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
** MacKinnon-Haug-Michelis (1999) p-values				

Table 5.12: Normalized cointegrating coefficients for Akbank and Şekerbank (standard errors in parentheses and c represents intercept coefficient)

Akbank	Şekerbank	c
1	-3.404745 (0.66968)	-0.416574 (1.54808)

Yapi Kredi and Garanti stock prices have no cointegration relationship with any of the banking stock prices. In Figure 5.3, time series graphs of these stocks are presented in the same plot.

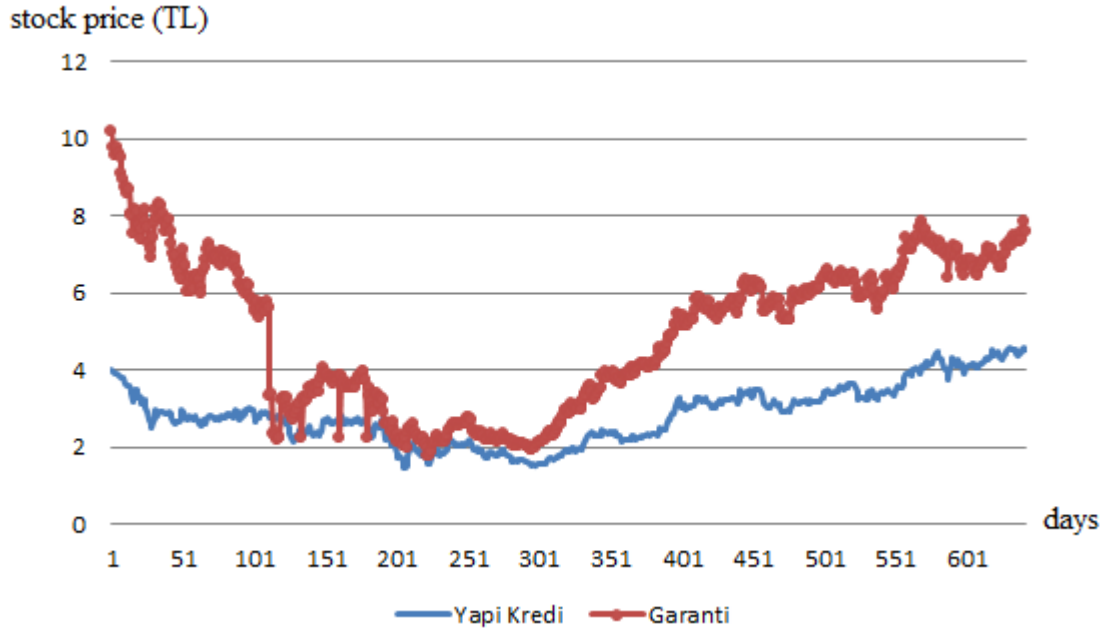


Figure 5.3: Time series graphs of Yapi Kredi Bank and Garanti Bank stock prices

Revealing the cointegrated stock prices, we can continue with time series model analysis for the stationary series formed by the cointegrated pairs. According to the time series analyses we have conducted, five series turned out to fit to  $AR(1)$  without intercept and

one of them, which is the cointegration series of Fortis and Vakıfbank, appears to fit to  $AR(2)$  model. Because most of the pairs fit to  $AR(1)$  model, we pretend that the cointegrated combinations of Turkish banking stock prices follow  $AR(1)$  processes, and we implement our method to Turkish banking stocks case with this assumption. In addition, we estimate the parameters of those  $AR(1)$  models in order to see most common parameter values in the banking sector, which are regression coefficient, i.e.  $\emptyset$ , and the standard error of regression, i.e.  $\sigma$ , for  $AR(1)$  models, so that we can use when we test our method on these data as empirical work in the rest of the study. The estimated parameter values of  $AR(1)$  models are reported in Table 5.13. In addition, you can find the model outputs of the stationary series formed from all pairs in Appendix B.

Table 5.13: Parameter estimations of the  $AR(1)$  models

<b>Pair</b>	$\emptyset$	$\sigma$
Fortis - Akbank	0.942	0.058
Fortis - Şekerbank	0.985	0.088
Vakıfbank - Şekerbank	0.986	0.236
Akbank - Vakıfbank	0.903	0.245
Akbank - Şekerbank	0.987	0.259

We see that the estimated  $\emptyset$  values are between 0.903 and 0.987. For our empirical work, we set  $\emptyset$  values 0.95, 0.96, 0.97, 0.98, 0.99 and examine the results for these parameter values. In addition, the estimated standard error of regression values turn out to be between 0.058 and 0.259 for the cointegrating series of the stock pairs. We set standard error of regression values 0.05, 0.1 and 0.2 for the empirical study.

Although we have stated that those 5 pairs fit to  $AR(1)$  model, they cannot meet the necessary requirements of diagnostic checks for the residuals, such as homoscedasticity or normality of the residuals. Yet, we assume they do and we continue with the assumption that the cointegrated series in concern fit to  $AR(1)$  without intercept model,

and we apply our method to  $AR(1)$  models with  $\phi$  between 0.95 and 0.99, and standard error of regression values 0.05, 0.1 and 0.2.

## 5.2 Profit per Unit Time for $X_0 = 0$

We have the probabilities to construct the P matrix. Before we obtain optimal trade-in values, i.e. optimal  $a$  and  $-a$  points, let us first calculate long-run expected profit per unit time if the process starts from point 0, i.e.  $a / EC(0)$  for the inputs of  $\sigma, \phi, \Delta b, b$  and  $a$ . The reason we start with the case  $X_0 = 0$  is that 0 is the farthest point to both trade-in points; thus, this case shows the longest time for the process to trade-in. In other words, we start constructing the method for the worst case. However, recall that our aim is to find the optimal, i.e. the most profitable,  $a$  points for  $a/EC(x)$  function. Thus, we will extend our study in order to find the most profitable  $a$  points wherever the process starts at.

We take standard error of regression values, i.e.  $\sigma$  values as 0.05, 0.1, 0.2, and regression coefficient values, i.e.  $\phi$  values, as 0.95, 0.96, 0.97, 0.98, 0.99. Remember that those  $\sigma$  and  $\phi$  values are decided from  $AR(1)$  model estimations of the banks' data in the previous section. We take  $\Delta b$  values, which represent the step sizes of the process, at 0.005, 0.01 and 0.02 for our estimations. It is obvious that estimating at different step sizes negatively affects our comparisons. However, we cannot calculate them all at the same  $\Delta b$  value since we do the calculations at MATLAB R2007b and as  $\sigma$  and  $\phi$  values increase MATLAB cannot finish computations in a reasonable time amount unless we also increase step sizes, i.e.  $\Delta b$  values.

Let us illustrate the results in Figure 5.4, where  $\sigma$  is taken as 0.05. In the figure, the x-axis shows the trade-in points  $a$  and the y-axis shows the long-run expected profit per unit time values. The step size for  $\sigma = 0.05$  is taken as 0.005. Having estimated the optimal objective values, i.e.  $a/EC(0)$ , for all parameter pairs  $(\phi, \sigma)$ , we see that the results do not change significantly for different  $\phi$  values at the same  $\sigma$  value. Thus, even if we fail to estimate the  $\phi$  value correctly for some  $\sigma$ , we still obtain an objective value

close to the one for the correct  $\phi$  value. Also note that the objective value as a function of the trade-in value seems to be concave for these  $\sigma$  values. Moreover, long-run expected profit per unit time decreases as  $\phi$  increases. That is, for  $\phi = 0.95$  the long-run expected profit per unit time is the biggest and for  $\phi = 0.99$  the long-run expected profit per unit time is the smallest for the  $\sigma$  values we concern. The intuition behind this relationship can be explained as

$$X_t - X_{t-1} = (\phi - 1)X_{t-1} + \varepsilon_t \quad (27)$$

$(\phi - 1)$  shows the change, or speed, of the  $AR(1)$  process. Since  $\phi$  takes value between  $-1$  and  $1$ ,  $(\phi - 1)$  takes values between  $-2$  and  $0$ . Then, as  $\phi$  increases,  $(\phi - 1)$  approaches to  $0$ . So, the process slows down.

In addition, note that as  $\phi$  increases, optimal  $a$  value also increases. Optimal trade-in points are as followings:  $a^* = 0.135$  for  $\phi = 0.95$ ,  $a^* = 0.145$  for  $\phi = 0.96$ ,  $a^* = 0.165$  for  $\phi = 0.97$ ,  $a^* = 0.190$  for  $\phi = 0.98$  and  $a^* = 0.245$  for  $\phi = 0.99$ , respectively.

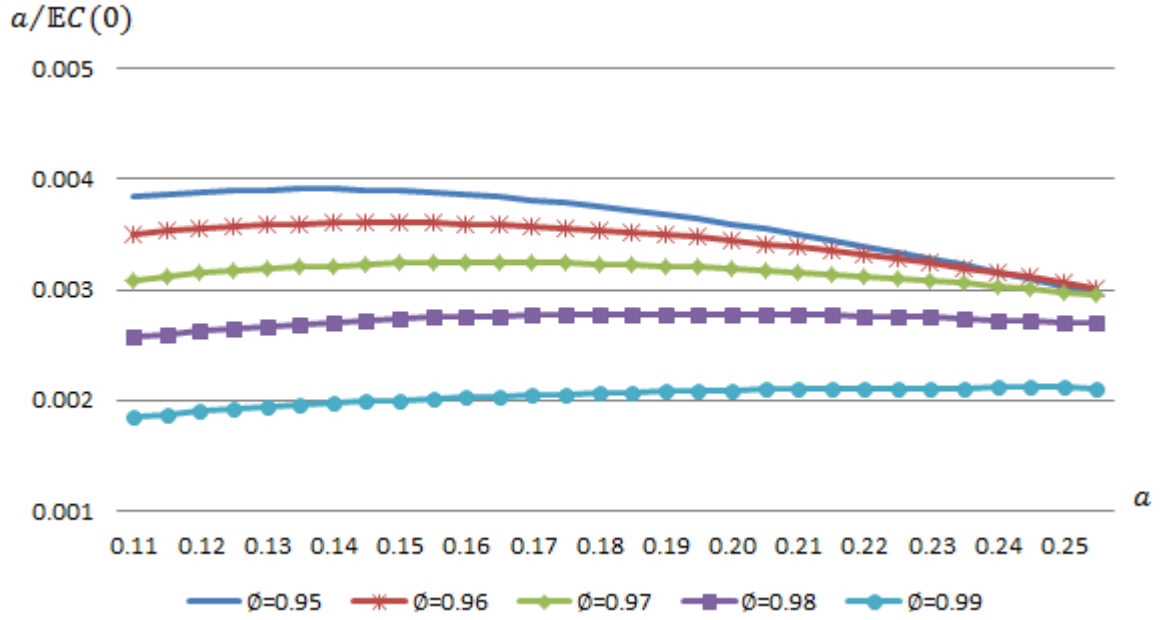


Figure 5.4: Profit per unit time graph for  $X_0 = 0$  and  $\sigma = 0.05$

The graphs illustrating profit per unit time for different trade-in points are included in Appendix C for  $\sigma$  values 0.1 and 0.2.

### 5.3 Profit per Unit Time for a Random Starting Point

Now, let us remove the restriction we consider in the previous section that the starting point of the trading process is 0, and let the process start at any point between  $-a$  and  $a$ . Because we do not know at which point the process starts, we need the distribution of the starting point  $X_0$ .

Note that, although we do not have random starting points for the process, we have trade-out points when the process starts at point 0. We will use this information in order to find the distribution of starting points. Let us conduct a simulation such that we check if the trade-out points come from the same distribution wherever the starting point is. The reason we want to check this assumption is that if it is revealed that the trade-out points come from a stationary distribution, then the starting points also come from this stationary distribution since each trade-out point is a starting point for the next trade



cycle. We can make this assumption since we work on a nonstop trading strategy. Because the trade goes on forever, if it exists, starting points must come from a stationary “long-run starting point” distribution. The purpose of this section is to check existence and to calculate, if it exists, that stationary distribution.

We conducted simulation at MATLAB for this investigation. Because we do not know the long-run starting point of the process, the first trading cycle starts at 0. So, we start with simulating the case for  $X_0 = 0$  and let the process trade-in and out for 100 times. Then, we record the 100<sup>th</sup> trade-out point, and we repeat this for 1000 times. We write down those 1000 trade-out points.

Secondly, since the trading should continue indefinitely, we take those 1000 trade-out points as possible starting points. Then, we simulate the case when one of those 1000 possible starting points is chosen as  $X_0$  at random, and again let the process trade-in and out until the 100<sup>th</sup> trade-out and record that trade-out point. We repeat that process 1000 times and record those trade-out points as trade-out points when the process starts at a random point. Now, we have two samples, the one with 1000 trade-out points for  $X_0 = 0$  and the one with 1000 trade-out points for a random  $X_0$ .

As a next step, we want to check if these two samples come from the same distribution for each parameter pair  $(\emptyset, \sigma)$  to find out if the trade-out points of our process has a stationary distribution. We conduct those simulations for the  $\emptyset$  values 0.95 and 0.99, and for  $\sigma$  values 0.05, 0.1, 0.2 at their optimal trade-in points based on long-run expected profit per unit time when the process starts at point 0, i.e.  $a/EC(0)$ , that we have computed before.

We use Two-Sample Kolmogorov-Smirnov Tests (K-S Test) and Q-Q Plots at R and MATLAB in order to check this hypothesis.

The null hypothesis of Kolmogorov-Smirnov Tests is that the samples are drawn from the same distribution. Kolmogorov-Smirnov Tests and Q-Q Plots confirmed our

assumption that the two samples of trade-out points come from the same distribution for all parameter estimations. So, the trade-out points have stationary distribution wherever the process starts at. Thus, we can also conclude that the starting points of our process have also stationary distribution.

Figures 5.5 and 5.6 present density and Q-Q plots for  $\sigma = 0.05$ ,  $a = 0.135$  and  $\emptyset = 0.95$ . In Appendix D, D statistics and p-values of Two Sample Kolmogorov-Smirnov Test, the probability and cumulative density graphs, and the Q-Q Plots for all  $(\emptyset, \sigma)$  pairs are reported.

As an example, let us examine the density graph and Q-Q Plot for  $\emptyset = 0.95$ ,  $a = 0.135$  and  $\sigma = 0.05$ . Figure 5.5 is the histogram plot for the two series, the series of the trade-out points when the trading process starts from 0, and the series of the trade-out points when the process starts from a random point. According to the graphs, those two series seem fit to the same distribution. Q-Q plot in Figure 5.6 also supports our hypothesis that they come from the same distribution. We see that the Q-Q follows the  $45^\circ$  line  $y=x$ , which suggests the identical distributions according to this Probability Plot. The p-value for the Two Sample K-S Test is 0.9689, which fails to reject the null hypothesis of the test that the two samples come from the same continuous distribution. Although, the density functions in Figure 5.5 resemble the Gaussian distribution, the one sample K-S Test reveals that they are different than Gaussian distribution. Because we only need to know that the two samples come from the same distribution, our study is not affected.

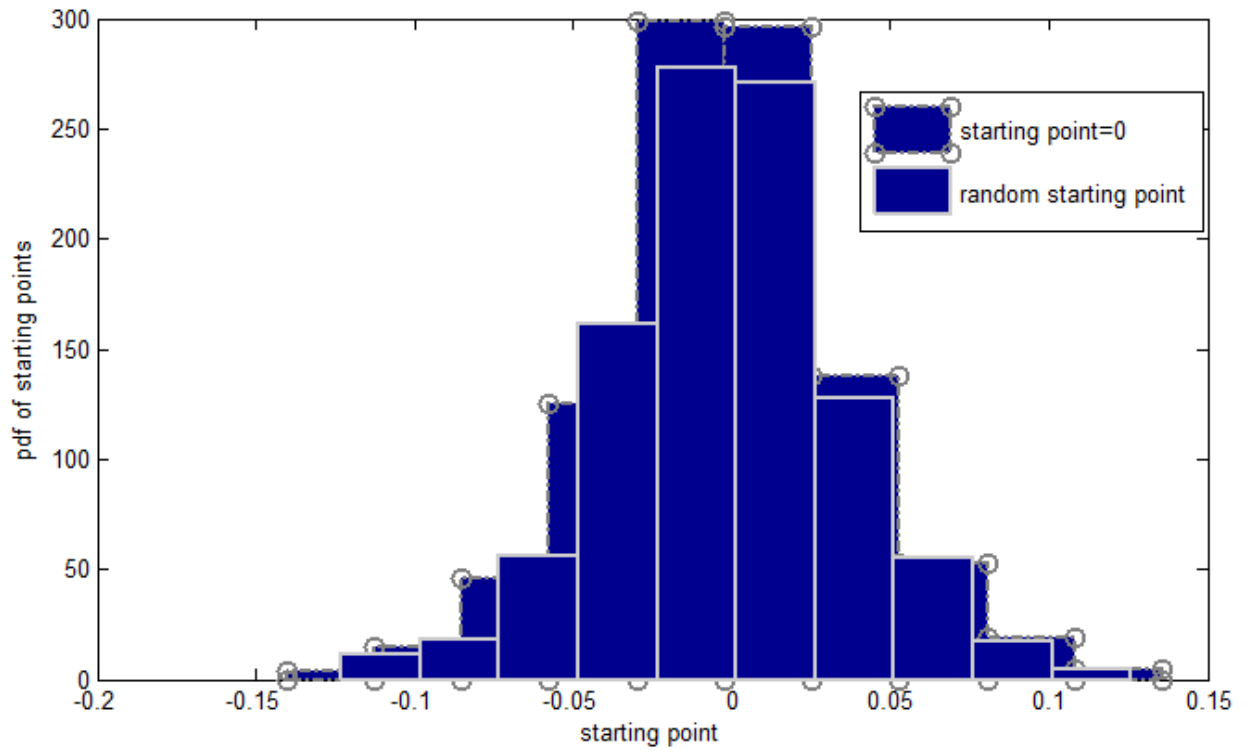


Figure 5.5: Density graph for  $\phi = 0.95$ ,  $a = 0.135$  and  $\sigma = 0.05$

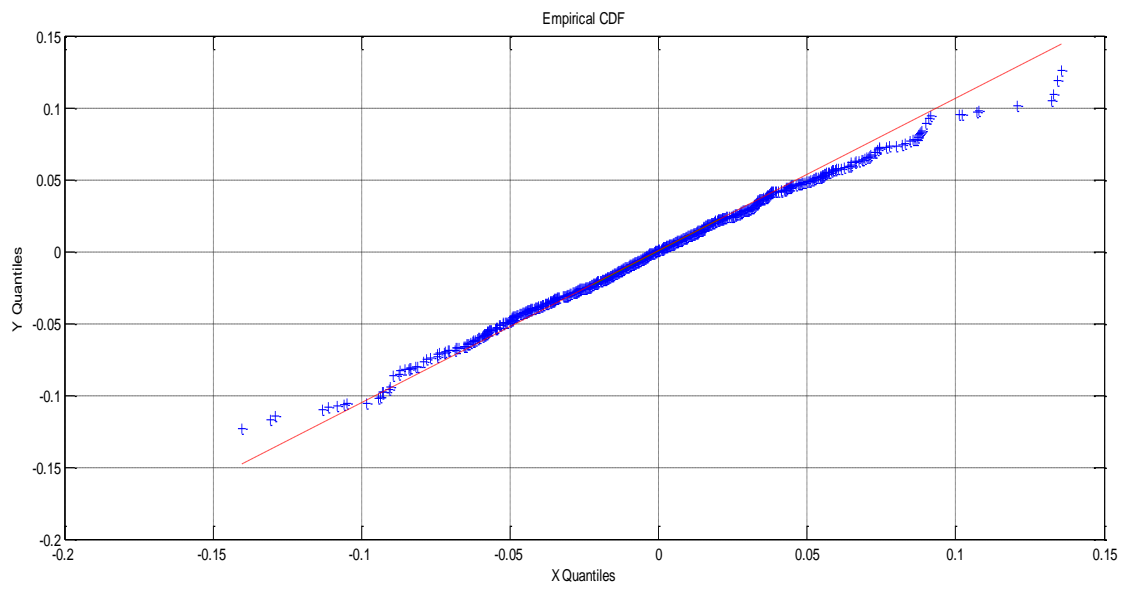


Figure 5.6: Q-Q plot for  $\phi = 0.95$ ,  $a = 0.135$  and  $\sigma = 0.05$

Now we are ready to find the long-run expected profit per unit time values when the process starts at a random point from the long-run starting point distribution. Then, we need to find the long-run expected value of profit per unit time for a predetermined trade-in point.

Expected cycle duration, i.e.  $\mathbb{E}C(x)$ , is calculated by

$$\mathbb{E}C(X) = \sum_{-b < j < b} P(j)C(j) \quad (28)$$

where  $X_0$  is distributed according to the stationary distribution of trade-out point under given trade-in threshold. The calculations are again done at MATLAB R2007b. Firstly,  $C$  vector, representing the cycle length, is calculated for the given inputs. Then, a simulation takes place to record 1000 trade-out points for the process. Recall that trade-out points also act as starting points of the next process in our nonstop process. Then we use those 1000 points to calculate the probability distribution of starting points, i.e.  $P(j)$ . We calculate the probabilities using the frequencies of the points out of 1000 points. Finally, profit per unit time values  $a / \mathbb{E}C(X)$  are calculated in order to find out the optimal  $a$  values.

Figure 5.7 shows the graph of  $a / \mathbb{E}C(x)$  values on the vertical axis against trade-in values on the horizontal axis for  $\sigma = 0.05$ . The step size for  $\sigma = 0.05$  is taken as 0.005. We see that the long-run expected profit per unit time graphs are also concave and smooth, and long-run expected profit per unit time also decreases as  $\emptyset$  increases as in the expected profit per unit time graphs for  $X_0 = 0$ . For  $\emptyset = 0.95$ , the long-run expected profit per unit time is the biggest; and for  $\emptyset = 0.99$ , the long-run expected profit per unit time is the smallest for the  $\sigma$  values we concern. In addition, as  $\emptyset$  increases, optimal  $a$  value also increases. Additionally, those two graphs are very similar to each other for each parameter pair  $(\emptyset, \sigma)$ . Optimal trade-in points are reported in Table 5.14.

Table 5.14: Optimal trade-in points for  $\emptyset$  values between 0.95 and 0.99 for a random starting point process

$\emptyset$	$a^*$
0.95	0.135
0.96	0.140
0.97	0.155
0.98	0.190
0.99	0.245

Note that the optimal trade-in points for  $a / EC(X)$  function are very close to the optimal trade-in points for  $a / C(0)$  function. You can find  $a / EC(X)$  graphs for all  $(\emptyset, \sigma)$  values are reported in Appendix E.

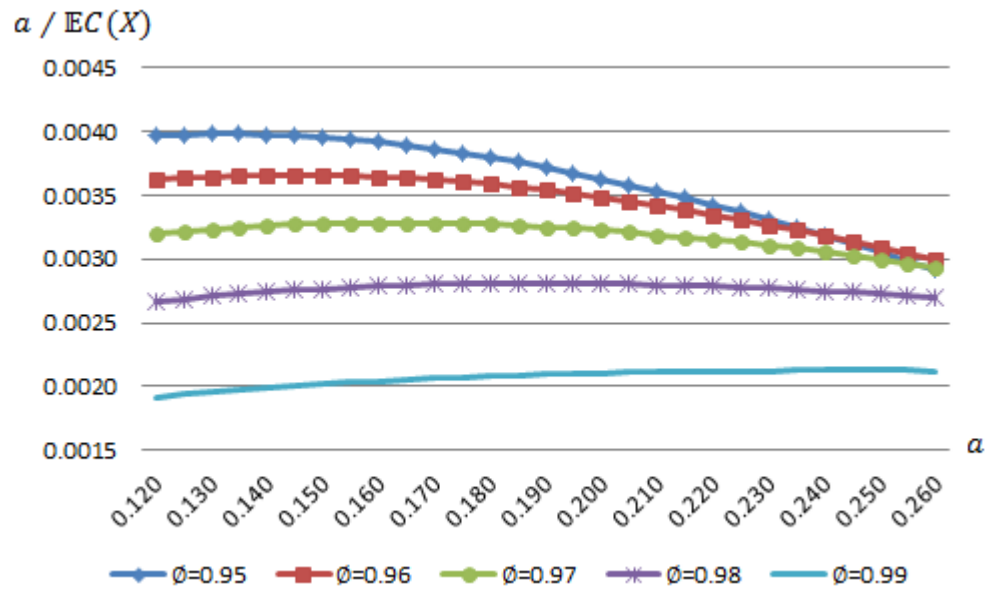


Figure 5.7: Profit per unit time graph for a random starting point and  $\sigma = 0.05$

# Chapter 6

## Numerical Results

We now investigate the performance of our method on simulated and real data in order to measure its performance.

### 6.1 Performance Measurement Using Simulated Data

In this section we conduct a simulation at MATLAB R2007b to see the performance of the method we propose. We generate  $AR(1)$  without constant series,  $X_t$ , having 0 mean, representing the stationary process in our study as (29).

$$X_t = \phi X_{t-1} + \varepsilon_t, \tag{29}$$

where  $\varepsilon_t \sim Gsn(0,1)$ . Recall that an  $AR(1)$  series is stationary as long as  $|\phi| < 1$ . We use the estimated  $\phi$  values we have found out in Chapter 5 as the regression coefficient values for the simulated  $AR(1)$  series. In other words, we simulate five  $AR(1)$  series having  $\phi$  values 0.95, 0.96, 0.97, 0.98 and 0.99. So, stationarity holds for the simulated time series. We do not set a  $\sigma$  value for the simulation, because we simulate  $AR(1)$  series having an error term  $\varepsilon_t$ , following  $Gsn(0,1)$ . In other words, we set  $\sigma$  value to 1. In the simulation process, we control the parameters the regression coefficient  $\phi$ , the trade-in point  $a$ , and the starting point  $X_0$  for the simulation process. We simulate five

$AR(1)$  series using the  $\emptyset$  values we have found out for each  $a$  value. Although we do not have to control starting point, we do in order to fairly compare the profit amounts for different  $\emptyset$  values. We set the starting point of the process randomly from Standard Normal Distribution.

We start the process with a 1000-step simulation. We use those 1000 points as daily price values  $X_t$  of our portfolio, and let the process trade-in when  $X_t \geq |a|$  and trade-out when  $X_t$  crosses over point 0. When one cycle is completed, the profit is calculated by

$$\text{Profit per cycle} = (\text{The absolute value of trade-in point of the cycle}) + (\text{The absolute value of trade-out point of the cycle}) \quad (30)$$

Recall that, since we work on discrete state space, although we indicate the trading-out rule as passing point 0, the value of the process at the first time the process crosses the point 0 is different than 0. Then the absolute value of this trade-out point contributes to our total profit as well. Also, although we set an  $a$  value as preset trade-in point, often, the first time the process passes point  $\pm a$  is reached at a number different than  $\pm a$ . Thus, the absolute value of this trade-in point contributes to our total profit.

At the end of 1000 steps, if the process has traded in at any point but has not traded out yet, we force it to trade-out, that is we add the trade-in point to the profit (since we take short position at trade-in point) and subtract the 1000<sup>th</sup> point of the simulated data  $X_{1000}$  from the profit (since we take long position at trade-out point). Finally, we divide the total profit to 1000 in order to find the profit per unit time. We record profit per unit time values with the optimal trade-in points for each  $\emptyset$  value.

The results are consistent with the analyses of Chapter 5 in that as  $\emptyset$  value increases, maximum long-run profit per unit time decreases. However, in the simulation analyses we have confronted with fluctuating lines of profit per unit time functions on the contrary of the previous analyses. The fluctuations are the result of sampling errors because we work on only one sample for each  $\emptyset$  value in concern. Since we obtain



consistent results with the previous chapter, we go on to the next step without repeating the samples to overcome the sampling error.

Profit per unit time functions are shown in Figure 6.1. The horizontal axis in the graph shows the trade-in points and the vertical axis shows the profit per unit time values corresponding to those trade-in points. The most profitable trade-in points for all  $\emptyset$  values are reported in Table 6.1.

Table 6.1: Optimal trade-in points for  $\emptyset$  values between 0.95 and 0.99 for simulated data

$\emptyset$	$a^*$
0.95	0.73
0.96	0.50
0.97	0.46
0.98	0.64
0.99	0.52

Despite the fluctuations, we can still tell that as  $\emptyset$  increases, profit per unit time value decreases. Due to the sampling errors  $a$  does not increase consistently as  $\emptyset$  increases in Figure 6.1 unlike in the Figures 5.4 and 5.7.

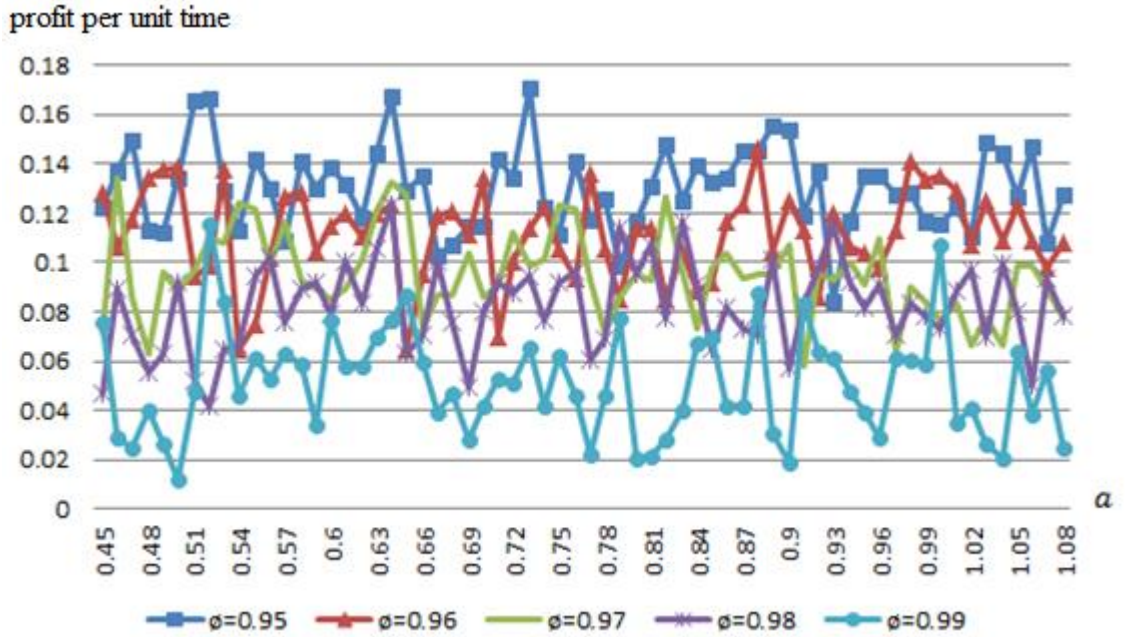


Figure 6.1: Profit per unit time graph using simulated data

## 6.2 Performance Comparison between Two Methods

In this section, we want to compare our trading strategy with a very basic alternative method based on historical data. We create a scenario that an investor observes a cointegrated pair of stock series for a period of time and decides optimal trade-in point for those series using the past observations. Then, he invests using pairs trading according to those optimal points. The investor does not use a statistical model to decide the optimal trade-in point. He only observes and find out at which point the portfolio pays the most for the historical data. Then, he uses this point as trade-in point for present data. On the other hand, we model the past data with an  $AR(1)$  series and, we estimate the parameter values  $\phi$  and  $\sigma$ . We then select the optimal trade-in point we have found in Chapter 5 for this parameter values. If we estimate  $\phi$  and  $\sigma$  values different than the ones we used in the previous chapter, we interpolate.

We assume that the investor uses point 0 as trade-out point as we do. We make this assumption because we want to measure the performance of our strategy which depends

on deciding the trade-in points correctly. Thus, we set trade-out point to the same value for both strategies, the strategy of the investor and our strategy.

Firstly, we simulate  $AR(1)$  series as in (29). We simulate this  $AR(1)$  series for all five  $\emptyset$  values we have used before, i.e. 0.95, 0.96, 0.97, 0.98 and 0.99, as in Section 6.1. In addition, we set  $\sigma = 1$  as in Section 6.1. These series become the historical data representing the cointegrated equation series, i.e. the portfolio values. The investor observes the most profitable trade-in point of those series. On the other hand, we estimate  $\emptyset$  and  $\sigma$  values for those series. Note that, although we set  $\emptyset$  and  $\sigma$  values in the simulation process, we pretend that we do not know or observe them and estimate them from the simulated data. After we estimate  $\emptyset$  and  $\sigma$  values, we pick the most profitable trade-in values according to our strategy described in Section 5.3.

Then, we simulate another  $AR(1)$  series for each  $\emptyset$  value. We use these data as future data that we intend to invest. We apply a pairs trading to these series for both set of trade-in points, the one we have decided according to our rule using  $\emptyset$  estimations and the one the investor has decided using the historical data in the previous step. Then, we compare the profit per unit time values gained from these rules.

In Figure 6.2, the investor's investment scenario is illustrated, in which the investor firstly observes and analyzes the historical data, and then invests using the optimal point he has decided in the previous step. In the figure,  $t+1$  shows the present time, which is the point of time that the investor starts investing; and  $t+n$  shows the point of time that the investor stops trading.

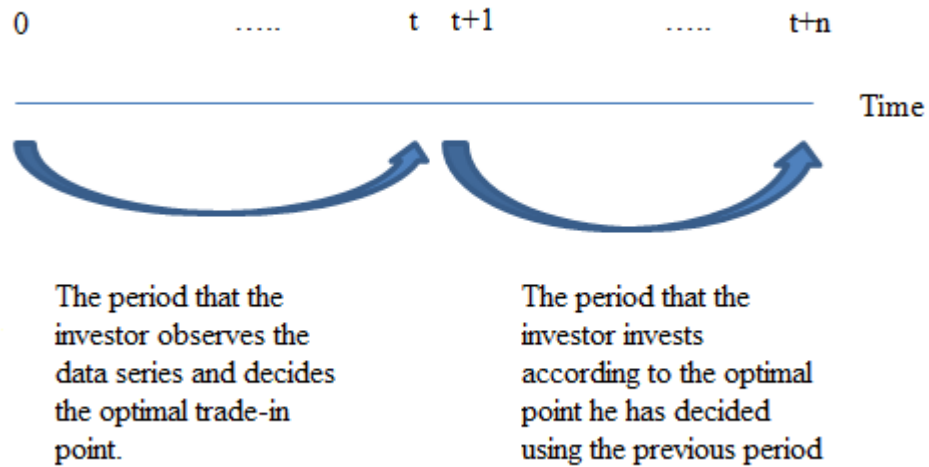


Figure 6.2: The timeline for the investor's process

Now, let us elaborate on the analysis we have mentioned briefly above. Firstly, we conduct a simulation of 10,000 steps according to an  $AR(1)$  without constant process as in (29). We assume that these simulated values are historical values an investor observes and we write down the profit per unit time values that the investor would earn for all possible trade-in points, if he invested in this period. In addition, we check if the series fits to  $AR(1)$  series and, if it does, we estimate the parameter values  $(\phi, \sigma)$  of this simulated series.

We repeat this historical data process with 20 independent samples. That is, we conduct a total of 20 simulations for the series in (29) of the same characteristics, i.e. for the same  $\phi$  value. We record the profit per unit time for each run of the investor's strategy for all possible trade-in ( $a$ ) points and we check the parameter estimations of these 20 samples for our strategy. The reason we repeat this process for 20 times is that we want to overcome the sampling error when we measure the profits. When 20 samples are gathered, we take the mean values of the 20 profit per unit time values that the investor recorded for each trade-in point, and decide the most profitable trade-in point according

to the investor's method. Those 20 samples of those values for all  $\emptyset$  values are reported in Appendix F. For the trade-in point of our method, we estimate the parameter values of the series for each sample and take the mean value of them. In Table 6.2, those estimated values and their mean and standard deviation values for  $\emptyset = 0.95$  are reported for the 20 samples. For  $\emptyset$  equals 0.96, 0.97, 0.98 and 0.99, estimated parameter tables are included in Appendix G. The mean of the estimated  $\emptyset$  values equals  $0.947 \approx 0.95$ . Thus, we pick the most profitable  $a$  value for  $\emptyset = 0.95$ . In addition, the mean value of the estimated standard error of regression ( $\sigma$ ) values equals 0.98. Note that we do not have optimal  $a$  values for  $\sigma > 0.2$ . So, we pick the most profitable  $a$  value closest to the parameter estimations  $\emptyset = 0.95$  and  $\sigma = 0.2$ .

Table 6.2: Estimated parameter values for  $\emptyset = 0.95$

Sample No	Estimated $\emptyset$ value	Estimated $\sigma$ value
1	0.936	1.029
2	0.956	0.981
3	0.943	1.013
4	0.945	1.001
5	0.952	0.975
6	0.947	1.038
7	0.957	1.121
8	0.934	1.120
9	0.961	0.998
10	0.954	1.089
11	0.938	1.041
12	0.955	0.997
13	0.949	0.982
14	0.942	1.011
15	0.960	1.081
16	0.951	1.031
17	0.941	0.974
18	0.939	1.021
19	0.929	1.241
20	0.941	1.008
mean	0.947	1.038
std.dev.	0.009	0.064

Having chosen the trade-in values by both the investor's observation and our method, we now conduct another simulated  $AR(1)$  without constant series of 10,000 steps using the same  $\emptyset$  value as the previous one and observe the series to trade-in and out according to the optimal points decided from both rules, and we record the profit per unit time values for both methods. This process is, again, repeated for 20 times in order to overcome possible sampling errors. Then, we take the mean of the profits per unit time of the 20 samples for both techniques and compare the results.

The 20 samples of profit per unit time values for the present data for all  $\emptyset$  values in are reported in Appendix H. The mean and standard deviation values of those profit per unit time values for 20 samples are also reported.

Let us examine the results for  $\emptyset = 0.95$ . In Appendix F.1, we check the profits per unit time for trade-in points between 0.45 and 0.64 to find the optimal trade-in point, because the optimal trade-in point decided by the rule of our study is 0.54 for  $(\emptyset, \sigma) = (0.95, 0.2)$ ; and thus, we think it is a good idea to check the interval around this point to find the best profitable point of the historical data. After we gather the 20 samples, we take the mean of the profits per unit time for all trade-in points as we mention before, and pick the point giving the biggest mean profit per unit time as the optimal trade-in point of the related  $\emptyset$  value. For  $\emptyset = 0.95$ , the optimal trade-in point according to the method of the investor is revealed to be 0.48. Meanwhile, we find the mean values of  $(\emptyset, \sigma)$  values as  $(0.947, 1.038)$  as Table 6.1 presents. Thus, we decide optimal trade-in point equal 0.54, corresponding to the optimal trade-in point for the closest parameter values we have worked on, i.e.  $(\emptyset, \sigma) = (0.95, 0.2)$ . Then, we observe the process to trade-in and out for both of the trade-in points, which are 0.48, the most profitable point according to the investor's strategy, and 0.54, the most profitable point according to the method we have proposed. Appendix H.1 displays the result of this comparison. We earn the same amount of profit per unit time using the two points. That is, the rule that the investor follows using historical data gives 0.1268 while the rule we have proposed gives 0.1268 per unit time.

On the other hand, for  $\emptyset = 0.97$  we see that our method gives more profit than the investor's method. Our method earns 0.0977 while the other one gives 0.0976 per unit time.

As a consequence, we see that our method pays approximately the same profit per unit time as the method of historical data. However, note that an investor should repeat his observation series before deciding the trade-in point, because his decision depends only

on observation without any statistical methods. This is a big waste of time and money in real life. On the other hand, the disadvantage of our method is that our method's success depends on estimating the parameters accurately. So, making estimations with mean value of samples is also necessary in order not to be affected from sampling errors. Yet, because we have not used the exact  $\sigma$  value but the closest value to the values we have calculated before, performance of our method will be better when we use the  $a$  values for the exact parameter values. Thus, we can state that, if we can estimate the parameters correctly, our method becomes a good asset for predicting stock series fitting  $AR(1)$  model.

### **6.3 Performance Measurement Using BIST Banking Stock Data**

As final part of our study, we want to apply our method on the historical cointegrated stock series. For this implementation, we use three stock price series we have used in the beginning of our study, namely Akbank, Şekerbank and Vakıfbank stocks from Turkish Stock Exchange, i.e. BIST, between years 2008 and 2010. We obtain three cointegrated pairs from those three stocks. Note that using these series will not positively affect the performance results by overestimating the profits per unit time, since we have used these data only to have an idea on the type the common time series models of the cointegrated stocks in BIST Banking Field, and their  $\emptyset$  and standard error of regression values. Since the method have not seen the data series before, the profit level will not be more than of any other data series fitting to our requirements, which are fitting to  $AR(1)$  model with  $\emptyset$  between 0.95 and 0.99 and standard error of regression values between 0.05 and 0.2.

We have three cointegrated pairs Vakıfbank-Şekerbank, Vakıfbank-Akbank and Şekerbank-Akbank, whose normalized cointegration equations are shown in Tables 5.8, 5.10 and 5.12, respectively.

Let us present the performance results of the pair Şekerbank-Akbank. The series formed from the cointegration equation of Şekerbank and Akbank, i.e.  $X_t$ , seems to fit  $AR(1)$  model with regression coefficient  $\emptyset = 0.987$  and standard error of regression  $\sigma =$



0.259. However, as we have mentioned at the beginning of our study, the pairs violate the diagnostic rules for  $AR(1)$  models, such as homoscedasticity or normality of the residuals. However, since it is very rare for a real life data series to fit a time series model meeting all the requirements, this analysis is still important as a real life implementation.

We take 0.98 as trade-in point which corresponds to the optimal trade-in point for  $\emptyset = 0.99$  and  $\sigma = 0.2$  according to our method. If we use 0.98 as our trade-in point, we earn 0.031TL/day from one unit of the portfolio for these data. We see that this trade-in point earns us more than every possible trade-in point which proves the success of our method for this series.

Remember that we have mentioned about the sensitivity issue of our method to correct estimations before. To be specific, the optimal trade-in points turn out to be very sensitive to estimated  $\sigma$  values. That is, when we fix  $\sigma$  values, we see that optimal trade-in points do not extremely change for different  $\emptyset$  values. However, when we fix  $\emptyset$  values, we see that optimal trade-in points do extremely change for different  $\sigma$  values. Let us illustrate this situation with Şekerbank and Akbank pair. If we were mistaken to estimate the  $\emptyset$  value accurately and we estimated it as 0.97, but estimated  $\sigma$  value truly as 0.2, then we would take the trade-in point as 0.66, which makes us earn 0.028TL/day. However, if we were mistaken to estimate the  $\sigma$  value truly and we estimate it 0.1, but we estimated the  $\emptyset$  value truly as 0.99, then we would take the trade-in point as 0.49, which makes us earn 0.029TL/day. Although the profit change does not give a big alert in this example, it may be adversely affected in another situation because of the huge shift in optimal trade-in point.

Now, let us summarize the performance of our method on the next stock pair we concern, which is the Akbank-Vakıfbank pair.

The series formed from the cointegration equation of Akbank and Vakıfbank seems to fit AR(1) model with  $\phi = 0.903$  and  $\sigma = 0.245$ . Note that the  $\phi$  value is lower than all  $\phi$  values we have concerned before, so we will take the closest  $\phi$  value to 0.903 we concern in our study which is 0.95. In addition, we will take the closest  $\sigma$  value to 0.245 we concern in our study which is 0.2. Then, we take 0.54 as our trade-in point which is the optimal trade-in point corresponding to  $\phi = 0.95$  and  $\sigma = 0.2$  according to our method. If we invest in this pair using our method over 2008-2010 and if we use 0.54 as our trade-in point, then we earn 0.021TL/day for one unit of portfolio. Although this point does not give the biggest profit per day for this series, we see that it gives us the second best profit. If we used 0.52 as our trade-in point, we would earn 0.022TL/day which is the biggest profit for this series. Although our method does not provide the most profitable trade-in point in this case, the trade-in points and the daily profit values are very close to each other for the two cases. Besides, we have assumed the estimated  $\phi$  value bigger than what it actually is in this case. So, we are still sure about the success of our method.

Let us continue with another cointegrated stock pair, Vakıfbank and Şekerbank. The  $X_t$  series, which is formed from the cointegration equation of Şekerbank and Vakıfbank, seems to fit AR(1) model with  $\phi = 0.986$  and  $\sigma = 0.236$ . Then, we take 0.98 as our trade-in point which is the optimal trade-in point for  $\phi = 0.99$  and  $\sigma = 0.2$  according to our method. When we use 0.98 as our trade-in point, we earn 0.020TL/day for one unit of portfolio if we invest over 2008-2010 years. We realize that this point gives again a very close profit to the best profit we can earn from this series. That is, if we chose 1.1 as the trade-in point, we could earn 0.021TL/day for one unit of portfolio. Although this profit is the second best, we see that it is very close to the optimal. For example, if we chose point 3 as our trade-in point, we would earn 0.013TL/day. This example shows how close our profit level to the optimal value.

This analysis shows that the method we propose has a good performance on real life data. The importance of measuring the performance of our method on real data is that real data rarely meet the diagnostic assumptions required for autoregressive series. So, this analysis shows that our method is successful not only on data meeting the required assumptions but also on real life data.

# Chapter 7

## Conclusion and Future Research

In this thesis, pairs trading strategy is studied. Pairs are selected using cointegration relationship. The aspect of this study different from other pairs trading studies is that we maximize the long-run expected profit per unit time while other studies are interested in the maximum total profit over finite horizon.

Our problem reduces to deciding on a trade-in point. That point is chosen so as to maximize the long-run expected profit per unit time for a pairs trading on the series in concern. We discretize stock price processes and model them as Markov Chains. We use the renewal reward theorem to calculate the long-run expected profit per unit time. We have worked with six real stocks in banking field from BIST and implemented our method on these real life data. Cointegration tests revealed six pairs from those six banking stocks. We have analyzed the series obtained from the cointegration equations of all six pairs and noticed that all those series follow  $AR(1)$  model. Then, we have used the parameter estimations of those series to simulate data to apply our method on. We have analyzed the cases that starting point of the process is known, which we set to 0 in the study, and that starting point of the process is unknown, which we assume as a random variable. We have revealed that random starting point has a stationary

distribution. In addition, both studies show that long-run expected profit per unit time functions are concave functions of trade-in points.

Then, we have simulated series from  $AR(1)$  models whose error terms follow  $Gsn(0,1)$ . We use the parameter values we have estimated from BIST stock price data for the simulated data. When we let the simulated series trade in and out, we observe the same result as in the case study: as the regression coefficient of the  $AR(1)$  model increases, the profit per unit time value decreases. We have simulated other series fitting to  $AR(1)$  model whose error terms follow  $Gsn(0,1)$ . We have compared our method with a basic method in which investors observe stock prices for a while and decide on the best trade-in point for the series by a simple search, and then invest using that trade-in point in the future. On the other hand, to implement our method, we observe the series for a while and estimate the parameters of that series, and use the trade-in point that maximizes the long-run expected profit per unit time. Because in the case study we have worked using the parameter estimations we acquired from the historical stock series, we do not have trade-in points for all values. Yet, we use the closest estimated parameter value we have to decide the trade-in point. The results show that two methods result in the similar long-run profit per unit time amounts to each other. Yet, considering we have not used exact trade-in points but used the one for the closest parameter estimations, we conclude that the method we propose has good performance. Finally, we have implemented our method on the historical data we used in the case study. We analyzed the results on three pairs. The results show that the performance of the proposed method depends on estimating the parameter values correctly.

We have built our method on a no transaction cost world. Thus, one future direction is to consider transaction costs of trades. Additionally, a method with more than one trade-in point can be studied. Waiting the stock price to pass “one” trade-in point may cause a cycle to last very long. As a solution, two trade-in points may be set so that investor trades in with a preset proportion of his portfolio when stock price reaches the smaller trade-in point; and trades the rest of his portfolio when the stock price reaches to the

bigger trade-in point. Note that after the investor trades at the smaller trade-in point, he may trade-out before the price level reaches to the bigger trade-in point because the price level may reach to point 0 before reaching to the bigger trade-in point. Although those trading processes at smaller trade-in point do not bring big amount of profit, it may be better to wait until the best and one trade-in point without trading. Investigation of this hypothesis will be a good future research direction.

# Bibliography

- [1] Avellaneda, M., and Lee J.H., "Statistical Arbitrage in the U.S. Equities Market," *Social Science Research Networks*, 2008.
- [2] Bala, A., and Mukund K. S., "Interrelationship between Indian and US stock markets," *Journal of Management Research*, vol.1, no. 3, pp.141-148, 2001.
- [3] Banerjee, A., Dolado J.J, Galbraith J.W., and Hendry D., "Co-integration, error correction, and the econometric analysis of non-stationary data," *OUP Catalogue*, 1993.
- [4] Chan, C. K., Benton G.E., and Min S.P., "International stock market efficiency and integration: a study of eighteen nations," *Journal of Business Finance and Accounting*, 24(6), July, 0306-686X
- [5] Do, B., Faff R., and Hamza K., "A new approach to modeling and estimation for pairs trading." *Proceedings of 2006 Financial Management Association European Conference*, 2006.
- [6] Do, B., and Faff R. W., "Does simple pairs trading still work?," *Financial Analysts Journal*, vol. 66, no. 4, pp. 83-95, 2010.
- [7] Dolado, J.J, Gonzalo J., and Marmol F., "Cointegration," *A Companion to Theoretical Econometrics*, pp. 634-654, 2001.
- [8] Ehrman, D.S., *The handbook of pairs trading: strategies using equities, options, and futures*, John Wiley & Sons, vol. 240, 2006.
- [9] Enders, W., *Applied econometric time series*, John Wiley & Sons, 2008.

- [10] Engelberg, J., Gao P., and Jagannathan R., "An anatomy of pairs trading: the role of idiosyncratic news, common information and liquidity," *Common Information and Liquidity (February 22, 2009)*, 2009.
- [11] Engle, R.F., and Granger C.W., "Co-integration and error correction: representation, estimation, and testing," *Econometrica: journal of the Econometric Society*, pp. 251-276, 1987.
- [12] Gatev, E., Goetzmann W.N., and Rouwenhorst K.G., "Pairs trading: performance of a relative-value arbitrage rule," *Working Paper 7032, National Bureau of Economic Research, Washington D*, 1999.
- [13] Gillespie, T., and Ulph C., "Pair trades methodology: a question of mean reversion," *Proceedings of International Conference on Statistics, Combinatorics and Related Areas and the 8th International Conference of Forum for Interdisciplinary Mathematics, NSW*, 2001.
- [14] Granger, C.W., "Some properties of time series data and their use in econometric model specification," *Journal of econometrics*, vol. 16, no.1, pp. 121-130, 1981.
- [15] Hong, G., and Susmel, R., "Pairs-trading in the asian ADR market," *Working Paper, University of Houston*, 2003.
- [16] Karlin, S., and Taylor. H. E., "A first course in stochastic processes," *Elsevier*, Chapter 3, Section 7, pp. 169-175, 1975.
- [17] Khandani, A., and Lo A.W., "What happened to the quants in August 2007?," *Available at SSRN 1015987*, 2007.
- [18] Lin, Y.X., McCrae M., and Gulati C., "Loss protection in pairs trading through minimum profit bounds: A cointegration approach," *Advances in Decision Sciences*, 2006.
- [19] Mudchanatongsuk, S., Primbs J. A., and Wong W., "Optimal pairs trading: A stochastic control approach," *In American Control Conference, 2008*, pp. 1035-1039, 2008.



- [20] Narayan, P. K., "The saving and investment nexus for China: evidence from cointegration tests." *Applied economics*, vol. 37, no. 17, pp. 1979-1990, 2005.
- [21] Nath, P., "High frequency pairs trading with us treasury securities: Risks and rewards for hedge funds," *SSRN eLibrary*, 2003.
- [22] Nelson, C.R., and Plosser C.R., "Trends and random walks in macroeconomic time series: some evidence and implications," *Journal of monetary economics*, vol. 10, no. 2, pp. 139-162, 1982.
- [23] Newbold, P., and Granger C.W., "Experience with forecasting univariate time series and the combination of forecasts," *Journal of the Royal Statistical Society. Series A (General)*, pp. 131-165, 1974.
- [24] Papadakis, G., and Wysocki P., "Pairs trading and accounting information," *Boston University and MIT Working Paper*, 2007.
- [25] Peskin, M., and Boudreau B., "Why hedge funds make sense," 2000, <http://www.thehfa.org/articles/1.pdf>.
- [26] Puspaningrum, H., "Pairs trading using cointegration approach," *Ph.D. Thesis, School of Mathematics and Applied Statistics, University of Wollongong*, 2012.
- [27] Puspaningrum, H., Lin Y. X., and Gulati C., "Finding the optimal pre-set boundaries for pairs trading strategy based on cointegration technique," *Centre for Statistical and Survey Methodology Working Paper Series*, 2009.
- [28] Vidyamurthy, G., "Pairs trading: quantitative methods and analysis," *Wiley*, vol. 217, 2004.
- [29] Walls, W. D., "A cointegration rank test of market linkages with an application to the US natural gas industry," *Review of Industrial Organization*, vol. 9, no. 2, pp. 181-191, 1994.
- [30] Wong, W. K., A. Agarwal, and Du J., "Financial integration for Indian stock market, a fractional cointegration approach," *National University of Singapore Working Paper No. WP0501*, 2005.

# Appendix A

## Time Series Plots of the Banking Stock Pairs

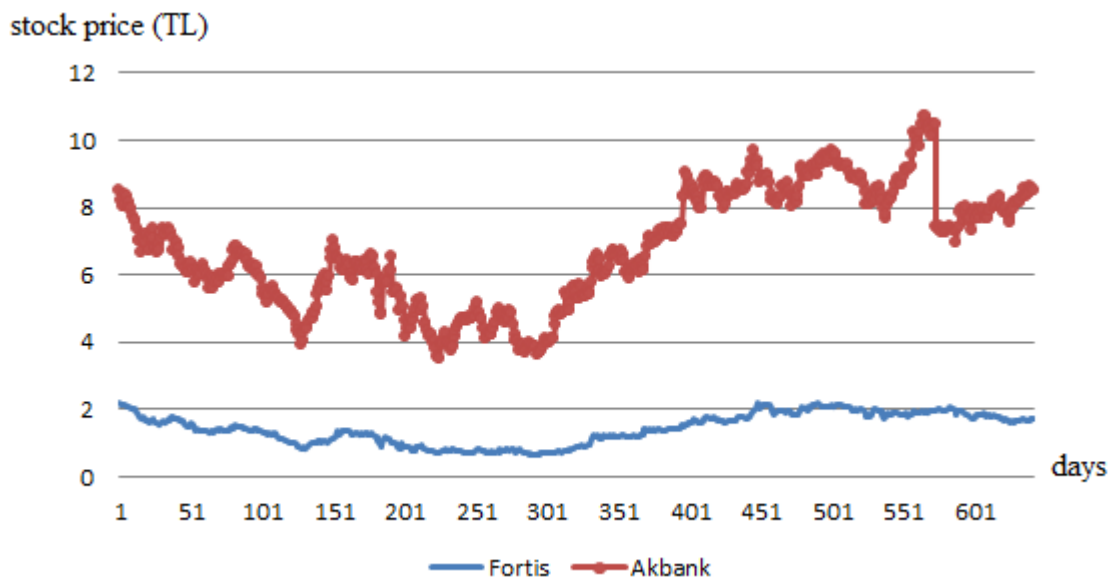


Figure A.1: Time series graphs of Fortis and Akbank

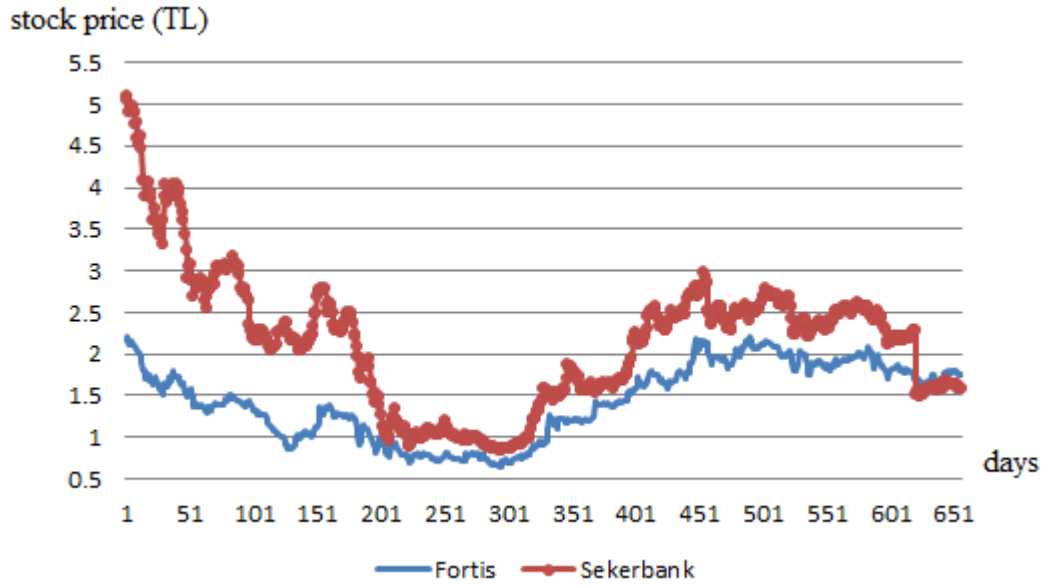


Figure A.2: Time series graphs of Fortis and Şekerbank

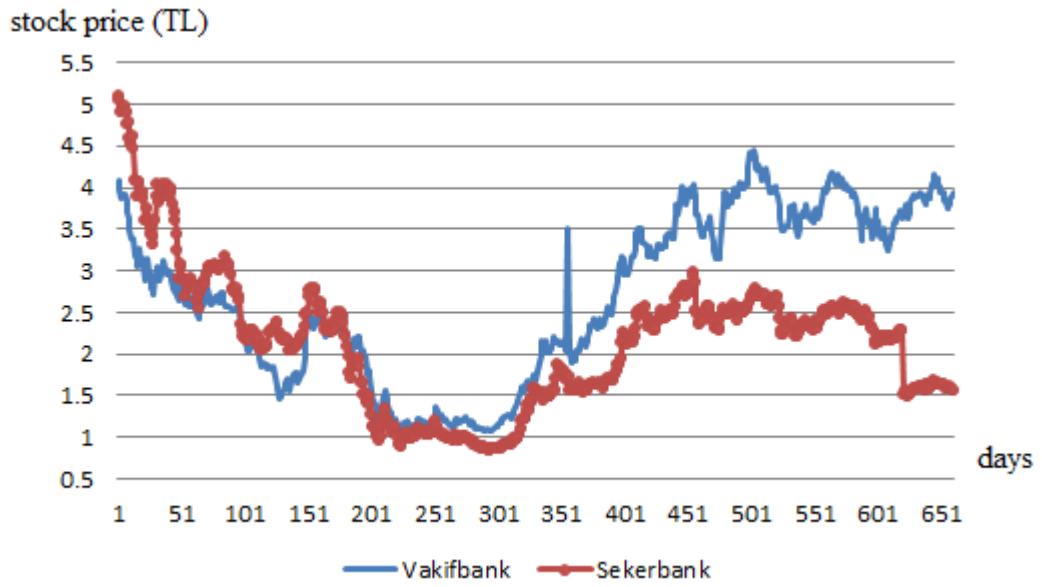


Figure A.3: Time series graphs of Vakıfbank and Şekerbank

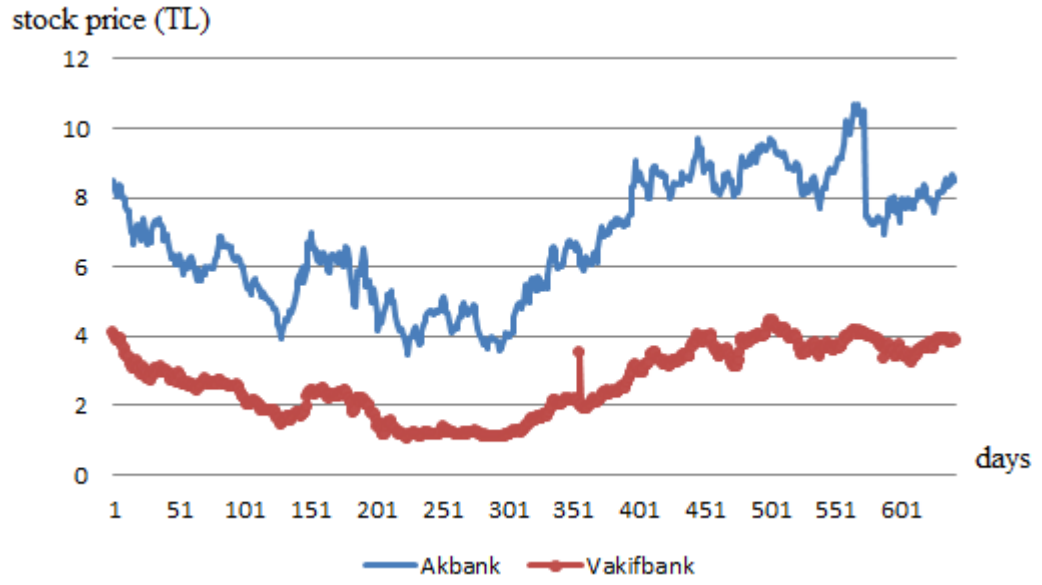


Figure A.4: Time series graphs of Akbank and Vakıfbank

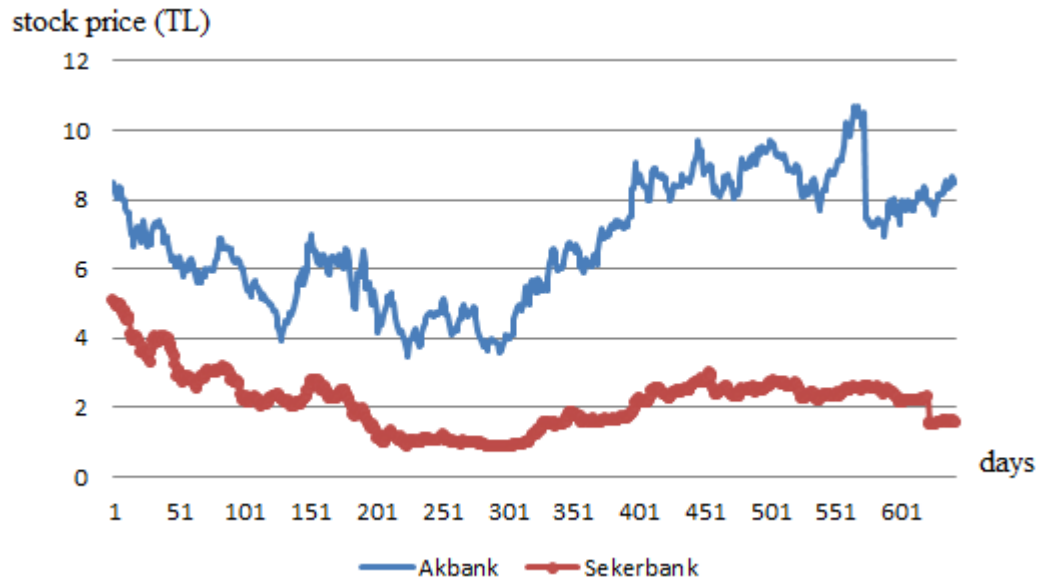


Figure A.5: Time series graphs of Akbank and Şekerbank

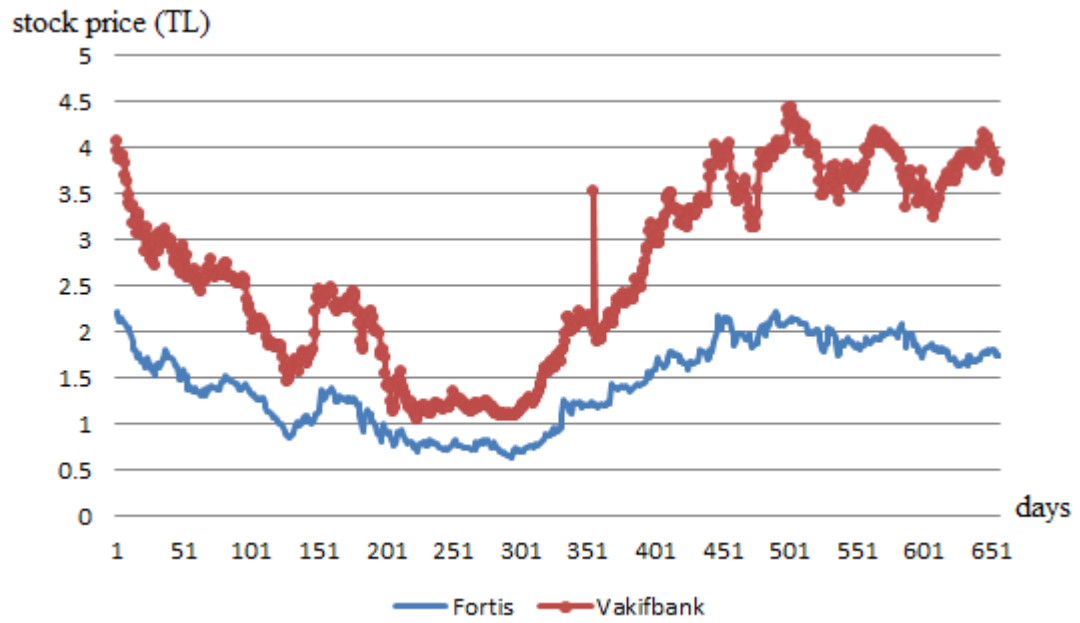


Figure A.6: Time series graphs of Fortis and Vakifbank

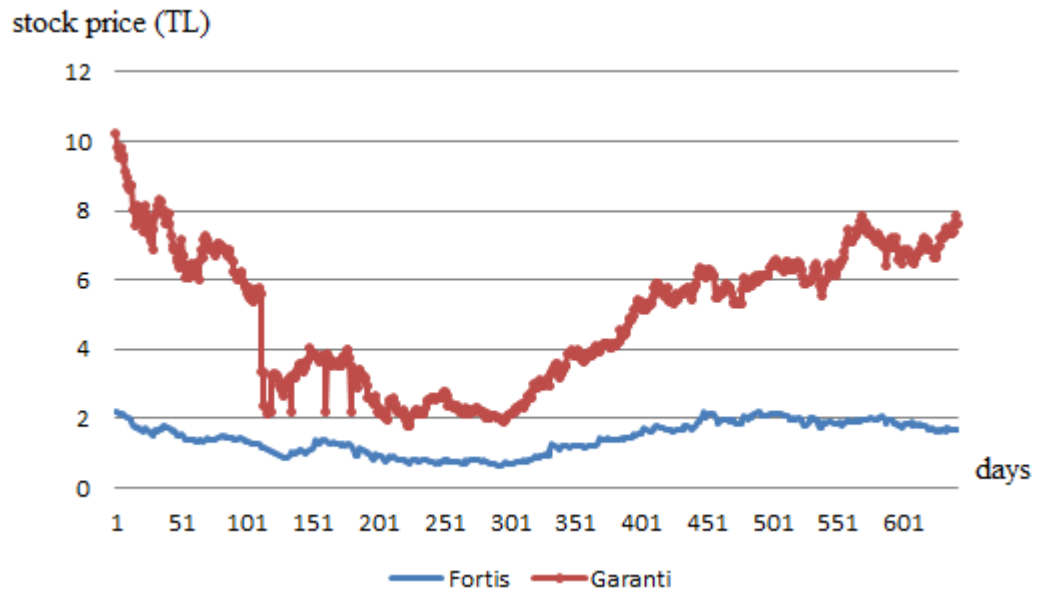


Figure A.7: Time series graphs of Fortis and Garanti Bank

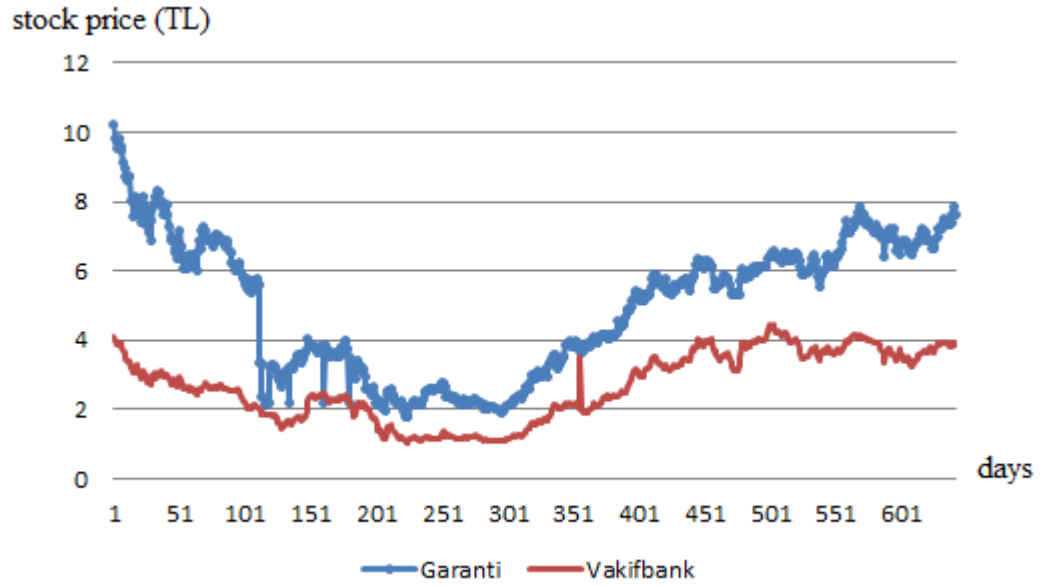


Figure A.8: Time series graphs of Garanti Bank and Vakıfbank

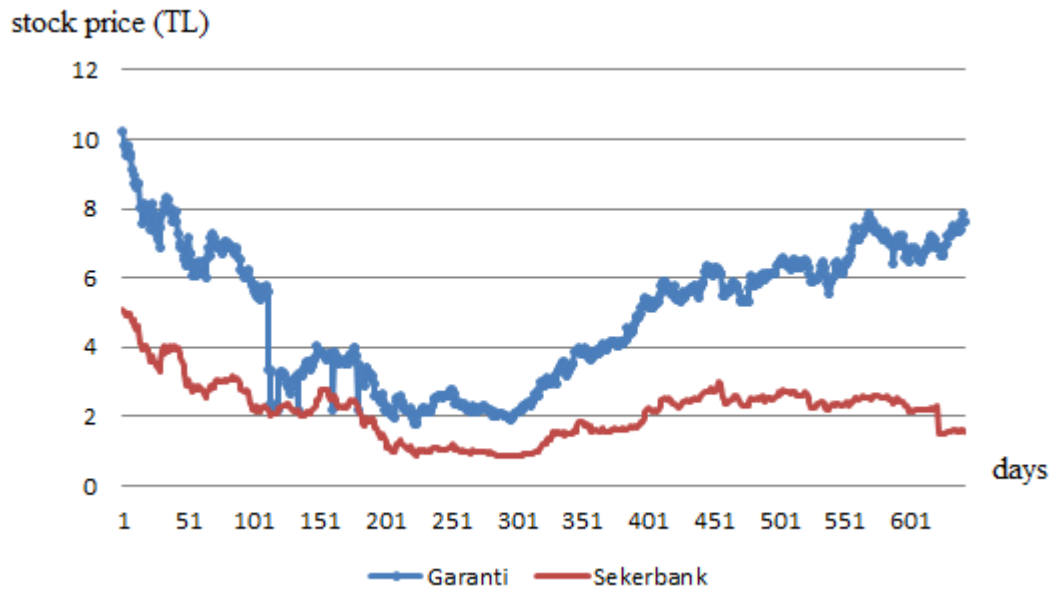


Figure A.9: Time series graphs of Garanti Bank and Şekerbank

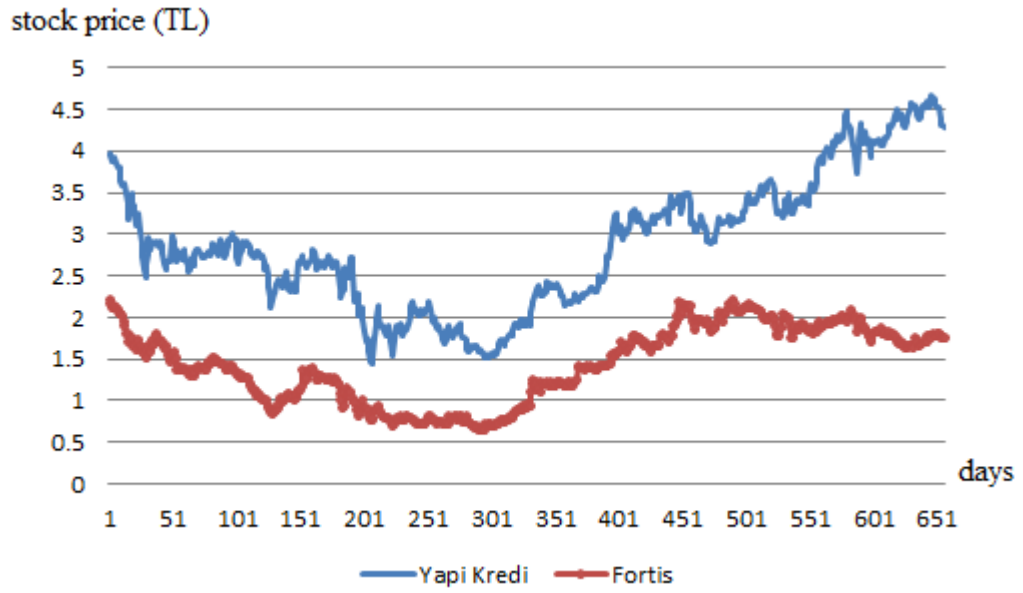


Figure A.10: Time series graphs of Yapi Kredi Bank and Fortis

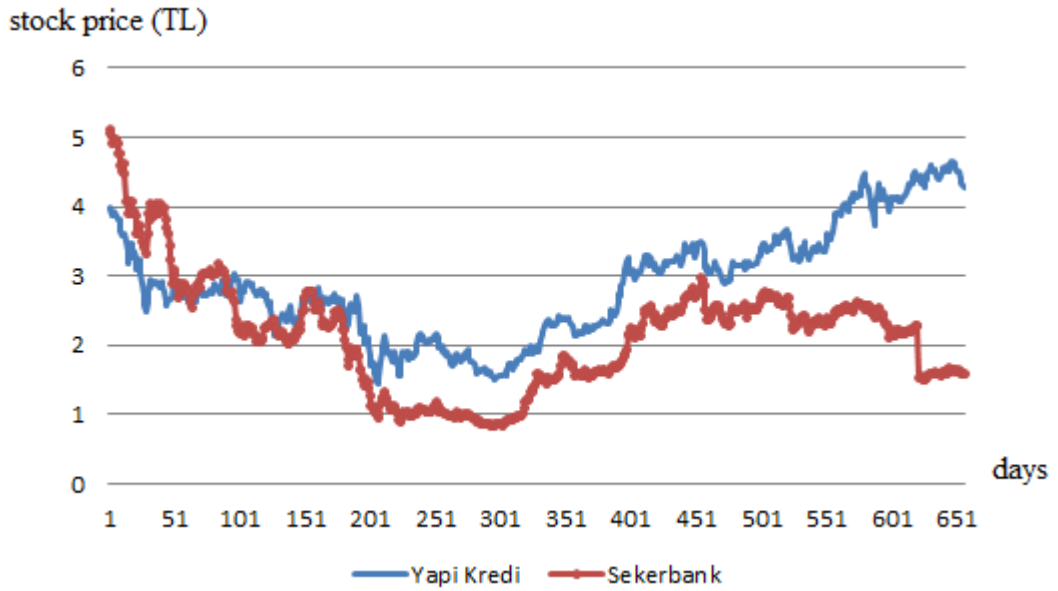


Figure A.11: Time series graphs of Yapi Kredi Bank and Şekerbank

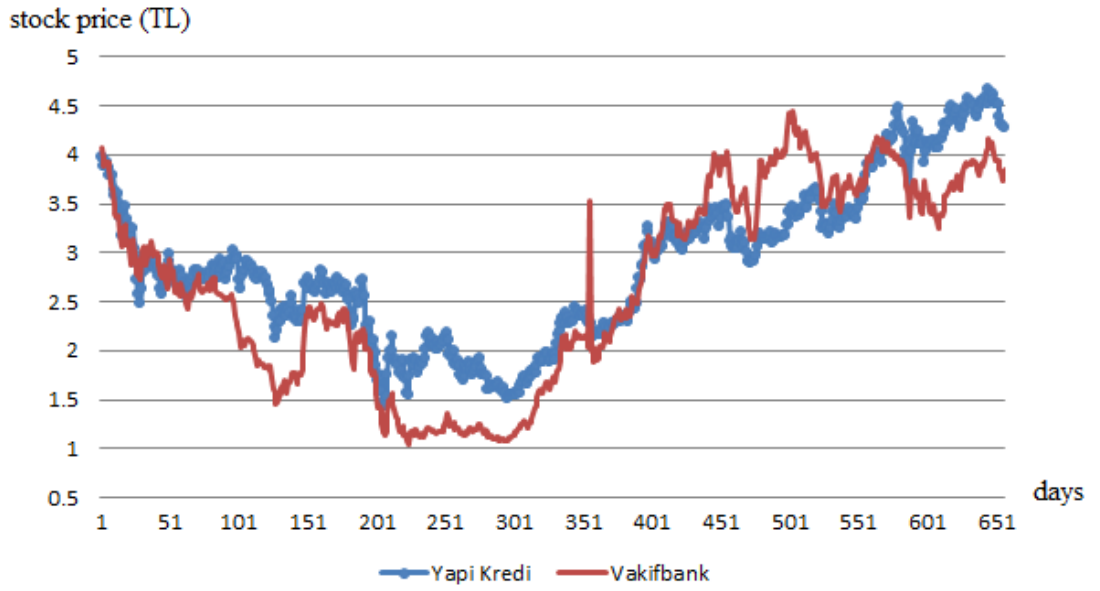


Figure A.12: Time series graphs of Yapi Kredi Bank and Vakifbank

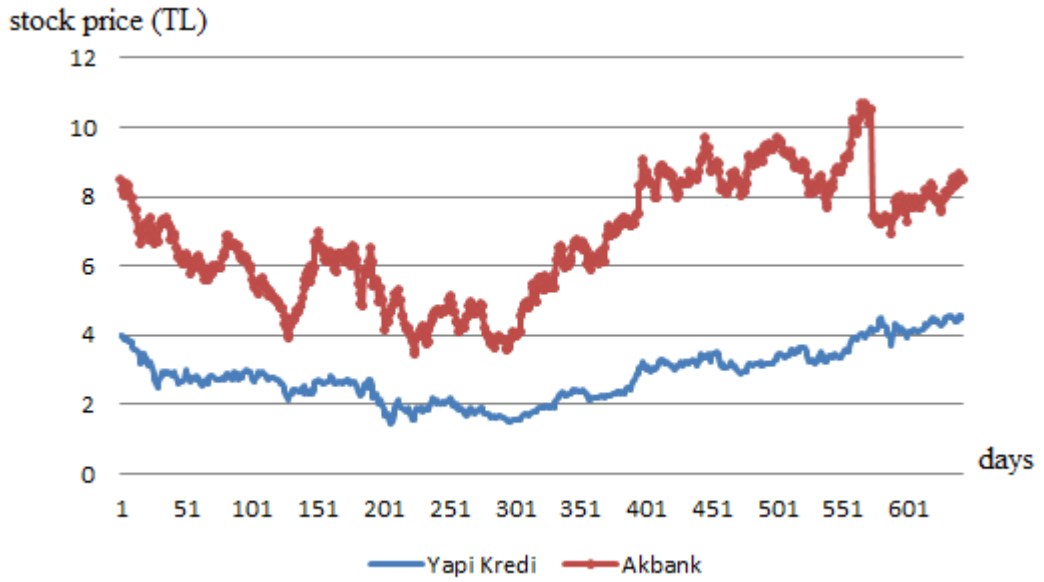


Figure A.13: Time series graphs of Yapi Kredi Bank and Akbank



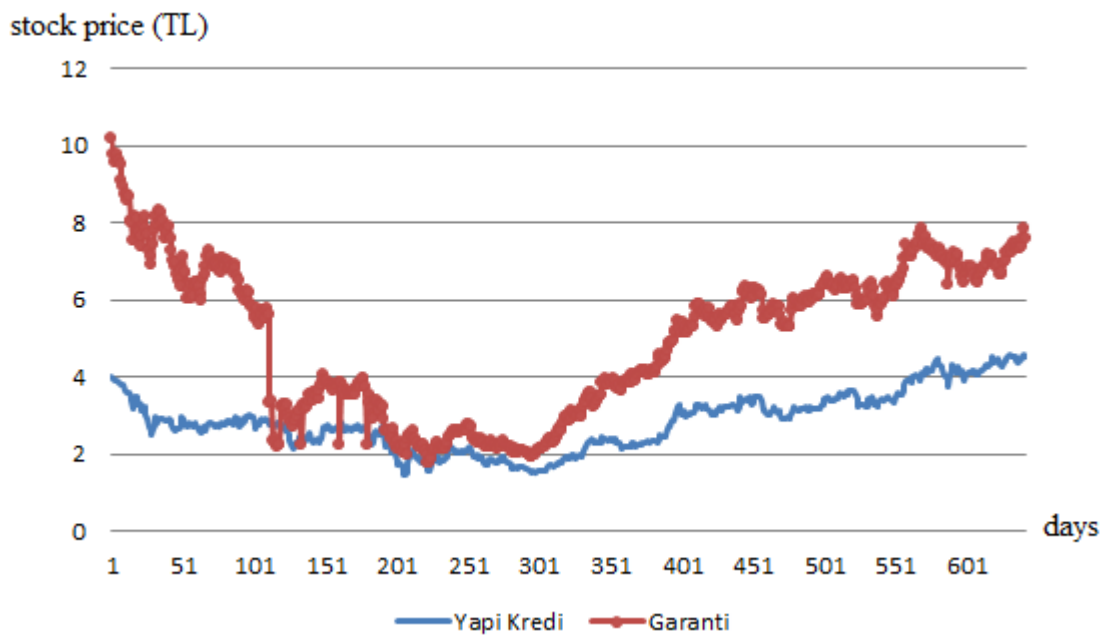


Figure A.14: Time series graphs of Yapi Kredi Bank and Garanti Bank

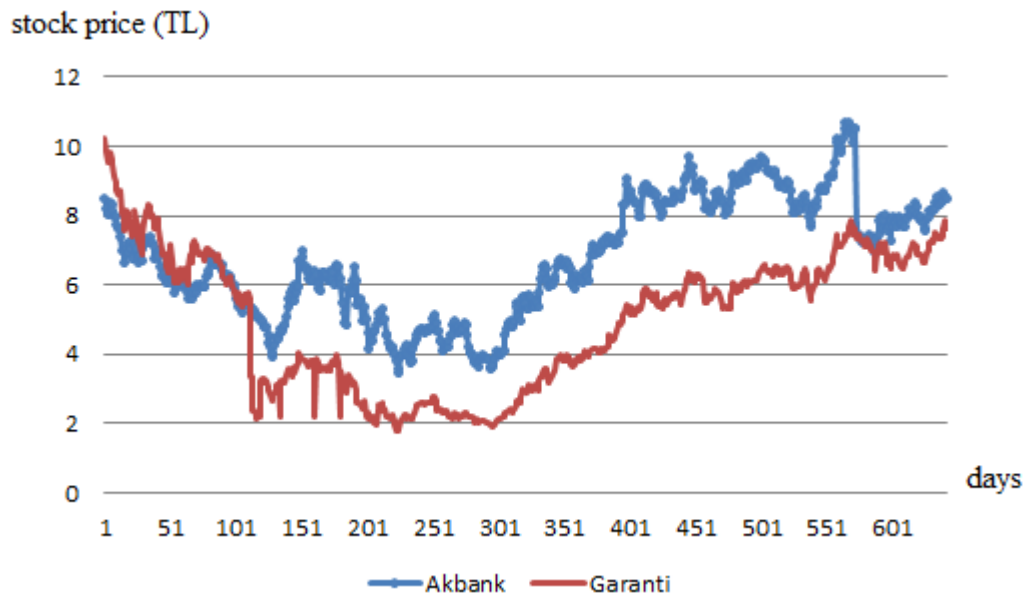


Figure A.15: Time series graphs of Akbank and Garanti Bank

# Appendix B

## Model Outputs of the Cointegration Equations

Test is done under the null hypothesis:  $H_0$ : The regression coefficient ( $\emptyset$ ) of the variable is 0.

Table B.1: Model outputs of the pair Fortis-Akbank

Dependent Variable: Fortis-Akbank  
Method: Least Squares  
Date: 09/24/10 Time: 11:54  
Sample (adjusted): 2 642  
Included observations: 641 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	9.06E-05	0.002298	0.039423	0.9686
Fortis-Akbank (-1)	0.942390	0.013033	72.30990	0.0000
R-squared	0.891099	Mean dependent var		0.014588
Adjusted R-squared	0.890929	S.D. dependent var		0.175525
S.E. of regression	0.057969	Akaike info criterion		-2.854714
Sum squared resid	2.147271	Schwarz criterion		-2.840789
Log likelihood	916.9359	Hannan-Quinn criter.		-2.849310
F-statistic	5228.722	Durbin-Watson stat		1.913889
Prob(F-statistic)	0.000000			

Table B.2: Model Outputs of the pair Fortis-Şekerbank

Dependent Variable: Fortis-Sekerbank  
 Method: Least Squares  
 Date: 07/25/13 Time: 18:08  
 Sample (adjusted): 2 657  
 Included observations: 656 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001292	0.003685	0.350594	0.7260
Fortis-Sekerbank (-1)	0.985278	0.004075	241.7855	0.0000
R-squared	0.988937	Mean dependent var		-0.333844
Adjusted R-squared	0.988920	S.D. dependent var		0.830909
S.E. of regression	0.087464	Akaike info criterion		-2.032140
Sum squared resid	5.003044	Schwarz criterion		-2.018462
Log likelihood	668.5418	F-statistic		58460.22
Durbin-Watson stat	1.825625	Prob(F-statistic)		0.000000

Table B.3: Model Outputs of the pair Vakıfbank-Şekerbank

Dependent Variable: Vakifbank-Sekerbank  
 Method: Least Squares  
 Date: 07/25/13 Time: 18:40  
 Sample (adjusted): 2 661  
 Included observations: 660 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.003726	0.009951	0.374429	0.7082
Vakifbank-Sekerbank (-1)	0.985465	0.004119	239.2509	0.0000
R-squared	0.988635	Mean dependent var		-0.910456
Adjusted R-squared	0.988618	S.D. dependent var		2.212463
S.E. of regression	0.236039	Akaike info criterion		-0.046616
Sum squared resid	36.66001	Schwarz criterion		-0.033003
Log likelihood	17.38317	F-statistic		57240.97
Durbin-Watson stat	2.069032	Prob(F-statistic)		0.000000

Table B.4: Model Outputs of the pair Akbank-Vakıfbank

Dependent Variable: Akbank-Vakıfbank

Method: Least Squares

Date: 07/25/13 Time: 17:16

Sample (adjusted): 2 642

Included observations: 641 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000393	0.009683	-0.040583	0.9676
Akbank-Vakıfbank (-1)	0.903026	0.016924	53.35808	0.0000
R-squared	0.816700	Mean dependent var		-0.008871
Adjusted R-squared	0.816413	S.D. dependent var		0.572095
S.E. of regression	0.245126	Akaike info criterion		0.029026
Sum squared resid	38.39543	Schwarz criterion		0.042952
Log likelihood	-7.302967	F-statistic		2847.085
Durbin-Watson stat	2.393919	Prob(F-statistic)		0.000000

Table B.5: Model Outputs of the pair Akbank-Şekerbank

Dependent Variable: Akbank-Sekerbank

Method: Least Squares

Date: 07/25/13 Time: 17:50

Sample (adjusted): 2 642

Included observations: 641 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.005275	0.011091	0.475549	0.6346
Akbank-Sekerbank (-1)	0.986514	0.004283	230.3156	0.0000
R-squared	0.988097	Mean dependent var		-0.976598
Adjusted R-squared	0.988078	S.D. dependent var		2.374316
S.E. of regression	0.259242	Akaike info criterion		0.141004
Sum squared resid	42.94483	Schwarz criterion		0.154929
Log likelihood	-43.19182	F-statistic		53045.27
Durbin-Watson stat	1.827689	Prob(F-statistic)		0.000000

Table B.6: Model Outputs of the pair Fortis-Vakıfbank

Dependent Variable: Fortis-Vakıfbank

Method: Least Squares

Date: 07/25/13 Time: 18:26

Sample (adjusted): 3 657

Included observations: 655 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.13E-05	0.001945	0.021213	0.9831
Fortis-Vakıfbank (-1)	0.663943	0.038046	17.45109	0.0000
Fortis-Vakıfbank (-2)	0.240691	0.038045	6.326402	0.0000
R-squared	0.776792	Mean dependent var		0.006940
Adjusted R-squared	0.776107	S.D. dependent var		0.104914
S.E. of regression	0.049642	Akaike info criterion		-3.163373
Sum squared resid	1.606769	Schwarz criterion		-3.142832
Log likelihood	1039.005	F-statistic		1134.522
Durbin-Watson stat	2.016356	Prob(F-statistic)		0.000000

# Appendix C

## Profit per Unit Time Graphs for $X_0 = 0$

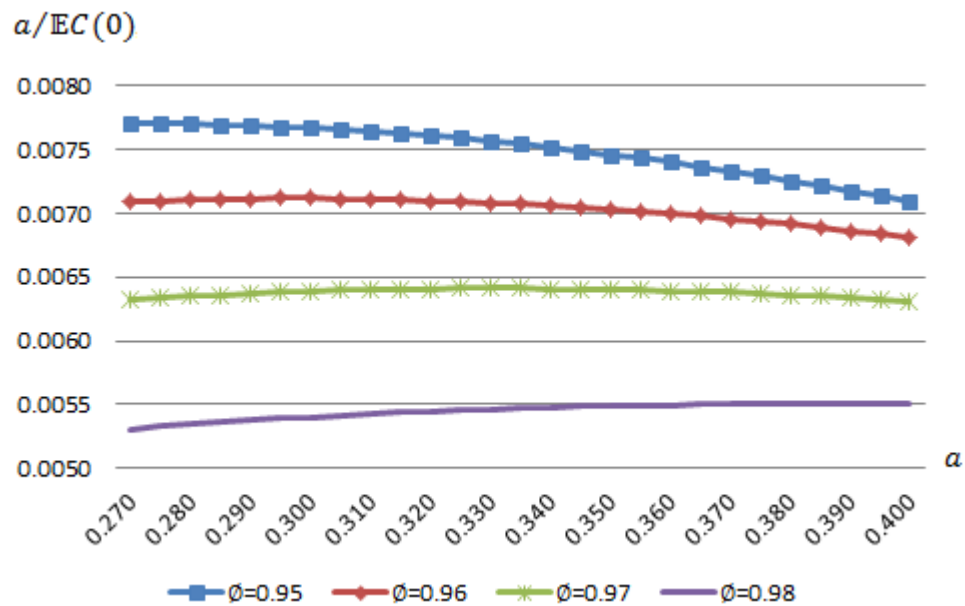


Figure C.1: Profit per unit time graph for  $\phi$  in  $[0.95, 0.98]$ ,  $\sigma = 0.1$  and  $\Delta b = 0.005$

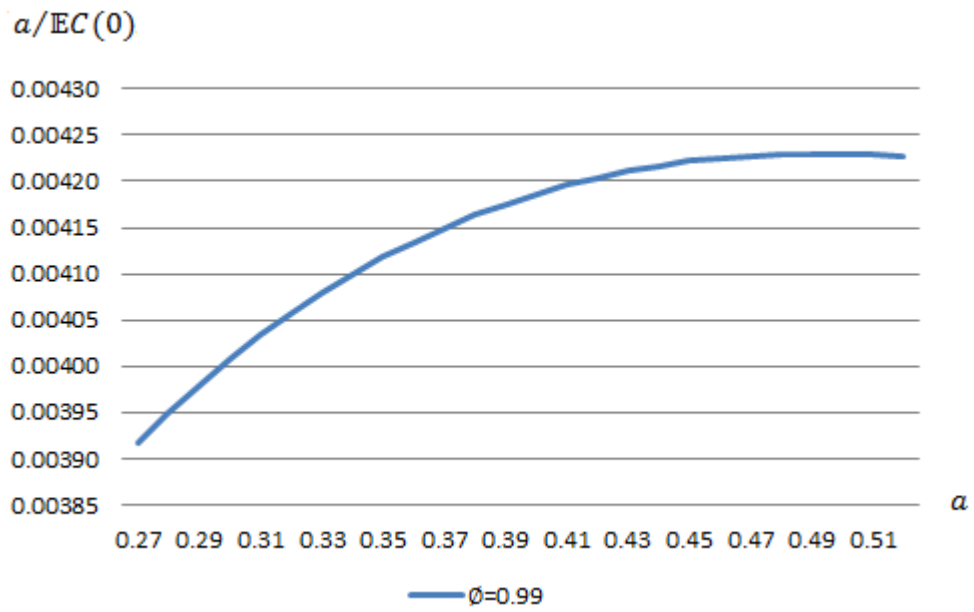


Figure C.2: Profit per unit time graph for  $\phi = 0.99$ ,  $\sigma = 0.1$  and  $\Delta b = 0.01$

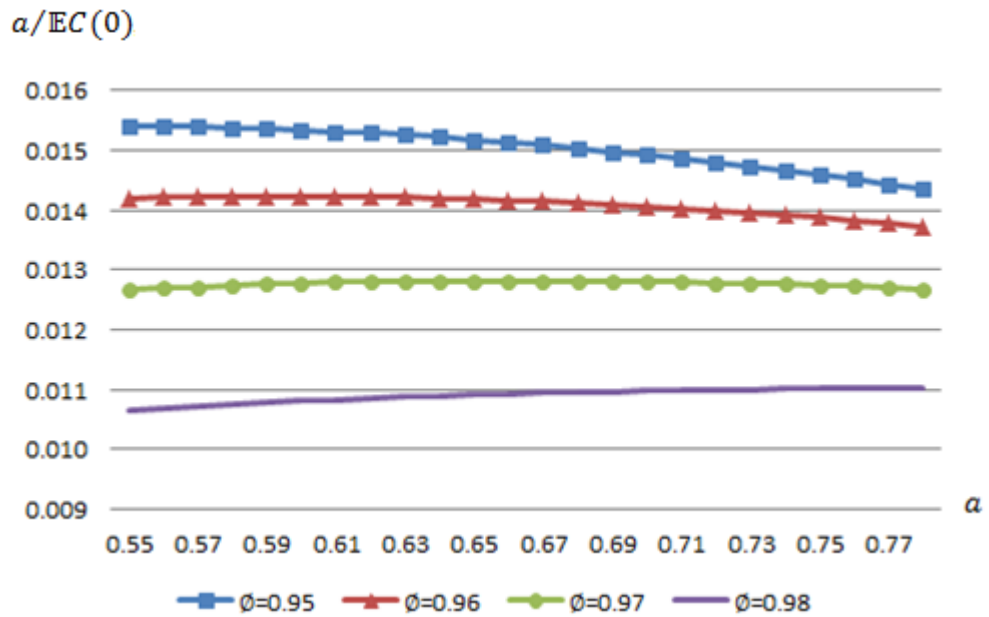


Figure C.3: Profit per unit time graph for  $\phi$  in  $[0.95, 0.98]$ ,  $\sigma = 0.2$  and  $\Delta b = 0.01$

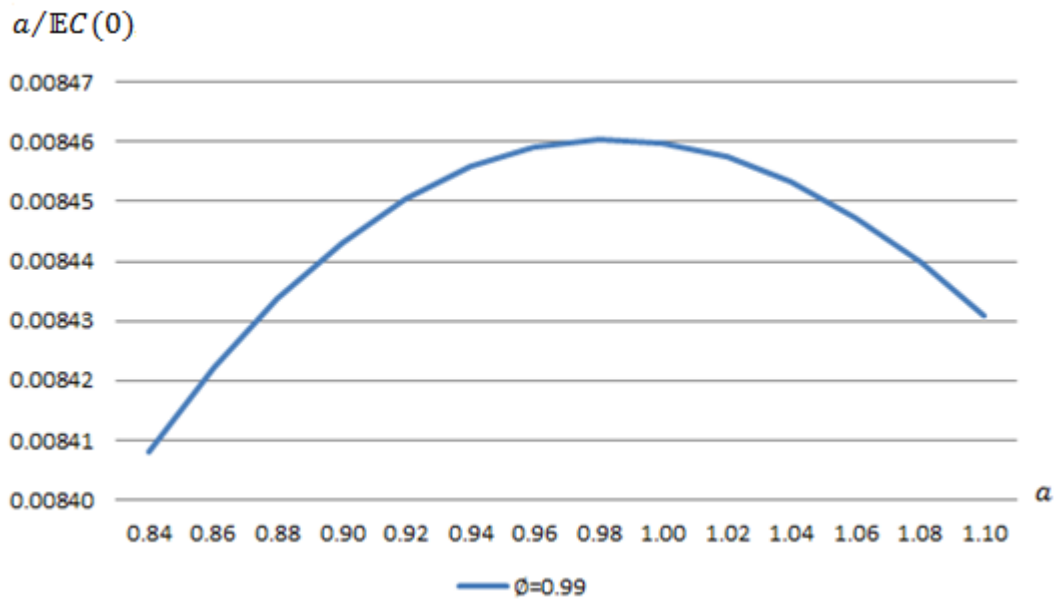


Figure C.4: Profit per unit time graph for  $\phi = 0.99$ ,  $\sigma = 0.2$  and  $\Delta b = 0.02$



# Appendix D

## Density, Empirical CDF and Q-Q Plots of Starting Points

Kolmogorov-Smirnov Two Sample Test is applied to the samples with the below hypothesis.

$H_0$ : The two samples come from the same continuous distribution

$H_1$ : The two samples do not come from the same continuous distribution

**Kolmogorov-Smirnov Two Sample Test for  $\phi = 0.95, a = 0.135$  and  $s = 0.05$**

$D = 0.022$

p – value = 0.9689

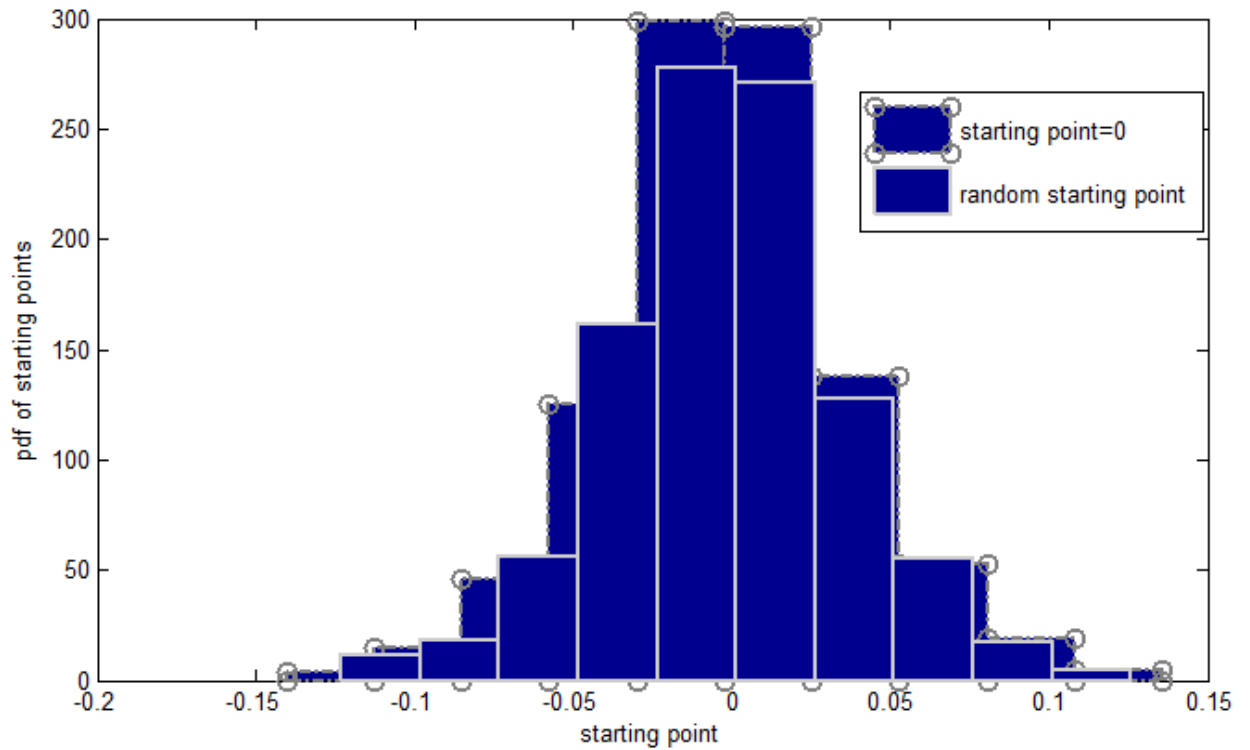


Figure D.1: Density graph for  $\phi = 0.95, a = 0.135$  and  $s = 0.05$

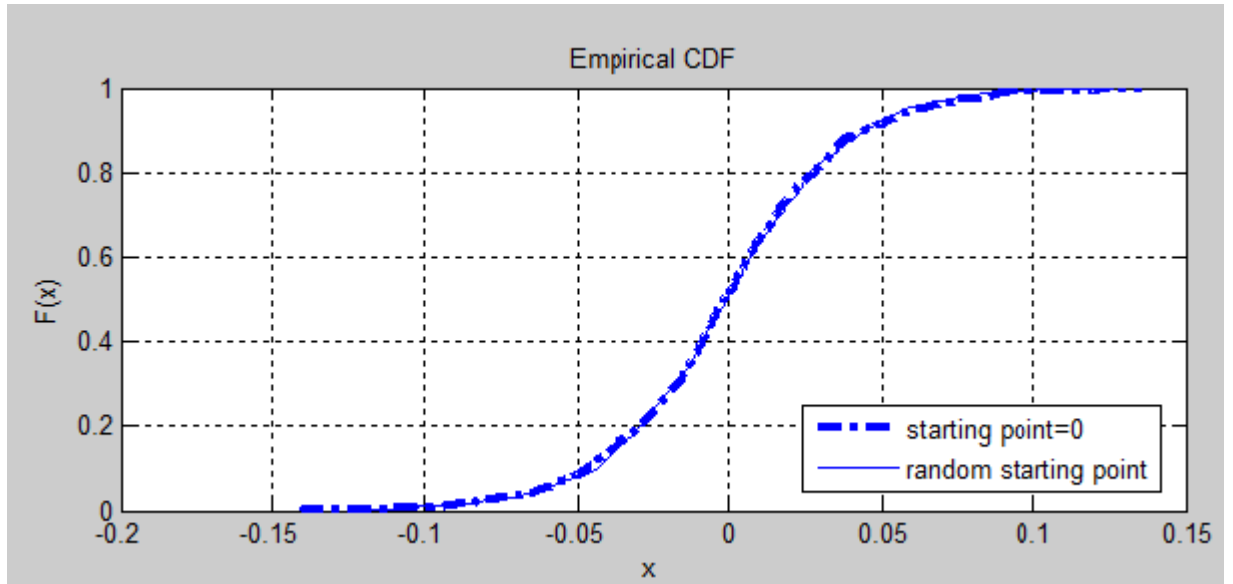


Figure D.2: Empirical CDF graph for  $\phi = 0.95, a = 0.135$  and  $s = 0.05$

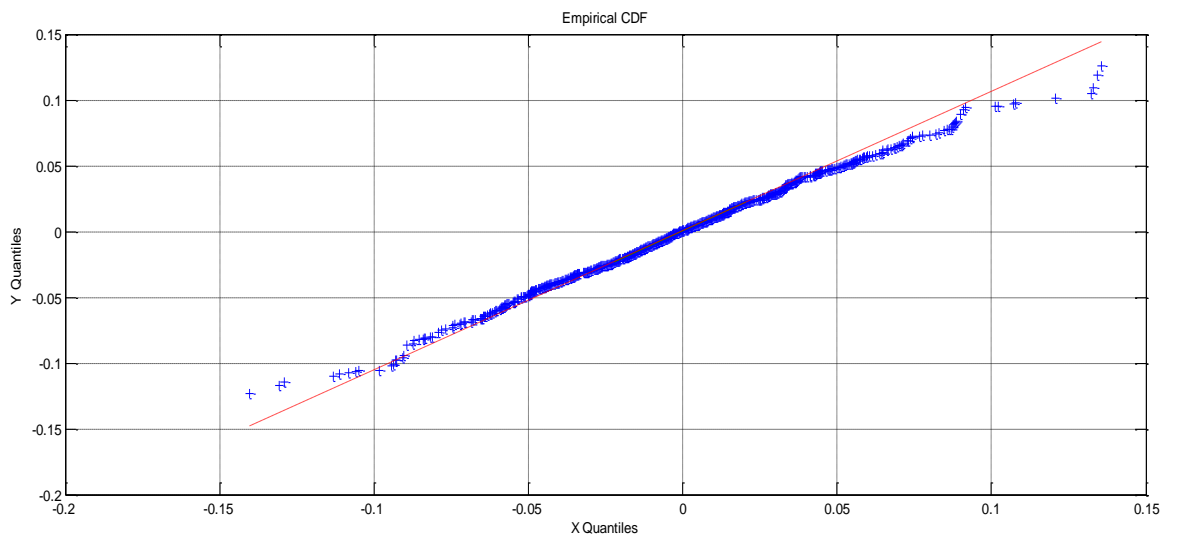


Figure D.3: Q-Q Plot for  $\phi = 0.95, a = 0.135$  and  $s = 0.05$

**Kolmogorov-Smirnov Two Sample Test for  $\phi = 0.99$ ,  $a = 0.245$  and  $s = 0.05$**

$D = 0.057$

p – value = 0.07762

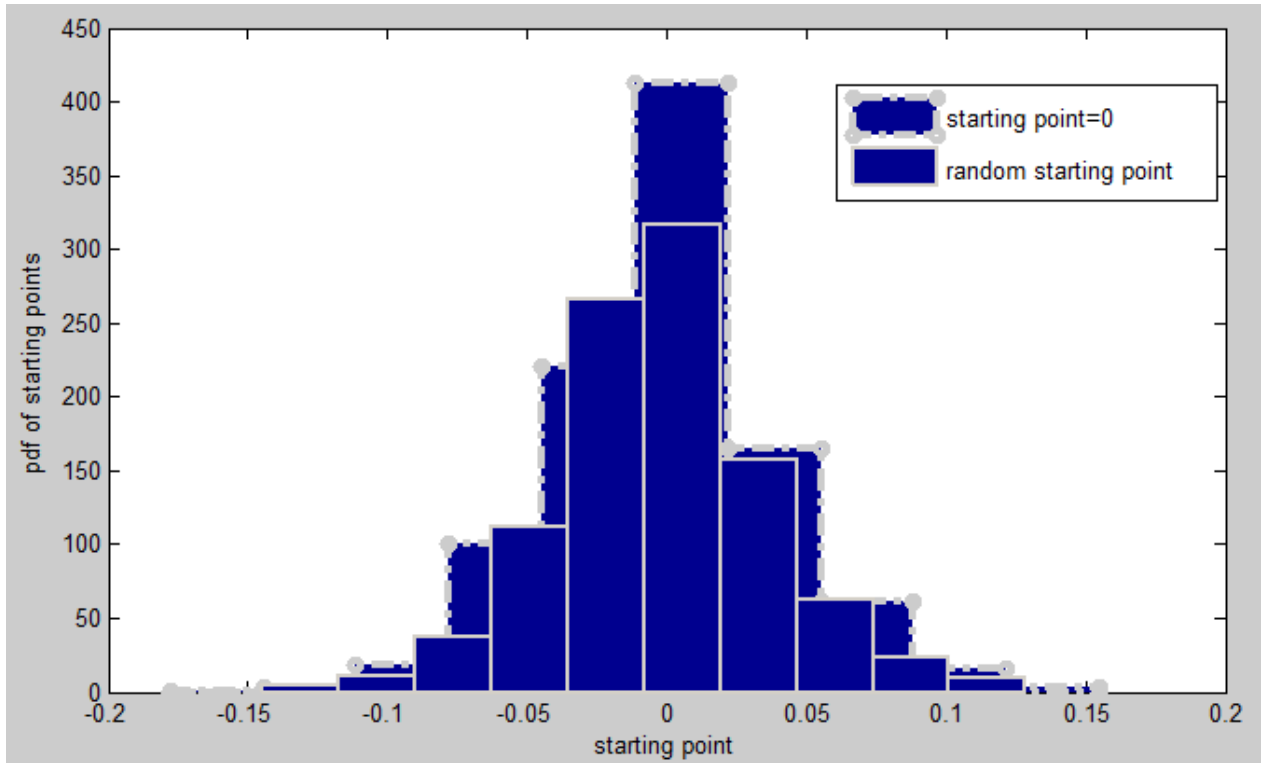


Figure D.4: Density graph for  $\phi = 0.99$ ,  $a = 0.245$  and  $s = 0.05$

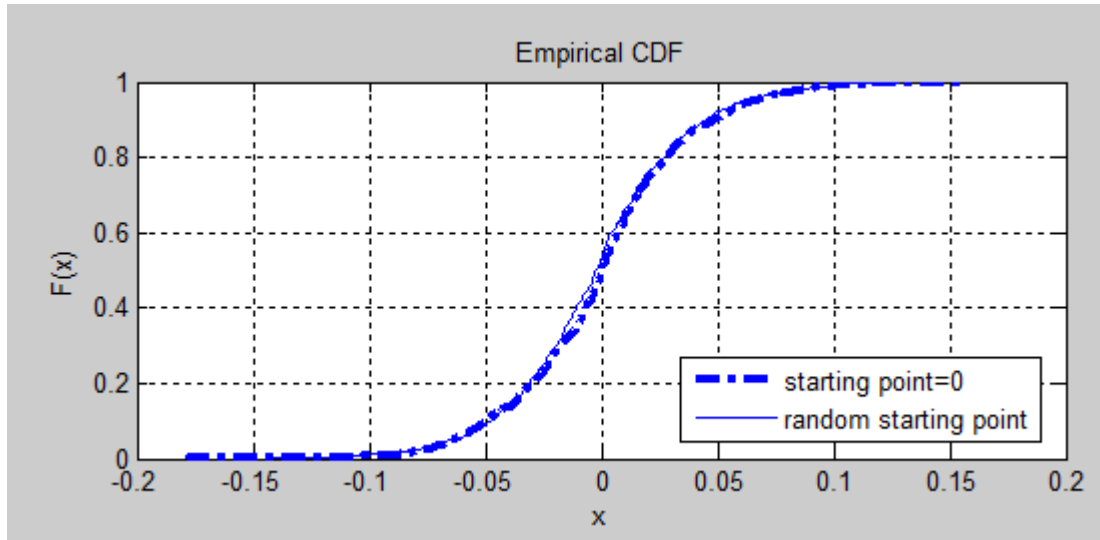


Figure D.5: Empirical CDF graph for  $\phi = 0.99, a = 0.245$  and  $s = 0.05$

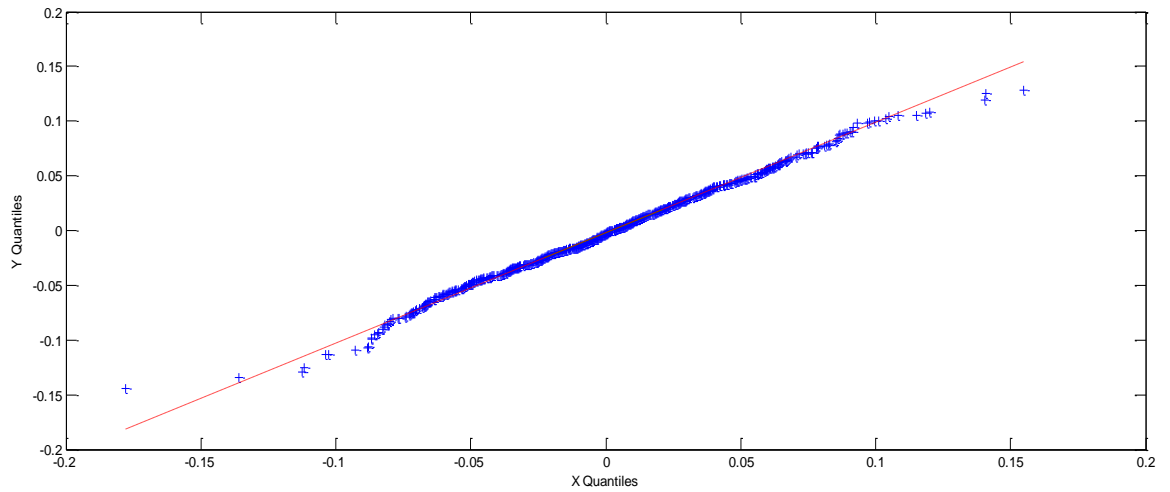


Figure D.6: Q-Q Plot for  $\phi = 0.99, a = 0.095$  and  $s = 0.05$

**Kolmogorov-Smirnov Two Sample Test for  $\phi = 0.95, a = 0.275$  and  $s = 0.1$**

$D - statistic = 0.04$

$p - value = 0.4005$

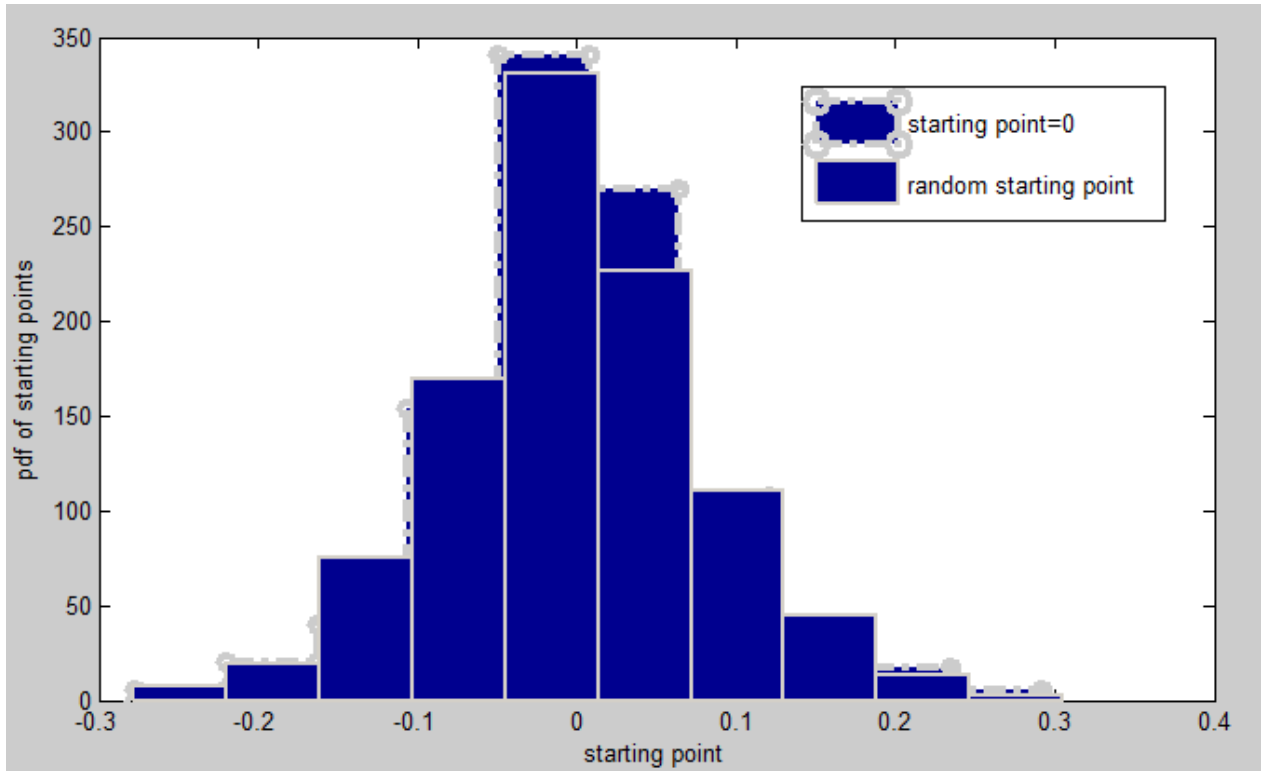


Figure D.7: Density graph for  $\phi = 0.95, a = 0.275$  and  $s = 0.1$

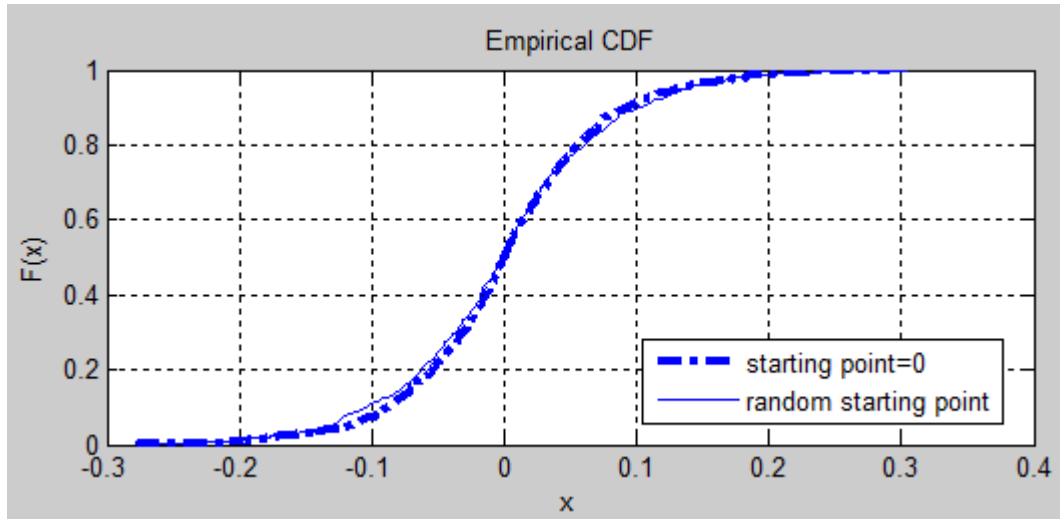


Figure D.8: Empirical CDF graph for  $\phi = 0.95, a = 0.275$  and  $s = 0.1$

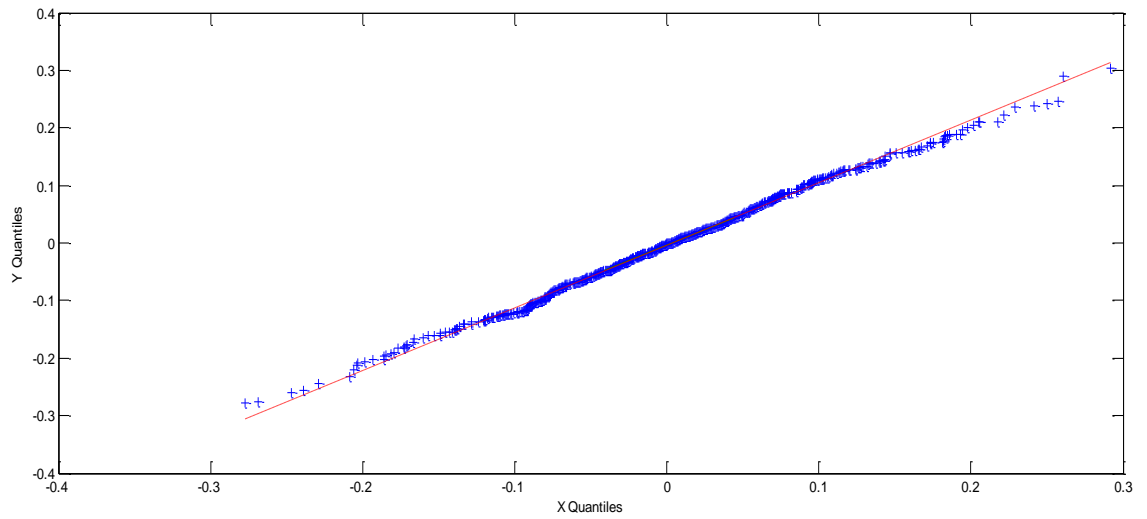


Figure D.9: Q-Q Plot for  $\phi = 0.95, a = 0.275$  and  $s = 0.1$

**Kolmogorov-Smirnov Two Sample Test for  $\phi = 0.99, a = 0.49$  and  $s = 0.1$**

$D - statistic = 0.043$

$p - value = 0.3136$

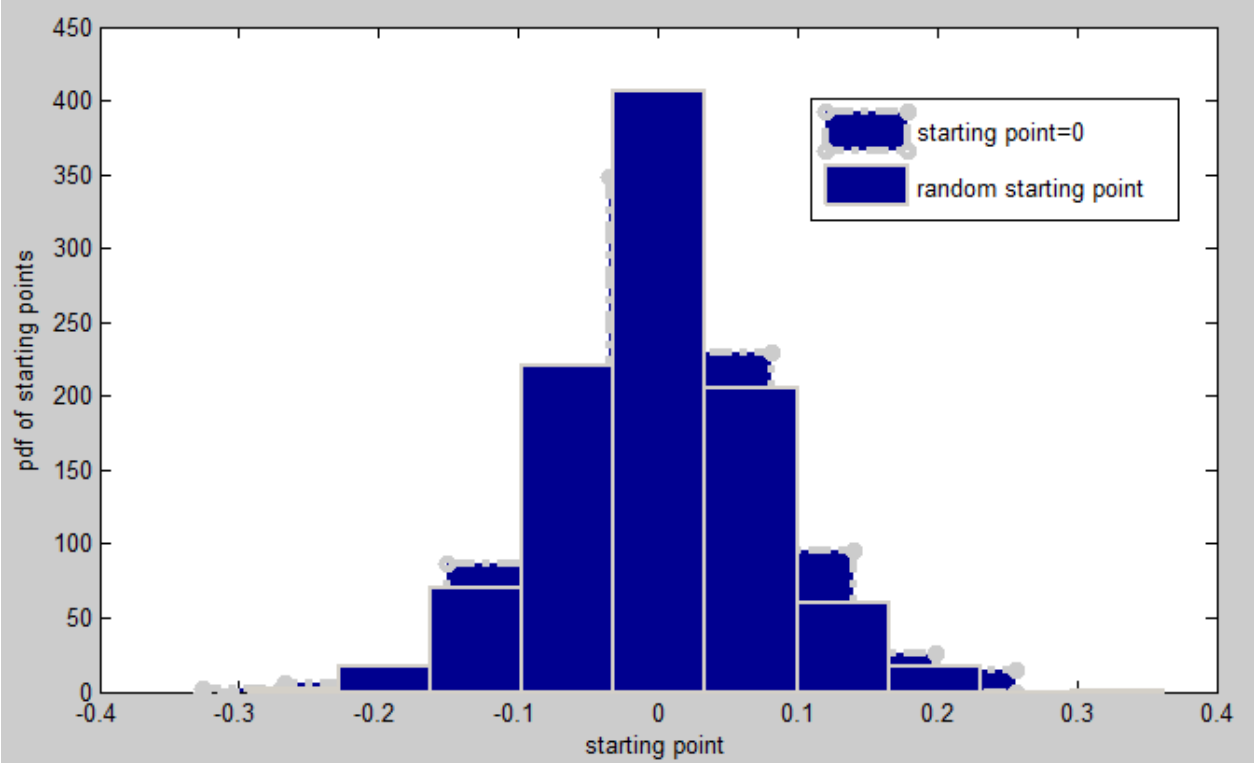


Figure D.10: Density graph for  $\phi = 0.99, a = 0.49$  and  $s = 0.1$



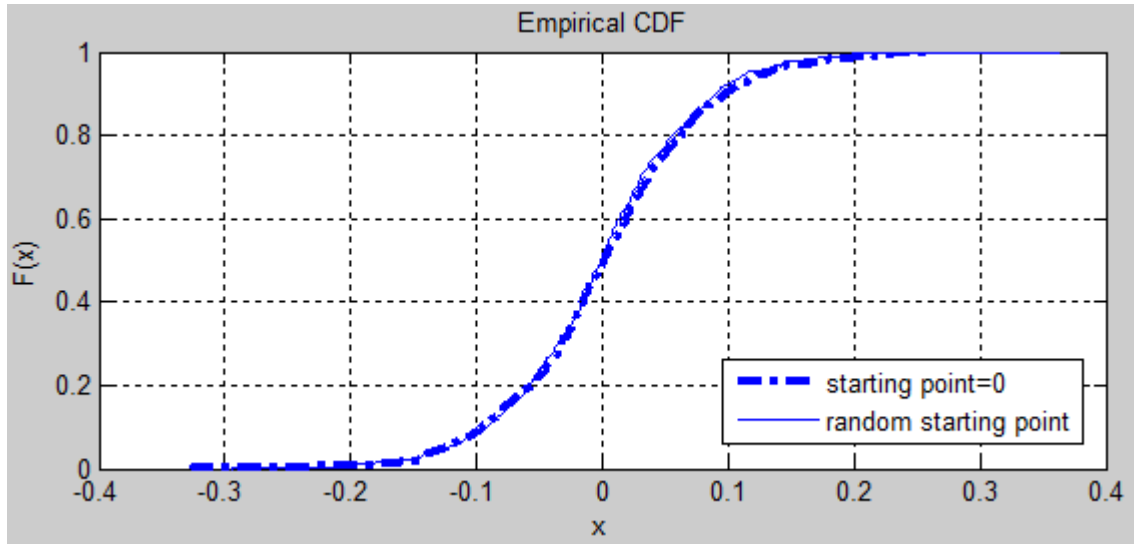


Figure D.11: Empirical CDF graph for  $\phi = 0.99, a = 0.49$  and  $s = 0.1$

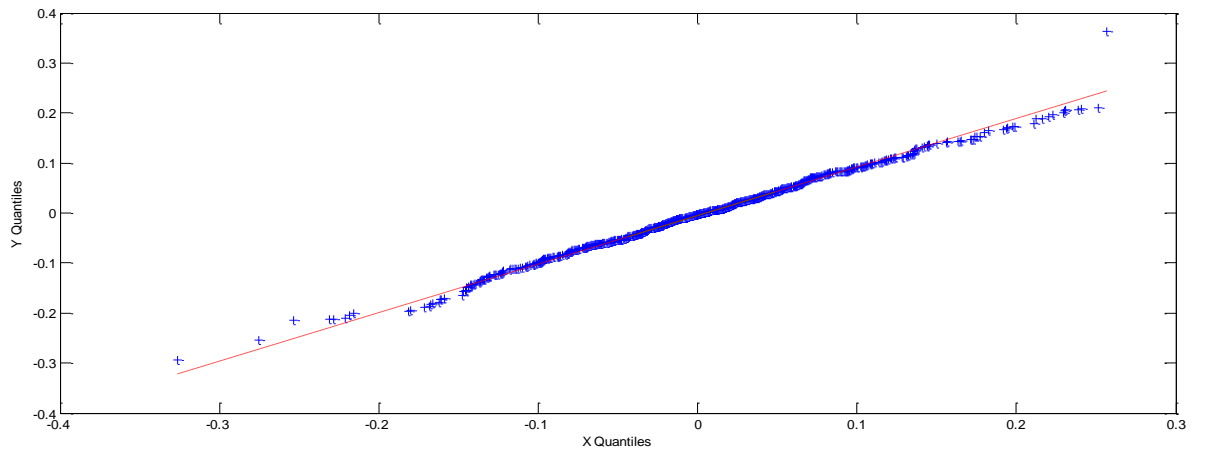


Figure D.12: Q-Q Plot for  $\phi = 0.99, a = 0.49$  and  $s = 0.1$

**Kolmogorov-Smirnov Two Sample Test for  $\phi = 0.95, a = 0.55$  and  $s = 0.2$**

$D - statistic = 0.028$

$p - value = 0.828$

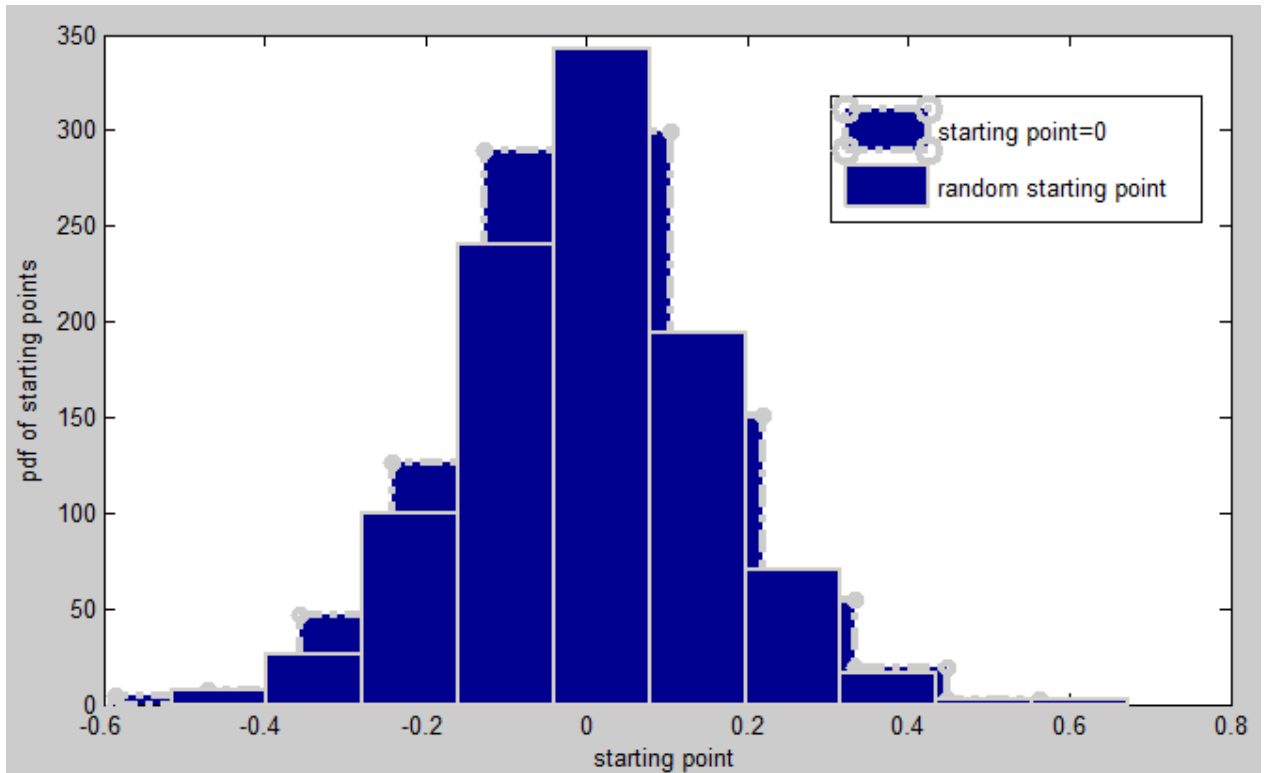


Figure D.13: Density graph for  $\phi = 0.95, a = 0.55$  and  $s = 0.2$

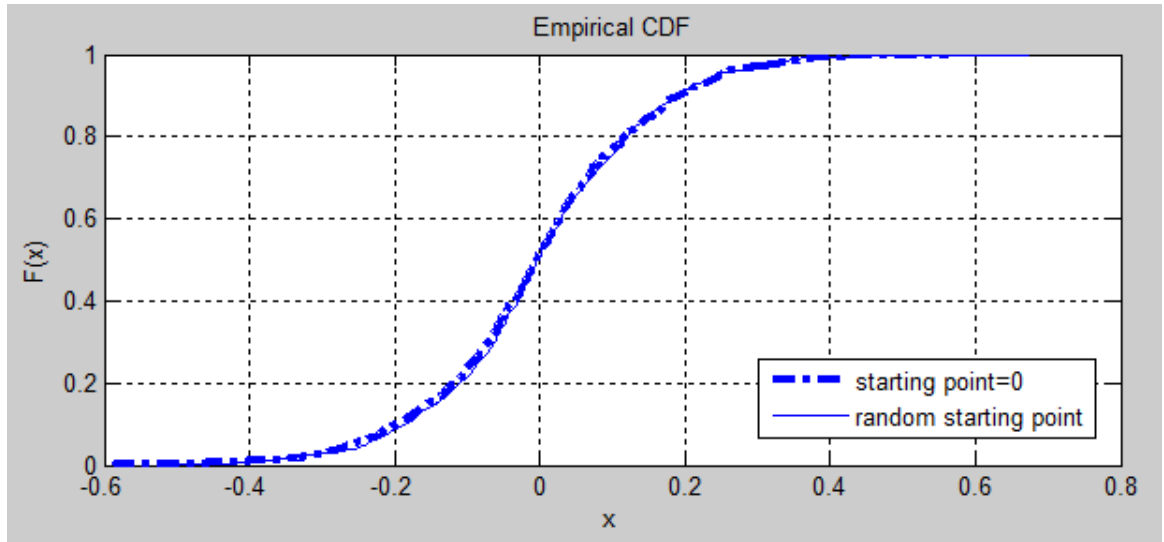


Figure D.14: Empirical CDF graph for  $\phi = 0.95, a = 0.55$  and  $s = 0.2$

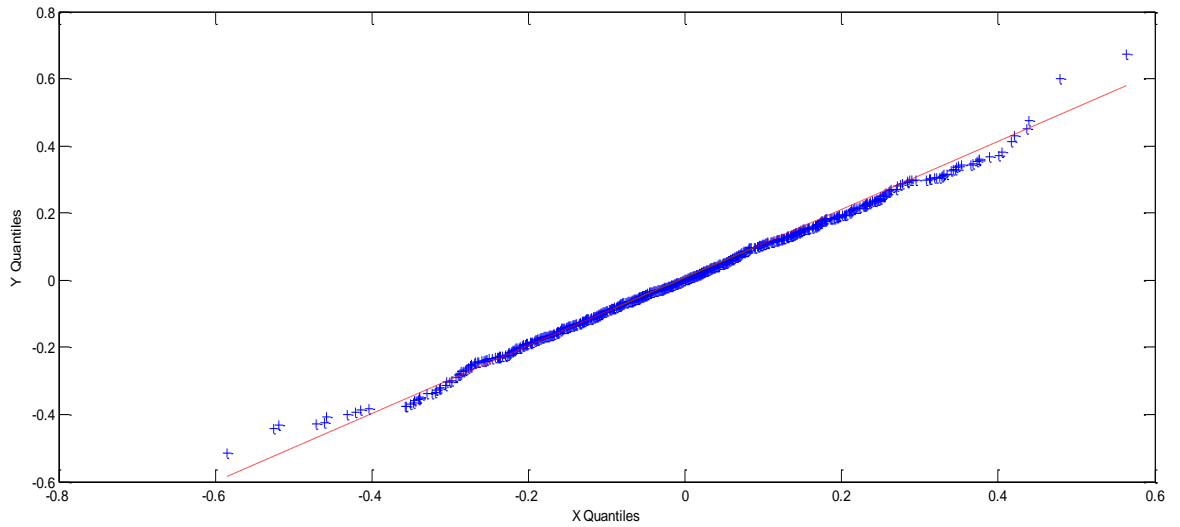


Figure D.15: Q-Q Plot for  $\phi = 0.95, a = 0.55$  and  $s = 0.2$

**Kolmogorov-Smirnov Two Sample Test for  $\phi = 0.99, a = 0.98$  and  $s = 0.2$**

$D = 0.05$

p – value = 0.1641

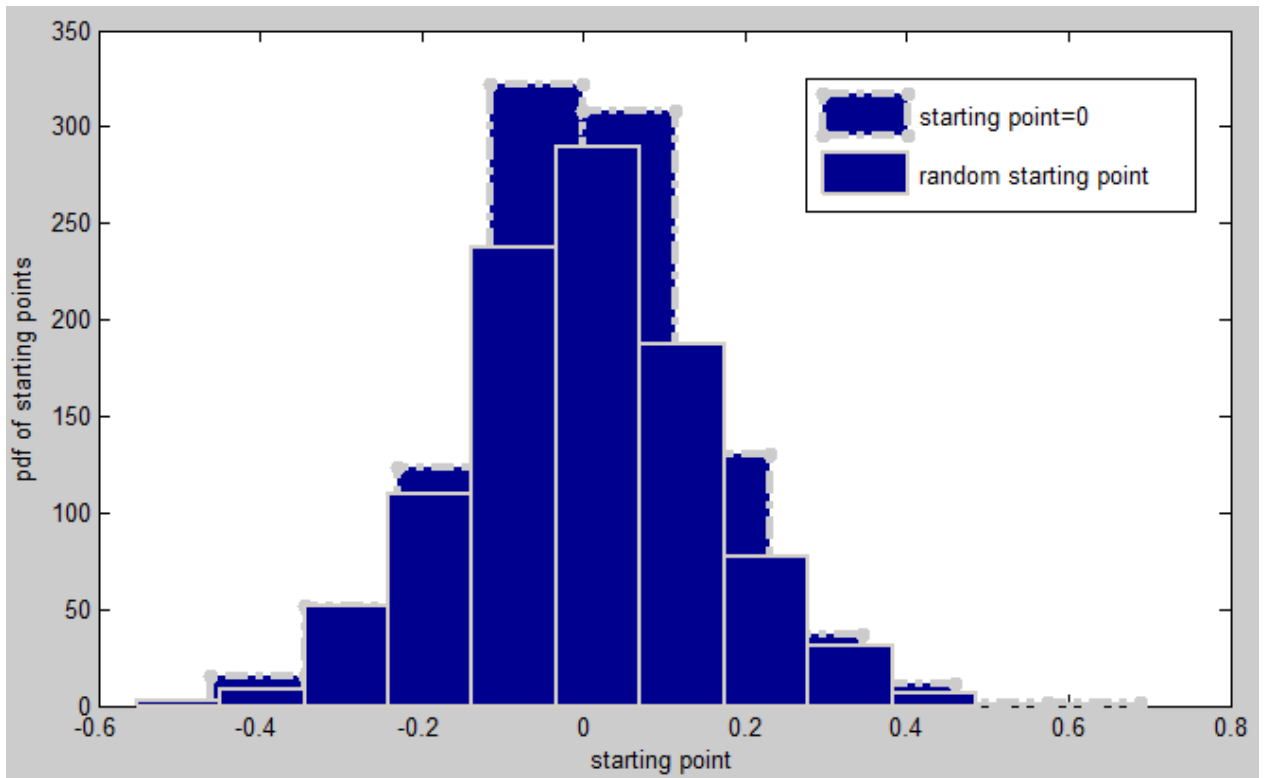


Figure D.16: Density graph for  $\phi = 0.99, a = 0.98$  and  $s = 0.2$

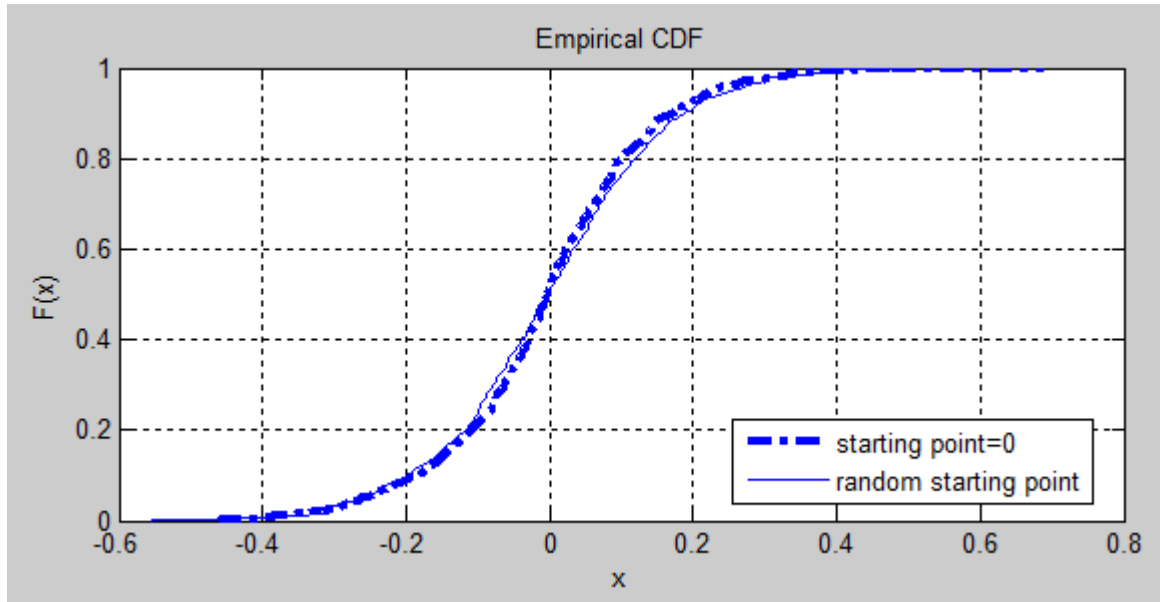


Figure D.17: Empirical CDF graph for  $\phi = 0.99, a = 0.98$  and  $s = 0.2$

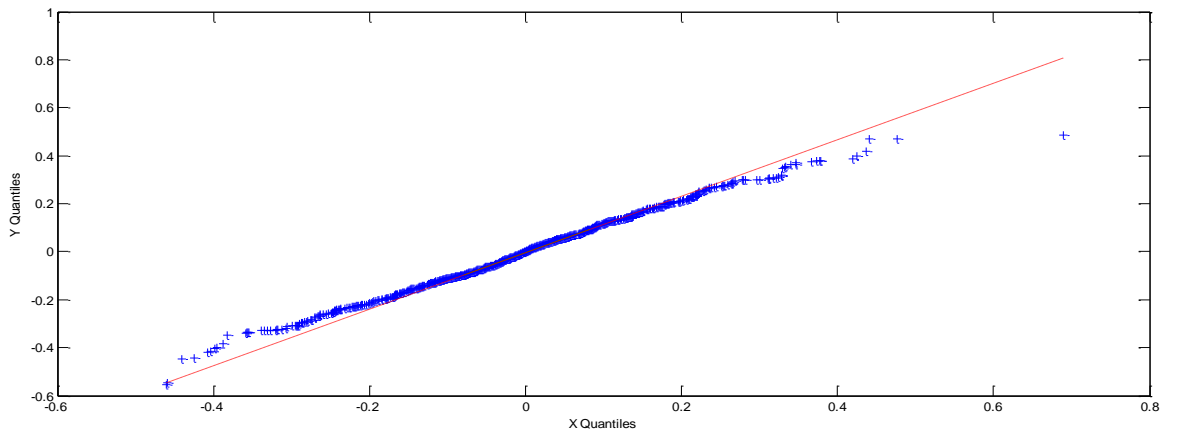


Figure D.18: Q-Q Plot for  $\phi = 0.99, a = 0.98$  and  $s = 0.2$

# Appendix E

## Expected Profit per Unit Time for Random Starting Point

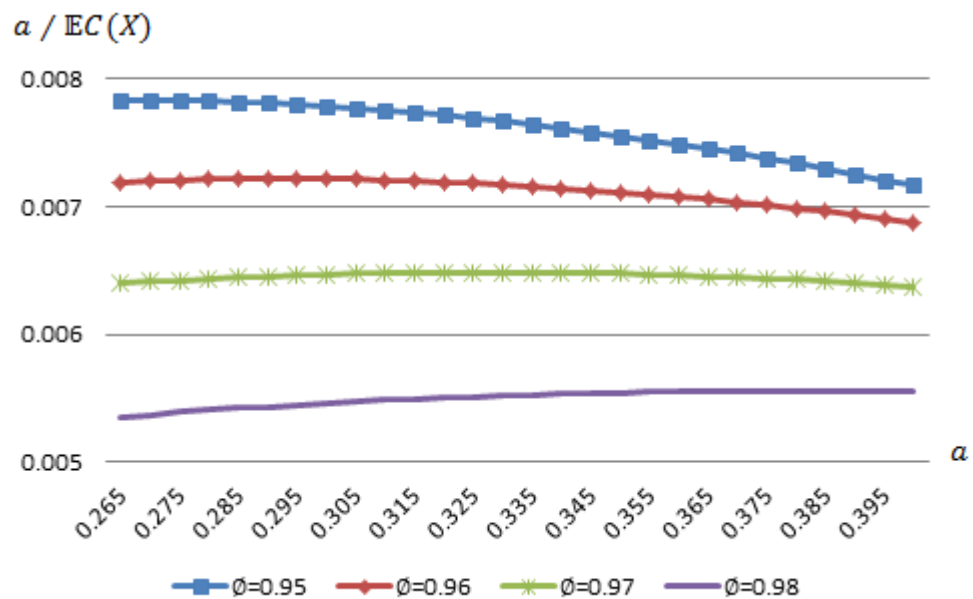


Figure E.1: Expected profit per unit time graph for  $\phi$  in  $[0.95, 0.98]$ ,  $\sigma = 0.1$  and  $\Delta b = 0.005$

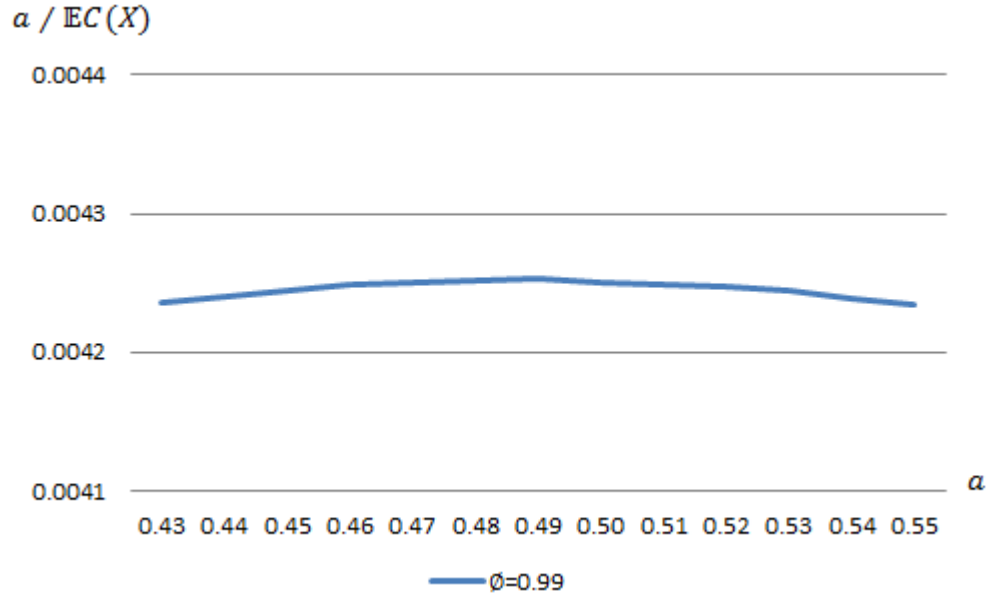


Figure E.2: Expected profit per unit time graph for  $\phi = 0.99$ ,  $\sigma = 0.1$  and  $\Delta b = 0.01$

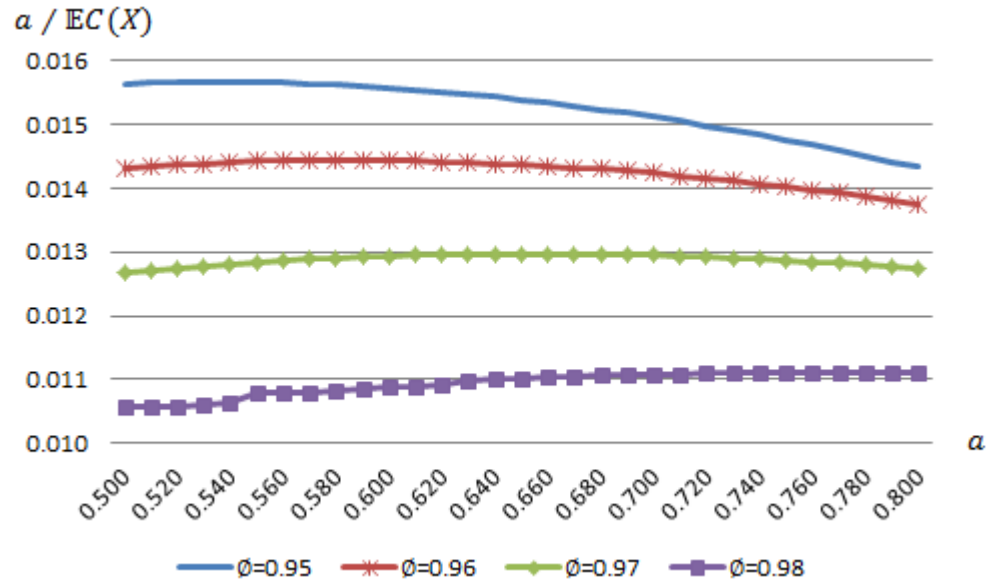


Figure E.3: Expected profit per unit time graph for  $\phi$  in  $[0.95, 0.98]$ ,  $\sigma = 0.2$  and  $\Delta b = 0.01$

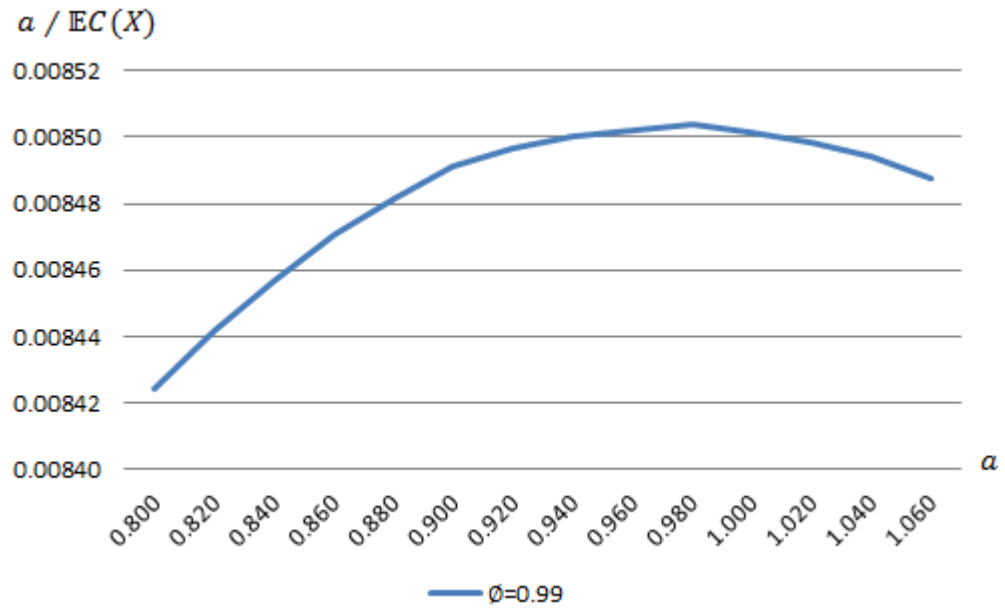


Figure E.4: Expected profit per unit time graph for  $\phi = 0.99$ ,  $\sigma = 0.2$  and  $\Delta b = 0.02$



# **Appendix F**

## Profit per Unit Time Values for Simulated Samples

Table F.1: Profit per unit time values for 20 samples for  $\emptyset = 0.95$  series

Trade in Point	sample 1	sample 2	sample 3	sample 4	sample 5	sample 6	sample 7	sample 8	sample 9	sample 10	sample 11	sample 12	sample 13	sample 14	sample 15	sample 16	sample 17	sample 18	sample 19	sample 20	Mean Profit
0.45	0.1369	0.1261	0.1315	0.1314	0.1254	0.1333	0.1243	0.1331	0.1316	0.1335	0.1322	0.131	0.1176	0.1236	0.1203	0.1276	0.1273	0.1163	0.1222	0.1206	0.1272900
0.46	0.1274	0.1204	0.1325	0.1276	0.1225	0.1276	0.1223	0.1257	0.1241	0.1264	0.1377	0.1158	0.1344	0.138	0.1177	0.1302	0.122	0.1284	0.1261	0.128	0.1261750
0.47	0.1224	0.1308	0.1207	0.1237	0.1253	0.1291	0.1237	0.1299	0.133	0.1292	0.1289	0.1365	0.1246	0.1262	0.1272	0.1133	0.1304	0.1261	0.1382	0.1269450	
0.48	0.1335	0.1196	0.1263	0.1261	0.1272	0.1424	0.1244	0.1287	0.1378	0.1272	0.1234	0.1244	0.127	0.1242	0.1265	0.1309	0.1345	0.1233	0.1308	0.1189	0.1279900
0.49	0.1279	0.1296	0.1221	0.1311	0.1133	0.1216	0.1373	0.1332	0.1356	0.131	0.1344	0.1208	0.1201	0.1295	0.1164	0.1212	0.1272	0.1236	0.1174	0.1267	0.1260000
0.5	0.129	0.1284	0.13	0.1227	0.1314	0.1307	0.1241	0.1153	0.1182	0.1283	0.121	0.124	0.1291	0.1228	0.1253	0.1278	0.1196	0.1176	0.1111	0.1309	0.1248650
0.51	0.1264	0.1292	0.1256	0.1321	0.1275	0.1279	0.1226	0.1285	0.1193	0.1381	0.1271	0.1218	0.1251	0.1381	0.1244	0.1244	0.1218	0.1279	0.1352	0.1339	0.1276100
0.52	0.1273	0.1272	0.1186	0.121	0.1326	0.1168	0.1154	0.1249	0.1211	0.1205	0.1283	0.1267	0.1354	0.1178	0.1211	0.1268	0.1239	0.1174	0.1354	0.1237	0.1240950
0.53	0.1263	0.1221	0.1321	0.1201	0.128	0.1175	0.1183	0.1299	0.1279	0.1263	0.1289	0.1259	0.1244	0.1268	0.1265	0.1244	0.123	0.1241	0.1157	0.1217	0.1249950
0.54	0.1212	0.1272	0.1163	0.1208	0.132	0.129	0.1193	0.1286	0.1329	0.1251	0.125	0.1309	0.1179	0.1204	0.1168	0.1327	0.1153	0.1305	0.1291	0.1245	0.1247750
0.55	0.1231	0.1327	0.1302	0.1205	0.1225	0.1236	0.1235	0.1269	0.1221	0.1195	0.1301	0.1334	0.117	0.129	0.1306	0.1296	0.1284	0.1248	0.1285	0.1219	0.1258950
0.56	0.1341	0.1184	0.1204	0.1202	0.1256	0.1286	0.1175	0.1318	0.1231	0.1163	0.1259	0.1212	0.1244	0.1176	0.123	0.1273	0.1255	0.1253	0.1214	0.1267	0.1237150
0.57	0.1275	0.1283	0.1387	0.1226	0.1234	0.1228	0.1272	0.1267	0.1303	0.1215	0.1172	0.1237	0.1327	0.1352	0.1262	0.13	0.1402	0.1299	0.1271	0.1171	0.1274150
0.58	0.1249	0.1267	0.1213	0.1217	0.1321	0.1199	0.1327	0.1209	0.1349	0.118	0.1222	0.1283	0.1289	0.1286	0.1236	0.1294	0.1281	0.1203	0.1347	0.1387	0.1267950
0.59	0.1313	0.124	0.1246	0.125	0.1268	0.1283	0.1319	0.1172	0.1418	0.1149	0.1353	0.1207	0.1168	0.1349	0.1267	0.1262	0.1301	0.1202	0.1312	0.1107	0.1259900
0.6	0.1263	0.1262	0.1199	0.1248	0.1285	0.1265	0.1297	0.1297	0.124	0.1319	0.1333	0.137	0.1319	0.1268	0.1212	0.1279	0.1226	0.1209	0.1276	0.1207	0.1268700
0.61	0.1216	0.1197	0.1353	0.1229	0.1219	0.1157	0.1175	0.1233	0.1158	0.1299	0.123	0.1213	0.1261	0.1233	0.1159	0.1263	0.1295	0.1237	0.1232	0.12	0.1227850
0.62	0.1258	0.134	0.1362	0.1297	0.1234	0.1159	0.1198	0.1285	0.1234	0.12	0.1269	0.1269	0.1309	0.1182	0.1268	0.1174	0.129	0.1275	0.1295	0.1259750	
0.63	0.1167	0.1203	0.1383	0.1246	0.1233	0.1309	0.1196	0.1298	0.1357	0.1328	0.132	0.1244	0.1294	0.1306	0.1359	0.1243	0.127	0.1147	0.1211	0.1335	0.1270950
0.64	0.1121	0.1286	0.1204	0.1185	0.1204	0.1234	0.1207	0.1338	0.1172	0.1215	0.1207	0.1153	0.1313	0.1292	0.1241	0.1213	0.1168	0.1277	0.118	0.1312	0.1230550

Table F.2: Profit per unit time values for 20 samples for  $\emptyset = 0.96$  series

Trade in Point	Profit per Unit Time																				Mean Profit
	sample 1	sample 2	sample 3	sample 4	sample 5	sample 6	sample 7	sample 8	sample 9	sample 10	sample 11	sample 12	sample 13	sample 14	sample 15	sample 16	sample 17	sample 18	sample 19	sample 20	
0.47	0.0988	0.1046	0.1126	0.1184	0.1045	0.1131	0.117	0.114	0.1174	0.1218	0.1136	0.1148	0.1152	0.1213	0.1136	0.1171	0.1071	0.1141	0.1177	0.108	0.1132350
0.48	0.1181	0.1117	0.1219	0.1083	0.1099	0.1187	0.1115	0.1042	0.1202	0.1088	0.0961	0.1041	0.0985	0.1132	0.1261	0.1081	0.1037	0.1159	0.1128	0.1169	0.1114350
0.49	0.1109	0.1054	0.0984	0.1233	0.1244	0.1096	0.1165	0.1159	0.1119	0.1104	0.1162	0.1126	0.1098	0.112	0.1169	0.1058	0.112	0.1083	0.1012	0.1187	0.1120100
0.5	0.1052	0.1118	0.1132	0.1139	0.1196	0.1195	0.1187	0.1044	0.1137	0.1044	0.1296	0.1122	0.1052	0.1018	0.1086	0.1133	0.1033	0.1123	0.1057	0.1112	0.1113800
0.51	0.1051	0.1203	0.1167	0.1108	0.1038	0.1143	0.1152	0.1131	0.1241	0.1063	0.1309	0.1143	0.1117	0.1163	0.1107	0.1089	0.115	0.1025	0.1126	0.1133000	
0.52	0.1166	0.1187	0.1128	0.1149	0.1127	0.1176	0.1262	0.1084	0.1208	0.1103	0.1123	0.1186	0.1114	0.1202	0.1162	0.1018	0.1181	0.1185	0.1156	0.1218	0.1156750
0.53	0.103	0.1088	0.1121	0.1118	0.1182	0.1199	0.1014	0.1174	0.1082	0.1145	0.1089	0.1172	0.1187	0.1125	0.1241	0.1119	0.1133	0.1127	0.1119	0.1063	0.1128400
0.54	0.1149	0.1185	0.1238	0.1162	0.1186	0.108	0.1139	0.1143	0.1091	0.1116	0.1038	0.1119	0.1176	0.1182	0.1112	0.1079	0.1153	0.1089	0.1065	0.115	0.1132600
0.55	0.1123	0.1135	0.1242	0.1053	0.1115	0.1094	0.1119	0.1125	0.1145	0.1187	0.1112	0.1164	0.1023	0.1143	0.1158	0.0991	0.1131	0.1134	0.1209	0.1097	0.1125000
0.56	0.1077	0.1069	0.1134	0.113	0.1181	0.1176	0.107	0.1134	0.1067	0.1211	0.1082	0.1155	0.1167	0.1191	0.1015	0.1097	0.1078	0.116	0.115	0.1142	0.1122550
0.57	0.1073	0.1058	0.1145	0.1119	0.0962	0.1144	0.1118	0.1114	0.1185	0.1086	0.1148	0.1074	0.1124	0.1179	0.1236	0.1172	0.1175	0.1117	0.112	0.1153	0.1125100
0.58	0.1159	0.1094	0.1085	0.1113	0.1177	0.1019	0.1142	0.1105	0.1055	0.1189	0.1167	0.1182	0.1109	0.1123	0.1235	0.1091	0.1019	0.1129	0.1108	0.1072	0.1118650
0.59	0.1067	0.1086	0.1046	0.1062	0.1094	0.1114	0.1161	0.1109	0.1234	0.1077	0.1033	0.1009	0.1106	0.112	0.1113	0.1132	0.1054	0.1075	0.1116	0.1302	0.1105500
0.6	0.1185	0.1044	0.1199	0.1168	0.1182	0.1208	0.106	0.1095	0.1197	0.1164	0.1169	0.1089	0.1156	0.117	0.1009	0.1073	0.1154	0.1151	0.1067	0.1245	0.1139250
0.61	0.1251	0.1091	0.1019	0.1162	0.115	0.1232	0.1151	0.108	0.1062	0.1024	0.1095	0.1132	0.1184	0.1058	0.1033	0.121	0.1146	0.1163	0.1247	0.113	0.1131000
0.62	0.1085	0.1216	0.1049	0.1161	0.1166	0.1058	0.1101	0.1115	0.1105	0.1143	0.1054	0.1147	0.1161	0.1108	0.106	0.1097	0.1158	0.123	0.1144	0.1137	0.1124750
0.63	0.1032	0.1072	0.1206	0.1121	0.1216	0.1232	0.1103	0.1126	0.1103	0.1072	0.1054	0.1059	0.1099	0.1164	0.1131	0.1146	0.1111	0.119	0.116	0.1068	0.1123250
0.64	0.1181	0.1188	0.1155	0.1145	0.1035	0.1059	0.1168	0.1265	0.1221	0.1117	0.1194	0.1199	0.106	0.1036	0.1061	0.1143	0.1133	0.1018	0.1126	0.1226	0.1136500
0.65	0.101	0.1161	0.1276	0.1031	0.1132	0.12	0.1045	0.1087	0.1051	0.1045	0.1089	0.1117	0.1142	0.1009	0.1058	0.1093	0.1042	0.1143	0.1243	0.1039	0.1103300
0.66	0.1092	0.1121	0.1129	0.1103	0.1139	0.1142	0.1141	0.1156	0.1104	0.1062	0.1184	0.1117	0.1105	0.1051	0.1104	0.1153	0.1074	0.1063	0.1152	0.1155	0.1120000
0.67	0.1166	0.1028	0.1205	0.1106	0.1227	0.1237	0.1116	0.1185	0.1214	0.0985	0.1086	0.1068	0.0996	0.1151	0.1146	0.106	0.1184	0.111	0.1056	0.109	0.1120800
0.68	0.1234	0.1156	0.1063	0.1131	0.1137	0.1165	0.1118	0.1138	0.1177	0.1082	0.1042	0.1074	0.1103	0.1118	0.1173	0.1174	0.1103	0.1102	0.1055	0.1067	0.1120600

Table F.3: Profit per unit time values for 20 samples for  $\emptyset = 0.97$  series

Trade In Point	sample 1	sample 2	sample 3	sample 4	sample 5	sample 6	sample 7	sample 8	sample 9	sample 10	sample 11	sample 12	sample 13	sample 14	sample 15	sample 16	sample 17	sample 18	sample 19	sample 20	Mean Profit
0.55	0.0925	0.0951	0.0977	0.0937	0.095	0.1032	0.1004	0.098	0.0986	0.0996	0.1019	0.0916	0.0983	0.1016	0.0995	0.1013	0.108	0.0922	0.0999	0.0946	0.0981350
0.56	0.0969	0.1023	0.0984	0.0974	0.0845	0.0953	0.0967	0.1002	0.1008	0.0866	0.1071	0.0997	0.1061	0.0931	0.0995	0.1142	0.099	0.083	0.1033	0.1002	0.0982150
0.57	0.0913	0.0862	0.0965	0.0981	0.0959	0.088	0.0997	0.0958	0.099	0.1063	0.0812	0.0936	0.091	0.1079	0.0969	0.0918	0.0904	0.0951	0.0925	0.0959	0.0946550
0.58	0.0961	0.0971	0.0922	0.0928	0.098	0.093	0.0954	0.0863	0.097	0.0972	0.0925	0.0893	0.0934	0.0938	0.1062	0.1005	0.094	0.0886	0.0945	0.0939	0.0949000
0.59	0.1061	0.0988	0.0959	0.0958	0.1026	0.0886	0.0947	0.1006	0.0894	0.1037	0.0923	0.1014	0.0928	0.0876	0.1	0.0978	0.096	0.1057	0.1002	0.0896	0.0969800
0.6	0.1058	0.0971	0.1011	0.0952	0.0864	0.0838	0.1019	0.0968	0.0973	0.0875	0.0914	0.1021	0.1015	0.0998	0.0997	0.087	0.0851	0.1065	0.0957	0.1019	0.0961800
0.61	0.1	0.097	0.097	0.1121	0.0978	0.1018	0.084	0.0958	0.0962	0.1078	0.0926	0.1018	0.0909	0.1035	0.1036	0.103	0.0955	0.0961	0.1011	0.0928	0.0985200
0.62	0.0941	0.1004	0.0828	0.0917	0.1123	0.0948	0.0879	0.1001	0.1004	0.0985	0.0975	0.1006	0.1054	0.0945	0.0957	0.1012	0.1022	0.0923	0.0976	0.1044	0.0977200
0.63	0.103	0.0995	0.1001	0.0891	0.0897	0.084	0.0925	0.0992	0.0963	0.1033	0.0942	0.0964	0.1018	0.0963	0.0983	0.0986	0.0939	0.1021	0.1012	0.1009	0.0970200
0.64	0.0866	0.0974	0.1002	0.1086	0.1075	0.1034	0.0926	0.0934	0.0851	0.0946	0.1055	0.0965	0.0968	0.0951	0.108	0.1055	0.0971	0.1001	0.0886	0.1	0.0981300
0.65	0.0976	0.101	0.1035	0.0986	0.0978	0.0975	0.0982	0.0956	0.106	0.0954	0.1048	0.1072	0.0913	0.1022	0.105	0.0948	0.0949	0.0991	0.1076	0.0922	0.0995150
0.66	0.1047	0.1005	0.0959	0.0995	0.0916	0.11	0.1032	0.0953	0.0764	0.0872	0.0869	0.1057	0.087	0.0939	0.1025	0.0918	0.0923	0.1067	0.0981	0.1019	0.0965550
0.67	0.0952	0.1005	0.0986	0.0985	0.0956	0.0988	0.0988	0.0939	0.1004	0.0915	0.0927	0.0984	0.1003	0.0911	0.1085	0.0927	0.0959	0.0945	0.0985	0.1035	0.0973950
0.68	0.0991	0.0943	0.0879	0.0974	0.1061	0.1005	0.0859	0.1088	0.1041	0.1004	0.1002	0.0964	0.1093	0.0972	0.098	0.0967	0.1043	0.0992	0.097	0.0964	0.0989600
0.69	0.0949	0.1015	0.1037	0.0943	0.1171	0.1001	0.0921	0.0929	0.0958	0.096	0.0995	0.0848	0.0953	0.0988	0.1018	0.0924	0.1014	0.0977	0.1016	0.097	0.0979350
0.7	0.1057	0.1009	0.0915	0.1119	0.0918	0.0938	0.0924	0.098	0.0951	0.0975	0.0916	0.1017	0.1005	0.1054	0.0869	0.1036	0.1031	0.1044	0.1026	0.0985	0.0988450
0.71	0.092	0.1004	0.0974	0.0965	0.1009	0.0971	0.0948	0.1016	0.0893	0.0969	0.1004	0.0961	0.0887	0.1035	0.0847	0.0914	0.0973	0.0928	0.0921	0.0922	0.0963050
0.72	0.0907	0.1074	0.0917	0.104	0.1037	0.0888	0.1021	0.0983	0.0995	0.0976	0.0923	0.0883	0.0957	0.0961	0.0963	0.0953	0.0999	0.0982	0.0986	0.0906	0.0967550
0.73	0.1144	0.0964	0.0961	0.0999	0.0962	0.0945	0.0889	0.1029	0.0994	0.0935	0.0956	0.097	0.0943	0.0966	0.0859	0.0931	0.0924	0.0998	0.0881	0.1092	0.0967100
0.74	0.1009	0.0969	0.1007	0.1004	0.0909	0.1032	0.0959	0.0969	0.095	0.1039	0.1045	0.1055	0.0934	0.0937	0.0927	0.0865	0.0948	0.1012	0.0996	0.0954	0.0976000
0.75	0.1063	0.0896	0.1022	0.1002	0.0961	0.0912	0.0945	0.0948	0.0984	0.1011	0.0966	0.0899	0.0976	0.0988	0.0962	0.0971	0.0878	0.0973	0.1067	0.0991	0.0970750

Table F.4: Profit per unit time values for 20 samples for  $\emptyset = 0.98$  series

Trade in Point	Profit per Unit Time																				Mean Profit
	sample 1	sample 2	sample 3	sample 4	sample 5	sample 6	sample 7	sample 8	sample 9	sample 10	sample 11	sample 12	sample 13	sample 14	sample 15	sample 16	sample 17	sample 18	sample 19	sample 20	
0.65	0.0685	0.0753	0.0848	0.0917	0.0812	0.0876	0.0753	0.0846	0.0843	0.0939	0.0798	0.0817	0.0807	0.0699	0.0833	0.0842	0.0809	0.0869	0.0838	0.0822	
0.66	0.0726	0.0765	0.0855	0.0776	0.0766	0.0763	0.0832	0.0765	0.0735	0.0843	0.0832	0.0717	0.0797	0.0789	0.086	0.0859	0.0791	0.0743	0.081	0.0829	
0.67	0.0799	0.0785	0.0824	0.0947	0.086	0.0906	0.0891	0.0854	0.0902	0.0739	0.0813	0.081	0.0811	0.0784	0.0785	0.0854	0.0737	0.0762	0.0795	0.085	
0.68	0.0809	0.0702	0.0874	0.0811	0.0807	0.0791	0.0838	0.0784	0.0744	0.0787	0.0816	0.0733	0.0858	0.0819	0.08	0.0772	0.0892	0.0897	0.07	0.0823	
0.69	0.0828	0.0836	0.0754	0.0814	0.0767	0.0778	0.0807	0.0815	0.0818	0.0843	0.0761	0.0712	0.0739	0.0735	0.0875	0.0815	0.0706	0.0782	0.0749	0.087	
0.7	0.0835	0.0851	0.0674	0.0802	0.0809	0.0794	0.0746	0.0745	0.0735	0.0847	0.0709	0.0884	0.0861	0.0764	0.0848	0.0723	0.085	0.083	0.0806	0.093	
0.71	0.086	0.0763	0.0654	0.0904	0.0712	0.0754	0.0758	0.0807	0.0853	0.0797	0.082	0.0623	0.0757	0.08	0.0886	0.0877	0.0852	0.0823	0.0801	0.0891	
0.72	0.0844	0.0805	0.0836	0.0747	0.0752	0.0664	0.0852	0.0695	0.0902	0.0821	0.0794	0.0828	0.0764	0.0825	0.0837	0.0835	0.071	0.0847	0.0691	0.0787	
0.73	0.0744	0.0744	0.0816	0.0774	0.0783	0.0788	0.0811	0.0784	0.0817	0.0813	0.0853	0.0764	0.076	0.0838	0.0803	0.0792	0.0735	0.0708	0.0806	0.0798	
0.74	0.0779	0.085	0.0775	0.0884	0.0957	0.0759	0.073	0.0782	0.0751	0.0818	0.0859	0.0785	0.0901	0.088	0.0779	0.075	0.0743	0.0748	0.079	0.0786	
0.75	0.0685	0.0737	0.0668	0.0807	0.0782	0.0826	0.067	0.0813	0.0822	0.0754	0.0873	0.0754	0.087	0.0829	0.0799	0.0755	0.0882	0.0805	0.0682	0.0753	
0.76	0.0786	0.0692	0.0742	0.0795	0.0811	0.074	0.0801	0.0844	0.0738	0.069	0.0767	0.0806	0.0817	0.0842	0.0838	0.0914	0.0925	0.0744	0.0867	0.077	
0.77	0.0821	0.077	0.0739	0.0836	0.0875	0.0817	0.0763	0.0831	0.0821	0.0879	0.0841	0.082	0.0739	0.0768	0.0702	0.0811	0.0822	0.0801	0.0788	0.0803	
0.78	0.0776	0.0756	0.0836	0.0841	0.0846	0.0845	0.0865	0.0818	0.0685	0.0806	0.0928	0.0748	0.0864	0.0651	0.076	0.0774	0.0843	0.0899	0.083	0.081	
0.79	0.0773	0.0789	0.0794	0.0799	0.0856	0.077	0.0748	0.0918	0.0671	0.0691	0.0802	0.0742	0.0799	0.0896	0.0794	0.0763	0.079	0.0789	0.0873	0.0843	
0.8	0.0814	0.082	0.081	0.0696	0.072	0.0735	0.0759	0.0698	0.0771	0.0766	0.0741	0.0836	0.0763	0.0847	0.0798	0.0761	0.0837	0.0731	0.0778	0.0768	
0.81	0.088	0.0714	0.0914	0.075	0.0857	0.0773	0.0815	0.0823	0.0796	0.0736	0.0793	0.0685	0.0705	0.0925	0.0794	0.0883	0.0809	0.0801	0.0748	0.0735	
0.82	0.0651	0.0832	0.0743	0.0727	0.0782	0.0799	0.0856	0.0823	0.0809	0.0767	0.0828	0.0832	0.0861	0.0805	0.0723	0.0824	0.0816	0.0797	0.0747	0.0804	
0.83	0.0802	0.0834	0.0747	0.0792	0.0823	0.081	0.0916	0.0794	0.0796	0.0739	0.0895	0.0746	0.0774	0.0807	0.0825	0.0823	0.0874	0.0868	0.0755	0.0809	
0.84	0.0786	0.0846	0.0835	0.0832	0.08	0.0716	0.0842	0.0776	0.0855	0.0738	0.0815	0.079	0.0819	0.0725	0.076	0.0881	0.0776	0.0782	0.0799	0.0882	
0.85	0.0799	0.0829	0.0874	0.0875	0.0728	0.0848	0.096	0.0821	0.0869	0.0791	0.0757	0.0865	0.0865	0.0781	0.0725	0.0807	0.0844	0.0813	0.0865	0.0769	

Table F.5: Profit per unit time values for 20 samples for  $\emptyset = 0.99$  series

Trade in Point	Profit per Unit Time																				Mean Profit
	sample 1	sample 2	sample 3	sample 4	sample 5	sample 6	sample 7	sample 8	sample 9	sample 10	sample 11	sample 12	sample 13	sample 14	sample 15	sample 16	sample 17	sample 18	sample 19	sample 20	
0.88	0.0621	0.0673	0.0561	0.0604	0.0563	0.0431	0.0571	0.0535	0.0565	0.0548	0.0543	0.0609	0.0502	0.0602	0.0652	0.062	0.0461	0.0694	0.0455	0.0479	0.05644
0.89	0.0656	0.056	0.0669	0.0497	0.0499	0.0636	0.0616	0.0598	0.0609	0.057	0.0521	0.0582	0.0604	0.0625	0.0603	0.0558	0.055	0.0489	0.053	0.0486	0.05709
0.9	0.0626	0.0584	0.0591	0.0383	0.0593	0.0474	0.0517	0.0548	0.0591	0.0536	0.0489	0.0575	0.0633	0.0584	0.0556	0.0467	0.0523	0.057	0.0588	0.0646	0.05537
0.91	0.0522	0.0611	0.0546	0.049	0.0565	0.0619	0.0529	0.0517	0.058	0.0599	0.0564	0.0556	0.061	0.047	0.0415	0.0646	0.0442	0.0431	0.0448	0.0593	0.053765
0.92	0.0623	0.0509	0.056	0.0558	0.058	0.0528	0.0504	0.0537	0.0525	0.0609	0.073	0.0391	0.0653	0.0535	0.0486	0.0626	0.0579	0.0608	0.0683	0.0566	0.05695
0.93	0.0541	0.0606	0.0656	0.0686	0.0589	0.0563	0.0548	0.0557	0.0531	0.0493	0.0603	0.0569	0.0511	0.0452	0.059	0.0503	0.0407	0.0529	0.0557	0.0561	0.05526
0.94	0.0623	0.0522	0.0654	0.0582	0.0564	0.0578	0.0502	0.0551	0.046	0.0528	0.0592	0.0497	0.051	0.0534	0.0637	0.0587	0.059	0.0484	0.0559	0.0558	0.05556
0.95	0.0541	0.063	0.0508	0.0488	0.0638	0.0534	0.0543	0.0544	0.067	0.0632	0.0578	0.0529	0.0529	0.0556	0.0514	0.0513	0.0473	0.0638	0.0513	0.0489	0.0553
0.96	0.0504	0.0526	0.053	0.0428	0.0519	0.0533	0.0594	0.0586	0.0596	0.0561	0.0633	0.0501	0.0444	0.0569	0.0632	0.0546	0.0532	0.0545	0.0517	0.0557	0.05476
0.97	0.0543	0.0504	0.05	0.0549	0.0622	0.0576	0.0624	0.0532	0.0566	0.044	0.0562	0.0501	0.058	0.0453	0.0648	0.0466	0.0525	0.0535	0.0544	0.0562	0.05416
0.98	0.0637	0.0558	0.0554	0.0612	0.0602	0.0564	0.061	0.0552	0.0448	0.0559	0.0603	0.0458	0.0581	0.0545	0.0539	0.0583	0.0532	0.062	0.0652	0.0598	0.057035
0.99	0.0623	0.0464	0.0584	0.0717	0.0539	0.068	0.055	0.0637	0.054	0.0678	0.0546	0.0677	0.0613	0.0585	0.0553	0.0586	0.0663	0.0483	0.049	0.0589	0.058985
1	0.0533	0.0569	0.0605	0.0454	0.0462	0.0611	0.0534	0.0544	0.0746	0.0597	0.0611	0.0616	0.0448	0.0576	0.0525	0.0418	0.0469	0.0467	0.0485	0.0546	0.05408
1.01	0.0548	0.0539	0.06	0.0625	0.0507	0.0474	0.0668	0.0487	0.0521	0.0504	0.0446	0.054	0.0523	0.0606	0.0482	0.0532	0.048	0.0537	0.0559	0.0626	0.05412
1.02	0.0629	0.0641	0.0645	0.0535	0.0518	0.0529	0.0517	0.061	0.0504	0.0631	0.061	0.0607	0.0522	0.0627	0.0635	0.0585	0.06	0.0639	0.0556	0.056	0.0585
1.03	0.0522	0.0699	0.0548	0.0631	0.0595	0.0645	0.0526	0.0664	0.0516	0.0553	0.0468	0.0534	0.0572	0.0431	0.0551	0.0641	0.0589	0.0503	0.0513	0.0564	0.056325
1.04	0.0612	0.0711	0.0628	0.0515	0.0543	0.0495	0.0546	0.0595	0.0467	0.0595	0.0611	0.0545	0.0546	0.0445	0.0663	0.0634	0.0551	0.0569	0.0437	0.0523	0.056155
1.05	0.0506	0.0533	0.0644	0.0562	0.0547	0.0496	0.054	0.0516	0.0884	0.0552	0.0572	0.0527	0.0585	0.0664	0.054	0.056	0.0486	0.0564	0.0608	0.0647	0.056665
1.06	0.0524	0.0509	0.0517	0.0621	0.0639	0.0567	0.0531	0.0432	0.0543	0.0513	0.0539	0.0504	0.0607	0.0599	0.0531	0.0601	0.0622	0.0492	0.0611	0.0564	0.05533
1.07	0.0588	0.0543	0.0608	0.0705	0.0491	0.0524	0.0584	0.0595	0.0616	0.0574	0.0562	0.0544	0.0568	0.0641	0.0608	0.055	0.0706	0.0546	0.0605	0.0612	0.05885
1.08	0.0694	0.0582	0.0619	0.0534	0.0655	0.0599	0.0547	0.0522	0.0575	0.0555	0.054	0.0515	0.056	0.0672	0.0542	0.06	0.0628	0.0644	0.0566	0.0577	0.05863

# **Appendix G**

Estimated Parameter Values of  
Simulated Data

Table G.1: Estimated parameter values of simulated data for  $\phi = 0.96$

Sample No	Estimated $\phi$ value	Estimated $\sigma$ value
1	0.951	0.900
2	0.958	0.881
3	0.964	0.967
4	0.952	1.200
5	0.956	0.910
6	0.962	1.134
7	0.955	0.981
8	0.961	0.972
9	0.973	0.997
10	0.959	0.942
11	0.971	0.987
12	0.969	1.023
13	0.966	0.979
14	0.956	0.984
15	0.956	1.054
16	0.967	1.122
17	0.976	0.989
18	0.972	0.886
19	0.968	1.034
20	0.974	0.993
mean	0.963	0.997
std.dev.	0.008	0.080



Table G.2: Estimated parameter values of simulated data for  $\emptyset = 0.97$

Sample No	Estimated $\emptyset$ value	Estimated $\sigma$ value
1	0.972	0.983
2	0.971	0.979
3	0.969	0.990
4	0.976	0.971
5	0.977	0.993
6	0.981	1.056
7	0.979	0.989
8	0.981	1.210
9	0.976	0.979
10	0.964	1.190
11	0.971	1.014
12	0.982	0.968
13	0.972	0.973
14	0.978	0.981
15	0.984	0.995
16	0.968	1.159
17	0.972	0.997
18	0.966	0.984
19	0.970	0.992
20	0.975	0.987
mean	0.974	1.020
std.dev.	0.005	0.073

Table G.3: Estimated parameter values of simulated data for  $\varnothing = 0.98$

Sample No	Estimated $\varnothing$ value	Estimated $\sigma$ value
1	0.982	0.991
2	0.989	0.989
3	0.961	1.097
4	0.998	0.999
5	0.957	0.988
6	0.969	0.984
7	0.990	1.171
8	0.978	0.981
9	0.981	1.089
10	0.986	1.010
11	0.991	0.981
12	0.989	1.068
13	0.964	1.023
14	0.991	1.082
15	0.980	1.105
16	0.979	1.019
17	0.982	0.993
18	0.996	0.992
19	0.981	0.998
20	0.978	0.982
mean	0.981	1.027
std.dev.	0.011	0.053

Table G.4: Estimated parameter values of simulated data for  $\emptyset = 0.99$

Sample No	Estimated $\emptyset$ value	Estimated $\sigma$ value
1	0.989	0.996
2	0.997	0.994
3	0.986	0.989
4	0.998	1.152
5	0.994	1.146
6	0.979	0.988
7	0.981	0.992
8	0.986	1.019
9	0.990	0.995
10	0.986	1.087
11	0.993	0.985
12	0.976	1.039
13	0.998	0.982
14	0.980	1.029
15	0.998	1.098
16	0.978	1.019
17	0.995	1.109
18	0.996	0.992
19	0.979	0.998
20	0.987	1.095
mean	0.988	1.035
std.dev.	0.007	0.056

# Appendix H

## Comparison of Two Methods

$a_1$  shows the most profitable trade-in point of historical based method and  $a_2$  shows the most profitable trade-in point of the proposed method.

Table H.1: Profit per unit time values of 20 samples for  $\phi = 0.95$ ,  $a_1 = 0.28$   
and  $a_2 = 0.54$

	Profit per Unit Time for Historical Based Method	Profit per Unit Time for Proposed Method
Sample 1	0.1227	0.1216
Sample 2	0.1410	0.1395
Sample 3	0.1384	0.1378
Sample 4	0.1237	0.1237
Sample 5	0.1239	0.1232
Sample 6	0.1285	0.1292
Sample 7	0.1249	0.1260
Sample 8	0.1236	0.1245
Sample 9	0.1238	0.1239
Sample 10	0.1297	0.1298
Sample 11	0.1134	0.1130
Sample 12	0.1171	0.1159
Sample 13	0.1308	0.1322
Sample 14	0.1312	0.1303
Sample 15	0.1254	0.1256
Sample 16	0.1300	0.1296
Sample 17	0.1171	0.1182
Sample 18	0.1247	0.1251
Sample 19	0.137	0.1378
Sample 20	0.1302	0.1292
Mean Profit	0.1268	0.1268

Table H.2: Profit per unit time values of 20 samples for  $\phi = 0.96$ ,  $a_1 = 0.60$   
and  $a_2 = 0.58$

	Profit per Unit Time for Historical Based Method	Profit per Unit Time for Proposed Method
Sample 1	0.112	0.1121
Sample 2	0.1115	0.1116
Sample 3	0.1133	0.1126
Sample 4	0.1204	0.1203
Sample 5	0.104	0.1035
Sample 6	0.1044	0.105
Sample 7	0.1146	0.1138
Sample 8	0.1148	0.1156
Sample 9	0.1124	0.1122
Sample 10	0.117	0.1171
Sample 11	0.1015	0.1014
Sample 12	0.1162	0.1162
Sample 13	0.1197	0.1191
Sample 14	0.1049	0.1052
Sample 15	0.1075	0.1072
Sample 16	0.114	0.1143
Sample 17	0.1169	0.1166
Sample 18	0.1121	0.1125
Sample 19	0.1051	0.1053
Sample 20	0.1137	0.1139
Mean Profit	0.1118	0.1118

Table H.3: Profit per unit time values of 20 samples for  $\phi = 0.97$ ,  $a_1 = 0.65$   
and  $a_2 = 0.66$

	Profit per Unit Time for Historical Based Method	Profit per Unit Time for Proposed Method
Sample 1	0.1051	0.105
Sample 2	0.0991	0.0995
Sample 3	0.0933	0.0938
Sample 4	0.0911	0.091
Sample 5	0.0952	0.0954
Sample 6	0.1031	0.103
Sample 7	0.1002	0.1004
Sample 8	0.0916	0.0916
Sample 9	0.0999	0.1004
Sample 10	0.0924	0.0929
Sample 11	0.1038	0.1038
Sample 12	0.0943	0.0949
Sample 13	0.1004	0.1002
Sample 14	0.1094	0.1101
Sample 15	0.1004	0.1001
Sample 16	0.0991	0.0992
Sample 17	0.0996	0.0996
Sample 18	0.096	0.0957
Sample 19	0.0849	0.0849
Sample 20	0.0929	0.0928
Mean Profit	0.0976	0.0977

Table H.4: Profit per unit time values of 20 samples for  $\phi = 0.98$ ,  $a_1 = 0.67$   
and  $a_2 = 0.76$

	Profit per Unit Time for Historical Based Method	Profit per Unit Time for Proposed Method
Sample 1	0.0825	0.0822
Sample 2	0.0756	0.0753
Sample 3	0.0763	0.076
Sample 4	0.0764	0.0753
Sample 5	0.0724	0.073
Sample 6	0.0691	0.0685
Sample 7	0.0826	0.081
Sample 8	0.0844	0.0843
Sample 9	0.0796	0.0793
Sample 10	0.0709	0.0708
Sample 11	0.0753	0.0753
Sample 12	0.0766	0.0766
Sample 13	0.0786	0.0779
Sample 14	0.0873	0.0865
Sample 15	0.0804	0.0814
Sample 16	0.0762	0.0762
Sample 17	0.0756	0.077
Sample 18	0.0746	0.0736
Sample 19	0.075	0.0757
Sample 20	0.0839	0.0835
Mean Profit	0.0777	0.0775



Table H.5: Profit per unit time values of 20 samples for  $\phi = 0.99$ ,  $a_1 = 0.99$   
and  $a_2 = 0.98$

	Profit per Unit Time for Historical Based Method	Profit per Unit Time for Proposed Method
Sample 1	0.0632	0.0632
Sample 2	0.0578	0.0578
Sample 3	0.057	0.057
Sample 4	0.056	0.0558
Sample 5	0.0619	0.0621
Sample 6	0.0532	0.0532
Sample 7	0.0568	0.0568
Sample 8	0.0621	0.0621
Sample 9	0.0638	0.0643
Sample 10	0.0704	0.0705
Sample 11	0.0667	0.0664
Sample 12	0.0615	0.0616
Sample 13	0.0623	0.0625
Sample 14	0.0503	0.0502
Sample 15	0.052	0.052
Sample 16	0.0519	0.0517
Sample 17	0.063	0.063
Sample 18	0.0591	0.059
Sample 19	0.0481	0.0479
Sample 20	0.0496	0.0495
Mean Profit	0.0583	0.0583