145283

Modelling Stock Market via Fuzzy Rule Based Systems

Dissertation submitted to the

Graduate Institute of Social Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Management

by

Hakan Aksoy

Boğaziçi University

2004

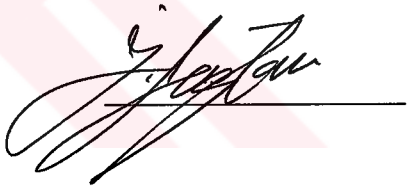
The dissertation of Hakan Aksoy

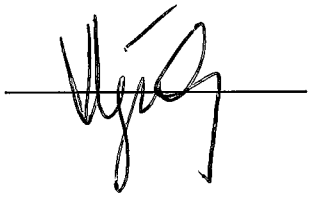
is approved by

Prof. Dr. Vedat Akgiray (Committee Chairperson) 

Assoc. Prof. Dr. Celal Aksu 

Prof. Dr. Özer Ertuna 

Assoc. Prof. Dr. İsmail Sağlam 

Prof. Dr. Mine Uğurlu 

June 2004

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Prof. Dr. Vedat Akgiray, Assoc. Prof. Dr. İsmail Sağlam and Prof. Dr. Mine Uğurlu for their guidance and insight throughout the research. I also would like to extend my deep gratitude to Prof. Dr. Muzaffer Bodur, Prof. Dr. Özer Ertuna and Prof. Dr. Mine Uğurlu for their valuable lectures in the PhD Program. The suggestions and comments of Assoc. Prof. Dr. Celal Aksu, the technical assistance of Assoc. Prof. Dr. Erdem Başçı and Prof. Dr. Kemal Leblebicioğlu, and the physical and the moral support of Koç Portfolio have also been major contributions worth acknowledging.

Finally, this thesis is dedicated to my wife Gülnihal, for her patience during my PhD study, and to my mother Münevver and my father Faruk, who have played a unique role in the formation of my vision.

VITA

HAKAN AKSOY

EDUCATION

- Department of Management

  BOĞAZİÇİ UNIVERSITY İstanbul, Turkey 2000-2004 *(exp.)*

  Ph.D. in Finance

- Electrical and Electronics Engineering Department

  METU Ankara, Turkey 1996-1997, 1998-2001

  Master of Science in EE Engineering

- Department of Economics

  BİLKENT UNIVERSITY Ankara, Turkey 1995-1997

  Master of Arts in Economics

- Electrical and Electronics Engineering Department

  BİLKENT UNIVERSITY Ankara, Turkey 1991-1995

  Bachelor of Science in EE Engineering

WORK EXPERIENCE

- Koç Asset Management İstanbul, Turkey 2002-*Present*

  Senior Portfolio & Fund Manager

- İktisat Asset Management İstanbul, Turkey 1999-2002

  Fund Manager & Strategist,

- İktisat Securities Inc. İstanbul, Turkey 1998-1999

  Strategist & Economist,

- Bilkent University Ankara, Turkey 1995-1997

  Research and Teaching Assistantship. Department of Economics.

PAPERS

- *"A Model of Boundedly Rational Learning in Dynamic Games", August 1997,* MA Thesis in Economics, Bilkent University Ankara/Turkey.
- *"Modelling the ISE100 Index and the Movements of Four Different Types of Players: Large-cap, Corporate, Mid-cap and Foreign Investors", September 2001,* MS Thesis in Electrical Engineering, METU Ankara/Turkey.
- *"Social Recommendations rather than Social Values" with Erdem Başçı, April 2003,* Modelling and Control of Economic Systems 2002, Proceedings of the IFAC (Int. Federation of Automatic Control) Sym. on Modelling and Control of Economic Systems, Sep. 2001, Elsevier Pergamon Press, Ed. by Reinhard Neck, pp. 425-431.
- *"A New Anomaly Observation in the ISE by Using Classifier Systems" (in Turkish) with İsmail Sağlam, Jan.-Mar. 2001,* Maliye Yazıları, 70, pp 50-63.
- *"Modelling the ISE100 Index by Using Neural Fuzzy Systems" (in English) with Kemal Leblebicioğlu, Apr.-Jun. 2001,* Maliye Yazıları, 71, pp 26-65.
- *"A Discriminant Function Calculation Method of the Turkish Industrial Firms and Testing its Validity" (in English), Oct.-Dec. 2001,* Maliye Yazıları, 73, pp 23-49.

PAPERS IN PROGRESS

- *"Fuzzy Logic Estimation of the ISE100 Index" with Kemal Leblebicioğlu.*
- *"Bankruptcy Prediction for Turkish Industrial Firms by Using Factor and Discriminant Analysis" with Mine Uğurlu.*
- *"Patience Extracts Sugar from a Lemon: Buy and Hold Strategy with A Classifier System in the Istanbul Securities Exchange " with İsmail Sağlam.*
- *"Recommendations rather than Values: Social Influence in Accepting A Commodity as A Medium of Exchange" with Erdem Başçı.*

CONFERENCE PARTICIPATION

- International Federation of Automatic Control (IFAC) Symposium on Modelling and Control of Economic Systems, University of Klagenfurt, Austria, Sep. 6-8, 2001. *"Social Recommendations rather than Social Values".*
- International Conference in Economics VII, METU, Turkey, Sep. 6-9, 2003. *"Social Influence in Accepting A Commodity as A Medium of Exchange".*
- International Conference in Economics VII, METU, Turkey, Sep. 6-9, 2003. *"Patience Extracts Sugar from a Lemon: Buy and Hold Strategy with A Classifier System in the Istanbul Stock Exchange".*
- Traditional Finance Symposium "Financial Integration at International Markets", Institute of Banking and Insurance, Marmara University, Turkey, May 27-28, 2004. *"Modelling Stock Market via Fuzzy Rule Based Systems".*

PERSONAL INFORMATION

Native Turkish, Fluent English, Intermediate German

Computer Usage of C++, Turbo Pascal, Matlab, Gauss, SAS, SPSS, Excel, Powerpoint

BÜVAK Scholarship for Conference Participations Sep.2003, Sep.2001

Ranked 157th in Grad. Exam (LES) among more than 30 thousand cand. Dec. 1998

Advisor in Bilkent University 1995-1997

VAKSA Scholarship 1990-1995

Ranked 18th in ÖYS among more than 1 million candidates Aug. 1990

Success Certificate from Ministry of Education Jun. 1990

TÜBİTAK Mediterranean Region Biology 2nd Price Apr. 1990

TÜBİTAK Mediterranean Region Physics 4th Price Apr. 1989

Milliyet Primary School Competition 3rd Price in Burdur Apr. 1983

ABSTRACT

Modelling Stock Market via Fuzzy Rule Based Systems

by

Hakan Aksoy

A rule based fuzzy logic model is implemented to forecast the monthly return of the Istanbul Securities Exchange 100 (ISE100) Index by combining common stock market data analysis techniques. These are technical analysis, financial analysis and macroeconomic analysis.

Starting with the technical analysis, an index level observation by using classifier systems is used as a long term input for the rule based modelling of the fuzzy logic. The negative correlation of the level of the ISE100 Index and the daily returns is obtained. The observation is statistically tested and found to be significant by using bootstrap method. A basic technical analysis rule using moving average is also modified as a short period input for the model.

In the financial analysis that borrows from the methodology of the Altman's bankruptcy prediction analysis, a model for predicting next period's stock return is developed with the logistic regression by using 81 financial ratios before the factor analysis. It is shown in the ANOVA analysis that the output of the analysis is statistically significant as "good" companies perform better than "bad" companies.

Thus, the weighted average of the logit score of the stocks is used to forecast the next period return of the index in the fuzzy logic model.

After the calculation of the technical and financial inputs for the model, macroeconomic data is gathered in three main groups: real economy, FX market and TL market. The data is ruled and modeled within the period from 1996 to 2002 and optimized with steepest descent learning algorithm. The R-Square of the model is obtained to be better than those of other stock market return estimation models in the literature. The reasons are: a) the ISE is not a developed and an efficient market, thus, predicting future prizes are easier than the other developed markets; b) rule based fuzzy logic modelling with large enough data set improves the explanation of the future movements.

Furthermore, the model is also tested for optimal investment decision in 2003. The algorithm is as follows: If the next month's predicted return of the index is positive, the model suppests to invest all of the existing money in the stock market. Otherwise, it is optimal to invest in repo only. The performance of the investor is compared according to the returns of the ISE100 Index, USD, repo and TL in the same period. Consequently the statistical comparison of the results by using the bootstrap method is promising and the model's suggestion performs better than the return of the repo and the ISE100 Index in 2003.

Keywords: Stock Market, Banktruptcy Prediction, Logistic Regression, Learning and Fuzzy Logic.

ÖZET

Hisse Senedi Piyasasının

Bulanık Kurala Dayalı Sistemler ile Modellenmesi

Hakan Aksoy

İstanbul Menkul Kıymetler Borsası 100 (İMKB100) Endeksi'nin aylık getirilerinin tahmin edilebilmesi için kurallara dayalı bulanık mantık modelinin genel hisse senedi data analizi teknikleri kullanılarak uygulaması yapıldı. Hisse senedi data analizi teknikleri olarak teknik analiz, mali analiz ve makroekonmik analiz kullanıldı.

Endeks seviyesi üzerine sınıflandırma sistemlerinin kullanılması ile hazırlanan bir gözlem uzun vadeli teknik analiz değişkeni olarak bulanık mantıkta kurala dayalı öğrenme modeli için düşünüldü. İMKB100 Endeks seviyesi ile günlük getirilerin arasındaki bir negatif korelasyon bir gözlem olarak ortaya çıktı. Bootstrap metodu ile gözlemin sonuçları istatistiki olarak test edildi ve anlamlı bulundu. Modelin kısa vadeli değişkeni için hareketli ortalama kullanılarak basit bir teknik analiz kuralı eklendi.

Temel analiz tarafında ise Altman'ın iflas tahmini analizindeki metoda benzer şekilde bir sonraki dönem hisse senedi getirisini tahmin için lojistik regresyon analizi yapıldı. Bunun için faktör analizi öncesinde 81 mali rasyo kullanıldı. ANOVA analizi kullanılarak "iyi" şirketlerin "kötü" şirketlerden daha iyi performans sağladığı istatistiki

olarak anlamlı bulundu. Sonuç olarak hisse senetlerinin logit skorlarının ağırlıklı ortalaması bir sonraki dönem endeks tahmini yapan bulanık mantık modeli için temel analiz değişkeni olarak kullanıldı.

Model için hesaplanan teknik ve mali analiz değişkenleri sonrasında, makroekonomik data üç ana grup altında toplandı: reel ekonomi, FX piyasası ve TL piyasası. 1996 ve 2002 yılları arasındaki datanın kuralları çıkarıldı ve model için "steepest descent" öğrenme algoritması ile optimize edilerek kullanıldı. Modelin açıklanabilirlik performansı literatürdeki diğer hisse senedi getiri tahmin modellerine göre başarılı bulundu. Neden olarak: a) İMKB'nin gelişmiş ve etkin bir piyasaya sahip olmaması öne sürüldü, ki bu yüzden gelecekteki fiyatları tahmin etmek diğer gelişmiş piyasalara oranla daha kolay olabileceği düşünüldü; b) kurala dayalı bulanık mantık ile yapılan modellemeler yeteri derecede büyük bir data seti ile yapılması durumunda gelecekteki hareketlerin açıklanabilirliğini ilerletebileceği sonucuna varıldı.

Bunun yanında, model 2003 yılında optimal yatırım kararı için test edildi. Algoritma gereği, bir sonraki ayın endeks tahmini getirisinin pozitif olması durumunda yatırımcı bütün portföyü ile hisse senedi piyasasına yatırım yapmaya karar verdi. Getirinin negatif olması durumunda ise yatırımcı sadece repo'yu tercih etti. Yatırımcının performansı ile aynı dönemdeki İMKB100 Endeksi'nin, ABD Doları'nın, repo'nun ve TL'nin getirisi karşılaştırıldı. Sonuçların bootstrap methodu ile istatistiki olarak karşılaştırılması umut verici bulundu ve modelin tavsiyesinin 2003 yılı içinde İMKB100 Endeksi ve repo getirisinden daha iyi performans gösterdiği gözlendi.

Anahtar Kelimeler: Hisse Senedi Piyasası, İflas Tahmini, Lojistik Regresyon, Öğrenme ve Bulanık Mantık.

TABLE OF CONTENTS

TITLE..... i

APPROVAL.....ii

ACKNOWLEDGEMENT..... iii

VITA.....iv

ABSTRACT.....vii

ÖZET..... ix

TABLE OF CONTENTS.....xi

LIST OF FIGURES..... xiv

LIST OF TABLES.....xv

CHAPTER

1. INTRODUCTION..... 1

1.1 Literature on Predictability of the Stock Market... 1

1.2 System View of the Turkish Economy..... 15

2. AN INDEX LEVEL OBSERVATION BY USING CLASSIFIER SYSTEMS IN THE ISE..... 19

2.1 Observing Stock Market..... 19

2.2 Another Observation in the ISE for the Long Time Horizon.....21

2.3 Discussion and Conclusion 29

3. DISCRIMINATING FUNDAMENTALLY GOOD AND BAD STOCKS IN THE ISE..... 31

3.1 Fundamental Analysis.....31

3.2 Factor Analysis Results.....36

3.3 Discriminant Analysis Results..... 40

3.4 Logistic Regression Results..... 49

3.5 Discussion for the Discriminant Analysis..... 56

3.6 Discussion for the Logistic Regression Analysis... 61

3.7 Getting Information from Financial Data..... 65

4. A FUZZY LOGIC MODELLING OF THE ISE BY USING TURKISH ECONOMY.....68

4.1 Introduction to Rule Based System: Fuzzy Logic.... 68

4.2 Investment Theory and the Stock Market..... 73

4.3 Dependent Variables..... 75

4.3.1 Technical Analysis..... 77

4.3.2 Fundamental Analysis.....80

4.3.3 Macroeconomy..... 81

4.4 Implementation of Fuzzy Logic.....93

4.4.1 Prepositional Logic.....95

4.4.2 Membership Function..... 96

4.4.3. Rule Generation and Fuzzification..... 98

4.4.4 Defuzzification..... 99

4.5 Modelling.....103

4.6 Learning.....111

4.6.1 Steepest Descent Algorithm..... 112

4.6.2 Results of the Optimization..... 114

4.6.3 Testing the Model..... 119

4.7 Performance Expectations..... 121

4.7.1 Comparing the Results with OLS..... 124

4.7.2 Testing the Results with Bootstrap Method 124

5. CONCLUSION AND FUTURE WORK..... 129

REFERENCES..... 134

APPENDIX-A Program Codes of The Classifier System (Prepared with Gauss 3.1.1.)..... 152

APPENDIX-B Financial Ratios.....161

APPENDIX-C Outlier Companies for Discriminant Analysis and Logistic Regression..... 165

APPENDIX-D Program Codes for Developing The Fuzzy Logic Model (Prepared with Matlab 5.3.)..... 166

APPENDIX-E Program Codes for Testing The Fuzzy Logic Model (Prepared with Matlab 5.3.)..... 186

APPENDIX-F Program Codes for Testing the Model's Output With Bootstrap Re-Sampling Method (Prepared with Matlab 5.3.) 188

LIST OF FIGURES

FIGURES

2.1 Prices of the ISE100 Index in USD Terms..... 23

2.2 Returns of the ISE100 Index in USD Terms..... 24

3.1 Discriminated Companies.....48

3.2 Analysis of the ISE100 Index According to Weighted Z Score..... 60

3.3 Analysis of the ISE100 Index Acc. to Logistic Regression Results..65

4.1 Membership Functions.....97

4.2 Defuzzification..... 102

4.3 Fuzzy Logic Model for the ISE100 Index..... 109

4.4 Actual and Estimated Output Before Optimization..... 111

4.5 Convergence of the Steepest Descent Method.....114

4.6 Actual and Estimated Output After Learning..... 115

4.7 Correlation of Actual and Estimated Output..... 116

4.8 Membership Functions of the Fundamental Analysis Variable.....116

4.9 Membership Functions of the Technical Analysis Variable.....117

4.10 Membership Functions of the Macroeconomics Variable.....118

4.11 Membership Functions of the Output 118

4.12 Actual and Estimated Output (Overlapping)..... 119

4.13 Correlation of Actual and Estimated Output (Overlapping)..... 120

4.14 Actual and Estimated Output (Monthly Frequency)..... 120

LIST OF TABLES

TABLE

2.1 Classification of the Data.....25

2.2 20-Working Day Example..... 27

2.3 251-Working Day Example.....28

2.4 753-Working Day Example.....28

3.1 Grouping Similar Ratios..... 38

3.2 Reduction of Unnecessary Data.....39

3.3 Stepwise Selection.....41

3.4 Correct Classification.....44

3.5 Wilks' Lambda.....45

3.6 Function Coefficients..... 47

3.7 Coefficient of the Variables.....51

3.8 Test Results.....54

3.9 Classification..... 55

3.10 Anova Table of the Discriminant Analysis 58

3.11 Correlation of the ISE100 Index and the Weighted Z Score 59

3.12 Anova Table of the Logistic Regression.... 62

3.13 Correlation of the ISE100 Index and the Weighted Score..... 63

4.1 Fuzzification of the Variables.....97

4.2 Rules..... 99

4.3 Variables Used in the Model 105

4.4 Statistical and Membership Range of the Variables.....106

4.5 Rules of Technical Analysis for the Modelling of the ISE100 Index.107

4.6 Rules of Macroeconomy for the Modelling of the ISE100 Index.....108

4.7 Rules of Final Output for the Modelling of the ISE100 Index..... 109

4.8 Statistical Results of the Hypothesis 128

CHAPTER 1. INTRODUCTION

1.1. LITERATURE ON PREDICTIBILITY OF THE STOCK MARKET

One of the earliest and most enduring questions related to financial economics pertains to whether financial asset prices are predictable. Numerous scientists have mobilized their skills towards forecasting stock market prices via various modelling tactics. After Keynes's (1936) famous adage that "most investors' decisions can only be summarized as a result of animal spirits of a spontaneous passion to action rather than inaction, and not as the outcome of a weighted average of benefits multiplied by quantitative probabilities", many researchers have analyzed the efficiency of stock market price movements. However, results of empirical testing of predictability of the stock returns on various stock markets are often inconclusive. The existence and the non-existence of the predictability are studied in the papers of Fama (1965), Fama (1976), Fama (1991), and Granger (1992). It is generally claimed that if returns are not predictable, the market is efficient. Fama (1970) summarizes this judgment as follows: "A market in which prices fully reflect available information is called efficient".

The simplest way to model the stock market may be to forecast future price movements using past price changes. Although predicating the forecasts on past price changes may seem too restrictive, a simplified problem can yield surprisingly rich insights into the behaviour of stock market prices. The history of empirical research starts with works on distributional properties of returns in a market. In this aspect, the random walk and the martingale methods are the most popular and general theories to explain the stock market prices.

The martingale, whose origin dates to the "history of games" by John von Neumann, is the earliest method employed in modelling the stock market. Various equilibrium models have been used in this literature but the most widely used one is the fair game model, which is based on the assumption that the market equilibrium is stated in terms of expected returns. The sequence of excess returns in a fair game has zero expected mean in an efficient market. The fair game model stipulates that the conditions of market equilibrium can be explained in terms of expected returns, though making little mention the details of the stochastic process generating these returns. There are two special cases of the model that provide great help for empirical researchers. These are the sub-martingale model and the random walk model. In sub-martingale model, expected value of the current price one period ahead must be equal to or greater than the current price based on the information set. Martingale is a special case of this model where the equality sign holds (fair game or zero sum game). The model postulates that there can be no trading rule that can beat a buy-and-hold strategy. Various trading systems have been suggested both within and outside the academic world to beat the buy-and-hold strategy. Some technical analysis tools (such as Bollinger bands, moving average) are also used to beat the buy-and-hold strategy in the stock market. Random walk is the most popular and widely accepted model in the academic area. According to this, there is a random walk in the return series when successive one-period returns are independent and identically distributed. This means that forecasts based on today's price could not be improved by also using the information in the previous prices.

By definition, the distribution must be normal distributed; however, they must also be independent of each other, which means zero auto-correlation. It is essential to note that zero auto-correlation does not necessarily mean that the probability distribution is normal (Akgiray, 1989). However if the process has normal distribution,

it can be said that there is independence, which means zero auto-correlation. The returns must also be identically distributed. We then need a stationarity test before starting to test the data for the stock market. However, it is important to note that the stock market data in Turkey (ISE) is not stationary because of the high inflation rate. There are also various definitions of the random walk model in this context. According to Taylor (1988), if the process passes the linearity test, the returns are considered to be a strict white noise, otherwise a white noise. If the data passes the tests for stationarity, normality, independence and linearity, we can infer that there is a random walk and prediction of the process is impossible.

If the return generating process is proved not to exhibit a random walk, then another model must be developed that would fit the data. The models can be linear, non-linear or non-parametric but they should depend all on the past data, which provides some information for the future. In the vast literature dealing with the non random walk theory, Lo and MacKinlay (1988) suggests a simple volatility based specification test, called variance ratio test, in order to refute the random walk hypothesis. The variance ratios as test statistics are intuitively successful but the tests are not clearly established yet in the literature. Besides, in the context of emerging financial markets, the results more strongly favour the hypothesis of predictability because of the vast literature, which proves that emerging financial markets are not efficient. Malkiel (2003) makes a brief criticism of the efficient market hypothesis and states that pricing irregularities and even predictable patterns in stock returns can appear and persist for some periods.

Another reason for stock returns to deviate from random walk model pertains to the so-called calendar anomalies effects. If the prices do not reflect the value of the assets in the stock market, there are some ways to make better profits via trading compared to a buy-and-hold strategy. The anomalies can be defined as the abnormal

returns in securities that cannot be explained by any of the asset pricing theories. Most of the anomalies in the literature can be summarized as intra-day effect, day of the week effect, month of the year effect, size effect, P/E ratio effect, custody effect, holiday effect etc. French and Roll (1986) finds negative as well as positive auto-correlations in daily returns of stocks. For a long-term analysis (annual or more), Schiller (1984), Summers (1986), Fama and French (1988) and Poterba and Summers (1988) find negative auto-correlation between the future and past returns of stocks. In McQueen and Thorley (1991), low (high) annual returns are followed by high (low) returns. Fama and French (1988) finds that returns differ by statistically significant amounts during different periods of time such as day-of-the-week, week-of-the-month and month-of-the-year. The valuation ratios in the stock market can be used to find other miscellaneous observations by comparing different time periods or even the data in different countries. By using 12 different countries' valuation ratios, Campbell and Schiller (1998) shows the relation between the dividend yields, the P/E and the returns. However, Shen (2000) finds no relation between the profitability of the company and the return of its stock but the relation between the returns and the P/E ratios. Basu (1977) finds that P/E ratios, because of the exaggerated investor expectations, might be indicators of future investment performance. As another example of solid fundamental analysis, Fama and French (1995) studies whether the behaviour of stock prices in relation to size and book-to-market-equity reflects the behaviour of earnings. These two variables capture much of the cross section of average stock returns. Fama and French (1996) also explains the strong patterns in returns observed when portfolios are formed on earnings/price, cash flow/price and sales growth. Finally, Fama and French (1998) finds that value stocks have higher returns than growth stocks in markets around the world. For the Turkish version of this price movement analysis, Aydoğan and Güney

(1997) shows the relation between P/E ratios and the returns on the ISE. Muradoğlu and Ünal (1994) studies the distributional property of the returns of the ISE stocks in the January 1988 - December 1991 period and finds that the ISE was not efficient in the weak form. Thus, abnormal returns are possible by using the past data in the ISE. Muradoğlu and Oktay (1993) and Metin, Muradoğlu and Yazıcı (1997) show the weekend effect in the ISE. Çetiner (1993) finds the winner and loser portfolio effect in the ISE. Özmen (1997) searches the calendar anomalies in the ISE. All these studies imply that the ISE is not efficient.

Additionally, technical analysis, in other words charting, has also been part of financial practices for many decades by comparing specific technical indicators and formations such as head-and-shoulders or double bottoms. Burton Malkiel (1996) states in his book, A Random Walk Down Walk Street, "Under scientific scrutiny, chart reading must share a pedestal with alchemy". No one has consistently outperformed the usage of a buy-and-hold strategy. Thus, technical trading will come to the rescue during the course of trading. Lo, Mamaysky and Wang (2000) finds that over the 31-year sample period, several technical indicators provide incremental information and have some predictive value by comparing the unconditional empirical distribution of daily stock returns to the conditional distributions driven by these common technical analysis indicators' results. On the other hand, one of the major criticisms levelled against technical analysis is that it ignores the transaction costs. Lo, Mamaysky and Wang (2001) proposes a dynamic equilibrium model of asset prices and trading volume with heterogeneous agents facing fixed transactions costs. They show that even small fixed costs can give rise to large no-trade regions for each agent's optimal trading policy and a significant illiquidity discount in asset prices. In an intertemporal capital asset pricing model with multiple assets and heterogeneous investors, Lo and Wang (2001) shows

that their portfolio has considerable forecasting power in predicting the future returns of the market portfolio. Guo (2002) states that simple trading strategies based on documented predictability tend to generate mean returns with lower volatility than the buy-hold-strategy. Rapach and Wohar (2002) also finds that certain financial variables display significant in-sample and out-of-sample forecasts with respect to stock returns.

Besides, there are some theories on the stock market price movements. Generally, the theories are conglomerate on the momentum and contrarian movements of the prices. Jegadeesh (1990) finds that stock returns tend to show momentum; the price of the stocks that have increased over the previous few months continue to increase over the next month. In contrast, stocks that have low returns in recent months tend to maintain a poor performance for the subsequent months. Jegadeesh and Titman (1993) also shows that the momentum lasts for more than just one month. Their study indicates that the momentum is stronger for firms that have poor recent performance. The tendency of recent good performance to continue is weaker. Thus, in these momentum studies, long-term losers outperform long-term winners and short-term winners outperform short-term losers. For the contrarian movements, the mean reversion theory is a good example. Mean reversion is the change of the market return in the direction of a reversion level as a reaction to a prior change in the market return. After a positive change in the returns, mean reversion causes a negative subsequent change and after a negative change in the returns, it causes a positive subsequent change. This can happen with different speeds. In a stock market with mean reversion in returns, the players will develop expectations about the speed of the reversion. For a positive change in returns, the reaction will depend on the expectation of the reversion speed. High returns are expected to disappear quickly. Then, the reaction will be to sell in order to realize the high returns. After a positive change, the expectation of a fast

reversion leads to higher selling pressure than the expectation of a slow reversion. The mean reversion expectations can be deduced from the sales. High sales after a positive change in returns indicate a fast expected reversion. Black (1988) shows that misperceptions in the development of mean reversion expectations can cause stock market crashes when players understand their error. The index should be analyzed in both long and short periods in order to discern its degree of effectiveness. Fama (1991) notes that the predictability of the stock returns over time is among the most controversial issues on the stock market efficiency. The controversy has led to various explanations on the possibility and the sources of abnormal profits of contrarian and momentum strategies by analyzing the speed of the mean reversion at the same time.

Contributing to the literature on stock market inefficiency, Hall (2000) argues that the stock market moves in low frequency waves in relation to other aggregate measures, such as GDP. An examination of cash flows earned by corporations suggests that their level is broadly consistent with the principle of rational markets that securities' values reflect the present discounted values of cash earnings. As a result, to estimate the index level, the cash earnings of a country may be worth analyzing. Besides, the increase in GNP, industrial production or capacity utilization should be consistent with the level of the index for a given country. The reason underlying this argument is the risk premium taken by the investors. If the macroeconomic fundamentals are strong in an economy, the risk premium of the investor is low and the prices are high. Rigobon (2001) also uses an identification technique based on the heteroskedasticity of stock market returns to identify the reaction of the monetary policy to the stock market. The results indicate that monetary policy significantly affects the stock market movements. Additionally, Taylor (1998) shows the historical analysis of monetary policy rules in the U.S. and states the interrelations of the macroeconomic

variables. As for similar versions of these studies in Australia, Edey and Romalis (1996) designs an econometric model to quantify the role of monetary policy, while Brischetto and Voss (1999) constructs a structural vector autoregression model.

Furthermore, the studies in stock returns and inflation show that there is a negative relationship in the ex post data. Fama (1981) attempts to explain the relationship between expected inflation and stock returns as an induced consequence of a negative relationship between inflation and real activity. This relation is also a basic consequence of Fisherian quantity theory. Firth (1979) examines the relationship between stock market returns and inflation in the U.K. In addition, Hasbrouck (1984) shows the empirical relationships between stock returns, inflation, and economic activity by using survey evidence.

In the literature dealing with model testing and robustness, Pesaran and Timmermann (1995) examines the robustness of the evidence on predictability of the U.S. stock returns, and argues the issue of whether this predictability could have historically been exploited by investors to earn profits in excess of a buy-and-hold strategy in the market. This study finds that the predictive power of models tends to change with the volatility of returns. Pesaran and Timmermann (2002) proposes a two-stage approach for forecasting of financial return series because of the general assumption of a time-invariant relationship between state variables and returns. If the relationship between stock returns and state variables such as interest rates and dividend yields can change with sudden shifts, this result could lead to serious misallocations. Their findings bring into question the common practice of assuming the existence of a stable prediction model for asset returns. Kandel and Stambaugh (1996) shows the stock return prediction in an asset allocation perspective and argue the importance of investors' conditional Bayesian portfolio decision. Besides, Cremers (2002) analyzes

the stock return prediction in a Bayesian framework and concludes that posterior probabilities are more supportive of stock return predictability than prior probabilities.

As a good example for the modelling and testing the goodness of fit, Keim and Stambaugh (1986) constructs a linear model for the return prediction. To predict the monthly returns of a chosen stock portfolio, they use (a) the yield on under Baa-rated bonds less one-month T-bill yield, (b) the logarithm of the ratio given by the level of the S&P 500 Index deflated by the consumer price index and the average of the year-end real S&P 500 Index over the 45 prior years, and (c) minus the logarithm of share price, averaged equally across the quintile of smallest market value firms on the New York Stock Exchange. Their sample consists of monthly data during the period January 1928 through November 1978. In their studies, the R-square values range from 0.1% to 1.6%, depending on their equations. Campbell (1987) uses (a) the one-month T-Bill rate, (b) the two-month less one-month T-Bill rate, (c) the six-month less one-month T-Bill rate, and (d) one lag of the two-month less one-month T-Bill rate as regressors. He tests his model involving one stock portfolio return series with 11.2% R-square value during the period May 1959 through August 1979. Harvey (1989) predicts the returns of eleven stock portfolios using (a) the excess return on equal-weighted market return, (b) junk bond premium, (c) the yield on the S&P 500 stock index minus yield on a one-month T-Bill, (d) the difference in returns for holding a 90-day bill and a 30-day bill for one month, and (e) a January dummy variable. For the sample period September 1941 through December 1987, he finds in-sample R-square values for the portfolios from 6.7% to 17.9%. Ferson and Harley (1991) reports for the fifteen stock portfolio return prediction models fitted using monthly data for the period 1964 through 1986. In their model, all of the variables in the model of Harvey (1989) are used and one more variable, the one-month nominal T-bill rate, is included. The R-square values range

from 5.8% to 13.7%. Finally, Campbell, Chan and Viceira (2003) develops a multivariate model by using VAR. The model predicts excess stock returns by using (a) ex post real Treasury bill rate, (b) excess log stock return, (c) value weighted return and (d) excess log bond return during the second quarter of 1952 through the fourth quarter of 1999. The Rsquare results of the models for predicting excess stock return in both the quarterly and the annual sample are 8.6% and 5.0%, respectively.

In order to model the density functions on stock returns, Hsieh (1991) claims that the model should have nonlinear dependence and concludes that stock returns are not identically independent distributed. There are also numerous research papers claiming that applying nonlinear models (Chenoweth and Obradovic 1996; Bak, Paczuski and Shubik 1997; Gencay 1998, Leung, Daouk, and Chen 2000) such as neural networks, genetic algorithms and fuzzy logic can provide successful predictions. For neural network application, Kimoto and Asakawa (1990) uses several neural networks to train the relationships between the past values of technical and economy indices for forecasting the returns of the Tokyo Stock Exchange Index. The used variables are (a) an indicator of market momentum, (b) turnover, (c) interest rate, (d) foreign exchange rate, and (e) the Dow Jones Industrial Average. The model sets a system of buy/sell and this system performs much better than a buy-hold strategy. Komo, Chang and Ko (1994) uses a radial basis function network to predict the future index values by using past data of the index, the transaction costs, and bond market values. The radial basis function network has two hidden layers and gradient descent is used for learning.

Modelling macroeconomic theory is done under restrictive information regarding people's knowledge level, wisdom and learning ability. In order to quantify the vague rules, as characteristic of fuzzy systems, an expert, data set and a model are needed. Fuzzy logic can be perceived as the bridge between the knowledge of humans

and their expectations by using the information set with mathematical calculations. Thus, fuzzy rules represent the people's vague knowledge in the system. Transformation methods are applied to train this knowledge. As a result, fuzzy sets, introduced by Zadeh (1965) as a mathematical way to represent vagueness in linguistics, are different from the classical set theory. In a classical nonfuzzy set, an element of the universe either belongs to or does not belong to the set, in other words the membership of an element is crisp. A fuzzy set is a generalization of an ordinary set in that it allows the degree of membership for each element in a unit interval. There exist two major types of fuzzy controllers: Mamdani (1974) and Takagi-Sugeno (1985). Mamdani (1974) uses fuzzy sets and Takagi-Sugeno (1985) uses linear functions. Fiordaliso (1998) finds appropriate rules in the framework of chaotic time series forecasting. He proposes a Takagi-Sugeno fuzzy controller to evaluate the rules in the inference process by using gradient descent algorithm. He compares the learning strategies in terms of interpretability and accuracy. The results of his study point out the advantage of using a nonlinear "combining forecast" method. Lee and Antonsson (2000) also works on the Takagi-Sugeno Fuzzy System with an evolutionary computation approach and find a locally Pareto optimal set of solutions. Their new approach posts good results after using two different data sets. Fuzzy membership functions require adaptation for purposes of fine-tuning in the input data. Thus, Nomura, Hayashi and Wakami (1992) works on this optimization process. Membership functions are optimized by steepest descent algorithm to minimize the error. After this optimization study, there needs to be an efficient rule working mechanism. Sun (1994) proposes a rule combination mechanism to build a simplified version of the original rule base. He introduces a data structure to organize rules so that the rule base can be matched against input signals with logarithmic efficiency. This can be used in various

situations of data compression, such as in image coding or in hierarchical pattern recognition. For the problems of structure identification of a fuzzy model, Sugeno and Kang (1988) suggests an algorithm for identifying a structure by applying an example in their studies. Additionally, Patel (2004) works on the twelve different fuzzy reasoning methods along with numerical comparisons by using transformation functions in fuzzy systems. His method reduces enormously the number of mathematical operations and amount of memory to compute the defuzzified control output.

One of the initial applications of estimation using fuzzy systems is Economakos (1979). This study estimates the demand for electrical power by using linguistic terms of the needed electricity. Then, Tanaka, Uejima and Asai (1982) develops the fuzzy linear regression analysis. Shnaider and Kandel (1989) constructs a model to forecast corporate income tax revenue by using the moving averages of the corporate tax revenue and per capita GNP for Florida. Song and Chissom (1993) analyzes the theoretical part of the fuzzy time series modelling. This study fuzzifies the data, develops the time series model, and calculates the output. Song and Chissom (1994) applies a time invariant forecasting model by using different defuzzifying methods. Meanwhile, Chen (1996) analyzes this model and changes the arithmetic operations in rule generation part. Macroeconomic modelling of fuzzy logic can be seen in recent studies. Kooths (1999) analyzes a business cycle simulation model in neuro fuzzy systems. He builds an experimental macroeconomic model by using money supply, interest rate, exchange rate, currency reserves, net capital imports, price increase, unemployment rate and technical progress. Linn and Tay (2001) analyzes the fuzzy inductive reasoning and nonlinear dependence in security returns by using the outputs of artificial stock market model. The model provides an explanation for nonlinearities observed in the US stock return by observing how individuals actually reason.

Since the rule generation for complex methods is very important and the available rules may not be adequate for modelling, there can be other experimental environments how to learn the mechanism of the process in the model. In order to understand the price movements of the stocks with effect and reaction mechanism, a laboratory environment can be designed as in the models of computational economics. Laboratory asset markets are a useful tool for understanding relative performance because they permit the controlled manipulation of the rules and procedures. Many hypotheses can be investigated under laboratory environment. The Santa Fe stock market experiment investigates market efficiency and price convergence with rational expectation asset pricing model. It is one of the most common complex artificial markets. It is described in detail in Arthur, Holland, LeBaron, Palmer & Tayler (1997) and LeBaron, Arthur & Palmer (1999). This model attempts to integrate the training mechanism into a well-defined economic structure, along with inductive learning using a classifier-based system. One feature of this model is that it allows agents to explore a fairly wide range of possible forecasting rules (Chen and Yeh, 2001). Agents have flexibility in using and changing their trading strategies. A major attribute as a major restriction in the market is the assumption of homogeneity among agents. The agents are quite identical, regardless of their information set, signal processing ability and learning mechanism. Moreover, the model adopts the Walrasian auctioneer mechanism, which provides a market-clearing price after the aggregation of the traders' demand and supply. The auctioneer announces a market-clearing price where aggregate demand equals aggregate supply. Then, traders determine their optimal demand or supply according to this price. There is no trading allowed outside the equilibrium price. The results on this market show that the artificial stock market is able to replicate certain empirical problems observed in real markets. These are predictability, excess volatility,

volatility persistence, and positive excess returns. When the heterogeneity of traders increase, the market is driven by more noisy factors than just fundamental trading. It becomes more difficult for traders to predict each other's forecasts. As in the case of the Turkish stock market, when the number of different types of traders increases, volatility increases and prices do not converge.

Consequently, stock prices have specific properties that make the prediction more complex because of the strong resemblance to random walk processes. A successful prediction algorithm does not have to provide predictions for all points in the time series. The important measures of success are the generated profit, related to which the algorithm produces predictions, and the explanatory power of the model above the standards. In spite of some academic support for efficient market hypothesis, lots of traders in the market all over the world believe and claim that they predict the prices in a way in order to make some profits for themselves and their customers. Almost all professional traders rely to some extent on either technical or macroeconomic and fundamental analyses of past information. They increase their risky position when the market is bullish and decrease their risky position when the market is bearish. Thus, it can be assumed that there is a direct correlation between the current situation of market prices and the future prices. Additionally, different time perspectives with different traders in the market create certain arbitrage opportunities. As an example, a majority stockowner will react quite differently when compared with a day trader when a stock suddenly loses value. As a result, it may be possible to identify these movements and predict future changes. Most arguments against the efficient market hypothesis refer to a time delay between the point when new information enters the system and the point when the information has been evaluated globally and a new equilibrium price has been established. From this perspective, the controversy revolves mainly around how the

word "immediately" should be interpreted in the efficient market hypothesis definition. Being an academician as well as a trader, I believe this is related with the fact that traders generally react faster to developments in the market when compared with academics.

1.2. SYSTEM VIEW OF THE STOCK MARKET AND THE ECONOMY

Based on the conclusion on the previous section, this thesis proposes a model to forecast the future movements of the stock market by using some rules. Thus, the designed model should work much faster than the market wherein the price adjustments occur "immediately" at the equilibrium level. Furthermore, because of the mechanical advantages of the fuzzy logic system, rule based modelling can be an alternative approach to develop a model of the stock market. As result, the thesis tries to model the Istanbul Stock Exchange by using the financial variables in the Turkish economy.

The Istanbul Stock Exchange is the only stock market in Turkey whereby stocks may be bought and sold. Like every investment, investing in stocks entails some degree of risk. There are two types of risk: systematic risk and erroneous risk. The erroneous risk can be overcome by a sound investment strategy, called diversification. However, by using the variables in the macroeconomic system, systematic risk can be minimized if not eliminated. Thus, the stock market can be estimated by using different variables such as fundamental analysis of the stocks and the Turkish economy and technical analysis. While forecasting the stock market that is the most risky market, the systematic risk should be reduced by using the variables and by developing new rules in the system of the economy. In this thesis, classifier systems, discriminant analysis,

logistic regression and fuzzy logic architectures are constructed to forecast the returns of the ISE100 Index by using financial variables. These financial variables are technical analysis of the ISE100 Index, companies' fundamental analysis by using logistic regression, gross national product, industrial production index, capacity utilization index, balance of payments, FX reserves, FX rate, inflation rate, budget deficit, repo rate and bond price index. The next chapters can be summarized as follows:

Chapter 2 introduces a classifier system to detect an observation in the Istanbul Stock Exchange. A classifier system is used to calculate expected return and risk at various levels of the ISE100 Index in the last 15 years. The classifier system groups the daily data of the ISE100 index into classes of fixed size, and computes the average return for each class. For a hold period of one month, the average return at almost all levels of the ISE100 index is found to be statistically zero. With the investment period extending to one year, the results drastically change; the average return becomes positive at low levels of the index while being negative at high levels. Furthermore, for a hold period of three years, average return becomes positive and statistically different from zero at all middle and lower classes of the index. Thus, it is observed that the expected return from and the confidence in the market portfolio are higher at the low levels of the market index, independent of the length of the hold period. The same classifier system also helps to estimate the optimal length of period to hold the market portfolio. Finally, by using "Bootstrap Method" the inverse correlation of the level of the ISE100 Index in USD terms and the daily returns are expressed as an observation in this chapter.

Chapter 3 attempts to identify the prediction of the equity returns by using the data from the Istanbul Stock Exchange in both discriminant analysis and logistic regression. A comparative analysis of discriminant analysis and logistic regression

methods are analyzed together and reveals that both methods have almost similar findings for the data used between 1996 and 2002. In the first step, factor analysis is used to scale down the 81 financial ratios documented by the previous studies. In the second step, discriminant analysis and logistic regression are used to improve the prediction accuracy through a comparative analysis of each method. In both the discriminant analysis and the logistic regression, the independent variable is the next (3 months) period returns of the industrial companies in the ISE. Consequently, cash ratio is the most important factor in the analyses. EBIT and the sales factors of the companies are also comparably important. However, the stock market reacts to the crises quickly whereas the financial data follow with a lag. Because of the delay, price volatility and the market value of the equity are adjusted on time and can be important variables. Besides, in the testing period of the forecastibility power of the both discriminant analysis and the logistic regression analysis, it is seen that the failed companies are more than thirty times riskier than the nonfailed companies. Consequently, these ANOVA results show the predicting power of both analyses used in this section for detecting the failed companies or for estimating the underperformed companies. Since there is a considerably high enough correlation between the weighting scores of both analyses and the ISE100 Index, these scores should be used in the fuzzy logic model as a long-term fundamental analysis indicator.

Chapter 4 presents general concepts about fuzzy logic, membership functions, rule generation and the description of a financial time series modelling for the ISE100 Index, followed by the conclusion and future studies to be undertaken in the future in Chapter 5. The fuzzy logic model is divided into three main parts: the technical analysis, the fundamental analysis and the macroeconomic analysis to forecast the monthly return of the ISE100 Index. For the technical analysis part, the result of the Chapter 2, an index

level observation is used for the rule based modelling of the fuzzy logic model. Besides, a basic technical analysis rule by using moving average is also modified as a short period input for the fuzzy logic model. For the fundamental analysis part, the result of the logistic regression analysis in Chapter 3 is used as a financial data input. Finally, macroeconomic data for Turkey is gathered in three main groups: real economy, FX market and TL market. The data are ruled and modelled by using fuzzy logic. Furthermore, the final model by using technical analysis, stocks' financial analysis and macroeconomic analysis is constructed for the period starting from the beginning of 1996 to the beginning of 2003 and is optimized to perform better with steepest descent learning algorithm. The R-Square of the model, in other words the ratio of the explained sum of squares to the total sum of squares, to estimate the monthly return of the ISE100 Index is calculated as 11.8%. The model is also used for investment in the period between the beginning of 2003 to the beginning of 2004. In this testing period, the suggestion of the model performs statistically better than the return of the ISE100 Index, the return of the USD and the return of the repo market. However, we also state that the testing period of one year is not enough to feel confident on this out of sample results and there is no guarantee for the future performances of the model. There is a risk of underestimation and overestimation in all of the forecasting models. To minimize this risk, forecasts should use best mathematical tools and/or econometric models. Here, in this study, we use fuzzy logic in order to forecast the monthly return of the ISE100 Index. However, we should emphasize that our aim is not to give valuable advices or insights to investors in the stock market. The only aim in this modelling exercise is to test a new methodology in return forecast and to compare its power with those of other mathematical methods.

CHAPTER 2. AN INDEX LEVEL OBSERVATION BY USING CLASSIFIER SYSTEMS IN THE ISE

2.1. OBSERVING STOCK MARKET

Random walk model is the restrictive case of efficient markets hypothesis. In the random walk model, the return of the next period is independent of the previous periods. If the prices do not reflect the value of the assets in the stock market, there are some ways to make better profits with trading than buy-and-hold strategy. What kind of strategies or rules are these?

Consequently, by using some observations in the stock market, an investor can have abnormal returns on securities and these observations cannot be explained by any of the asset pricing theories. Most of the findings in the literature can be summarized as intra-day effect, day of the week effect, month of the year effect, size effect, P/E ratio effect, custody effect, holiday effect etc. Fama (1965) finds the positive correlation between the daily returns of stocks. Similar results are expressed in Fischer (1966) for monthly returns and in Lo and MacKinlay (1988) and Conrad and Kaul (1988) for weekly returns. French and Roll (1986) finds the negative and positive auto-correlations in daily returns of the stocks according to its company valuation.

For long-term analysis (annul or more) Schiller (1984), Summers (1986), Fama and French (1988) and Poterba and Summers (1988) find negative auto-correlation between the future and past returns of the stocks. A comparatively recent research by McQueen and Thorley (1991) shows that low (high) annual returns are followed by high

(low) returns. Fama and French (1989) finds that returns differ by statistically significant amounts during different periods of time such as day-of-the-week, week-of the month and month-of-the-year.

The valuation ratios in the stock market can be used to find other miscellaneous observations by comparing different time periods or even the data in different countries. By using 12 different countries' valuation ratios, Campbell and Schiller (1998) shows the relation between the dividend yields, the P/E and the returns. However, Shen (2000) finds no relation between the profitability of the company and the return to its stock whereas depicting a significant relationship between the returns and the P/E ratios. Basu (1977) finds that P/E ratios, because of the exaggerated investor expectations, might be indicators of future investment performance. Also, as a good fundamental analysis example, Fama and French (1995) studies whether the behaviour of stock prices in relation to size and book-to-market-equity reflects the behaviour of earnings. Size and book-to-market-equity variables capture much of the cross section of average stock returns. Fama and French (1996) also explains the strong patterns in returns observed when portfolios are formed on earnings/price, cash flow/price and sales growth. Finally, Fama and French (1998) finds that value stocks have higher returns than growth stocks in markets around the world. For the Turkish version in the literature, Aydoğan and Güney (1997) shows the relation between P/E ratios and the returns in the ISE. Because of the chronic high inflation in Turkey, prices are unstable and thus P/E ratios can represent the time series data better than the stock prices in TL terms in the long-run. Another popular method is to study the stocks' prices in US terms.

Muradoğlu and Ünal (1994) studies the distributional property of the returns of the ISE stocks in the period of January 1988 and December 1991 and finds that the ISE

was not efficient in the weak form. Thus, abnormal returns can be possible to estimate by using the past data in the ISE.

Muradoğlu and Oktay (1993) and Metin, Muradoğlu and Yazıcı (1997) show the weekend effect in the ISE. Çetiner (1993) finds the winner and loser portfolio effect in the ISE. According to the winner loser effect, if there are two portfolios with different return performances in the recent past, the one with a low return performance will deliver a considerably better return performance in the future than the one with high return performance. Özmen (1997) also searches the calendar anomalies detailed in the ISE.

2.2. ANOTHER OBSERVATION IN THE ISE FOR THE LONG TIME HORIZON

In this section a classifier system is used to calculate expected return and risk at various levels of the ISE100 Index in the last 15 years. It is observed that the expected return from and the confidence in the market portfolio are higher at the low levels of the market index, independent of the length of the hold period. The same classifier system also helps to estimate the optimal length of period to hold the market portfolio. Finally, by using Bootstrap Method, the winner-loser effect is statistically proved in the next section. The buy-and-hold strategy, which is known to be, on average, the optimal investment decision in stock markets with significantly growing annual net returns, may also yield, under the prescriptions of a classifier system, high returns in an extremely volatile and almost horizontal market, namely, the Istanbul Stock Exchange.

The simple classifier system in this study groups 15 years' daily data of the ISE100 index into classes of fixed size, and computes for every observation in each class the return over the successive T days.[1]

Next, the average return in each class is calculated. Confidence intervals for average returns are constructed using bootstrap re-sampling method, since the returns do not have a known distribution. It is observed that for an investment period of at least one year, average classified returns are very high at low levels of the ISE100 index.[2]

The data for this study are taken from the Istanbul Stock Exchange and the Central Bank of Turkey. Since Turkey is one of the emerging markets in the world and has a high inflationary data, the foreign, domestic and even corporate investors prefer to compare their returns in USD terms. Thus, it is almost a must to analyze the data in USD terms for the long time horizon (15-years is a very long investment period for Turkey). The study uses 3513 data observations of the ISE100 index in USD terms, which is also stationary, for the time period between July 1987 and February 2002. Figure 2.1 shows that the time series data fluctuates between 0.20 and 3.57 cents, with the median and the average being 0.88 and 1.03 cents, respectively.

---

[1] A classifier system, first introduced by Holland (1975), is a list of condition-action statements. The condition part looks for certain kind of obligations. When the conditions are satisfied, the action part of the classifier system is activated.

[2] There are a number of recent studies that show the inefficiency of the ISE. Sayın (1993) finds the existence of winner-loser effect in the ISE. Muradoğlu and Ünal (1994) establish that stock returns in the ISE deviate from random walk. Metin, Muradoğlu and Yazıcı (1997) show that the ISE exhibits the well known weekend effect.

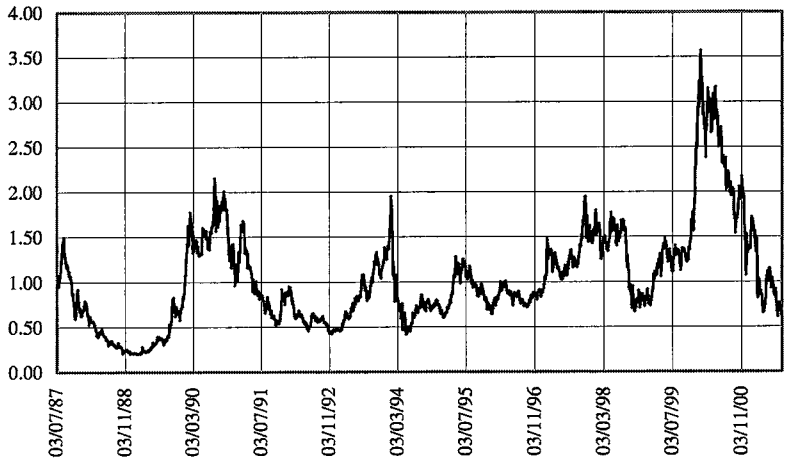


Figure 2.1. Prices of the ISE100 Index in USD Terms

The returns of the ISE100 index are given in Figure 2.2 As can be seen, maximum, minimum, average and median returns are 20%, -24%, 0.05% and -0.06%, respectively. This is to note that 49% of the returns are positive and 51% of the returns are negative. The average of the returns is higher than the median; the prices rise faster when there is an upward momentum and, inversely, the prices hike slower when there is a downward momentum.

As apparent from Figures 2.1 and 2.2, the ISE100 index does not rise in a regular pattern; it just fluctuates in a wide range. Thus, buy-and-hold strategy in the naivest form does not seem to work well in the Istanbul Stock Exchange; and investors generally prefer to use buy-and-sell strategy. Consequently, the other less risky markets, in particular bond market, are preferred more by investors. While in many developed countries the stock market is comparable in size to other security markets, the share of the stock market capacity in the overall security market capacity is 6.8% in Turkey.[3] This comparison alone shows that the stock market in Turkey is considerably underdeveloped. In fact, one can immediately establish by a quadratic programming that

[3] For example, the share of the stock market capacity in the overall security market is as high as 49% in U.S.A.

the optimal portfolio in Turkey consists of repo and bond holdings because of the higher variance and lower mean return of the stock market on the average. Nevertheless, the Istanbul Stock Exchange has also witnessed periods with incomparably high stock returns. Hence, investment in the Turkish stock market prompts a very thorough analysis of the behaviour of past prices and returns. The analysis is also aimed at finding out how the average return of the ISE100 index relates to the length of the investment period and the level of the index. Like the findings in Campbell and Shiller (1998) and Shen (2000), the valuation ratios can be used to classify the time periods with high returns and low risk. In this section, the relation between the returns of the ISE100 Index and the reliance (mean to standard deviation ratio) is investigated.

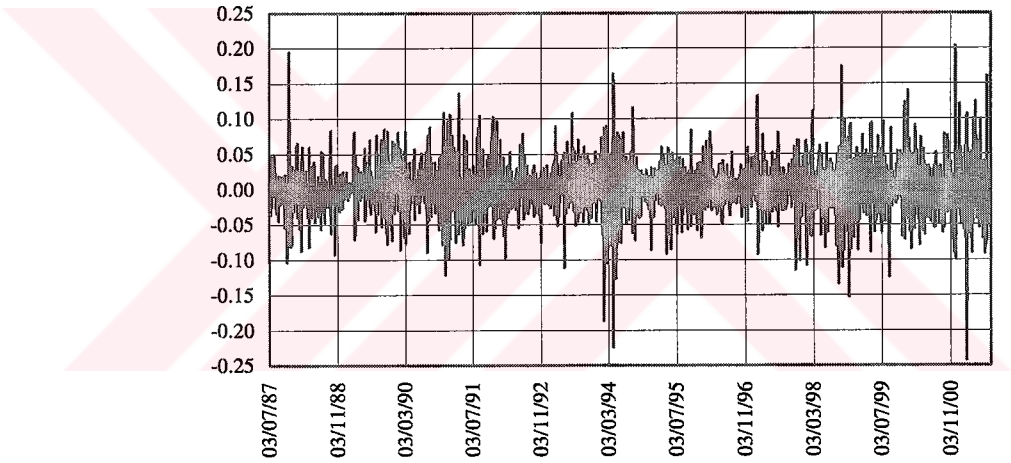


Figure 2.2. Returns of the ISE100 Index in USD Terms

We classify the ISE100 index over the last 175 months into intervals ($i = 1,2,...9$) of size 0.2 cent, and observe the number of data within each class as reported in Table 2.1.

For every observation in each class, we construct a consecutive T-day series (investment period) and compute its return. The average and standard deviation of the returns in class i are then calculated as μ_i^T and σ_i^T, respectively. We aim to test

whether the true mean $\hat{\mu}_i^T$ of returns in class i is zero. The test cannot be immediately performed, since the actual test statistic $Z_{\mu_i^T} = (\mu_i^T - \hat{\mu}_i^T)/\sigma_i^T$ does not have a known distribution with tables. So, we try to approximate the distribution of $Z_{\mu_i^T}$. By bootstrap re-sampling method, we simulate 20,000 values of pivotal test statistic $Z_{\mu_i^{T,B}} = (\mu_i^{T,B} - \hat{\mu}_i^T)/\sigma_i^{T,B}$ where $\mu_i^{T,B}$ and $\sigma_i^{T,B}$ are the average and the standard deviation of bootstrap samples of returns, respectively. By using the definition of quantiles, we then construct $1-\alpha$ confidence interval $C_{i,1-\alpha}^T$ for the pivotal test statistic. The null hypothesis of zero average return in class i is accepted with $1-\alpha$ confidence if the actual test statistic $Z_{\mu_i^T}$ lies within the interval $C_{i,1-\alpha}^T$.

Table 2.1. Classification of the Data

Class No	Index In USD	Observed Data
1	[0.2, 0.4)	317
2	[0.4, 0.6)	346
3	[0.6, 0.8)	734
4	[0.8, 1.0)	581
5	[1.0, 1.2)	389
6	[1.2, 1.4)	425
7	[1.4, 1.6)	258
8	[1.6, 1.8)	181
9	[1.8, +∞)	282
Total	[0.2, +∞)	3513

The classifier system and the bootstrap re-sampling method that we use are formally described below: (All computations are performed with the software Gauss 3.1.1.)

1. For $i = 1,2,...,9$, the data entering class i is recorded. The total number of data in class i is denoted by N_i. The data in class i is indexed by $j = 1,2,...,N_i$ and denoted by $X_{i,j}$.

2. The below steps are repeated for the values 20, 251 and 753 of T.
3. For $i=1,2,...,9$, $j=1,2,...,N_i$ and observation $X_{i,j}$ in class i, the consecutive T-day series $S_{i,j}^T=(X_{i,j},X_{i,j}+1,...,X_{i,j}+T-1)$ is constructed. Thus, we get $3513-(T-1)$ distinct time series over all classes.
4. For $i=1,2,...,9$, and $j=1,2,...,N_i$, the mean return $\mu_{i,j}^T$ of the series $S_{i,j}^T$ is calculated. The set $M_i^T=\{\mu_{i,1}^T,\mu_{i,2}^T,...,\mu_{i,N_i}^T\}$ is recorded.
5. For $i=1,2,...,9$, the average μ_i^T and the standard deviation σ_i^T of the set of returns M_i^T are calculated.
6. For $i=1,2,...,9$, the actual test statistic $Z_{\mu_i^T}=(\mu_i^T-\hat{\mu}_i^T)/\sigma_i^T$ for the postulated true average return $\hat{\mu}_i^T=0$ is calculated.
7. For $i=1,2,...,9$, a total of 20,000 samples denoted by $(B_{i,1}^T,B_{i,2}^T,...,B_{i,20000}^T)$ of size $ceiling[N_i/4]$ (i.e. the highest integer less than or equal to $N_i/4$) are sequentially drawn from the set M_i^T with replacement.
8. For $i=1,2,...,9$, and $k=1,2,...,20000$, the average $\mu_{i,k}^{T,B}$ and the standard deviation $\sigma_{i,k}^{T,B}$ of the returns in the bootstrap sample $B_{i,k}^T$ are calculated.
9. For $i=1,2,...,9$, and $k=1,2,...,20000$, the pivotal statistic $Z_{\mu_{i,k}^{T,B}}=(\mu_{i,k}^{T,B}-\mu_i^T)/\sigma_{i,k}^{T,B}$ is calculated.
10. For $i=1,2,...,9$, $(1-\alpha)$ confidence interval $C_{i,1-\alpha}^T=[z_{i,(\alpha/2)}^T,z_{i,(1-\alpha/2)}^T]$ for $Z_{\mu_{i,k}^{T,B}}$'s is obtained noting that Prob$(z_{i,(\alpha/2)}^T<Z_{\mu_{i,k}^{T,B}}<z_{i,(1-\alpha/2)}^T)=1-\alpha$.
11. For class $i=1,2,...,9$, the hypothesis $H_0:\hat{\mu}_i^T=0$ is accepted with $(1-\alpha)$ confidence if (and only if) $Z_{\mu_i^T}\in\mu_{i,1-\alpha}^T$.

We first run the classifier system setting the investment period length T to 20 (working days). The results are reported in Table 2.2. The average class returns are negatively correlated with the ISE100 index level. We cannot reject the null hypothesis of zero average returns for any classes except for classes 2 and 9. For these two classes, the respective (daily) average returns are 0.0023 and -0.0031, which are statistically different from zero at 98% confidence level. This is to note that the actual test statistics of a particular class can be interpreted as the average return per risk (or simply reliance) of the class. Since the standard deviation (risk) of returns is found to be not varying much over the classes, there is a close relation between the mean returns and the reliance. Therefore, from the viewpoints of both risk neutral (return maximizer) and risk averse (reliance maximizer) investors who regress on the past prices, the optimal (naive) prediction for the buy level of the ISE100 index to hold for 20 working days lies between 0.4 and 0.6 cents (class 2). The daily average return figure of %0.23 for class 2 implies a monthly average return as high as %4.7.

Table 2.2. 20-Working Day Example

i	μ_i^T	$\sigma_{\mu_i^T}$	$Z_{\mu_i^T}$	$C_{i,0.98}^T$	$C_{i,0.95}^T$	$C_{i,0.90}^T$	$C_{i,0.80}^T$	p	H_0
1	0.0011	0.006	0.19	[-0.32,0.28]	[-0.26,0.24]	[-0.22,0.20]	[-0.17,0.16]	0.89	Yes(0.90)
2	0.0023	0.007	0.34	[-0.28,0.27]	[-0.23,0.22]	[-0.20,0.19]	[-0.15,0.15]	1.00	NO
3	0.0011	0.006	0.19	[-0.19,0.19]	[-0.16,0.16]	[-0.13,0.13]	[-0.11,0.10]	0.98	Yes(0.98)
4	-0.0003	0.006	-0.06	[-0.23,0.20]	[-0.19,0.17]	[-0.16,0.15]	[-0.12,0.11]	0.28	Yes(0.80)
5	0.0008	0.006	0.13	[-0.26,0.27]	[-0.22,0.23]	[-0.19,0.19]	[-0.14,0.15]	0.78	Yes(0.80)
6	0.0006	0.005	0.11	[-0.26,0.25]	[-0.22,0.21]	[-0.18,0.18]	[-0.14,0.14]	0.76	Yes(0.80)
7	-0.0012	0.006	-0.20	[-0.32,0.34]	[-0.27,0.28]	[-0.23,0.24]	[-0.17,0.18]	0.17	Yes(0.90)
8	-0.0019	0.008	-0.25	[-0.42,0.39]	[-0.35,0.32]	[-0.29,0.27]	[-0.22,0.21]	0.17	Yes(0.90)
9	-0.0031	0.006	-0.47	[-0.28,0.35]	[-0.24,0.30]	[-0.20,0.24]	[-0.16,0.19]	0.00	NO

Test results for a hold period of one year (251 working days) are given in Table 2.3 below.

Table 2.3. 251-Working Day Example

i	μ_i^T	$\sigma_{\mu_i^T}$	$Z_{\mu_i^T}$	$C_{i,0.98}^T$	$C_{i,0.95}^T$	$C_{i,0.90}^T$	$C_{i,0.80}^T$	p	H_0
1	0.0035	0.0019	1.84	[-0.28,0.32]	[-0.24,0.27]	[-0.20,0.22]	[-0.16,0.17]	1.00	NO
2	0.0013	0.0019	0.68	[-0.27,0.30]	[-0.23,0.25]	[-0.19,0.21]	[-0.15,0.16]	1.00	NO
3	0.0006	0.0015	0.41	[-0.19,0.21]	[-0.16,0.18]	[-0.13,0.15]	[-0.11,0.11]	1.00	NO
4	0.0003	0.0015	0.22	[-0.23,0.23]	[-0.19,0.19]	[-0.16,0.16]	[-0.13,0.12]	0.97	Yes(0.98)
5	-0.0001	0.0017	-0.07	[-0.28,0.28]	[-0.23,0.24]	[-0.19,0.20]	[-0.15,0.16]	0.27	Yes(0.80)
6	0.0001	0.0013	0.04	[-0.25,0.28]	[-0.21,0.23]	[-0.18,0.19]	[-0.14,0.15]	0.63	Yes(0.80)
7	-0.0007	0.0011	-0.61	[-0.32,0.35]	[-0.27,0.29]	[-0.23,0.24]	[-0.18,0.18]	0.00	NO
8	-0.0011	0.0011	-1.07	[-0.41,0.40]	[-0.34,0.33]	[-0.28,0.27]	[-0.21,0.21]	0.00	NO
9	-0.0022	0.0006	-3.43	[-0.30,0.32]	[-0.25,0.27]	[-0.21,0.23]	[-0.17,0.18]	0.00	NO

Table 2.3 shows that the null hypothesis of zero average returns is rejected for all classes, except for the middle ones (4-6). The return is positive at lower classes (1-3), statistically zero at middle classes (4-6), and negative at upper classes (7-9). The optimal buy level of the ISE100 index to hold for 251 working days may be predicted between 0.2 and 0.4 cents (class 1) for both risk neutral and risk averse investors. The daily average return %0.35 for class 1 implies a yearly average return as high as %140.

Finally, test results for a hold period of three years (753 working days) are reported in Table 2.4.

Table 2.4. 753-Working Day Example

i	μ_i^T	$\sigma_{\mu_i^T}$	$Z_{\mu_i^T}$	$C_{i,0.98}^T$	$C_{i,0.95}^T$	$C_{i,0.90}^T$	$C_{i,0.80}^T$	p	H_0
1	0.0011	0.0003	3.88	[-0.32,0.29]	[-0.26,0.24]	[-0.21,0.20]	[-0.17,0.16]	1.00	NO
2	0.0009	0.0002	4.26	[-0.27,0.31]	[-0.23,0.25]	[-0.19,0.21]	[-0.15,0.16]	1.00	NO
3	0.0006	0.0003	2.16	[-0.19,0.22]	[-0.16,0.18]	[-0.13,0.15]	[-0.10,0.12]	1.00	NO
4	0.0004	0.0004	1.04	[-0.24,0.22]	[-0.20,0.18]	[-0.17,0.15]	[-0.13,0.12]	1.00	NO
5	0.0005	0.0004	1.24	[-0.29,0.31]	[-0.24,0.26]	[-0.20,0.22]	[-0.16,0.17]	1.00	NO
6	0.0003	0.0005	0.56	[-0.32,0.30]	[-0.27,0.25]	[-0.23,0.21]	[-0.17,0.17]	1.00	NO
7	-0.0000	0.0004	-0.08	[-0.38,0.35]	[-0.32,0.30]	[-0.26,0.25]	[-0.20,0.20]	0.32	Yes(0.80)
8	-0.0001	0.0003	-0.38	[-0.49,0.46]	[-0.40,0.39]	[-0.33,0.33]	[-0.26,0.25]	0.07	Yes(0.95)
9	-0.0002	0.0003	-0.60	[-0.85,0.61]	[-0.72,0.52]	[-0.61,0.43]	[-0.45,0.34]	0.11	Yes(0.90)

In Table 2.4, the null hypothesis is rejected for all but upper classes (7-9). The return is statistically zero at upper classes, while being positive at middle and lower

classes. The optimal buy level of the ISE100 index to hold for three years may be predicted between 0.2 and 0.4 cents (class 1) for risk neutral investors while being between 0.4 and 0.6 cents (class 2) for risk averse investors. The daily average returns %0.11 and %0.09 for classes 1 and 2 imply three-year returns of %129 and %97, respectively.

These findings indicate the similar results with Çetiner (1993) in the context of winner-loser effect in the ISE.

2.3. DISCUSSION AND CONCLUSION

This chapter introduces a classifier system to detect an observation in the Istanbul Stock Exchange. The classifier system groups the daily data of the ISE100 index into classes of fixed size, and computes the average return for each class. For a hold period of one month, the average return at almost all levels of the ISE100 index is found to be statistically zero. With the investment period extending to one year, the results drastically change; the average return becomes positive at low levels of the index while being negative at high levels. Furthermore, for a hold period of three years, average return becomes positive and statistically different from zero at all middle and lower classes of the index.

It is interesting to observe that the optimal buy level of the ISE100 index for both risk neutral and risk averse investors lies between 0.2 and 0.6 cents, irrespective of the length of the hold period. From a comparison between Tables 2.1, 2.2 and 2.3, it follows that risk neutral investors may optimally choose to invest in the ISE100 index for a year and whenever its level falls between 0.2 and 0.4 cents. On the other hand, risk

averse investors should better invest in the ISE100 index for three years and whenever its level is between 0.4 and 0.6 cents.

In a recent study, Bessembinder and Chan (1998) shows that technical trading rules possess significant return forecast power for the US equity market, which, they argue, is not inconsistent with market efficiency since the transaction costs are high enough to eliminate the difference between returns obtained from technical trading rules and those obtained from buy-and-hold strategy. Interestingly, in our paper, we have established that even the simple buy-and-hold strategy in the Turkish equity market implies abnormal returns in USD terms, as high as %140 for an annual investment, for example. The value of an emerging market may change from time to time because of the movements of hot money, creating inconsistent and huge fluctuations in the market.

Consequently, the forecast power of the classifier system clearly shows that huge returns are available to patient investors. Thus, the results should be used in the fuzzy logic model as a long-term technical analysis indicator. For the short-term, technical analysis indicator simple moving analysis can be used to catch the momentum of the equity market.

CHAPTER 3. DISCRIMINATING FUNDAMENTALLY GOOD AND BAD STOCKS IN THE ISE

3.1. FUNDAMENTAL ANALYSIS

With respect to the rules regarding the technical analysis, the fundamental analysis is a more realistic approach to estimate the index or price levels. To construct some rules from the balance sheets and income statements of the companies, Altman z-score discriminant function can be used as a fundamental modelling. The goal of the discriminant function analysis is to predict group membership from a set of predictors. The major purpose of discriminant analysis is to predict membership in two or more mutually exclusive groups from a set of predictors, when there is no natural ordering on the groups.

Furthermore, a discriminant function of the Turkish companies can be calculated and the average z-score of the companies will be a very good indicator for estimating the index level. This chapter attempts to predict equity returns with a statistical method. This method provides an improved predictive accuracy using data from the Istanbul Stock Exchange. A comparative analysis of discriminant and logit methods reveals that the methods end up in similar findings for the data used between 1996 and 2002.

Altman z score discriminant function for bankruptcy prediction has been one of the most challenging tasks in accounting after Fitzpatrick in 1930s and the early researches (Ramser and Foster 1931;Fitzpatrick 1932; and Winakor and Smith 1935) focused on the comparison of the values of financial ratios in failed and non-failed firms

and concluded that the ratios of the failed firms were poorer. Later, Altman's theoretical and empirical research has been successfully completed (Altman 1984). Two main studies were conducted for the fundamental analysis of healthy and poor companies. The first study was on the empirical search of financial ratios for lowest misclassification rates. The second one dealt with the search for statistical methods for improved prediction accuracy.

In the 1930s, there were no advanced statistical methods or computers available for researchers. Thus, the values of financial ratios in good and bad firms were compared with each other, with bad firms observed to be poorer (e.g. Fitzpatrick, 1932). In 1966, Beaver showed the method of discriminant analysis and in 1968, Altman expanded this study to multivariate analysis. In Altman (1968), multiple discriminant analysis (MDA) is used to classify financial ratios and to develop a bankruptcy prediction model. The discriminant analysis model for bankruptcy prediction results in accurate predictions up to two years prior to actual failure. However, accuracy rapidly diminishes after the second year.

Other researches for bankruptcy prediction by using multiple discriminant analysis will come up after this study (Altman, Haldeman and Narayanan 1977; and Blum 1974). In Deakin (1972), there are some modifications in the methodology of the discriminant analysis by calculating the probability of the bankruptcy. Taffler (1983) calculates performance score for the companies to modify the discriminant analysis. Performance score is used to grade the bankruptcy risk of the companies. After a detailed research on which financial ratios would better serve to differentiate healthy firms from ailing firms, selected market ratios were employed in the academic studies to analyze the financial strength of a company. In Levy (1978), beta plays no role in price determination and variance provides a better explanation for price behaviour in an

imperfect market. Besides, in 1980 Aharony, Jones and Swary found that variance of the mean returns for the failed firm is significantly larger than those for the control group for every week prior to bankruptcy week.

In Joy and Tollefson (1975) and Altman and Eisenbeis (1976), there are different aspects of the discriminant analysis in size and financial profile of business failures and financial reporting techniques. In 1977, Altman, Haldeman and Narayanan developed the prediction model, dubbed the ZETA credit risk model, to be applicable to non-manufacturing firms and private companies whose stocks are not publicly traded. The ZETA model is more valid in bankruptcy classification in five years compared to Altman's original model with the modification of the previous model in emerging market scoring.

However, in the method of discriminant analysis, the assumption of normality of the financial ratio distribution is almost impossible to realize. The discriminant analysis was the sole method used until 1980s. After this period, the logistic regression analysis was developed. In addition to the distributional normality problem in discriminant analysis, there is also the problem of matching the procedures used in this analysis.

Logistic regression allows one to predict a discrete outcome such as group membership from a set of variables that may be continuous or discrete. For both methods of the discriminant analysis and logistic regression, the categories in the outcome must be mutually exclusive. One of the ways to determine whether to use logistic regression or discriminant analysis is to analyze the assumptions. The logistic regression is much more relaxed and flexible in its assumptions than the discriminant analysis. Unlike the discriminant analysis, the logistic regression does not have the requirements of the independent variables to be normally distributed, linearly related, and equal variance within each group. The logistic regression analysis requires larger

sample size. At least 50 cases per independent variable might be required for an accurate hypothesis testing. However, when the assumption of the distribution is met, the discriminant function analysis is more powerful than the logistic regression. As a result, Ohlson (1980) firstly constructs the logit model by analyzing the bankruptcy of the companies. Zavgren (1985) develops the classification success of the logit model.

After the development of logistic regression, a novel method dubbed "artificial neural networks" started to be used in academic studies, with promising results in failure prediction, thanks to the help of computer technology in 1900s. (Back, Laitinen, Sere and Wezel, 1996)

In both the discriminant analysis and the logistic regression, financial ratios have been selected based on their ability to increase the prediction accuracy. There are some efforts to create theoretical constructions in failure prediction concept, but there are no significant developments to be construed as a basis for the theoretical ratio selection. A stepwise selection procedure is used to identify the empirical characteristics of the ratios on statistical grounds.

In the first step, factor analysis is used to search for predictors that make higher classification rates utilizing 81 financial ratios documented by the previous studies and listed in Appendix-B. In the second step, discriminant analysis and logistic regression are used to improve the prediction accuracy through a comparative analysis of each method using data from the Istanbul Stock Exchange.

The data span two important crises. The first one starts after the deep crises of the Asian countries, and then like a domino effect, the Russian Crisis comes into the scene with the impact of weak global oil prices. After the Russian Government's declaration of a moratorium in August 1998, international investors stampeded for the

exit in Turkish financial markets as well. The government implemented a tight monetary policy and the system was stabilized until the next storm broke out.

Unfortunately, another storm erupted in February 2001, with the collapse of the banking system in Turkey, which had a very deep effect on Turkish capital markets. The rapid downturn of the stock market and the stratospheric rise in interest rates with an enormous demand for foreign currencies wreaked havoc on the economic system, culminating in the devaluation of the Turkish lira against major currencies. Thus, the data span a period of deeply effective crises, which had severe ramifications on the corporations, leading several firms to bankruptcy. Fortunately(?), the data serve to pinpoint the critical variables in turbulent environments.

As a result, the quarterly balance sheets and income statements of the industrial companies traded in the four major sectors (textile, non-metallic mineral production, food and beverage, and metal sectors) of the Istanbul Stock Exchange during the 1996-2003 period are used as input data for the analysis.

The rest of the chapter is organized as follows: The next section gives a short description of the technique and the variable selection methods. After calculating discriminant function by using factor and discriminant analysis, the hypothesis is tested to determine the level of success of the discriminant function. To compare the results, similar steps are repeated for the analysis of the logistic regression. Since, the discriminant analysis and the logistic regression are conducted by using the balance sheet and income statement data for the period from 1996 to 2002 (inclusive), the final result depicting the failed and unfailed industrial stocks traded on the ISE, according to both the discriminant analysis and the logistic regression are tested by using the data for the year of 2003. Finally, the empirical results are presented and then the conclusion follows.

3.2. FACTOR ANALYSIS RESULTS

Discriminant analysis tries to derive the linear combination of independent variables that will discriminate best between a priori defined groups, which in our case are failing and non-failing companies. This is achieved by the statistical decision rule of maximizing the between group variance relative to the within-group variance. However, in logistic regression, maximum likelihood function is calculated and maximized by using derivation. For the database, we have 81 ratios to define the companies (see Appendix-B). Since such a broad number of ratios is excessive to calculate a general discriminant function or to run the logistic regression, the factor analysis is firstly used to eliminate the unnecessary ratios. In other words, the size of huge data is reduced without loss of information (or with a negligible loss of information) in factor analysis.

These 81 financial variables are grouped into twenty-two categories after the factor analysis, as seen in the table below, including profitability, liquidity, solvency, degree of economic distress, leverage, efficiency, variability and size. Factor analysis is conducted to identify the variables that seem to be doing the best job in producing the necessary database.

Consequently, 22 variables are identified by stepwise selection and are used in subsequent analyses to reveal those that are significant predictors of failure.

The results indicate that in case of total variance explanation, ratios related to the financial expense item, liquidity item and turnover item seem to play an important role as seen in the following table (definitions and the formula of the financial ratios are given in Appendix-B). This was done to find out if the variables in alternative models

are describing different financial dimensions, so that the selection of one variable into the model is not only a consequence of extremely small differences in the values of test statistics. The models that combine descriptiveness with simplicity are preferred.

Factor analysis rests on the quarterly financial statements of the firms. The number of observations per variable is approximately seventy, which is above the minimum ratio of five observations per variable. The variables are selected so as to have an absolute value of minimum of 0.5 loading. The observations that are +/- 3 standard deviations from the mean are regarded as outliers and are eliminated. The eliminated companies for the analysis are given in Appendix-C. The factors are labelled and the variable that has the highest loading on each factor is selected. The analysis reveals that 22 variables that are selected explain 76.76% of the variance and the addition of one more variable contributes to the explained variance by less than 1.25%.

Ten factors in the analysis have a 57.18% total variance, as presented in the second table. The first table presents the results of the varimax rotated factor patterns and the grouping ratios.

Table 3.1. Grouping Similar Ratios

Rotated Component Matrix[a]

	Component																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
DR	-.906																					
SETA	.906																					
STDTA	-.851																					
STFDTA	-.828																					
FDTA	-.823																					
WCSTD	.721																					
CR	.721																					
WCTA	.711				.527																	
RETA	.689																					
ROA	.689																					
EBTTA	.679																					
QR	.648								.528													
STFDSTD	-.614					-.508																
IETD	-.520																					
CLCA																						
EBITS		.926																				
WCS		.921																				
NIS		.928																				
FES		-.807																				
RDS		-.719																				
EBITDS		.708												.538								
FASE			.992																			
TFASE			.957																			
TDSE			.915																			
LTDSE			.889																			
RSSE			.887																			
LTDTD				.916																		
LTDTA				.903																		
LTFDTD				.886																		
STFEFE				-.555																		
CATA					.838																	
STARTA					.805																	
QATA					.787																	
APTL						.890																
APNPTA						.890																
APCL						.883																
APT						-.533																
TOAS							.974															
TAFAS							.938															
LTD							.886															
SALE							.760															
STFA								.975														
FAT								.958														
SEFA								.956														
CASH									.825													
LAMSTA									.777													
LAMSS									.622	.528												
QAS										.967												
CAS										.951												
SES		-.504								.738												
EBITPC											.742											
NIPC											.702											
PGTL											-.532											
SCA												.838										
TAT												.791										
ART												.513										
SSE																						
MVETL													.913									
MVETD													.865									
EBITDT														.774								
EBITTA														.536								
GPM															.744							
OPS																						
QAI																.807						
INVT																.774						
ROE																	.834					
EBITSE																	.810					
VOL																		.853				
PROFIT																		.538				
WCLTD																			.545			
FDLTD																						
FASELTD																						
EXS																						
SWC																				.717		
SLA																						
FEFL																					-.845	
TIE																						
IDIOOI																						
OIBOIA																						.847
CCL																						

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 15 iterations.

Table 3.2. Reduction of Unnecessary Data

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.700	14.625	14.625	11.700	14.625	14.625	9.487	11.858	11.858
2	6.106	7.632	22.257	6.106	7.632	22.257	4.782	5.977	17.836
3	4.894	6.117	28.374	4.894	6.117	28.374	4.286	5.357	23.193
4	4.473	5.591	33.966	4.473	5.591	33.966	3.806	4.758	27.950
5	4.243	5.304	39.270	4.243	5.304	39.270	3.738	4.673	32.623
6	3.736	4.670	43.939	3.736	4.670	43.939	3.577	4.471	37.093
7	3.312	4.141	48.080	3.312	4.141	48.080	3.472	4.340	41.434
8	2.626	3.283	51.363	2.626	3.283	51.363	3.132	3.916	45.349
9	2.536	3.170	54.533	2.536	3.170	54.533	3.123	3.904	49.253
10	2.116	2.645	57.178	2.116	2.645	57.178	2.929	3.661	52.914
11	2.020	2.525	59.703	2.020	2.525	59.703	2.703	3.379	56.293
12	1.661	2.076	61.779	1.661	2.076	61.779	2.372	2.965	59.258
13	1.594	1.992	63.771	1.594	1.992	63.771	1.963	2.454	61.712
14	1.463	1.829	65.600	1.463	1.829	65.600	1.825	2.281	63.993
15	1.283	1.604	67.204	1.283	1.604	67.204	1.591	1.988	65.981
16	1.239	1.548	68.753	1.239	1.548	68.753	1.531	1.914	67.895
17	1.181	1.477	70.229	1.181	1.477	70.229	1.519	1.898	69.794
18	1.120	1.400	71.629	1.120	1.400	71.629	1.198	1.498	71.291
19	1.050	1.313	72.942	1.050	1.313	72.942	1.189	1.486	72.778
20	1.044	1.305	74.247	1.044	1.305	74.247	1.089	1.361	74.139
21	1.011	1.264	75.510	1.011	1.264	75.510	1.082	1.352	75.491
22	1.001	1.251	76.762	1.001	1.251	76.762	1.017	1.271	76.762
23	.996	1.245	78.007						
24	.982	1.227	79.234						
25	.970	1.212	80.446						
26	.965	1.207	81.653						
27	.921	1.152	82.805						
28	.887	1.109	83.914						
29	.869	1.087	85.000						
30	.830	1.038	86.038						
31	.805	1.006	87.045						
32	.753	.941	87.985						
33	.692	.865	88.851						
34	.655	.818	89.669						
35	.636	.795	90.464						
36	.604	.755	91.219						
37	.587	.733	91.952						
38	.564	.705	92.657						
39	.516	.645	93.302						
40	.492	.615	93.917						
41	.456	.570	94.486						
42	.427	.534	95.020						
43	.422	.527	95.548						
44	.391	.489	96.036						
45	.364	.454	96.491						
46	.341	.426	96.917						
47	.312	.390	97.306						
48	.290	.363	97.669						
49	.267	.334	98.004						
50	.180	.225	98.228						
51	.170	.213	98.441						
52	.142	.178	98.619						
53	.133	.166	98.785						
54	.119	.149	98.935						
55	.100	.125	99.060						
56	9.819E-02	.123	99.182						
57	9.200E-02	.115	99.297						
58	8.022E-02	.100	99.398						
59	6.943E-02	8.679E-02	99.484						
60	6.363E-02	7.953E-02	99.564						
61	5.802E-02	7.253E-02	99.636						
62	4.915E-02	6.144E-02	99.698						
63	4.034E-02	5.043E-02	99.748						
64	3.291E-02	4.113E-02	99.789						
65	3.209E-02	4.012E-02	99.830						
66	2.911E-02	3.638E-02	99.866						
67	2.722E-02	3.402E-02	99.900						
68	2.379E-02	2.974E-02	99.930						
69	2.172E-02	2.715E-02	99.957						
70	1.311E-02	1.639E-02	99.973						
71	9.191E-03	1.149E-02	99.985						
72	5.890E-03	7.363E-03	99.992						
73	3.037E-03	3.797E-03	99.996						
74	2.712E-03	3.389E-03	99.999						
75	5.798E-04	7.248E-04	100.000						
76	4.608E-12	5.760E-12	100.000						
77	2.213E-16	2.767E-16	100.000						
78	1.620E-16	2.026E-16	100.000						
79	-2.13E-16	-2.664E-16	100.000						
80	-1.10E-15	-1.372E-15	100.000						

Extraction Method: Principal Component Analysis.

3.3. DISCRIMINANT ANALYSIS RESULTS

To start the discriminant analysis, the stepwise selection begins with the variables, which are chosen according to their highest factor scores in each of the 22 factors, given in the rotated component matrix in the previous table. Since multicollinearity is a serious problem for the discriminant analysis, some of the unnecessary variables should be eliminated by using factor analysis from the data set with 81 ratios. In each of the factors, only one variable is selected and this variable has the highest factor score in this factor. Finally, before starting with the discriminant analysis, factor analysis has eliminated 59 of the ratios and one factor is represented by only one variable.

Disriminant analysis does very well, provided that the variables in every group follow a multivariate normal distribution and the covariance matrices for each group are equal. However, empirical experiments have shown that especially failing firms violate the normality condition. In addition, the equal variances condition is also violated. Moreover, multicollinearity among independent variables is often a serious problem, especially when stepwise procedures are employed (Hair et al., 1998). However, empirical studies have proved that the problems connected with normality assumptions were not weakening its classification capability, but its prediction ability.

In both the discriminant analysis and the logistic regression, the independent variable should be selected before running the program. The next period returns of the industrial companies in the ISE are calculated. Since the balance sheets and the income statements are announced quarterly, the testing period (next period) for the returns are calculated in the next 3 months to follow the announcements of the companies. However, the pure returns are not enough. In order to compare them, the return of the

ISE100 Index in the same period should be excluded from the pure returns by using the following formula:

$$\hat{\mu}_{Company_i}^{Next_Period} = [(1+\mu_{Company_i}^{Next_Period})/(1+\mu_{ISE100}^{Next_Period})]-1$$

The calculated real returns for each company are used as independent variable for the analysis. If the real returns are positive, the categorical independent variable is recorded 1. Otherwise, it is recorded 0.

The discriminant analysis rests on the data generated from the quarterly financial statements of the firms and is carried out to identify the most important predictors in return prediction using the variables identified by the factor analysis. The stepwise procedure is used and a significance criterion of 5% or less is utilized. The stepwise selection identifies five variables as significant predictors. The following table shows the entering variables step by step into the model.

Table 3.3. Stepwise Selection

Variables in the Analysis

Step		Tolerance	F to Remove	Wilks' Lambda
1	CASH	1.000	11.420	
2	CASH	.995	9.962	.997
	VOL	.995	9.644	.997
3	CASH	.994	9.327	.995
	VOL	.994	10.147	.995
	STFA	.998	7.560	.994
4	CASH	.983	10.627	.994
	VOL	.991	10.921	.994
	STFA	.997	7.805	.993
	MVETL	.984	4.575	.992
5	CASH	.979	9.688	.993
	VOL	.991	10.922	.993
	STFA	.997	7.692	.992
	MVETL	.978	5.289	.991
	TOAS	.989	4.774	.991

As seen in the table, at the final step there are five variables:

A) Cash Ratio = (Cash + Marketable Securities)/ Current Liabilities. Liquidity is seen as the most important financial ratio because of the frequently erupting crises in Turkey. Since the funding of debt increases with soaring interest rates in the crisis environment, the liquidity problem is aggravated and firms write losses.

B) Annual Volatility in Quarterly Period. Before upward or downward movements are noted in any given stock, a rise in volatility is observed in the stock.

C) Sales/ Tangible Fixed Assets. Since the turnover ratio decreases during crisis, firms increase the profitability margin in order to survive. In recession times, prices move upwards and government's inflation target was abandoned. Turnover factor is strongly related to the efficiency of the economy. In the crisis period, firms' goods cannot be marketed and sold. By using the increasing returns to scale assumption in the economy, firms increase the profit margins to fund the fixed costs. Thus, efficiency and turnover of firms decline in crisis periods.

D) Market Value of Equity/ Book Value of Total Liabilities. This variable is a measure of long-term liquidity as well as solvency. It shows by how much the firm's assets can decline in value before liabilities can exceed the assets and the firm becomes insolvent. Consequently, the lower mean of MVETL ratio for the firms reflects unfavourable expectations of the capital market investors and/or the relatively higher amounts of liabilities in the capital structures of these firms.

E) Total Assets/ WPI. This variable is a proxy of size and reflects the total asset values divided by the wholesale price index (1994=100). This adjustment in asset values eliminates the adverse effects of two-digit inflation in Turkey on assets that are recorded at historical costs. Interestingly, the financially distressed firms are significantly larger than the firms in the other group supporting Altman et al. (1977). This variable has been selected as a significant predictor of failure in discriminant and logit analyses but its coefficient is small relative to other predictors in both models.

The Wilks' lambda is the ratio of the within group sum of squares to total sum of squares. At each step, the variable that is included is the one with the smallest Wilks' lambda after the effect of variables already in the discriminant function is removed. Since the Wilks' lambda can be approximated by the F-ratio, the rule is tantamount to entering the variable that has the highest partial F-ratio. Because Wilks' lambda is equal to:

Lambda = Within Group Sum of Squares / Total Sum of Squares.

In the discriminant analysis, at each step, if the variable that contributes least to the discriminatory power of the model measured by Wilks' lambda fails to meet the criterion to stay, than it will be removed. The variable not in the model that contributes the most to the discriminatory power of the model is entered. When all variables in the model meet the criterion to stay and none of the other variables meets the criterion to enter, the stepwise selection process stops. In some contexts, it has been stated that the disadvantage of stepwise procedure is that the economic importance of variables is ignored and the statistical grounds are stressed.

Consequently, the 22 variables for the discriminant analysis were chosen using stepwise selection. Variables were chosen to enter or leave the model using the significance level of an F test from an analysis of covariance, where the already selected variables act as covariates and the variable under consideration is the dependent variable. In our analysis, we selected the significance level for adding or retaining variables in the model to be 0.05.

The discriminant analysis yields a correct classification rate of 54.4%. With 55.7% among 2124 samples, bad performed companies and with 52.8% among 1819 samples, good performed companies are correctly classified. 54.2% of cross validated grouped cases are also correctly classified for validation. The results are given in the following table.

Table 3.4. Correct Classification

Classification Results[b,c]

			Predicted Group Membership		
		PRIORI	0	1	Total
Original	Count	0	1183	941	2124
		1	858	961	1819
	%	0	55.7	44.3	100.0
		1	47.2	52.8	100.0
Cross-validated[a]	Count	0	1179	945	2124
		1	862	957	1819
	%	0	55.5	44.5	100.0
		1	47.4	52.6	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 54.4% of original grouped cases correctly classified.

c. 54.2% of cross-validated grouped cases correctly classified.

In the following table, the minimized value of Wilks' lambda is given with a significant Chi-square value of 37.9 with 5 degrees of freedom.

Table 3.5. Wilks' Lambda

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.990	37.920	5	.000

Wilks' Lambda

Step	Number of Variables	Lambda	df1	df2	df3	Exact F			
						Statistic	df1	df2	Sig.
1	1	.997	1	1	3772	11.420	1	3772.000	7.341E-04
2	2	.994	2	1	3772	10.545	2	3771.000	2.782E-05
3	3	.992	3	1	3772	9.562	3	3770.000	2.866E-06
4	4	.991	4	1	3772	8.322	4	3769.000	1.177E-06
5	5	.990	5	1	3772	7.619	5	3768.000	4.203E-07

Variables Entered/Removed[a,b,c,d]

Step	Entered	Wilks' Lambda							
						Exact F			
		Statistic	df1	df2	df3	Statistic	df1	df2	Sig.
1	CASH	.997	1	1	3772.000	11.420	1	3772.000	.001
2	VOL	.994	2	1	3772.000	10.545	2	3771.000	.000
3	STFA	.992	3	1	3772.000	9.562	3	3770.000	.000
4	MVETL	.991	4	1	3772.000	8.322	4	3769.000	.000
5	TOAS	.990	5	1	3772.000	7.619	5	3768.000	.000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 44.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

As seen in the above tables, the Wilks' Lambda is not too much less than 1 (which is not expected). However, the F statistics of the Wilks' Lambdas for each variable are significant because of the large number of data set. The reason for high

Wilks' Lambda can be explained by the high volatility in the stock market. This is also seen easily in the figure of the discriminating companies and the standard deviations (Figure 3.1). Because of the high volatility, there may be some distortions but the mean difference in the discriminating companies shows that the discriminant function is working. Thus, the result is also proven by the significant F statistics.

Box's test of equality of covariance matrices has lower significance than the critical level of 0.01(p=0.00), so the assumption of identical variance-covariance across groups is violated. When the two group sizes are equal, the violation of this assumption has minimal impact. However, the violation of normality assumption may weaken the prediction power of the model.

In the following table, the structured matrix of the variables are given and it is understood that only four of the variables are selected to be used in the discriminant function by using the within-groups correlation and by minimizing the Wilks' lambda.

Thus, the next table gives the variables used in the discriminant function with the coefficients (after removing the unnecessary variables and including the variables for the maximized Wilks' lambda). In the standardized canonical discriminant function, all of the coefficients are significantly different than zero. Thus, as seen in the following table, the coefficients of the canonical discriminant functions are significant. The coefficient of variable TOAS is seen as zero because the variable of the data is significantly large, but it does not affect the analysis negatively. Consequently, the final discriminant function is:

$$Z=-1.668+(-0.0000059TOAS)+(0.019STFA)+(0.597MVETL)$$
$$+(2.038\ VOL)+(-0.956CASH)$$

as seen in the table.

Table 3.6. Function Coefficients

Standardized Canonical Discriminant Function Coefficients

	Function
	1
TOAS	-.358
STFA	.452
MVETL	.378
VOL	.540
CASH	-.511

Canonical Discriminant Function Coefficients

	Function
Variables	1
TOAS	-.0000059
STFA	.0187812
MVETL	.5971799
VOL	2.0379370
CASH	-.9561915
(Constant)	-1.6681085

Unstandardized coefficients

Finally, the comparison of the means for the failed and nonfailed companies can be analyzed to validate the results of discriminant function's accuracy. As seen in the graphics below, the value discriminant function is higher for the unfailed companies (Priori=0) and lower for the failed companies (Priori=1). With a total of 2124 data, the unfailed companies have an average value of 0.1 with 1.12 standard deviation. Besides, with a total of 1819 data, the failed companies have an average value of –0.1 with 0.95 standard deviation.

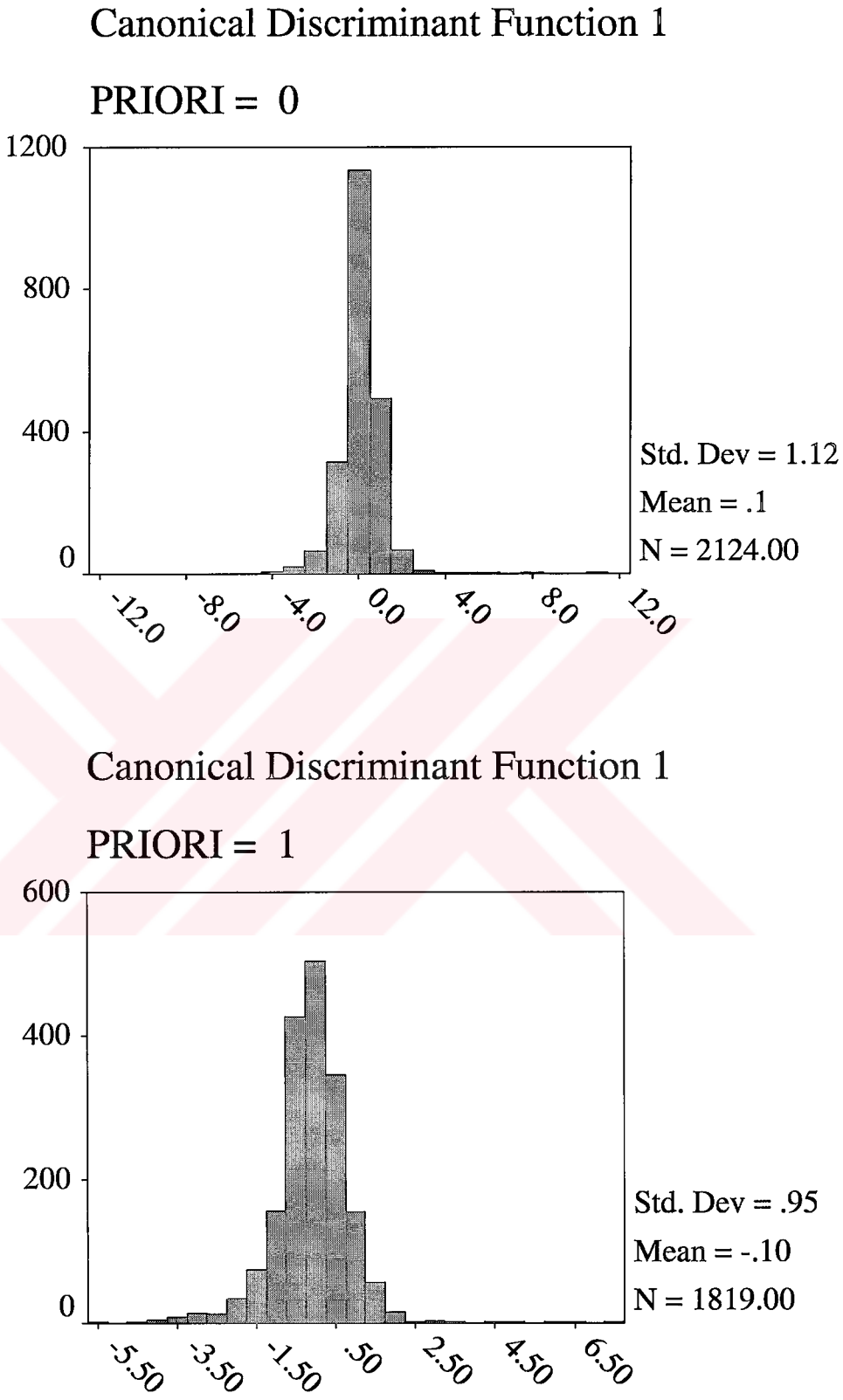


Figure 3.1. Discriminated Companies

Consequently, with 0.1 and –0.1 means of discriminant functions around 1 standard deviation for the failed and the nonfailed companies, respectively, the model is criticized to be not adequately descriptive. In this kind of model, we should not expect a high significance level because of the noisy financial environments especially in emerging markets, which entail frequent episodes of crises. In the next stage, a logistic regression is used to develop another model for the return prediction in the ISE.

3.4. LOGISTIC REGRESSION RESULTS

In logistic regression, there is no true R^2 value as there is in OLS regression. However, because deviance is analogous to "Mean Square Error" in regression analysis, one can approximate an R^2 based on lack of fit indicated by the deviance (-2LogLikelihood).

$$R^2_{\text{logistic}} = 1 - \frac{-2LL_{reduced} - 2LL_{full}}{-2LL_{reduced}} = 1 - \frac{G}{-2LL_{reduced}}$$

$$R^2_{OLS} = 1 - \frac{SS_{total} - SS_{residual}}{SS_{total}} = 1 - \frac{SS_{regression}}{SS_{total}}$$

While the reduced model is the logistic model with just the constant, the full model contains all the predictors in the model. In the other classical models, goodness of fit is calculated by using addition and subtraction. However, in logistic regression, because of the loglikelihood function, which is exponential, goodness of fit is calculated by multiplication and division.

The most common assessment of overall model fit in logistic regression is the goodness of fit test, which is simply the chi-square difference between the null model (the model with the constant only) and the model containing one or more predictors. This is one use of the likelihood ratio test between two nested models. It is an assessment of the improvement of fit between the predicted and observed values on Y by adding the predictor or predictors.

In logistic regression, the variables are identified through stepwise selection procedure as in discriminant analysis case. There is no multicollinearity among the independent variables since they are selected using factor analysis. The logistic regression model reveals four variables as predictors of the next period real returns. Tests of statistical significance and overall model fit for the logistic regression model, as summarized in the following tables, reveal that it is a better method to be used when normality assumption is violated. The correct classification rate of this model is 54.4%, which is equal to the correct classification power of the discriminant analysis.

Table 3.7. Coefficient of the Variables

Iteration History[a,b,c,d,e]

Iteration		-2 Log likelihood	Coefficients																					
			Constant	TOAS	VOL	EBITDT	GPM	CASH	MVETL	FEFL	STFA	SETA	LTDTD	APNPTA	WCLTD	CATA	EBITS	FASE	ROE	EBITPC	QAS	OIBOIA	SCA	SWC	QAI
Step	1	5176.503	.318	.000	-.341	.213	-.120	.178	-.114	.000	-.001	-.093	-.208	.119	-.004	-.210	.078	-.004	.013	.028	-.005	.000	.033	.001	.000
1	2	5175.548	.348	.000	-.356	.234	-.145	.180	-.115	.000	-.001	-.112	-.212	.138	-.004	-.234	.153	-.004	.014	.026	-.007	.000	.032	.002	.000
	3	5175.346	.353	.000	-.355	.220	-.163	.180	-.114	.000	-.001	-.122	-.211	.149	-.004	-.244	.216	-.004	.014	.025	-.009	.000	.034	.002	.000
Step	1	5176.522	.318	.000	-.341	.214	-.120	.178	-.114	.000	-.001	-.094	-.208	.120	-.004	-.211	.078	-.004	.013	.028	-.005		.033	.001	.000
2	2	5175.566	.349	.000	-.357	.235	-.146	.180	-.115	.000	-.001	-.112	-.212	.139	-.004	-.235	.153	-.004	.014	.026	-.007		.032	.002	.000
	3	5175.363	.354	.000	-.355	.220	-.163	.180	-.114	.000	-.001	-.123	-.212	.150	-.004	-.245	.217	-.004	.013	.025	-.009		.034	.002	.000
Step	1	5176.572	.319	.000	-.341	.214	-.120	.178	-.113		-.001	-.094	-.209	.119	-.004	-.210	.078	-.004	.013	.028	-.005		.033	.001	.000
3	2	5175.615	.350	.000	-.357	.235	-.145	.180	-.115		-.001	-.112	-.213	.137	-.004	-.234	.153	-.004	.014	.026	-.007		.032	.002	.000
	3	5175.413	.354	.000	-.355	.220	-.164	.180	-.114		-.001	-.123	-.212	.149	-.004	-.244	.217	-.004	.014	.025	-.009		.034	.002	.000
Step	1	5176.634	.343	.000	-.343	.216	-.126	.176	-.113		-.001	-.089	-.205	.146	-.004	-.223	.077	-.004	.014	.028	-.005			.001	.000
4	2	5175.686	.373	.000	-.360	.239	-.150	.179	-.115		-.001	-.107	-.209	.163	-.004	-.246	.151	-.004	.014	.027	-.007			.002	.000
	3	5175.491	.380	.000	-.358	.225	-.169	.178	-.113		-.001	-.117	-.209	.176	-.004	-.257	.213	-.004	.014	.026	-.009			.002	.000
Step	1	5176.791	.375	.000	-.344	.214	-.133	.177	-.112		-.001	-.112	-.230		-.004	-.216	.076	-.004	.014	.028	-.005			.002	.000
5	2	5175.881	.409	.000	-.360	.236	-.158	.179	-.113		-.001	-.133	-.237		-.004	-.237	.148	-.004	.014	.027	-.008			.002	.000
	3	5175.701	.418	.000	-.358	.223	-.176	.178	-.112		-.001	-.144	-.239		-.004	-.247	.207	-.004	.014	.026	-.009			.002	.000
Step	1	5177.030	.372	.000	-.344	.215	-.129	.176	-.112		-.001	-.110	-.228		-.004	-.214	.077	-.004		.028	-.005			.002	.000
6	2	5176.109	.406	.000	-.360	.237	-.154	.179	-.113		-.001	-.130	-.235		-.004	-.235	.150	-.004		.027	-.008			.002	.000
	3	5175.925	.415	.000	-.359	.224	-.173	.178	-.112		-.001	-.142	-.236		-.004	-.245	.209	-.004		.026	-.009			.002	.000
Step	1	5177.363	.375	.000	-.343	.214	-.124	.181	-.112			-.112	-.229		-.004	-.226	.079	-.004		.028	-.005			.002	.000
7	2	5176.436	.409	.000	-.359	.235	-.149	.184	-.113			-.133	-.236		-.004	-.248	.152	-.004		.027	-.008			.002	.000
	3	5176.250	.418	.000	-.357	.221	-.167	.183	-.112			-.145	-.238		-.004	-.258	.212	-.004		.026	-.009			.002	.000
Step	1	5177.808	.363	.000	-.345	.213	-.122	.182	-.111			-.107	-.229		-.004	-.216	.079			.028	-.005			.002	.000
8	2	5176.693	.397	.000	-.361	.234	-.147	.184	-.113			-.128	-.237		-.004	-.237	.152			.026	-.008			.002	.000
	3	5176.712	.406	.000	-.359	.221	-.165	.184	-.111			-.139	-.238		-.004	-.247	.211			.025	-.009			.002	.000
Step	1	5178.253	.293	.000	-.332	.196	-.136	.166	-.114				-.216		-.004	-.179	.076			.029	-.005			.001	.000
9	2	5177.473	.312	.000	-.345	.208	-.161	.166	-.116				-.221		-.004	-.192	.140			.028	-.007			.002	.000
	3	5177.339	.313	.000	-.342	.196	-.178	.164	-.115				-.221		-.004	-.197	.189			.027	-.008			.002	.000
Step	1	5178.947	.304	.000	-.337		-.130	.169	-.114				-.217		-.004	-.170	.078			.030	-.005			.001	.000
10	2	5178.072	.323	.000	-.350		-.156	.169	-.116				-.222		-.004	-.183	.149			.029	-.007			.002	.000
	3	5177.909	.324	.000	-.347		-.175	.167	-.115				-.222		-.004	-.189	.203			.028	-.009			.002	.000
Step	1	5179.295	.302	.000	-.340		-.147	.168	-.114				-.210		-.004	-.170	.080			.030				.001	.000
11	2	5178.653	.322	.000	-.356		-.173	.168	-.117				-.213		-.004	-.182	.136			.029				.002	.000
	3	5178.585	.322	.000	-.355		-.186	.167	-.116				-.212		-.004	-.185	.167			.029				.002	.000
Step	1	5180.103	.315		-.341		-.160	.171	-.111				-.186		-.004	-.175	.081			.030				.002	.000
12	2	5179.432	.334		-.356		-.187	.171	-.113				-.190		-.004	-.187	.138			.030				.002	.000
	3	5179.358	.334		-.355		-.201	.170	-.113				-.189		-.004	-.191	.171			.029				.002	.000
Step	1	5180.855	.241		-.341		-.162	.158	-.111						-.004	-.119	.078			.030				.001	.000
13	2	5180.181	.258		-.356		-.189	.158	-.113						-.004	-.129	.136			.029				.002	.000
	3	5180.104	.259		-.355		-.203	.156	-.113						-.004	-.134	.170			.029				.002	.000
Step	1	5181.292	.173		-.339		-.180	.164	-.109						-.004		.075			.026				.001	.000
14	2	5180.673	.185		-.354		-.207	.165	-.111						-.004		.129			.027				.002	.000
	3	5180.608	.183		-.353		-.221	.163	-.111						-.004		.160			.027				.002	.000
Step	1	5182.544	.120		-.335			.158	-.110						-.004		.072			.027				.002	.000
15	2	5182.154	.126		-.350			.159	-.112						-.004		.106			.027				.002	.000
Step	1	5184.235	.135		-.342			.159	-.111						-.004					.029				.002	.000
16	2	5184.058	.149		-.359			.161	-.114						-.004					.029				.002	.000
Step	1	5185.867	.137		-.344			.154	-.112						-.004					.029				.002	
17	2	5185.773	.150		-.361			.157	-.116						-.004					.029				.002	
Step	1	5187.994	.138		-.344			.155	-.112						-.004					.029					
18	2	5187.970	.153		-.361			.157	-.116						-.004					.029					
Step	1	5190.205	.132		-.341			.118	-.110											.029					
19	2	5190.183	.145		-.357			.121	-.113											.030					

a. Method: Backward Stepwise (Likelihood Ratio)

b. Constant is included in the model.

c. Initial -2 Log Likelihood: 5208.635

d. Estimation terminated at iteration number 3 because log-likelihood decreased by less than .010 percent.

e. Estimation terminated at iteration number 2 because log-likelihood decreased by less than .010 percent.

Consequently, logistic regression model that does not have the restrictions of the discriminant method seems more appropriate as a method of failure prediction in this study. As seen above the table, at the final step there are four variables. The first three variables are as same as the variables in the discriminant case. The forth variable in the logistic regression is EBIT/Paid in Capital. The other two variables in the discriminant analysis are, in order to their importance, Sales/Tangible Fixed Assets and Total Asset/WPI. The last variable of the logistic regression is similar to the fourth variable in the discriminant analysis Since the profitability and the sales are directly correlated to each other. Finally, four variables in the logistic regression are:

A) Annual Volatility in Quarterly Period. Before upward or downward price changes in the stock market, fluctuations are noted. In other words, volatility increases due to fact that stock market direction is unknown.

B) Cash Ratio = (Cash + Marketable Securities)/ Current Liabilities. Liquidity is one of the most important financial ratios because of the frequently occurring current account crises in Turkey. When the funding of debt increases with surging interest rates in an illiquid environment, firms write losses.

C) Market Value of Equity/ Book Value of Total Liabilities. A measure of long-term liquidity, in other words solvency, shows by how much the firm's assets can decline in value before the liabilities can exceed the assets and the firm becomes insolvent. The lower mean of MVETL ratio for the firms reflects unfavourable expectations of the capital market investors and/or the relatively higher amounts of liabilities in the capital structures of these firms.

D) EBIT/ Paid in Capital. Profit before interest payments and taxes over paid in capital gives generally the financial return of the equity for the given period. While the dependent variable in the logistic regression is the expected returns of the given equity with respect to the index, this variable gives the real life return of the company in the same period.

In some cases, the traditional goodness-of-fit test (G or the likelihood ratio test) may not be the best assessment of model fit. Some simulations suggest that the deviance statistic is not distributed as chi-square when the data are *sparse*. The term "sparse" refers to a circumstance in which there are few observed values (and therefore few expected values) in the cells formed by crossing all of the values of all of the predictors. The Hosmer-Lemeshow fit test is designed to correct for this (when there are continuous predictors). Note that they do not recommend the use of this test when there is a small n (less than 400; Hosmer & Lemeshow, 2000). The Hosmer-Lemeshow test is performed by dividing the predicted probabilities into deciles (10 groups based on percentile ranks) and then computing a Pearson chi-square that compares the predicted to the observed frequencies (in a 2 X 10 table). Lower values (and nonsignificance) indicate a good fit to the data and, therefore, good overall model fit. As seen in Table 3.8, the result of the test is insignificant with 4.07 chi-square value at the final step.

Table 3.8. Test Results

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	10.824	8	.212
2	12.802	8	.119
3	11.663	8	.167
4	10.932	8	.206
5	13.011	8	.111
6	21.764	8	.005
7	17.862	8	.022
8	15.030	8	.059
9	16.688	8	.034
10	9.268	8	.320
11	12.115	8	.146
12	13.993	8	.082
13	8.174	8	.417
14	18.669	8	.017
15	4.487	8	.811
16	9.708	8	.286
17	11.155	8	.193
18	10.667	8	.221
19	4.070	8	.851

The logistic regression analysis at the final step yields a correct classification rate of 54.4%, which is same as the result in discriminant analysis. The results are given in the following table.

Table 3.9. Classification

Classification Table[a]

			Predicted		
			PRIORI		Percentage
	Observed		0	1	Correct
Step 1	PRIORI	0	1772	263	87.1
		1	1457	282	16.2
	Overall Percentage				54.4
Step 2	PRIORI	0	1773	262	87.1
		1	1456	283	16.3
	Overall Percentage				54.5
Step 3	PRIORI	0	1774	261	87.2
		1	1456	283	16.3
	Overall Percentage				54.5
Step 4	PRIORI	0	1781	254	87.5
		1	1464	275	15.8
	Overall Percentage				54.5
Step 5	PRIORI	0	1786	249	87.8
		1	1471	268	15.4
	Overall Percentage				54.4
Step 6	PRIORI	0	1789	246	87.9
		1	1472	267	15.4
	Overall Percentage				54.5
Step 7	PRIORI	0	1790	245	88.0
		1	1474	265	15.2
	Overall Percentage				54.5
Step 8	PRIORI	0	1795	240	88.2
		1	1474	265	15.2
	Overall Percentage				54.6
Step 9	PRIORI	0	1792	243	88.1
		1	1466	273	15.7
	Overall Percentage				54.7
Step 10	PRIORI	0	1795	240	88.2
		1	1461	278	16.0
	Overall Percentage				54.9
Step 11	PRIORI	0	1800	235	88.5
		1	1467	272	15.6
	Overall Percentage				54.9
Step 12	PRIORI	0	1812	223	89.0
		1	1489	250	14.4
	Overall Percentage				54.6
Step 13	PRIORI	0	1821	214	89.5
		1	1503	236	13.6
	Overall Percentage				54.5
Step 14	PRIORI	0	1824	211	89.6
		1	1514	225	12.9
	Overall Percentage				54.3
Step 15	PRIORI	0	1837	198	90.3
		1	1531	208	12.0
	Overall Percentage				54.2
Step 16	PRIORI	0	1842	193	90.5
		1	1533	206	11.8
	Overall Percentage				54.3
Step 17	PRIORI	0	1844	191	90.6
		1	1540	199	11.4
	Overall Percentage				54.1
Step 18	PRIORI	0	1849	186	90.9
		1	1549	190	10.9
	Overall Percentage				54.0
Step 19	PRIORI	0	1879	156	92.3
		1	1565	174	10.0
	Overall Percentage				54.4

a. The cut value is .500

3.5. DISCUSSION FOR THE DISCRIMINANT ANALYSIS

In the results of the discriminant analysis section, we calculated the discriminant function as the following,

$$Z=-1.668+(-0.0000059TOAS)+(0.019STFA)+(0.597MVETL)+(2.038\ VOL)+(-0.956CASH).$$

By using the discriminant function above, the z-values can be calculated by using the balance sheets and the income statements data in the last quarter of 2002 and the first three quarters of 2003 of approximately 192 (number of stocks is changed quarter by quarter) industry firms traded on the ISE.

If the z-value of the company is lower for the given periods, the expected return for that company will be higher, too. To test this hypothesis, the stock returns of the four quarters, which are not used in the analyses, for approximately 192 industrial companies of the ISE are calculated and after dividing to the return of the ISE, we will get the real returns, as seen below.

$$\hat{\mu}_{Company_i}^{ith_Period} = [(1+\mu_{Company_i}^{ith_Period})/(1+\mu_{ISE100}^{ith_Period})]-1$$

When we analyze the total data set for the real returns, we see that the average real return of these companies for the last quarter of 2002 and the first three quarters of 2003 is –4.05% with 28.06% standard deviation. Thus, the ratio of the mean to standard deviation is –14.43%.

Finally, for each z value there is a real return. We can divide the total data set into two pieces according to the average of the z values for the total data set (the average of the z values of the total data set is 0.11). As a result, according to our data set for the industrial firms, if the z-value is greater than 0.11, we will have 334 companies. Similarly, if the z-value is smaller than 0.11, we will have the other 419 companies.

Consequently, the mean of the nonfailed companies is 0.10% with 33.77% standard deviation. However, the mean of the failed companies is –7.36% with 21.99% standard deviation. The standard deviations of the failed and nonfailed companies are slightly different from each other. The standard deviation of the nonfailed companies is higher as expected. Besides, the means are significantly different between the nonfailed and the failed companies. The difference is also seen from the ratio of the mean to the standard deviation. It is 0.30% for the nonfailed companies, while being -33.47% for failed companies. As a result, the risk in the nonfailed companies is incomparably higher than the risk in the failed companies.

Similarly, we can also do an Anova test in these two groups with a cut-off z-value of 0.11. As seen in the following table, the discriminant function is good enough to discriminate the good return expected companies to the bad return expected companies with 13.4 F value and 0.03% significant level. It is also seen from the Anova table that, these two groups are different as apparent from the ratio of the mean to standard deviation.

Table 3.10. Anova Table of the Discriminant Analysis

TOTAL DATA	
T Mean	-4.05%
Sample	753
Stdev	28.06%
Mean/Stdev	-14.43%

GROUP1				GROUP2			
Mean	Sample	Stdev	Mean/ Stdev	Mean	Sample	Stdev	Mean/ Stdev
0.10%	334	33.77%	0.30%	-7.36%	419	21.99%	-33.47%
Z Scores greater than			***0.11***	***Z Scores smaller than***			***0.11***

	TSS	df	MS	F-value	p-Value
Between Group	1.035	1	1.035	**13.361**	**0.03%**
Within Group	58.177	751	0.077		
Total	59.212	752			

Consequently, for today's estimations and future's forecasts, the discriminant function calculated in this paper can be very useful and effective in real life as seen from the empirical results. By using the book value of these stocks, weighted average of the z score can be calculated. As shown in the table above, if there is a correlation between the Istanbul Stock Exchange Index and this weighted z score, it can be concluded that the companies' balance sheet and income statement data can be really effective for the next period returns of the stock market index. As shown in Table 3.11, the correlation between weighted z score and the ISE100 Index is calculated as –28.6%, which attracts attention in stock market standards.

Table 3.11. Correlation of the ISE100 Index and the Weighted Z Score

Year	Quarter	Index	Weighted Z Score	ISE100 Return(t+1)
1996	1	1	0.549	5.14%
1996	2	2	0.097	2.78%
1996	3	3	-0.178	34.69%
1996	4	4	0.064	65.29%
1997	1	5	0.555	15.13%
1997	2	6	0.169	39.63%
1997	3	7	-0.057	33.09%
1997	4	8	0.406	-5.56%
1998	1	9	0.631	25.81%
1998	2	10	0.488	-44.73%
1998	3	11	0.949	14.65%
1998	4	12	0.546	75.29%
1999	1	13	0.564	8.70%
1999	2	14	0.699	22.65%
1999	3	15	0.612	150.52%
1999	4	16	1.148	4.67%
2000	1	17	1.406	-9.13%
2000	2	18	0.997	-21.54%
2000	3	19	0.983	-16.85%
2000	4	20	2.104	-14.98%
2001	1	21	1.533	39.65%
2001	2	22	1.546	-31.94%
2001	3	23	1.394	80.74%
2001	4	24	1.528	-15.27%
2002	1	25	0.931	-19.68%
2002	2	26	0.975	-5.74%
2002	3	27	0.862	17.28%
2002	4	28	1.435	-8.63%
2003	1	29	1.264	14.87%
2003	2	30	0.760	19.95%
2003	3	31	0.639	20.07%
Correlation between 1996/1 and 2002/3				**-28.6%**
Correlation between 2002/4 and 2003/3				**-82.2%**

As seen in the table above, a discriminant function is calculated between the periods 1996/1 and 2002/3 and in the calculation period the correlation between the periodic returns of the ISE100 Index and the weighted z score is –28.6%. However, the correlation jumps to –82.2% between the index and the weighted z score in the testing period. If the numbers are observed carefully, in the before and after crises periods there will be huge fluctuations and these fluctuations distort the correlation calculated in this period. Between the 1996/1 and 2002/3 period, two crises have hit the Turkish

economy: Russian Crisis in August 1998 and Liquidity Crisis in February 2001. Since between 2002/4 and 2003/3, there are no serious fluctuations in the prices, the correlation is significantly better than the correlation in the calculation period. Similar results will be observed for the logistic regression analysis in the next section.

Since there is one dependent variable and one independent variable in the above observations, regression analysis, in addition to correlation analysis, is also possible. If we regress the independent variable, the result will be as seen in the following picture:

Return_ISE100(t+1) = 0.3388-(0.2163*Weighted_Z_Score)

with 9.04% coefficient of determination value, which is high enough in stock market index forecast standards.

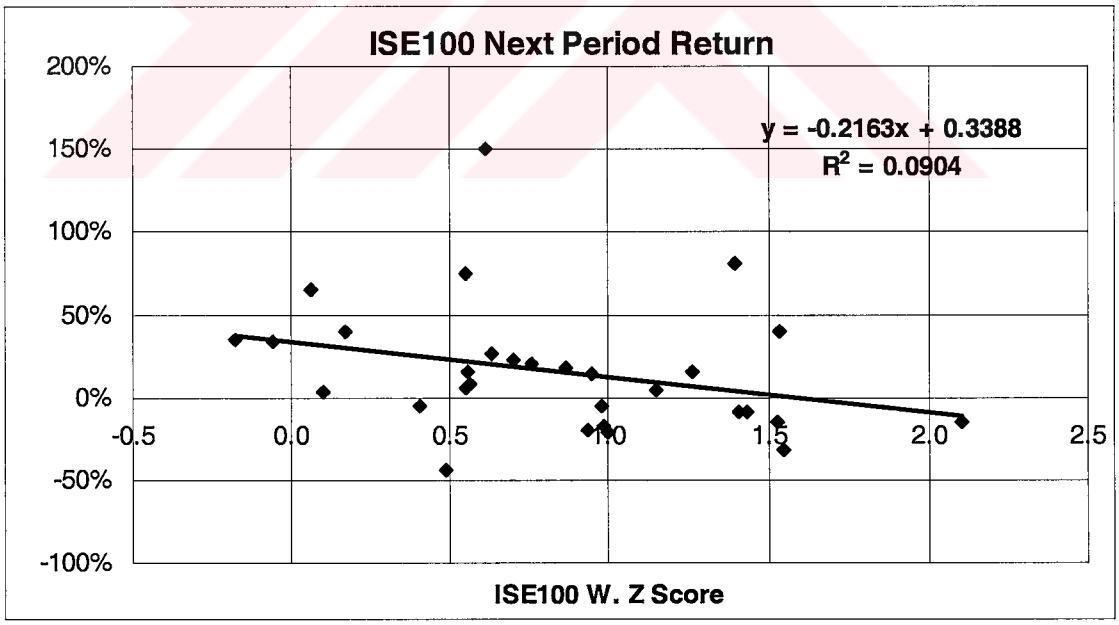


Figure 3.2. Analysis of the ISE100 Index According to Weighted Z Score

In discriminant analysis, the variables should be normally distributed. However, in this kind of return forecasts, the return of the prices cannot be normally but may log-normally be distributed. Thus, for index forecast of discriminant analysis, there is a serious assumption violation. However, in logistic regression, there is no assumption of normality and for this kind of analysis it is more flexible.

3.6. DISCUSSION FOR THE LOGISTIC REGRESSION

As in the discriminant analysis, the logistic regression results are tested by using the balance sheets and the income statements of approximately 192 industry firms traded on the ISE in the period between the last quarter of 2002 and the first three quarters of 2003.

If the logit-value of the company is higher for the given periods, the expected return for that company will be higher, too. To test this hypothesis, the company returns of the four quarters -- which are not used in the analyses -- for approximately 192 industrial companies of the ISE are calculated and after dividing with the return of the ISE, we end up with the real returns, as in the discriminant analysis case.

$$\hat{\mu}_{Company_i}^{ith_Period} = [(1+\mu_{Company_i}^{ith_Period})/(1+\mu_{ISE100}^{ith_Period})]-1$$

When we analyze the total data set for the real returns, we see that the average real return of these company for the last quarter of 2002 and the first three quarters of 2003 is –4.05% with 28.06% standard deviation. Thus, the ratio of the mean to the standard deviation is –14.43%, exactly the same as in the discriminant analysis.

Finally, for each logit value, there is a real return. The total data set can be divided into two pieces with respect to the average of this logit value (the average of the logit values of the total data set is 0.89). As a result, according to our data set for the industrial firms, if the logit value is greater than 0.89, we will have 368 companies. Similarly, if the logit value is smaller than 0.89, we will have the other 385 companies as in the following table.

Table 3.12. Anova Table of the Logistic Regression

TOTAL DATA	
T Mean	-4.05%
Sample	753
Stdev	28.06%
Mean/Stdev	-14.43%

GROUP1				GROUP2			
Mean	Sample	Stdev	Mean/ Stdev	Mean	Sample	Stdev	Mean/ Stdev
-7.93%	368	22.26%	-35.65%	-0.34%	385	32.25%	-1.05%
Logits greater than			***0.89***	***Logits smaller than***			***0.89***

	TSS	df	MS	F-value	p-Value
Between Group	1.086	1	1.086	**14.029**	**0.02%**
Within Group	58.126	751	0.077		
Total	59.212	752			

Consequently, the mean of the nonfailed companies is -0.34% with 32.25% standard deviation. However, the mean of the failed companies is –7.93% with 22.26% standard deviation. The standard deviation of the nonfailed companies is higher because of the higher the return higher the volatility assumption. Besides, the means are significantly different between the nonfailed and the failed companies. The difference is also seen at the ratio of the mean to the standard deviation. It is –1.05% for the nonfailed companies, while being –35.65% for failed companies. As a result, the risk in

the nonfailed companies is incomparably higher than the risk in the failed companies. The Anova test of these two groups with a cut-off logit value of 0.89 has enough evidence to discriminate the good return expected companies to the bad return expected companies with 14.0 F value and 0.02% significant level.

Table 3.13. Correlation of the ISE100 Index and the Weighted Score

Year	Quarter	Index	Weighted Score	ISE100 Return(t+1)
1996	1	1	0.889	5.14%
1996	2	2	0.948	2.78%
1996	3	3	0.995	34.69%
1996	4	4	0.955	65.29%
1997	1	5	0.897	15.13%
1997	2	6	0.952	39.63%
1997	3	7	0.994	33.09%
1997	4	8	0.922	-5.56%
1998	1	9	0.897	25.81%
1998	2	10	0.910	-44.73%
1998	3	11	0.846	14.65%
1998	4	12	0.913	75.29%
1999	1	13	0.918	8.70%
1999	2	14	0.913	22.65%
1999	3	15	0.933	150.52%
1999	4	16	0.900	4.67%
2000	1	17	0.848	-9.13%
2000	2	18	0.946	-21.54%
2000	3	19	0.931	-16.85%
2000	4	20	0.765	-14.98%
2001	1	21	0.773	39.65%
2001	2	22	0.842	-31.94%
2001	3	23	0.884	80.74%
2001	4	24	0.833	-15.27%
2002	1	25	0.868	-19.68%
2002	2	26	0.899	-5.74%
2002	3	27	0.909	17.28%
2002	4	28	0.845	-8.63%
2003	1	29	0.821	14.87%
2003	2	30	0.934	19.95%
2003	3	31	0.966	29.22%
Correlation between 1996/1 and 2002/3				**26.4%**
Correlation between 2002/4 and 2003/3				**70.6%**

As in the discriminant analysis case, by using the book value of the stocks, weighted average of the logit score can be calculated. If there is a correlation between the Istanbul Stock Exchange Index and the weighted logit score, it will be useful for the forecast of the ISE100 Index, studied in the next chapter. As seen in the table above, between the periods 1996/1 and 2002/3, the calculation period, the correlation between the 3 month period returns of the ISE100 Index and the weighted logit score is 26.4%. However, the correlation jumps to 70.6% between the index and the weighted logit score in the testing period. Similarly, in the results of the discriminant analysis, because of the Russia Crisis in August 1998 and Liquidity Crisis in February 2001, there are huge fluctuations and these fluctuations distort the correlation of the calculation period. Since between 2002/4 and 2003/3, there have been no major crises in Turkey, the economy is in a recovery process. Thus, the correlation in this period is significantly better than the correlation in the calculation period.

We have one dependent variable and one independent variable for the correlation analysis. If we regress the independent variable, the result will be as seen in the following picture:

$$Return_ISE100(t+1) = -1.5469+(1.9037*Weighted_Logit_Score)$$

with 7.39% coefficient of determination value, which is quite considerable for our analysis.

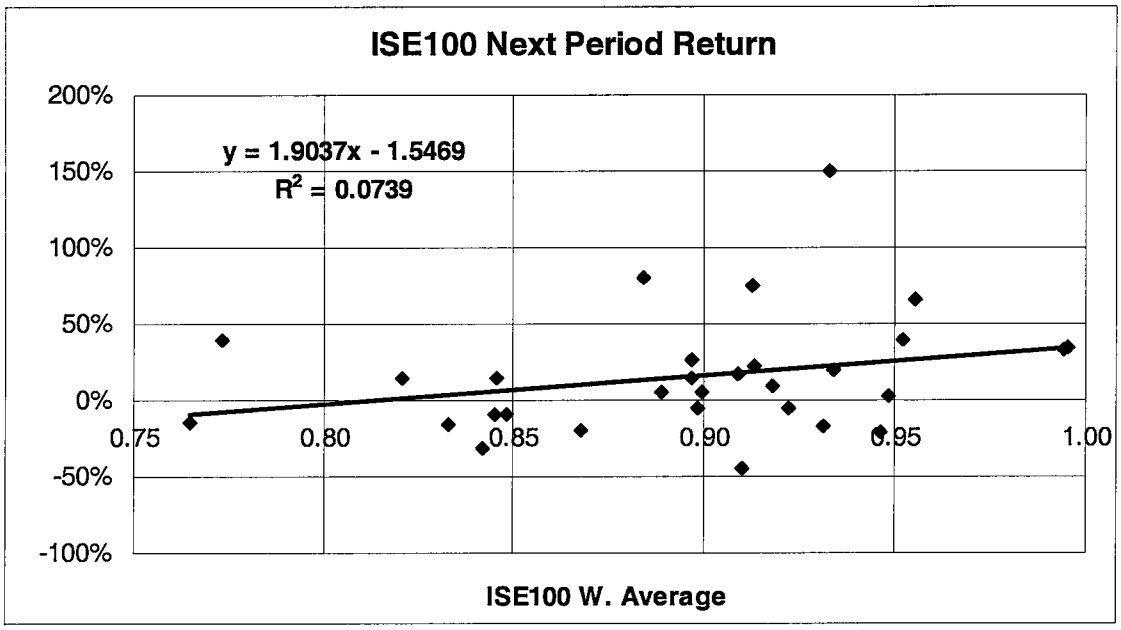


Figure 3.3. Analysis of the ISE100 Index According to Logistic Regression Results

3.7. GETTING INFORMATION FROM FINANCIAL DATA

In this chapter, a discriminant analysis and a logistic regression calculation method are presented for the elimination of the unsuccessful or the underperformed industrial firms in terms of the stock market performance and for the application of the ISE100 Index forecast. In other words, the failed company prediction for Turkish industrial companies, traded on the ISE was aimed for this section after the adaptation of Altman's famous bankruptcy prediction paper.

Despite being unfortunate occurrences, the fact that Turkey has had serious economic crises in recent years has rendered this era of the Turkish economy worthy of experimentation for academic purposes. Generally, it is known that a healthy company's shareholders' equity should be at least 40% of its total assets. Since the

average of this ratio decreases to around 20% for the post-crisis period (year1998 and year 2001), most of the companies have some liquidity and bankruptcy problems in Turkey.

As an example, the possible effects of the crises in August 1998 and February 2001 after the euphoria period of year 1997 and year 2000, respectively, are the following:

- Usage of money in the firms and the efficiency of the firms decreased.
- While the turnover ratio of the firms was decreasing, profitability margin was increasing.
- The overall active profitability increased because of the higher increase in profitability margin compared to the decrease in turnover ratio.
- Because of the high interest rates, the cost of debt or funding of debt increased incredibly.
- Thus, the amount of debt increased exponentially.
- Even active profitability was higher in euphoria period, because of the huge financial expenses, the firms faced high losses.
- High losses recorded by the firms caused them to perform worse in the stock market with respect to the stock market index, paving the way for losses in the firms' shareholders' equities and culminating in their bankruptcies.

As a result, cash ratio is the most important factor in the analyses. EBIT and the sales factors of the companies are also considerably important. However, the stock market reacts to the crises promptly but the financial data follow with a lag. Because of the delay, volatility of the prices and the market value of the equity are adjusted on time and can be important variables. The delay makes the correlation between the return of

the stock market index and the results of both analyses to decline sharply. In the long term, the correlation can be around 30%, which can attract the attention of the stock market investors, but in the short term, it fluctuates between 5% and 80%.

Besides, in the testing period of the forecastibility power of both the discriminant analysis and the logistic regression analysis, it is seen that the failed companies are more than thirty times riskier than the nonfailed companies. Consequently, these ANOVA results show the predicting power of both analyses used in this section for detecting the failed companies or for estimating the underperformed companies.

Since there is a considerably high correlation between the weighting scores of both analyses and the ISE100 Index, these scores should be used in the fuzzy logic model as a long-term fundamental analysis indicator.

CHAPTER 4. A MACROECONOMIC FUZZY LOGIC MODELLING OF THE ISE

4.1. INTRODUCTION TO RULE BASED SYSTEM: FUZZY LOGIC

In the past decade, fuzzy systems have been used with conventional techniques in many scientific applications and engineering systems, especially in system theory. Fuzzy sets, introduced by Zadeh (1965) as a mathematical way to represent vagueness in linguistics, are different from the classical set theory. In a classical nonfuzzy set, element of the universe either belongs to or does not belong to the set, in other words the membership of an element is crisp. A fuzzy set is a generalization of an ordinary set in that it allows the degree of membership for each element in a unit interval. In order to quantify the vague rules, as characteristic of fuzzy systems, an expert, data set and a model should be needed. Fuzzy logic can be perceived as the bridge between the knowledge of humans and their expectations by using the information set with mathematical calculations. Thus, fuzzy rules represent the people's vague knowledge in the system. Transformation methods are applied to train this knowledge.

In order to model the density functions on stock returns, Hsieh (1991) claims that the model should have nonlinear dependence and concludes that stock returns are not identically independent distributed. There are also numerous research papers claiming that applying nonlinear models (Chenoweth and Obradovic 1996; Bak, Paczuski and Shubik 1997; Gencay 1998, Leung, Daouk, and Chen 2000) such as

neural networks, genetic algorithms and fuzzy logic can provide successful predictions. For neural network application, Kimoto and Asakawa (1990) uses several neural networks to train the relationships between past values of technical and economy indices for forecasting the returns of the Tokyo Stock Exchange Index. Komo, Chang and Ko (1994) uses a radial basis function network to predict the future index values by using past data of the index, the transaction costs, and bond market values. Fuzzy logic and neural networks are complementary methods in designing advanced models. Each method has its good and bad aspects. Fukuda and Shibata (1994) presents the comparison of these techniques. Fuzzy logic is better than neural networks in mathematical modelling, knowledge representation and expert knowledge but worse than neural networks in learning and optimization ability.

The theory of fuzzy is widely used in control theory. There exist two major types of fuzzy controllers: Mamdani (1974) and Takagi-Sugeno (1985). Mamdani (1974) uses fuzzy sets and Takagi-Sugeno (1985) uses linear functions. Fiordaliso (1998) finds appropriate rules in the framework of chaotic time series forecasting. He proposes a Takagi-Sugeno fuzzy controller to evaluate the rules in the inference process by using gradient descent algorithm. The results of his study point out the advantage of using a nonlinear "combining forecast" method. Lee and Antonsson (2000) also works on the Takagi-Sugeno Fuzzy System with evolutionary computation approach and finds a locally Pareto optimal set of solutions. Nomura, Hayashi and Wakami (1992) works on the optimization process for fine-tuning of the input data. One of the acceptable ways of fine-tuning in fuzzy logic is the optimization by steepest descent algorithm to minimize the error. Not only the membership functions' working but also the rule generation in fuzzy logic system should work properly. Sun (1994) proposes a rule combination mechanism to build a simplified version of the original rule base. In his

study, a data structure is produced to organize rules so that the rule base can be matched against input signals with logarithmic efficiency. Sugeno and Kang (1988) suggests an algorithm for structure identification by applying an example in their studies. Additionally, Patel (2004) works on the twelve different fuzzy reasoning methods along with numerical comparisons by using transformation functions in fuzzy systems. According to his method, the number of mathematical operations and amount of memory to compute the defuzzified control output is enormously reduced.

Furthermore, some time series applications are also stated in the fuzzy logic theory literature. Economakos (1979) estimates the demand for electrical power by using linguistic terms of the needed electricity. Then, Tanaka, Uejima and Asai (1982) develops the fuzzy linear regression analysis. For the finance application, Shnaider and Kandel (1989) constructs a model to forecast corporate income tax revenue by using moving averages of the corporate tax revenue and per capita GNP for Florida. Besides, Song and Chissom (1993) analyzes the theoretical part of the fuzzy time series modelling. They analyze to fuzzify the data, develop the time series model, and calculate the output. Then, Song and Chissom (1994) applies a time invariant forecasting model by using different defuzzifying methods. Meanwhile, Chen (1996) works on the same model by changing the arithmetic operations in rule generation part.

As a detailed macroeconomic modelling example in fuzzy systems, Kooths (1999) develops a macroeconomic model to realize an alternative to conventional expectation hypothesis. The experience and rule based expectations can be used in forecasting behaviour that is characterized by explicit rule orientation (theory foundation), vague formulation (bounded rationality) and learning process (acquisition of experience). By using rules, he builds an experimental macroeconomic model by using money supply, interest rate, exchange rate, currency reserves, net capital imports,

price increase, unemployment rate and technical progress. Human reasoning can also be modelled, as if the thought process is described by the application of fuzzy logic (LeBaron, Arthur and Palmer 1999). Traders are capable of handling a large number of rules for mapping of market states into expectations. Linn and Tay (2001) shows this by allowing agents the ability to compress information into a few fuzzy notions, which they can in turn process and analyze with fuzzy logic. Their work also analyzes the fuzzy inductive reasoning and nonlinear dependence in security returns by using the outputs of artificial stock market environment.

Since the rule generation for complex methods is very important and the available rules may not be enough for modelling, there can be other experimental environments to learn the mechanism of the process, which may be constructed in the model. In order to understand the price movements of the stocks with effect and reaction mechanism, a laboratory environment should be designed as in the models of the computational economics. The laboratory asset markets permit the controlled manipulation of the rules and procedures. Many hypotheses can be investigated under laboratory environment. One of the famous experimental stock markets, the Santa Fe stock market investigates market efficiency and price convergence with rational expectation asset pricing model. The model is described in detail in Arthur, Holland, LeBaron, Palmer & Tayler (1997) and LeBaron, Arthur and Palmer (1999). This model attempts to integrate the training mechanism into a well-defined economic structure, along with inductive learning using a classifier-based system. One feature of this model is that it allows agents to explore a fairly wide range of possible forecasting rules (Chen and Yeh 2001). The results on this market show that the artificial stock market is able to replicate certain empirical problems observed in real markets.

Consequently, stock prices have specific properties that make the prediction more complex because of the strong resemblance to random walk processes (Lo and MacKinlay, 1988). A successful prediction algorithm does not have to provide predictions for all points in the time series. The important measure of the success is the generated profit where the algorithm produces predictions and the explanatory power of the model above the standards. This thesis supports the idea of the possibility to identify these movements and predict future changes. Our argument against the efficient market hypothesis is constructed because of the time delay between the point when new information enters and the point when the information has been evaluated and a new equilibrium price has been established. From this standpoint, the controversy revolves around how the word "immediately" should be interpreted in the efficient market hypothesis definition. Being an academician and a trader, I believe that the answer to this problem is a matter of perception and time differences of traders and the academicians. In other words, traders may be in general faster than academics in the market.

Consequently, to characterize the ISE100 Index, one may implement fuzzy models, explaining to some extent predictable behaviour patterns. The fuzzy models for the ISE100 Index in this thesis have been used for the first time in the literature. By using technical analysis indicators such as Bollinger bands and moving average, fundamental analysis indicators such as cash ratio, quick ratio and gross profit margin, and macroeconomic analyses such as repo rates, fx rates and gross national product, we can derive some rules. By using the rules in fuzzy logic method, the model can be constructed by an expert (myself) and optimized with steepest descent learning algorithm.

4.2. INVESTMENT THEORY AND THE STOCK MARKET

An investment is a commitment of money or value that is expected to receive additional money. Every investment entails some degree of risk; it requires a certain sacrifice at the present for an uncertain future benefit. Stocks, bonds, repos, options and futures are the various investment tools that provide attractive returns at different levels of risk. However, options and futures market has not been opened at the Istanbul Stock Exchange (ISE) yet. When we look at the risk return analysis of the remaining assets; stock, repo and bonds: Stock market has the highest level of risk. In Turkey, the average fluctuations of bond and repo market are almost similar and significantly less risky with respect to the stock market because the markets are affected by similar political and macroeconomic events.

Stock is an equity security (Modis, 1999), which implies ownership. If a company is represented by a stock traded in a stock exchange market, investors can own some percentage of the company. This is the easiest way for individuals to participate in the profits and losses of a business.

When a corporation is formed, nothing exists but the name. To raise money to buy materials with which to produce something or to run its operations, a corporation can sell shares of the stock for a determined price. When a corporation sells shares of the company, the person buying the shares becomes a shareholder of the corporation, someone who owns an interest in that corporation.

As large corporations were formed and the number of shareholders needed to finance large manufacturing ventures that began to rise, there came a need to find some

centralized place where people could buy and sell their shares in a corporation. Without a centralized meeting place, people would have to search for buyers and buyers would have to search for sellers.

Finally, a stock exchange can be simply explained as a place where people gather to exchange shares pertaining to their ownership in different corporations, either for shares of ownership in other corporations or for money. Over time, stock markets have grown to be the basis of market economies. Every major market country on the world has a stock exchange, although the rules and methods of operation may change from country to country.

In the stock exchanges, indices are composed in order to calculate price and return performances of all shares as well as on the basis of relative markets and sectors. Until the end of 1996, the ISE used to compute only the ISE100, financial and industrials price indices. As of 1997, the ISE began to calculate sector and sub-sector indices on the basis of prices and total return.

The ISE price indices are computed and published throughout the trading session while the return indices are calculated and published at the close of the session only. The ISE100 Index is used as a main indicator of the ISE.

The ISE100 is composed of market companies except investment trusts. The contents of the ISE100 Index are selected on the basis of predetermined criteria directed for the companies to be included in the indices as well as in consideration of their ability to represent relevant sectors. The market capitalization is weighted in the publicly-held portion of each constituent stock kept in custody (Takasbank) (except those kept in non-fundable accounts). The basic formula for calculating of the ISE100 index is as follows:

$$ISE100_t = \frac{\sum_{i=1}^{100}(P_{it} \quad N_{it} \quad Fl_{it})}{D_t}$$

where P_{it} is the closing price of stock i in period t , N_{it} is the total number of shares of the stock i in period t (paid-in capital/1,000), Fl_{it} is the flotation weight (publicly-held portion) of the stock i kept in custody (except those kept in non-fungible accounts) in period t , and D_t is the value of divisor in period t (adjusted base market value). In the calculation of the index only registered prices are taken into account.

As a result, the most extensive index in the Istanbul Stock Exchange is the ISE100 Index, which represents the Turkish stock market. The adjustment process for prices of the stocks and the index should be taken long enough. Thus, daily and weekly periods can be quite short for the adjustment of the fast movements in prices. Besides, macroeconomic data are generally released on a monthly basis in general. The independent variable can be taken as the monthly return of the ISE100 Index for the modelling. The dependent variables, used for the modelling will be presented in the next section.

4.3. DEPENDENT VARIABLES

Data selection must be performed judiciously to avoid the "garbage-in, garbage-out" syndrome often associated with computers. Performance of the fuzzy logic is highly dependent on the quality and appropriateness of its input data. If relevant data

inputs are not included, the performance will suffer needlessly. For this reason, it is important to have a solid understanding of the financial markets and intermarket relationships within the context of global investment.

Following a specific market analysis theory or analytic perspective on the markets when selecting input data for a financial fuzzy logic application has its own implications. I posit that the financial markets are nonlinear, and possibly chaotic, that market inefficiencies exist, can be discerned quantitatively, and persist long enough for investors to profit from them. Technical analysis suggests the use of only single-market technical data as inputs. Fundamental analysis focuses on economic factors that influence the target market. Neither of these traditional approaches alone is sufficient in today's global environment. Similarly, intermarket analysis that relies solely on the subjective and/or qualitative interpretation of several price charts cannot adequately model the simultaneous influence of multiple markets on a target market. Through the use of the fuzzy logic, the algorithm should combine these three methods of analysis into a multidimensional quantitative framework. Multiple data inputs reflecting a broad range of related markets and fundamental inputs can be used to discern the general tendencies and patterns between markets, and quantify these nonlinear relationships that affect a given market.

Investors need to adapt their investment strategies and methods, in order to survive and profit in this new world order. Today's limited single-market focus must yield a broader analytical framework that can quantify complex intermarket relationships. What is needed for successful investment is a broader concept of market analysis that recognizes the interrelatedness and interdependence of today's financial markets. Investors can no longer focus solely on individual markets. It is now imperative that they adopt an intermarket investing perspective and utilize analytical

tools capable of assessing complex intermarket relationships. To accomplish this, investors must expand their analytic perspective to include external factors in addition to internal factors that affect a given market. Technical market analysis must encompass the global context of today's financial markets, by incorporating the more narrowly defined schools of technical, fundamental, and intermarket analysis. Here, intermarket can be any market compatible with the invested market such as repo, bonds and bills market. In our analysis, these variables are grouped together under the category of macroeconomic variables. As a result, the dependent variables may be categorised broadly into the following three sections:

1. Technical Analysis
2. Fundamental Analysis
3. Macroeconomy.

4.3.1. Technical Analysis

The technical analysis approach is essentially a reflection of the idea that prices move in trends, which are determined by the changing attitudes of investors toward a variety of economic, monetary, political and psychological forces. The art of technical analysis is to identify trend changes at an early stage and to maintain an investment posture until the weight of the evidence indicates that the trend is reversed. The following points summarize technical analysis:

1. Technical analysis is based on published market data and focuses on internal factors by analyzing movements in the aggregate market, industry average, or stock. In contrast, fundamental analysis focuses on economic and political factors, which are external to the market itself.

2. The focus of technical analysis is identifying changes in the direction of stock prices, which tend to move in trends as the stock price adjusts to a new equilibrium level. These trends can be analyzed, and changes in trends detected, by studying the action of price movements and trading volume across time. The emphasis is on likely price changes.
3. Technicians attempt to assess the overall situation concerning stocks by analyzing breadth indicators, market sentiment, and momentum.

The volatility of the stock market prices in Turkey is very high. Since the technical analysis is the analysis of the prices as explained above, the analysis should be divided into time zones such as long term or short term. The momentum of the prices can be different according to taken periods. As a result, in the modelling part, the Technical Analysis Variable can be grouped with respect to the analysis of the periods into two factors:

1. Long Term Technical Analysis of the Istanbul Stock Exchange 100 Index,
2. Short Term Technical Analysis of the Istanbul Stock Exchange 100 Index.

1. Long Term Technical Analysis of the Istanbul Stock Exchange 100 (ISE100): This variable looks for the index level changes in the long term. Since most of the technical analysis indicators looked for the momentum of the market and this kind of momentums can only be in short or medium terms. There should be a broader analysis, which shows the price fluctuations in the long term. The analysis used in Aksoy and

Sağlam (2000) will be the basis for this variable as explained in detail in the second chapter. To sum up, the forecasting power of the classifier system clearly shows that sizeable returns are available to patient investors. Thus, the results should be used in the fuzzy logic model as a long-term technical analysis indicator. For the short-term, technical analysis indicator simple moving analysis can be used to catch the momentum of the equity market in the short run.

2. Short Term Technical Analysis of the Istanbul Stock Exchange 100 (ISE100): Short term technical analysis is based on the widely accepted premise that security prices are determined by the supply of and the demand for securities. The tools of the technical analysis are therefore designed to measure certain aspects of supply and demand. Typically, as explained above, technical analysts record historical financial data on charts and study these charts in search of meaningful patterns and they aim to use the patterns to predict future prices. The basic assumptions underlying technical analysis are as follows:

- Market value is determined by the interaction of supply and demand.
- Supply and demand are governed by numerous factors, both rational and irrational.
- Security prices tend to move in trends that persist for an appreciable length of time, despite minor fluctuations in the market.
- Changes in a trend are caused by the shifts in supply and demand.
- Shifts in supply and demand, no matter why they occur, can be detected sooner or later in charts of market transactions.
- Some chart patterns tend to repeat themselves.

As a result, with the 10 days moving average calculations of the ISE100 Index and by using the transaction volume of the ISE100 Index and the daily highest and lowest ISE100 Index values, a band is calculated for the short term technical analysis variable of the short term fuzzy logic model. If the price of the ISE100 Index is above the band, the short-term technical analysis variable will be the ISE100 Index return of the previous month. Otherwise, it will be the negative of the ISE100 Index return of the previous month.

4.3.2. Fundamental Analysis

Fundamental analysis is based on the premise that any security (and the market as a whole) has an intrinsic value, or the true value as estimated by an investor. This value is a function of the firms underlying variables, which combine to produce an expected return and an accompanying risk. By assessing these fundamental determinants of the value of a security, an estimate of its intrinsic value can be determined. This estimated intrinsic value can then be compared to the current market price of the security. Similar to the decision rules used for bonds, decision rules are employed for common stocks when fundamental analysis is used to calculate intrinsic value.

The analysis used in Aksoy (2003) will be the basis for this variable as explained in detail in Chapter 3. The failed company prediction for Turkish industrial companies, traded on the ISE was calculated after the adaptation of Altman's famous bankruptcy prediction paper.

As explained in Chapter 3, there is a considerably high correlation between the ISE100 Index and the weighting scores of both analyses, the logistic regression and the

discriminant analysis. Thus, these scores should be used in the fuzzy logic model as a long-term fundamental analysis indicator.

In discriminant analysis, the variables should be normally distributed. However, in this kind of return forecasts, the return of the prices cannot be normally but may log-normally be distributed. Thus, in the discriminant analysis, explained in Chapter 3, there is a serious assumption violation. However, in the logistic regression, there is no assumption of normality, which is more flexible for this kind of analysis.

Finally, the stock market variable used in the fuzzy logic model is produced from the weighting scores of the logistic regression, which summarizes all the ratios derived from balance sheets and income statements of the companies. As a result, the final results of Chapter 3 will be used as a variable in this section.

4.3.3. Macroeconomy

The contributions to the stock market modelling by macroeconomic variables are various in the literature. Hall (2000) argues that the stock market moves in low frequency in relation to other macroeconomic measures, such as GDP. Cash flows earned by corporations suggest that their level is broadly consistent with the rational markets that value of the securities reflects the present discounted value of the cash earnings. As a result of this argument, to estimate the index level the cash earning of the country may be meaningful. Besides, the increase in GNP, industrial production or capacity utilization should be consisted with the level of the index for a given country. Rigobon (2001) also uses an identification technique based on the heteroskedasticity of stock market returns to identify the reaction of monetary policy to the stock market. He concludes that monetary policy significantly affects the stock market movements. Additionally, Taylor (1998) also shows the historical analysis of monetary policy rules

in the U.S. and states the interrelations of the macroeconomic variables. Furthermore, Edey and Romalis (1996) designs an econometric modelling to quantify the role of monetary policy in Australia. Studies on stock returns and inflation have discovered a negative relationship using ex post data. Fama (1981) attempts to explain the relationship between expected inflation and stock returns as an induced consequence of a negative relationship between inflation and real activity as a basic consequence of Fisherian quantity theory. Firth (1979) examines the relationship between stock market returns and inflation in the U.K. In addition, Hasbrouck (1984) shows the empirical relationships between stock returns, inflation, and economic activity by using survey evidence.

Consequently, the macroeconomic variables can be grouped into three factors, where the last one, the TL market has high correlation with the ISE in the short run. Besides, the real economy variable is effective in the long run and the foreign exchange variable is useful to measure the resistance to foreign exchange based crises in Turkey. These three factors are grouped as follows:

1. Real Economy Indicators
    A. Gross National Product (GNP)
    B. Industrial Production Level
    C. Capacity Utilization Level
2. Foreign Exchange (FX) Market Indicators
    A. Balance of Payments
    B. FX Reserves
    C. FX Rate
3. TL Market Indicators

A. Primary Budget Deficit

B. Repo rates (after inflation)

C. Treasury Bills' and Government Bonds' Price Index in the ISE (DIBS) (after inflation).

1. Real Economy Indicators: Long-term macroeconomic performance in Turkey can be measured by three broad measures: GNP, industrial production index and capacity utilization level. News of these three variables makes the headlines because these issues affect our daily lives in the short run. They also dominate the research agenda in macroeconomics. On the other hand, the growth figures of the country affect the stock market in the long run because of the interactions between companies. Besides, firstly capacity utilization index data and secondly industrial production index are released earlier than GNP. By using these earlier figures, GNP figures can be calculated and fund managers can shape their strategies based on the results in the long run. To sum up, these three index variables as representatives of the real economy for Turkey can be summarized in the following:

1.A.Gross National Product (GNP): GNP is the value of all final goods and services produced in the economy in a given period (quarter or year). GNP is the basic measure of the economic activity. Nominal GNP measures the value of output at the prices prevailing in the period during which the output is produced, while real GNP measures the output produced in any one period at the prices of some base year. The growth rate of the economy is the rate at which real GNP is increasing. What causes real GNP to grow over time? The first reason is that the available amount of resources in the economy change. The resources are conveniently split into capital and labour. The labour force, consisting of people either working or looking for work, grows over time

and thus provides one source of increased production. The capital stock, including buildings and machines, likewise rises over time, providing another source of increased output. Increase in the availability of the factors of the production accounts for part of the increase in real GNP. The second reason, for real GNP to change is that the efficiency with which factors of production work may change. Over time, the same factors of production can produce more output. These increases in the efficiency of production result from changes in knowledge, including learning by doing, as people learn through experience to perform familiar tasks better.

When the growth rate is high, the production of goods and services is rising, making possible an increased standard of living. With the high growth rate typically go lower unemployment and the availability of more jobs. High growth is a target and hope of most securities. The growth rate of real GNP is the most important of all the macroeconomic indicators by which to judge the economy's long run performance. Thus, the value of the stocks and their performance expectations should be equal to the growth performances of the companies in the future and the growth figures of the companies represent the growth expectations in the country briefly. Consequently, the expectations for the growth figures, which can be calculated from pre-indexes such as capacity utilization and industrial production, and the realizations of the growth figures, which can be taken from the Turkish State Institute of Statistics, will be useful as indicators for the real economy.

In Turkey, quarterly real GNP values are available. Real increase in the data can be calculated and the increase is taken as the variable for the model.

1.B.Industrial Production Level: Industrial production is a government report that provides a fixed-weight measure of the physical output of the nation's factories, mines, and utilities. Manufacturing production, the largest component of the total, can be

accurately predicted using total manufacturing hours worked from the employment report. One of the bigger wildcards in this report is utility production, which can be quite volatile due to swings in the weather. Severe hot or cold spells can boost production as increased heating/cooling needs drive utility production up.

Besides, growth rates of total factor productivity differ widely across sectors. It also raises doubts about how effectively output is measured in some service industries, such as banks or the government. While analyzing the growth of a country, it should be better to consider the measurement of industrial productivity.

The State Institute of Statistics releases the industrial production data every month. Industrial Production Data is very useful to forecast the growth figures. Since GNP data is released quarterly and Industrial Production data is released monthly and two-three months earlier than the GNP data, the production data can also be used as a good indicator to gauge the status of the economy. Thus, the variable should be used in real economy section of the model with GNP figures. As a variable, because of the seasonal effects, the change in the ratio of the industrial production level with respect to the previous year's respective month is calculated.

1.C.Capacity Utilization Level: Capacity utilization for an individual manufacturing plant is the ratio of the plant's actual output level to its potential or capacity output level. Capacity output is a measure of the extent to which the manufacturing plant can produce goods, given its current technology and fixed factors of production (such as the capital stock).

In addition to industrial production, capacity utilization is also reported monthly and approximately one month earlier than the industrial production index. Though the rate of capacity utilization is seen as a critical gauge of the slack available in the economy, the market does not completely trust this measure. Capacity is very difficult

to measure. Economists closely watch capacity utilization for signs of inflation pressures. The data is typically expressed as an increase or decrease from the prior month.

Since the production depends on the amount of output produced in an economy to the inputs of factors of production and to the state of technical knowledge, any company in the economy that works in full capacity or less than full capacity depending on the external and internal effects, has a contribution to the growth figures of the country. Consequently, the capacity level of a company affects the production and the growth.

As a variable, the difference of the data with respect to the previous year's same month is used.

2.Foreign Exchange (FX) Market Indicators: Foreign Exchange, also known, as "Forex" or "FX", constitutes the largest financial market in the world, with a daily average turnover of approximately US$1.5 trillion. Foreign currency trading is the simultaneous buying of one currency and selling of another. Foreign currencies can be on a floating exchange rate regime or on pegged regime and are always traded in pairs; e.g., the Euro versus the Dollar or the Dollar versus the Lira.

Because of the heavy influences of the foreign investors in Turkish markets, it is important to analyze the variables of the foreign exchange (FX) market. Besides, resent crises in Turkey are the currency crises and the stress is increased with current account deficits in balance of payments. Not only foreigners but also Turkish investors analyze the foreign reserves in order to measure the resistance to possible foreign currency attacks. Foreign reserves and current account deficit are very useful to give an idea about the ability of the economy to withstand currency attacks.

On the other hand, in Turkey the currency regime was changed from the pegged regime to the floating exchange rate regime after the February 2001 crisis. Thus, the volatility of FX rate variable is increased after executing floating exchange rate regime in Turkey. Additionally, the volatility of the foreign reserves and the current account deficit variables are decreased after February 2001 crises. As a result, the variables, which measure the net FX amount in Turkey and the variable, which values the FX notes, should be analyzed together in general.

Consequently, these variables; current account deficit and fx reserves (as measuring fx amount) and fx rate (as measuring fx value) represent the foreign exchange market for Turkey and in the following, their employment in the model is summarized:

2.A.Balance of Payments: The balance of payments is the record of the transactions of the residents of a country with the rest of the world. The simple rule for balance of payments accounting is that any transaction that gives rise to a payment by a country's resident is a deficit item in that country's balance of payments. Thus, imports of cars or oil, use of foreign shipping, gifts to foreigners, purchase of land in the other country, or making a deposit in a Bank in Switzerland are all deficit items. Examples of surplus items, by contrast, would be sales of cars abroad, payment by foreigners, pension from abroad or foreign purchases of stock. When there is a net outflow, balance of payments is in deficit. To control the economy and attain a healthy growth level in a country, a surplus in the balance of payments is necessary. When foreign investors leave the country, it will turn into a deep deficit and disrupt the economy. As evidenced in the February 2001 crisis, high current account deficits increase the pressure on the economy. Fischer's recent research suggests that under certain circumstances, a large current account deficit may make the U.S. economy vulnerable to severe disruptions.

Recent theories and some data should be analyzed to understand how the current account deficit could be either an optimal outcome or a threatening one. Even if a large current account deficit can be justified from the perspective of macroeconomic theory, recent work by Obstfeld and Rogoff (2000) offers a reason for worry. Ostfeld and Rogoff note what might happen if an economy was forced to balance its current account suddenly. International investors might decide to park their funds elsewhere. Such an event could result from a variety of unforeseen shocks, such as a slowdown in the growth rate relative to other countries or a sudden fall in the stock market. Obstfeld and Rogoff argue that the effects of such a current account reversal would be especially potent for any economy. Using a simple model embodying their theoretical assumptions, the authors compute that a rapid reversal of the current account would generate a real exchange rate depreciation greater than 20% in magnitude.

Consequently, monthly current account deficit or surplus is calculated as a variable in the model for representing the fx market.

2.B.FX Reserves: The Central Bank of Turkey holds reserves that they sell in the market when there is an excess demand for dollars. Conversely, when there is an excess supply of dollars, they buy the dollars. The data is released in the weekly press bulletin of the Central Bank of Turkey.

Efficient management of foreign reserves is one important task of a central bank. Risk management contributes to the fulfilment of this task by strategically managing and controlling exposures to financial and operational risks. Most importantly, a risk management framework is needed to establish the appropriate strategic asset allocation for the institution. The framework will determine the ability to manage the reserves actively and to analyze performance and risk appropriately. Since the investors are also following the risk position of the central bank's foreign reserve,

the weaker positions constitute risk for the devaluation in the country. In cases of foreign attacks at times of weaker central bank positions, stock markets can also be deeply impacted.

Consequently, weekly change in fx reserves is calculated as a variable in the model for representing the fx market.

2.C.FX Rate: FX rate is the value of one unit of foreign money in TL terms. A rise in the exchange rate means that foreign prices have increased relative to the prices of goods produced here. Goods abroad have become more expensive relative to the prices of goods at home, which other things equal, implies people are likely to switch some of their spending to goods at home. This is often described as an increase in the competitiveness of our products, as our goods become cheaper relative to foreign goods. In this case of devaluation, the stock market will be cheaper and have comparative advantage in prices, while other things being constant.

A currency crash is defined as a large change of the nominal exchange rate that is also a substantial increase in the rate of change of nominal depreciation. Next comes the examination of the composition of the debt as well as its level, and a variety of other macroeconomic factors, external and foreign. Crashes tend to occur when: output growth is low; the growth of domestic credit is high; and the level of foreign interest rates is high. A low ratio of FDI to debt is consistently associated with a high likelihood of a crash. These can be summarized as the weakness of an economy for any foreign attack. As discussed above, fx rate is a valuation variable with respect to foreign currencies, while current account deficit and fx reserves are measuring the net flow in the economy.

Consequently, the difference of daily devaluation rate in United States Dollar is used as a variable in the model for the fx market section.

3.TL Market Indicators: In Turkey, there are other markets to invest in TL-terms. These are repo and bond market. Although their volatility is less than the ISE, they are influenced by similar news effects as well. On the other hand, these variables can be considered as alternative to the stock market, as they affect each other in an adverse manner.

In other words, if the stock market investor decides to sell his stock portfolio but stay invested in the country, the investor will either buy bonds or use his money in TL based instruments such as repo. However, if he decides to invest in another country, he can also buy foreign currency and leave the country. As a result, with respect to the developments and flash news in the economy, stock market, bond market and repo market move in the same directions but if the investors decide to decrease the stock portfolio and at the same time naturally increase repo and bond portfolio, they are considered alternative instruments. Thus, the correlation between each other is influenced by adverse effects. In spite of the fact that Turkey is commonly accepted as an emerging market, the liquid investments in the economy are influenced by flash news and developments more than the alternative instruments, while the movements of the stock market and the bond market are as positively correlated as in the developed economies. Besides, these TL markets are influenced by the government's primary deficit. Thus, in this group, budget deficit should be included as another dependent variable.

Since the TL's value declines due to inflation, the variables discussed above should be adjusted with respect to inflation. The return of the Istanbul Stock Exchange, used as a dependent variable in the model, is not divided to the rate of inflation because the dividends are not included in the indexes in the Istanbul Stock Exchange.

The inflation rate is the percentage rate of increase of the level of prices during a given period. The inflation is one of the main concerns of citizens, policy makers, and macroeconomists. During periods of inflation, the prices of normal and accepted goods are rising. Partly for this reason, inflation is unpopular, even if people's incomes rise along with the prices. Inflation is also unpopular because it is often associated with other disturbances to the economy such as the oil price increases that would make people worse off. Mainly, there are two types of inflation: wholesome price index (WPI) and consumer price index (CPI). These price indexes depend on different categorical variables. To sum up, CPI is more individual needs' price increase oriented. However, in an economy the price changes are generally measured by WPI because it gives the large amount of goods' price index.

Consequently, the variables; bond market index, repo rate and primary budget deficit after WPI adjustment can be used to represent the TL market indicators for Turkey and their methods of usage in the model are summarized in the following:

3.A.Primary Budget Deficit: Like private companies, the government has a budget too. When the government is spending more than it receives, the budget is in deficit. The size of the budget deficit is affected by the government's fiscal policy variables; such as government purchases, transfer payments, and tax rates. To finance the deficit, the government has to borrow debt from creditors or the budget should be in surplus by increasing revenues or decreasing expenditures in the next periods.

The debt ratio in Turkey increased during the currency crisis in February 2001, due to the accumulated duty losses of public banks as well as the burden caused by bankrupted private banks. As at the beginning of 2004, the debt ratio to GNP in Turkey is around 80%, which is considerably high with respect to other European countries. Besides, the interest of this debt has the highest portion of the budget expenditure in

Turkey. In order to decrease the public debt in Turkey, the difference between the revenues and the expenditures without interest in the budget, in other words primary budget deficit should be significantly positive. The positive expectations or realizations in primary budget deficit influence the TL market and decrease the interest rates in Turkey. With decreasing cost of carry, the companies in the economy have better opportunities and the valuations of the stocks increase.

As a result, the percentage increase in the total primary budget deficit after inflation (WPI) adjustment in the last 12 months is calculated and used in the model.

3.B.REPO Rates: Repurchase agreements (repos) are instruments used to help finance part of their inventories of marketable securities for one or few days. For instance, if an investor ends a day of trading with an increase of $1 million in its inventory of marketable securities, a repo may be sold to finance the $1 million inventory overnight. The investor is essentially making a short-term loan to the other investor with his inventory serving as collateral. Repos that last longer than overnight, called term repos, can span 30 days or even longer. Term repos are marketable securities, actively traded between the money market trading desks of large banks and brokerage houses. In another words, repo is a short-term loan instrument to find money to enter the stock market. The ISE data provides the monthly return of overnight repo rate after WPI adjustment and is used as a variable for cost of carrying money.

3.C.Treasury Bills' and Government Bonds' Price Index in the ISE (DIBS): A bond is a promise by a borrower to pay the lender a certain amount at a specified date and pay a given amount of interest per year. The interest rates on bonds issued by different borrowers reflect the differing risks of default. Default occurs when a borrower is unable to meet the commitment to pay interest or principal. Government Bonds represent the indebtedness of government. The owners of the bonds are creditors;

government is the debtor. Government Bonds are of such high quality that their yield is often used as an example of a risk free, default free or interest rate. On the ISE, government bonds are also traded. There are approximately 20-30 government bonds traded in the ISE and these bonds change according to the maturity date and interest rates in the meantime. The ISE calculates the weighted average price of the government bonds, the DIBS Index. As a variable, the monthly return of the DIBS Index after WPI adjustment is used for the model in the TL market section.

4.4. IMPLEMENTATION OF FUZZY LOGIC

Appropriate tools are needed to integrate data from related markets and quantify their influence. Fuzzy logic is very well suited to implementing synergistic analysis. Fuzzy Logic can be used to synthesize technical, fundamental, and macroeconomy analysis into one analytic framework as discussed earlier, capable of finding hidden patterns and complex relationships in the data. The application of this technology to financial market analysis has quickly become a hot subject in the financial industry, with dozens of articles and several books written on the subject in the past few years. Fuzzy logic is simply a mathematical tool. It is their ability to quantify relationships and find patterns in the data from related markets that allows fuzzy logic to play an important role in implementing synergistic market analysis in today's global markets.

Fuzzy logic has been applied practically in many areas where conventional model based approaches are difficult to implement for the design and the learning. However, as the system complexity increases, reliable fuzzy rules and membership functions used to describe the system behaviour become difficult to determine. In this

more complex environment, the fuzzy logic system needs more computational support to work with. In this case alternative conventional model based approaches have better performances.

Furthermore, due to the dynamic nature of economic and financial applications, rules and membership functions must be adaptive to the changing environment in order to continue to be useful. It has been considered as one of the most attractive strategies in identifying complex systems, particularly for nonlinear systems with imprecise and uncertain knowledge of system information and behaviour. In contrast to a regular knowledge base in expert system, the fuzzy rules, being structured knowledge, can be coded into relevant explicit numerical algorithms or a fuzzy model with fuzzy identification or fuzzy reasoning based on system non-fuzzy input or data by Lin and Lee (1996). Fuzzy logic can be applied to system modelling, estimation, optimal and optimization of control and adaptive control problems. Only system and input-output data are required. Fuzzy rule-based systems apply these methods to solve many types of "real-world" problems, especially where a system is difficult to model, is controlled by a human operator or expert, or where ambiguity or vagueness is common. A typical fuzzy system consists of a rule base, membership functions, and an inference procedure.

The benefits of fuzzy design are:

- Simplified and reduced development cycle,
- Ease of implementation,
- Provision of more "user-friendly" and efficient performance.

On the other hand, if the number of variables is higher, the number of rules will increase much faster. This will cause a rule explosion in the model. In order to prevent

the rule explosion, an expert should make a hierarchical modelling in fuzzy logic. In the beginning of this study, a hierarchical modelling is used too. In the model, there are also middle layers to combine related variables in groups. The membership functions of the input and middle variables are optimized. Finally, in this section a model for the stock exchange will be developed to estimate the monthly returns of the ISE100 Index with 12 financial input variables.

4.4.1. Prepositional Logic

In crisp logic, such as binary logic, variables are true or false, black or white, 1 or 0. An extension to binary logic is multi-valued logic, where variables have many crisp values. Prepositional logic on the other hand, is defined with uncertain terms. The next months' return of the ISE100 index is estimated as 6% increase. In some aspects it may be a good increase, in another it may be a very good increase. Thus, with some degree it is a good increase and a very good increase.

In the example given above, very bad decrease, bad decrease, neutral, good increase and very good increase are the linguistic terms, which can be converted from the real data with some degree. As a result, Fuzzy Logic is a departure from classical two-valued sets and logic. This logic uses "soft" linguistic (e.g. good, bad, normal) system variables and a continuous range of truth values in the interval [0,1], rather than strict binary (True or False) decisions and assignments. Formally, fuzzy logic is a structured, model-free estimator that approximates a function through linguistic input/output associations.

4.4.2. Membership Function

A relationship is defined to express the distribution of truth of a variable. For example, 'good increase' may be defined as a distribution around a value 'x'. Any value within the distribution may be interpreted as 'good increase', although with different degrees of truth or confidence.

Theoretically, a fuzzy set **F** of a universe of disclosure **X**={x} is defined as a mapping, by which each x is assigned a number in range [a,b]. This indicates the boundaries of the attribute F of x. If x is the monthly return of the ISE100 index, 'good increase' may be considered as a particular value of the fuzzy variable, and each x is assigned a number in the range of real numbers. **U**good inc.(x) is element of [a,b] that indicates the extent to which that x is considered to be good increase. **U**good inc.(x) is element of [a,b] is called a membership function.

Let **X** be a time-invariant set of objects x. A fuzzy set **F** in **X** may be expressed by a set of ordered pairs **F** = {(x,**U**(x)) | x is an element of **X**}, where **U** is the membership function that maps **X** to the membership space **M**, and **U**(x) is the grade of membership or degree of truth of x in **F**. In addition, when **M** contains the values 0 and 1 only, then **F** is non-fuzzy and **U** is the characteristic function of a non-fuzzy set. Thus, binary and multi-value logic are extreme sub-cases of fuzzy logic.

As an example, let the monthly returns of the ISE Index in 2000 be 2.1%, 9.4%, -4.5%, 16.7%, -2.3%, -5.3%, 2.8%, 2.5%, -11.3%, 1.5%, 2.2% and 4.1%, respectively. Increases around 10% are considered Very-Good-Increase, around 5% Good-Increase, around 1% Neutral, around -1% Bad-Decrease and -5% Very-Bad-Decrease. In the example, there are not more then two membership-functions for each x as seen in the below picture (While constructing the architectures of the models there are similarities).

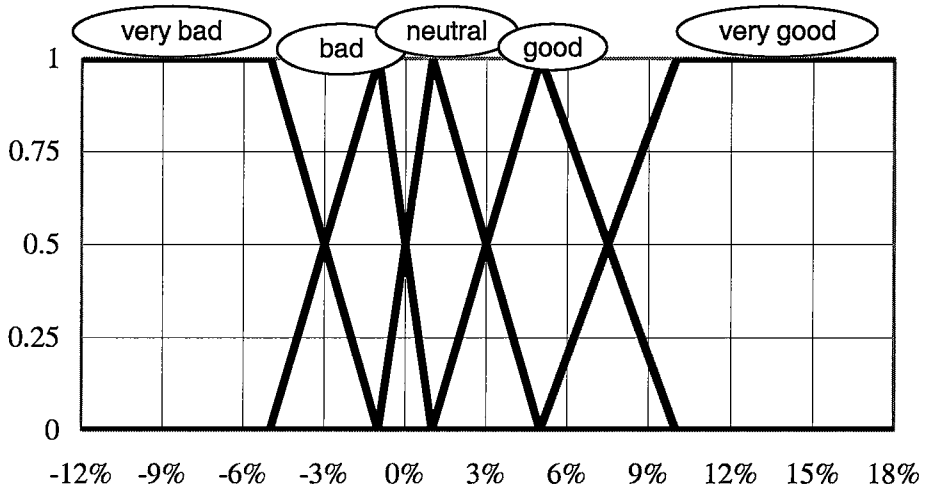


Figure 4.1. Membership Functions

As a result of the membership functions, the data can be converted easily. For simplicity, piecewise linear membership functions have been assumed. Membership functions can be continuous curves of many different shapes (Lin and Lee, 1996).

Table 4.1. Fuzzification of the Variables

Date	Monthly Return of ISE100	Very-Good-Increase	Good-Increase	Neutral	Bad-Decrease	Very-Bad-Decrease
00/01	2.10%	0%	0.275%	0.725%	0%	0%
00/02	9.40%	0.880%	0.120%	0%	0%	0%
00/03	-4.50%	0%	0%	0%	0.125%	0.875%
00/04	16.70%	1.000%	0%	0%	0%	0%
00/05	-2.30%	0%	0%	0%	0.675%	0.325%
00/06	-5.30%	0%	0%	0%	0%	1.000%
00/07	2.80%	0%	0.450%	0.550%	0%	0%
00/08	2.50%	0%	0.375%	0.625%	0%	0%
00/09	-11.30%	0%	0%	0%	0%	1.000%
00/10	1.50%	0%	0.125%	0.875%	0%	0%
00/11	2.20%	0%	0.300%	0.700%	0%	0%
00/12	4.10%	0.775%	0.225%	0%	0%	0%

Fuzzification of the five crisp variables as in the figure above, causes spreading of the variables with a distribution profile that suits the problem. As an example for the data, 2.1% in January 2000, the figure of the membership function should be used. In the figure, when x axis is equal to 2.1% and a vertical line is plotted crossing 2.1% at the x-axis, the line intersects with membership "good" and membership "neutral" at the point (2.1%, 27.5%) and (2.1%, 72.5%), respectively. Accordingly, all the vertical lines intersect in the membership function figure at one or two points. If it intersects at one point, the ordinate of the point is equal to 1 (The intersection is with outer membership functions in case of extreme data.). If it intersects at two points, the sum of the ordinates is equal to 1 (The intersection is with inner membership functions in case of normal data.).

4.4.3. Rule Generation and Fuzzification

In fuzzy problems, the rules are produced based on experiences. Concerning problems that deal with fuzzy engines or fuzzy control, all possible input-output relationships should be in fuzzy terms. To do this, there is a need for an expert. A fuzzy expert system is an expert system that uses a collection of fuzzy membership functions and rules, instead of Boolean logic, to reason about data. The input output relationships or rules are expressed with 'if then' statements such as:

If (a is in A1) and (b is in B1), then (c is in C5); or

If (a is in A1) and (b is in B2), then (c is in C1); or

If (a is in A2) and (b is in B5), then (c is in C3).

The A's and the B's are fuzzified inputs, and the C's are the actions for each variable. For 2 variables with 5 membership functions for each, total number of rules should be 25 rules (5x5) as seen in the below table.

Table 4.2. Rules

VAR. A	VAR. B: B1	B2	B3	B4	B5
A1	C5	C1	C2	C2	C5
A2	C4	C4	C1	C3	C3
A3	C2	C3	C4	C1	C5
A4	C1	C2	C4	C4	C1
A5	C1	C1	C3	C5	C2

The rules become more difficult to tabulate if the fuzzy statements are more in number, that is if there are 3 variables with 5 membership functions for each, total number of rules would increase to 125 rules (5x5x5). However, the need for fuzzy system is to simplify the environment and develop the most likely model. If the environment is not simplified as desired, the natural environment and the fuzzy environment can not be different and the fuzzy system may become useless. Consequently, the usage of fuzzy architecture depends on the capability of the simplification of the environment by fuzzy experts.

4.4.4. Defuzzification

Because of the nature of the fuzzy logic, the values are not crisp, and their fuzziness exhibits a distribution described by the membership function. If one tries to get two fuzzy variables, what will the output be? The question has been answered in Patel (2004) by various fuzzification and defuzzification techniques. He assumed that

the input variables are fuzzified with normal, convex fuzzy sets of any shape (triangular, bell etc) and that the output variable membership functions are convex, and continuous with trapezoidal or triangular membership functions. He presented the transformation functions for twelve fuzzy reasoning methods with their numerical comparisons among each other.

Consequently, by using the conclusion part of the rules, the result in one fuzzy subset is assigned to each output variable for each rule. Usually only "min" or "product" is used as inference rules. By the "min inference" of the rules' results, the output membership function is clipped off at a height corresponding to the rule premise's computed degree of truth (fuzzy logic and). In "product", the output membership function is scaled by the rule premise's computed degree of truth. Then, all of the fuzzy subsets assigned to each output variable are combined together to form a single fuzzy subset for each output variable. Again, usually "max" or "sum" is used. In "max", the combined output fuzzy subset is constructed by taking the pointwise maximum over all of the fuzzy subsets assigned to variable by the inference rule (fuzzy logic or). In "sum", the combined output fuzzy subset is constructed by taking the pointwise sum over all of the fuzzy subsets assigned to the output variable by the inference rule. Note that "sum" can result in truth values greater than one! Thus, "sum" is only used with normalization. In the composition subprocess, all of the fuzzy subsets assigned to each output variable are combined together to form a single fuzzy subset for each output variable.

The model in this thesis uses "Mamdani min" operator because the "Mamdani min" implication is a simplified and most intuitive version of "Zadeh max-min" proposed by Mamdani in the 1970s in connection with fuzzy control (Mamdani, 1977).

Furthermore, defuzzification is the process where the membership functions are sampled to find the grade of membership. The grade of the membership is used in the fuzzy logic equation and after an outcome region is defined, the output is deduced. There are some defuzzification methods. Mizumoto (1989) compared about ten defuzzification methods. Two of the more common techniques are the center of area and maximum methods. In the center of area method, the crisp value of the output variable is computed by finding the variable value of the center of gravity of the membership function for the fuzzy value. In the maximum method, one of the variable values at which the fuzzy subset has its maximum truth value is chosen as the crisp value for the output variable. There are several variations of the maximum method that differ only in what they do when there is more than one variable value at which this maximum truth value occurs. One of these, the mean of maximum method, returns the average of the variable values at which the maximum truth value occurs. Because of the widespread usage of center of area method, it is used in the defuzzification process of the model.

The result of an example is presented below with the data shown in the table above by using Mamdani min rule for fuzzy logic operation and center of area for defuzzification.

With the rules in Table 4.2, the output memberships will be:

if (a is in A4) and (b is in B1) then (c is in C1);

if (a is in A5) and (b is in B1) then (c is in C1);

if (a is in A4) and (b is in B2) then (c is in C2);

if (a is in A5) and (b is in B2) then (c is in C1).

The calculations show that:

UA4(a)=0.282,

UA5(a)=0.718,

UB1(b)=0.191 and

UB2(b)=0.809.

With Mamdani min operator:

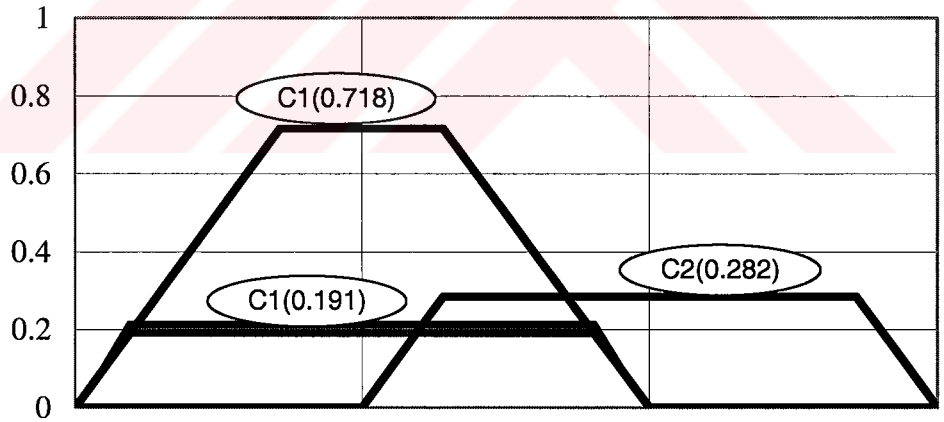
Min(**U**A4(a), **U**B1(b))=0.191 for (c is in C1);

Min(**U**A5(a), **U**B1(b))=0.191 for (c is in C1);

Min(**U**A4(a), **U**B2(b))=0.282 for (c is in C2);

Min(**U**A5(a), **U**B2(b))=0.718 for (c is in C1).

Since there is one question but four answers, the defuzzification process will be done via taking the center of area of the four rectangles as seen in the graphics below, and will give the value of the output.



Membership Fuctions: C1(0.191), C1(0.191), C1(0.718) and C2(0.282)

Figure 4.2. Defuzzification

Fuzzy system can express knowledge but cannot learn to adapt by itself. In many applications, partial understanding of the knowledge is available, but a complete

set of rules is not. However, input-output data may be available to determine the rest. What is needed is technology that can work with a partial knowledge base and can also learn from the additional data in order to perform the task correctly. Fuzzy logic handles the explicit knowledge, whereas optimization handles the knowledge implicit in the data. A fusion of both these into one provides a better way of resolving problems.

4.5. MODELLING

One of the main advantages of fuzzy logic is that it more closely models the kind of reasoning that a person engages in when dealing with issues or elements that are not precisely defined and that involve aspects of degree or judgment. However it may not be possible to think in terms of sharp boundaries for modelling. Thus, the disadvantages of fuzzy logic systems are the same as those found in traditional knowledge based systems: Someone has to write rules, which means that expert knowledge has to be available and formalized. These systems can not learn on their own, nor can they adapt to changing market conditions, except by manually rewriting the rules, adjusting the membership functions or other specific rule-finding methodologies.

According to professional investors, the stock markets are generally affected by four main criteria (Virtual Trading, 1995):

- Stock Market Factors
- Macro Economic Factors
- Risk Factors
- Inter-Market Factors.

In the model designed for the ISE100, the inter-market factors (effects of other markets) are included in macroeconomy part. Thus, the models in the following pages generally consist of three main parts; technical analysis, fundamental analysis and macroeconomy. The variables used in the model are discussed separately in the third section above. The design of the model is constructed according to the relations of the variables between each other.

Firstly, the technical analysis section index level observation, which is used to analyze the ISE100 Index in the long-run, and short term technical analysis, which is used to analyze the momentum of the ISE100 Index, are used as the variables to represent the first and dynamic section of the model.

Secondly, the fundamental analysis section is the results of the logit scores derived from data of the stock markets balance sheets and income statements by using 81 financial ratios after running logistic regression in SPSS. The final score is the weighted average of the companies' separated logit score, which represent the firms' financial statements in the medium and long run.

Finally, the macroeconomy section in the model has three main parts: Real Economy, FX Market and TL Market. Real economy shows the fundamental changes in Turkey. These are GNP, industrial production index and capacity utilization ratio. Capacity utilization ratio is an indicator for the industrial production and industrial production is an indicator for GNP. The FX Market consists of current account deficit, which is an important item of the balance of payments, FX reserve and FX rate. Current account deficit and FX reserve variables measure the amount of foreign exchange, while FX rate is used for the valuation of these assets. The aim of this part is to quantify the resistance of currency crises in the economy. TL Market consists of primary budget deficit, repo and DIBS after inflation adjustments. This last part as an alternative

investment section in TL terms other than the stock market, shows how good the TL market expectation is. Consequently, the 12 inputs used in the models are summarized in the table below.

Table 4.3. Variables Used in the Model

V1	Long Term Tech. Analysis	Long Term Risk (Aksoy&Saglam,2001)
V2	Short Term Tech. Analysis	Moving Average (for 10 working days)
V3	Fundamental Analysis	Discriminant Analysis for Turkish Companies
V4	GNP	% Change in Gross National Product (y-o-y)
V5	Industrial Production	% Change in Industrial Production (y-o-y)
V6	Capacity Utilization	% Change in Capacity Utilization (y-o-y)
V7	Balance of Payments	Current Account Deficit
V8	FX Reserves	Weekly change in FX Reserve of the Central Bank
V9	FX Rate	US Dollar/TL Parity
V10	Budget Deficit	Budget Deficit/Wholesome Prices (y-o-y)
V11	Repo Rates/Inflation	Monthly Average of Repo Rate over ex-ante Inflation
V12	DIBS/Inflation	Gov. Bills&Bonds Index over ex-ante Inflation

As shown in Figure 4.3, by using the first two of the variable, another variable will be produced, named the Technical Analysis Variable. The third variable will be used to estimate the output directly. From the fourth variable to the sixth variable, the Real Economy Variable will be produced, as the first middle variable. From the seventh variable to the ninth variable, the FX Market Middle Variable will be produced. From the tenth variable to twelfth variable, the TL Market Variable will be produced as the last middle one. By using these three middle variables, the Macroeconomy Variable will be produced. Finally, the output will be calculated with the Technical Analysis, Fundamental Analysis and Macroeconomy Variables.

For each variable the maximum, minimum, average and the standard deviations are given in the following table. By using these values with expert opinion, the groups of the fuzzy membership functions of the variables are given in the next table. All of the variables have five membership functions. As explained before, if there are two variables with five membership functions for each, number of rules will be 25. Number of rules will be 125 for three variables with five membership functions.

As seen in Figure 4.1, the inner three membership functions are triangles and the outer two membership functions are trapezoids. In Table 4.4; 1, 2 and 3 are the border points of the first trapezoid; 2, 3 and 4 are the border points of the first triangle; 3, 4 and 5 are the border points of the second triangle; 4, 5 and 6 are the border points of the last triangle and 5, 6 and 7 are the border points of the last trapezoid. The middle number for each triangle is the head and the others are the bottom numbers. The last number for the first trapezoid and the first number for the last trapezoid are the bottom numbers.

Table 4.4. Statistical and Membership Range of the Variables

	MAX	MIN	AVE	STDEV	1	2	3	4	5	6	7
Y	1.070	-0.505	0.055	0.188	1.070	0.732	0.393	0.055	-0.132	-0.318	-0.505
V1	3.583	0.463	1.241	0.598	3.583	2.802	2.022	1.241	0.982	0.722	0.463
V2	1.070	-0.681	0.050	0.189	1.070	0.730	0.390	0.050	-0.193	-0.437	-0.681
V3	2.104	-0.178	0.767	0.555	2.104	1.658	1.213	0.767	0.452	0.137	-0.178
V4	0.104	-0.123	0.025	0.072	0.104	0.078	0.052	0.025	-0.024	-0.074	-0.123
V5	0.213	-0.144	0.028	0.088	0.213	0.151	0.090	0.028	-0.029	-0.087	-0.144
V6	86.600	68.500	76.835	3.957	86.600	83.345	80.090	76.835	74.057	71.278	68.500
V7	1106.000	-1512.000	-167.816	586.616	1106.000	681.395	256.790	-167.816	-615.877	-1063.939	-1512.000
V8	5462.000	-5362.000	38.192	721.797	5462.000	3654.064	1846.128	38.192	-1761.872	-3561.936	-5362.000
V9	0.566	-0.155	0.004	0.021	0.566	0.378	0.191	0.004	-0.049	-0.102	-0.155
V10	0.682	-0.557	0.019	0.167	0.682	0.461	0.240	0.019	-0.173	-0.365	-0.557
V11	0.332	-0.082	0.009	0.038	0.332	0.225	0.117	0.009	-0.021	-0.051	-0.082
V12	0.122	-0.128	0.001	0.010	0.122	0.081	0.041	0.001	-0.042	-0.085	-0.128

After having the variables with membership functions, the rules should be produced for the last step of the modelling. As seen in Figure 4.3, there are 12 input

variables and the other 6 middle variables are produced from these 12 variables. One of the 6 middle variables is produced from two different fuzzified variables (25 rules for each variable). The other 5 variables are produced from three different fuzzified variables (125 rules for each variable).

Finally, there should be 650 rules for the overall model. The rules for each variable are given in the following three tables. In the first table, the technical analysis variable is driven by the variables: observation and short term technical analysis. In the second table, three middle variables are produced to generate the macroeconomy variable. Finally, three main variables are combined to produce final output in the last table.

Table 4.5. Rules of Technical Analysis for the Modelling of the ISE100 Index (in 2 by 2)

OUTPUT: TECHNICAL ANALYSIS (C)

		INPUT 2: Short Term Technical Analysis (B)				
		B1	B2	B3	B4	B5
INPUT 1: L. Term Technical Analysis (A)	A1	C1	C1	C2	C2	C2
	A2	C2	C2	C2	C3	C3
	A3	C2	C3	C3	C3	C4
	A4	C3	C3	C4	C4	C4
	A5	C4	C4	C4	C5	C5

Table 4.6. Rules of Macroeconomy for the Modelling of the ISE Index (in 3 by 3)

OUTPUT: REAL ECONOMY (D)

INPUT 3: Capacity Utilisation (C)
INPUT 2: Industrial Production (B)

INPUT 1: GNP (A)	C1					C2					C3					C4					C5				
	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5
A1	D1	D1	D2	D2	D2	D2	D2	D2	D3	D3	D2	D2	D3	D3	D3	D3	D3	D3	D4	D4	D3	D3	D4	D4	D4
A2	D1	D2	D2	D2	D3	D2	D2	D2	D3	D3	D2	D3	D3	D3	D4	D3	D3	D3	D4	D4	D3	D4	D4	D4	D5
A3	D1	D2	D2	D2	D3	D2	D2	D3	D3	D3	D2	D3	D3	D3	D4	D3	D3	D4	D4	D4	D3	D4	D4	D4	D5
A4	D2	D2	D2	D3	D3	D2	D2	D3	D3	D3	D3	D3	D3	D4	D4	D3	D3	D4	D4	D4	D4	D4	D4	D5	D5
A5	D2	D2	D2	D3	D3	D2	D3	D3	D3	D4	D3	D3	D3	D4	D4	D3	D4	D4	D4	D5	D4	D4	D4	D5	D5

OUTPUT: FX MARKET (D)

INPUT 3: FX Rate (C)
INPUT 2: FX Reserve (B)

INPUT 1: Balance of Payments (A)	C1					C2					C3					C4					C5				
	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5
A1	D1	D1	D2	D2	D2	D1	D2	D2	D2	D3	D1	D2	D2	D2	D3	D2	D2	D2	D3	D3	D2	D2	D2	D3	D3
A2	D2	D2	D2	D3	D3	D2	D2	D2	D3	D3	D2	D2	D3	D3	D3	D2	D2	D3	D3	D3	D2	D3	D3	D3	D4
A3	D2	D2	D3	D3	D3	D2	D3	D3	D3	D4	D2	D3	D3	D3	D4	D3	D3	D3	D4	D4	D3	D3	D3	D4	D4
A4	D3	D3	D3	D4	D4	D3	D3	D3	D4	D4	D3	D3	D4	D4	D4	D3	D3	D4	D4	D4	D3	D4	D4	D4	D5
A5	D3	D3	D4	D4	D4	D3	D4	D4	D4	D5	D3	D4	D4	D4	D5	D4	D4	D4	D5	D5	D4	D4	D4	D5	D5

OUTPUT: TL MARKET (D)

INPUT 3: FX Rate (C)
INPUT 2: FX Reserve (B)

INPUT 1: Primary Bud. Def. (A)	C1					C2					C3					C4					C5				
	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5
A1	D1	D1	D2	D2	D2	D2	D2	D2	D3	D3	D2	D2	D3	D3	D3	D3	D3	D3	D4	D4	D3	D3	D4	D4	D4
A2	D1	D2	D2	D2	D3	D2	D2	D2	D3	D3	D2	D3	D3	D3	D4	D3	D3	D3	D4	D4	D3	D4	D4	D4	D5
A3	D1	D2	D2	D2	D3	D2	D2	D3	D3	D3	D2	D3	D3	D3	D4	D3	D3	D4	D4	D4	D3	D4	D4	D4	D5
A4	D2	D2	D2	D3	D3	D2	D2	D3	D3	D3	D3	D3	D3	D4	D4	D3	D3	D4	D4	D4	D4	D4	D4	D5	D5
A5	D2	D2	D2	D3	D3	D2	D3	D3	D3	D4	D3	D3	D3	D4	D4	D3	D4	D4	D4	D5	D4	D4	D4	D5	D5

OUTPUT: MACROECONOMY (D)

INPUT 3: TL Market (C)
INPUT 2: FX Market (B)

INPUT 1: Real Economy (A)	C1					C2					C3					C4					C5				
	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5
A1	D1	D1	D1	D2	D2	D2	D2	D2	D2	D2	D2	D2	D2	D3	D3	D3	D3	D3	D3	D3	D3	D3	D3	D4	D4
A2	D1	D2	D2	D2	D2	D2	D2	D2	D2	D3	D2	D3	D3	D3	D3	D3	D3	D3	D3	D4	D3	D4	D4	D4	D4
A3	D2	D2	D2	D2	D2	D2	D2	D3	D3	D3	D3	D3	D3	D3	D3	D3	D3	D4	D4	D4	D4	D4	D4	D4	D4
A4	D2	D2	D2	D3	D3	D3	D3	D3	D3	D3	D3	D3	D3	D4	D4	D4	D4	D4	D4	D4	D4	D4	D4	D5	D5
A5	D2	D3	D3	D3	D3	D3	D3	D3	D3	D4	D3	D4	D4	D4	D4	D4	D4	D4	D4	D5	D4	D5	D5	D5	D5

Table 4.7. Rules of Final Output for the Modelling of the ISE100 Index (in 3 by 3)

OUTPUT: RETURN OF THE ISE100 INDEX (D)

		INPUT 3: Macroeconomy (C)																								
		INPUT 2: Fundamental Analysis (B)																								
		C1					C2					C3					C4					C5				
		B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5
INPUT 1: Technical Analysis (A)	A1	D1	D1	D1	D2	D2	D1	D2	D2	D2	D2	D2	D2	D2	D2	D2	D2	D2	D2	D3	D3	D2	D3	D3	D3	D3
	A2	D2	D2	D2	D2	D2	D2	D2	D2	D2	D3	D2	D2	D3	D3	D3	D3	D3	D3	D3	D3	D3	D3	D3	D3	D4
	A3	D2	D2	D2	D3	D3	D2	D3	D3	D3	D3	D3	D3	D3	D3	D3	D3	D3	D3	D4	D4	D3	D4	D4	D4	D4
	A4	D3	D3	D3	D3	D3	D3	D3	D3	D3	D4	D3	D3	D4	D4	D4	D4	D4	D4	D4	D4	D4	D4	D4	D4	D5
	A5	D3	D3	D3	D4	D4	D3	D4	D4	D4	D4	D4	D4	D4	D4	D4	D4	D4	D4	D5	D5	D4	D5	D5	D5	D5

Consequently, the final model for estimating the monthly returns of the ISE by using these 650 rules is shown in Figure 4.3.

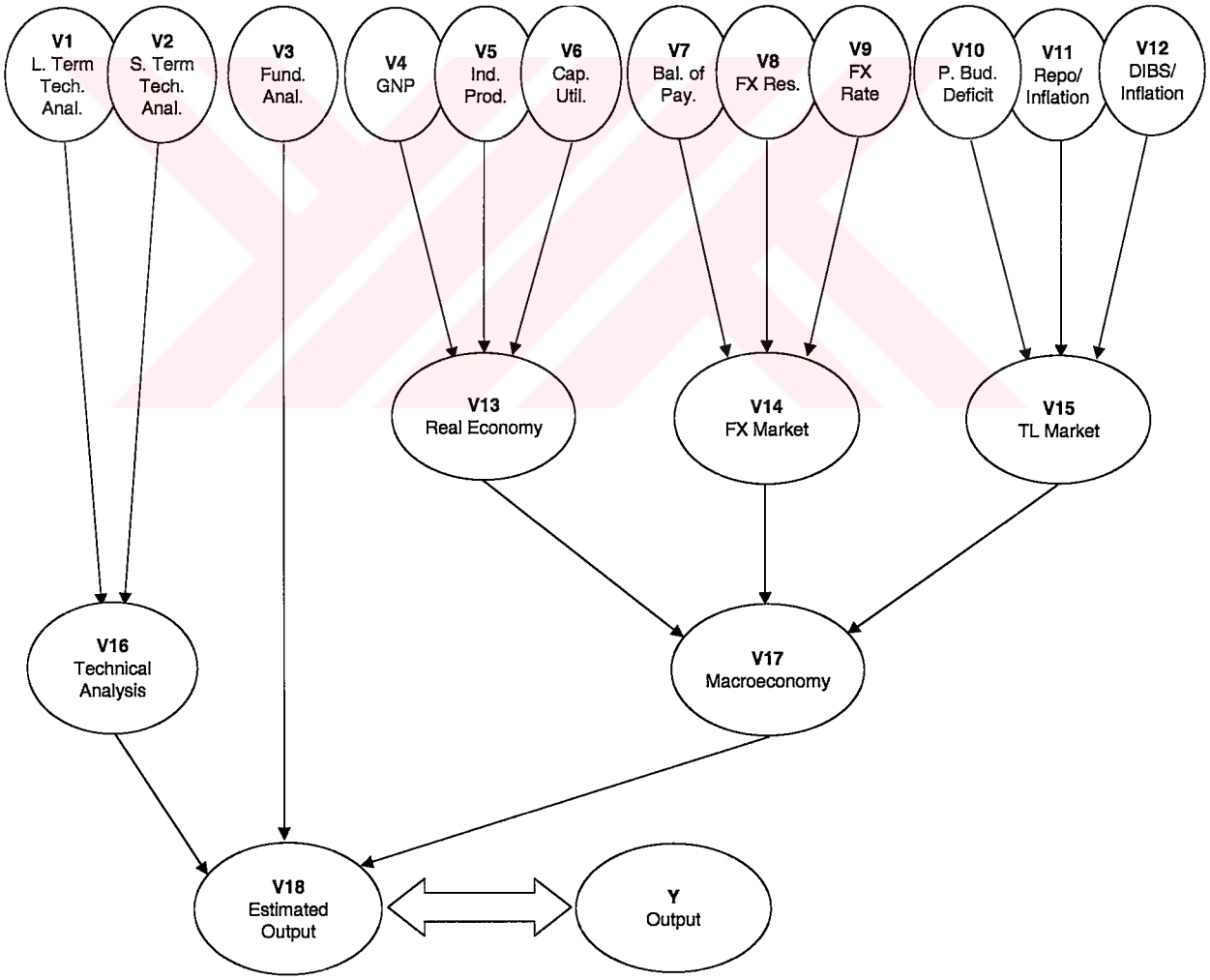


Figure 4.3. Fuzzy Logic Model for the ISE100 Index

For the structure of the data set, when the data is released, the variable is adapted according to the recent data in the next day. The monthly return of the stock market index is in daily frequency. Some others are not. The data set is constructed overlapping because the variables in the data set are announced daily, weekly, monthly and quarterly. Like the economists or strategists' points of view in portfolio management companies, when a variable in the data set is changed, the estimation of the model for the next day will be adapted to these changes.

If the rules and the model structure are carefully analyzed, the weightings of some variables for the output are more effective. For the technical analyses variable, the observation input is more effective than the short-term technical analysis input as seen in Table 4.5. As also seen in Table 4.6 and in Table 4.7, the rules are generated according to their priority with respect to the expert (myself)!

For the Capacity Utilization, the Industrial Production and the GNP inputs, the GNP data is released later than the Industrial Production data and the Industrial Production data is released later than the Capacity Utilization data. Thus, taken positions are realized after the announcement of the GNP. As a result, specific to this model, the Industrial Production and the Capacity Utilization data affect the model positively, while conversely the GNP data affect the model negatively. The other variables in the model are positively correlated with the model.

Besides, Turkey is a well-known emerging market and the importance of the flash news and the economic developments are considerably higher than accepting the DIBS and the repo as alternative instruments; thus, the DIBS and the repo inputs are also positively correlated with the model.

The result of the model by using the rules above with fuzzy logic algorithm is 5.02%, which is the sum of the squares of the difference of the real returns from the

estimated returns divided by the real returns of the stock market for the period between 1996 and 2003.

4.6. LEARNING

After having 5.02% R square value as seen in the figure below, the question should be: Is there any improvement in the accuracy? Since the data set covers the period starting from the beginning of 1996 and to the ending of 2003. Thus, the next work is to find the optimum stock market model, which is developed by changing the border points of the membership functions by using steepest descent algorithm.

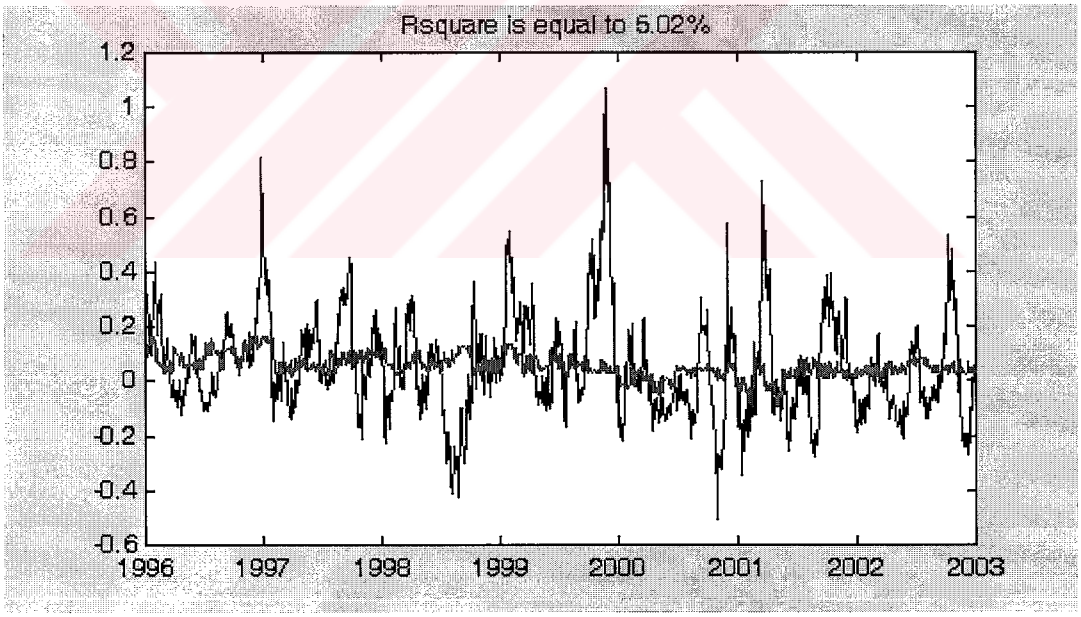


Figure 4.4. Actual and Estimated Output Before Optimization

4.6.1. Steepest Descent Algorithm

Various learning algorithms can be used when performing optimization, which provide methods of minimizing the overall error associated with the fuzzy logic output. This results in a trade-off between network performance and training time. Since algorithms are utilized that produce acceptable solutions within a reasonable time, one of the most common algorithms used in optimization is the gradient-descent algorithm. Starting at an initial point on the error surface, it determines the gradient that quantifies the slope, or steepness, of the curvature of the error surface at that point. The algorithm uses this information by moving in the exact opposite direction by an amount proportional to the learning rate. By utilizing this "downhill" movement along the error surface, the algorithm minimizes error. Other learning algorithms include conjugate-gradients, Newton-Raphson and Levenberg-Marquardt. However, no one algorithm is the best for all optimization problems. They all require a starting point somewhere in the vicinity of the optimal point, and some degree of smoothness of the performance landscape.

In the steepest descent method, the gradient of the objective function gives the direction $(-\nabla f)$ for the optimization. Here, the objective function is ratio of the mean square error of the difference of the output and the estimated output. The gradient will be:

$$\nabla f(\bar{x})^t = \lim_{\alpha \to 0^+} \frac{f(\bar{x} + \alpha \bar{d}) - f(\bar{x})}{\alpha},$$

where $\bar{d}$ is the direction. The steepest descent algorithm moves along the direction $\bar{d}$ with

$$\left\|\bar{d}\right\| = 1 \qquad \text{and} \qquad \bar{d} = \frac{-\nabla f(\bar{x})}{\left\|\nabla f(\bar{x})\right\|},$$

that minimizes the above limit. Here $\bar{d}$ represents the changes in the borders of the membership functions. As a result the algorithm will be;

1. Given x^0, set $k \to 0$.
2. $\bar{d}^k = -\nabla f(\bar{x}^k)$. If $\bar{d}^k = 0$, then stop.
3. By using one dimensional search, choose α^k.
   a. Set α^k=1.
   b. Calculate $f(\bar{x}^k)$, $f(\bar{x}^k + (\alpha^k/2)\bar{d}^k)$ and $f(\bar{x}^k + \alpha^k \bar{d}^k)$.
   c. If $\min\{ f(\bar{x}^k), f(\bar{x}^k + (\alpha^k/2)\bar{d}^k), f(\bar{x}^k + \alpha^k \bar{d}^k) \} = f(\bar{x}^k)$,
      i. Set counter,
      ii. Change $\alpha^k \to \alpha^k/2$,
      iii. If counter is less than 100, go to Step 3.c.,
      iv. Else go to Step 4.
   d. Elseif $\min\{ f(\bar{x}^k), f(\bar{x}^k + (\alpha^k/2)\bar{d}^k), f(\bar{x}^k + \alpha^k \bar{d}^k) \} = f(\bar{x}^k + (\alpha^k/2)\bar{d}^k)$,
      i. Calculate $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} (\alpha^k)^2 & \alpha^k & 1 \\ (\alpha^k/2)^2 & \alpha^k/2 & 1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} f(\bar{x}^k + \alpha^k \bar{d}^k) \\ f(\bar{x}^k + (\alpha^k/2)\bar{d}^k) \\ f(\bar{x}^k) \end{bmatrix}$,
      ii. Change $\alpha^k \to -b/2a$,
   e. Elseif $\min\{ f(\bar{x}^k), f(\bar{x}^k + (\alpha^k/2)\bar{d}^k), f(\bar{x}^k + \alpha^k \bar{d}^k) \} = f(\bar{x}^k + \alpha^k \bar{d}^k)$,
      i. Set counter,
      ii. Change $\alpha^k \to 2\alpha^k$,
      iii. If counter is less than 100, go to Step 3.e.,
      iv. Else go to Step 4.

4. Set $(\bar{x}^k + \alpha^k \bar{d}^k) \rightarrow \bar{x}^{k+1}$, $k \rightarrow k+1$.

5. Go to Step 2.

Consequently, the model learns how to use the dependent variables optimally with the steepest descent algorithm with one dimensional search.

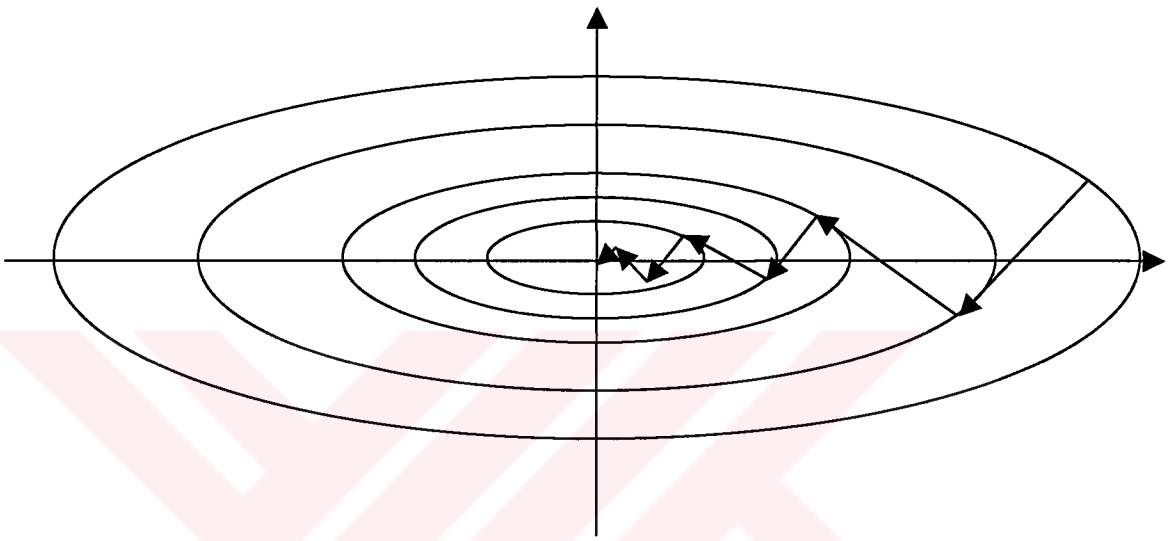


Figure 4.5. Convergence of the Steepest Descent Method

As seen in the figure above, the arrows are tangent to the inner circles when the gradient is equal to zero. The length of the arrows is calculated by the one dimensional search as α^k.

4.6.2. Results of the Optimization

By minimizing the square of the error, which is the difference of real stock market returns from estimated stock market returns, the inner five border points of the membership functions are changed. Firstly, this process is done for the last three variables (technical analysis, fundamental analysis and macroeconomy) and the output.

Secondly, the membership functions of the 12 input variables are changed according to the same algorithm. Lastly, the same process is done to the last three variables and the output again. Finally, the new model is a little different from the first model with membership functions of the 12 input variables, the last 3 variables and the output.

The learning of the model in Pentium IV computer after 12 hours and 18 minutes is with %11,8 R square value as seen in the following table, which is the sum of the squares of the difference of the real returns from the estimated returns divided by the real returns of the stock market. The result of the optimization will end with an improvement more than twice the first R square result in the modelling of the ISE100 Index. Since the fluctuation in stock market is significant, the estimation for peak points is not successfully estimated. However, the correlation analysis will be different as in Figure 4.7 and the model seems to be more successful with 34.4% correlation.

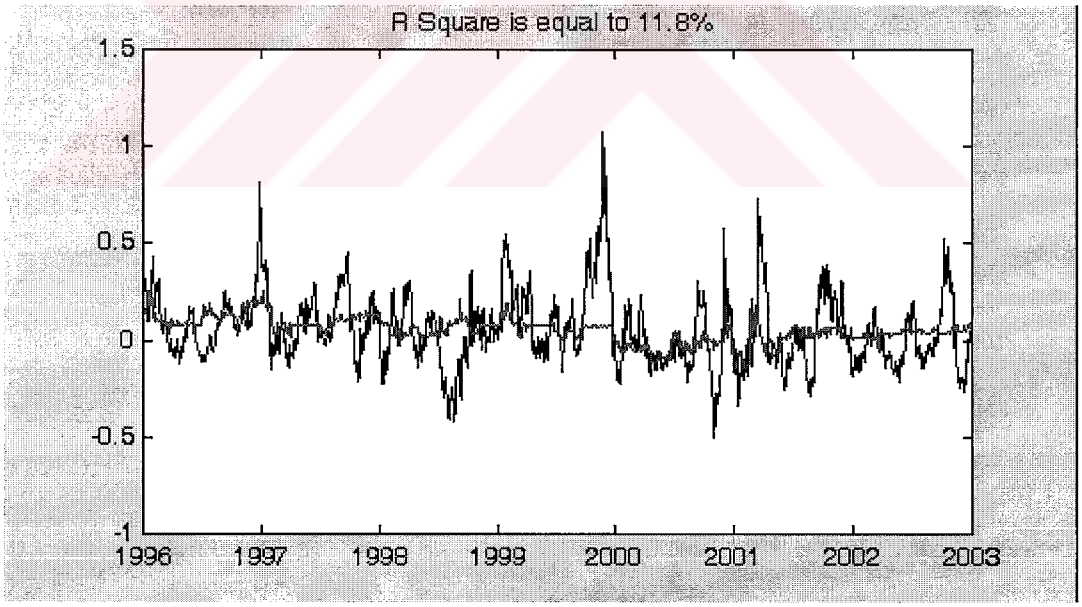


Figure 4.6. Actual and Estimated Output After Learning

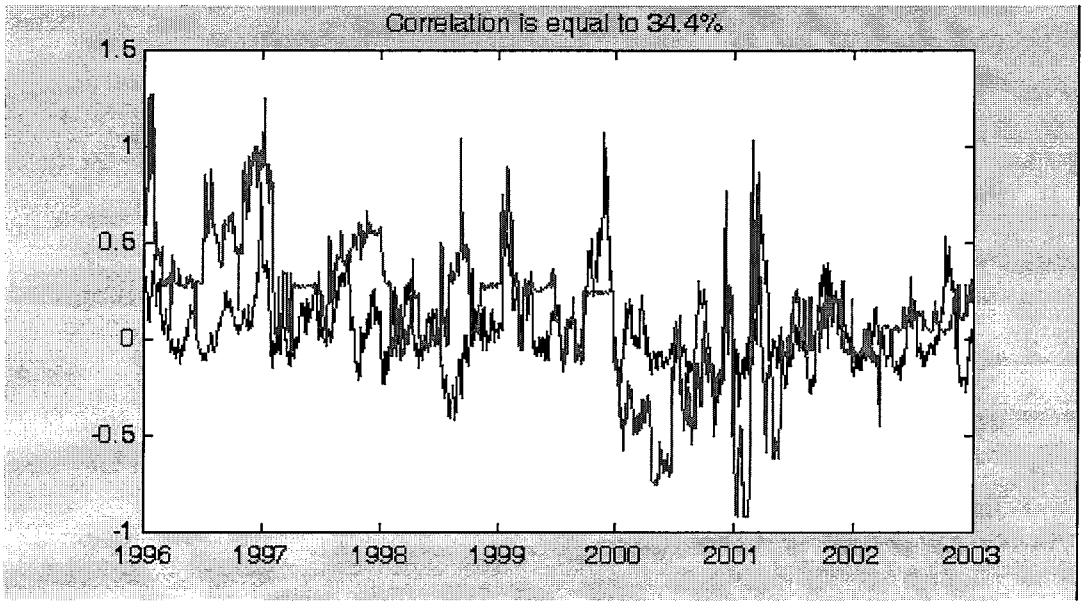


Figure 4.7. Correlation of Actual and Estimated Output

The membership of the output variable and the final intermediate variables are changed after learning process with the steepest descent algorithm. In the following four figures, the membership functions of the fundamental analysis, the technical analysis, the macroeconomy and the estimated output variables are demonstrated.

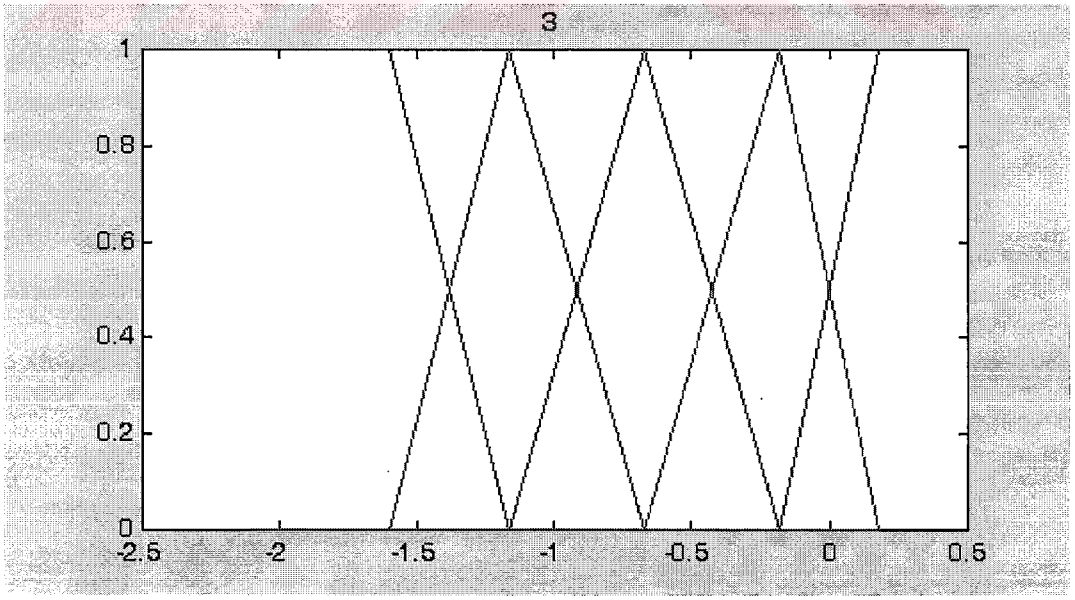


Figure 4.8. Membership Functions of the Fundamental Analysis Variable

When the membership of the fundamental analysis variable is analyzed, it is seen that the very good and very bad memberships' borders are enlarged, while the middle memberships contract, which means that the variable is sensitive.

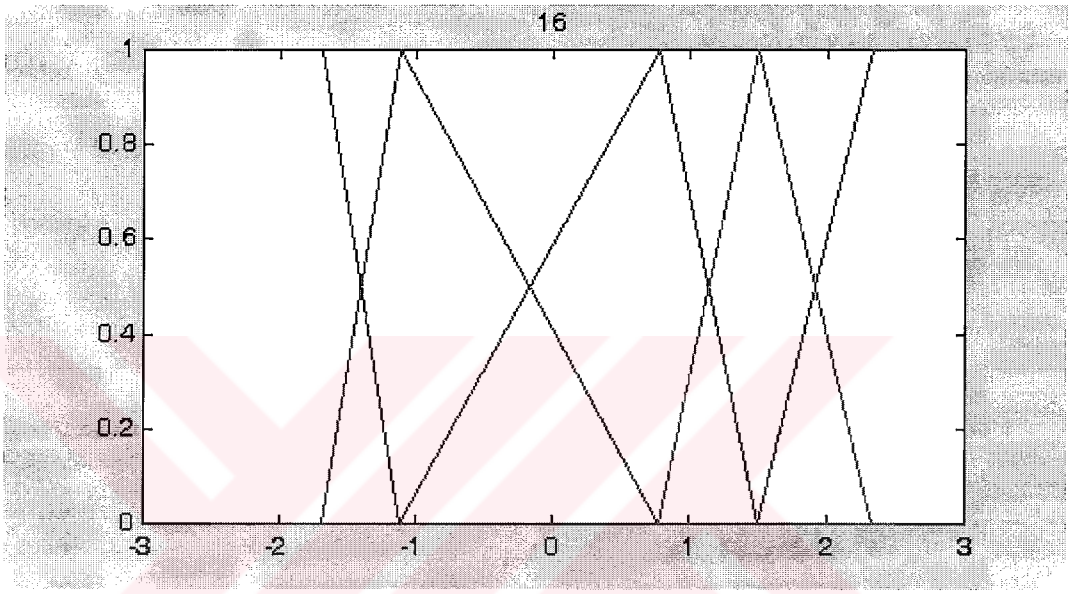


Figure 4.9. Membership Functions of the Technical Analysis Variable

In the memberships of the technical analysis variable, the bad and the neutral memberships are not symmetric and their borders are enlarged in the optimization period as result of insensitivity for these memberships.

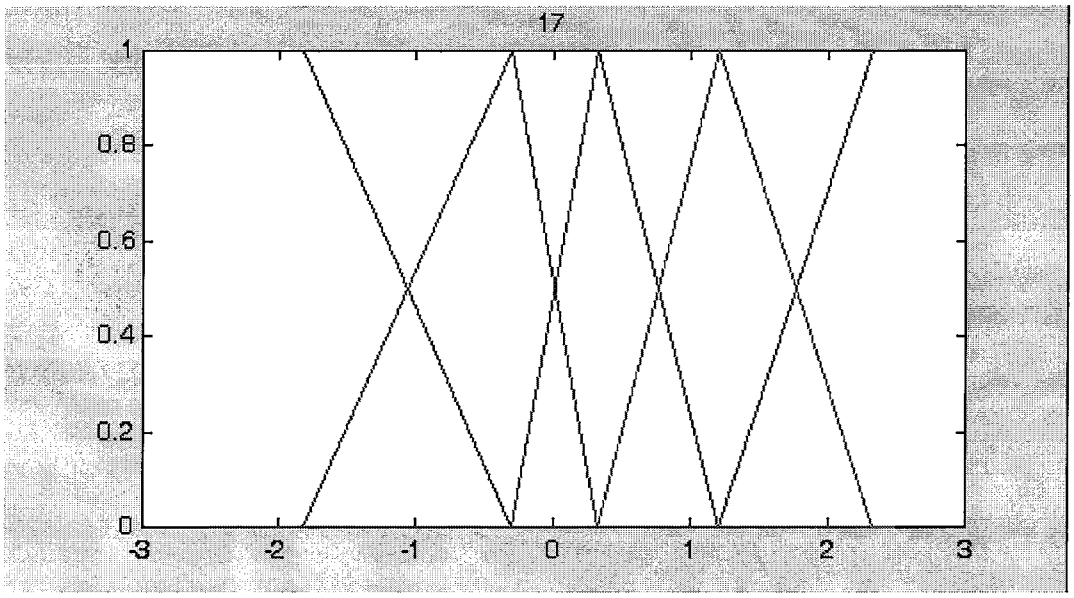


Figure 4.10. Membership Functions of the Macroeconomics Variable

The borders for the very bad membership of the macroeconomics variable are enlarged and the bad membership is not symmetric.

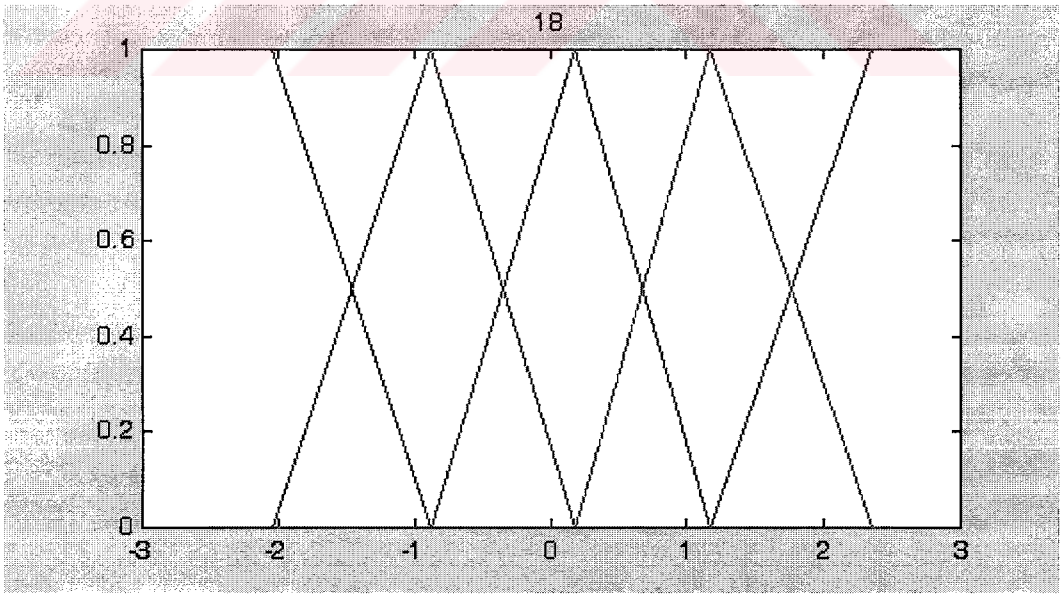


Figure 4.11. Membership Functions of the Output (Monthly Return of the ISE100)

The memberships for the output do not change so much. There is fine tuning for the output memberships at the borders since the membership functions of input and intermediate variables are changed.

4.6.3. Testing the Model

It will be better to compare these models with the data that are not used in the optimization process. By using the data in the period between 2003 and 2004, the accuracy of the model is better than the model development period with 17.6% R square as seen in the following figure.

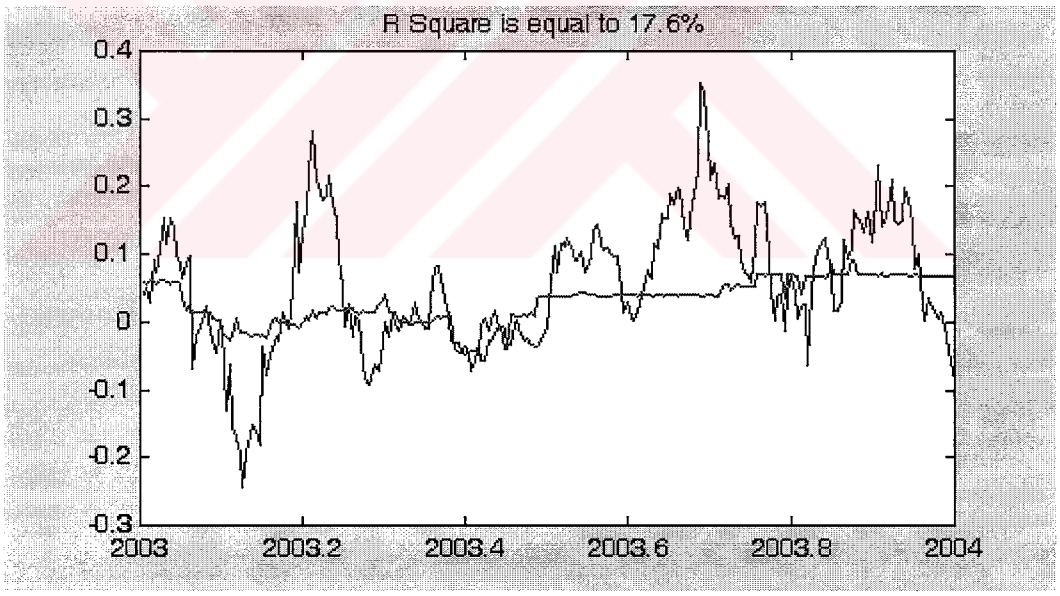


Figure 4.12. Actual and Estimated Output (Overlapping)

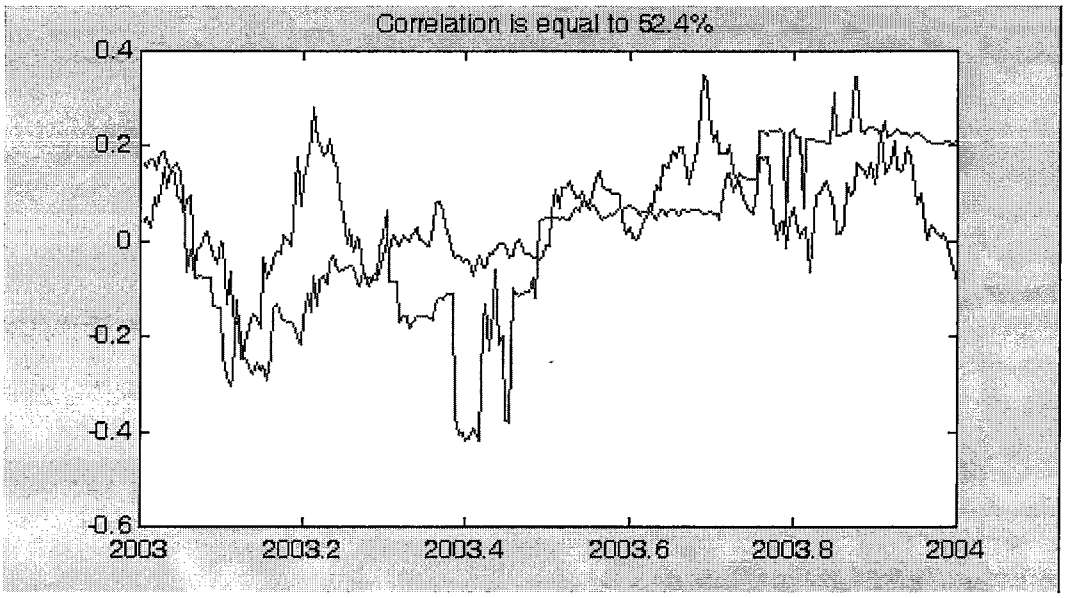


Figure 4.13. Correlations of Actual and Estimated Output (Overlapping)

As seen in the figure above, the correlation of the actual and estimated output is considerably high in terms of stock market results in the testing period. In the learning period it was 34.4% in Figure 4.7. Since the data is overlapping, the monthly frequency of the graphics is plotted below. The accuracy is improved to 18.7% R square value, as seen in Figure 4.14.

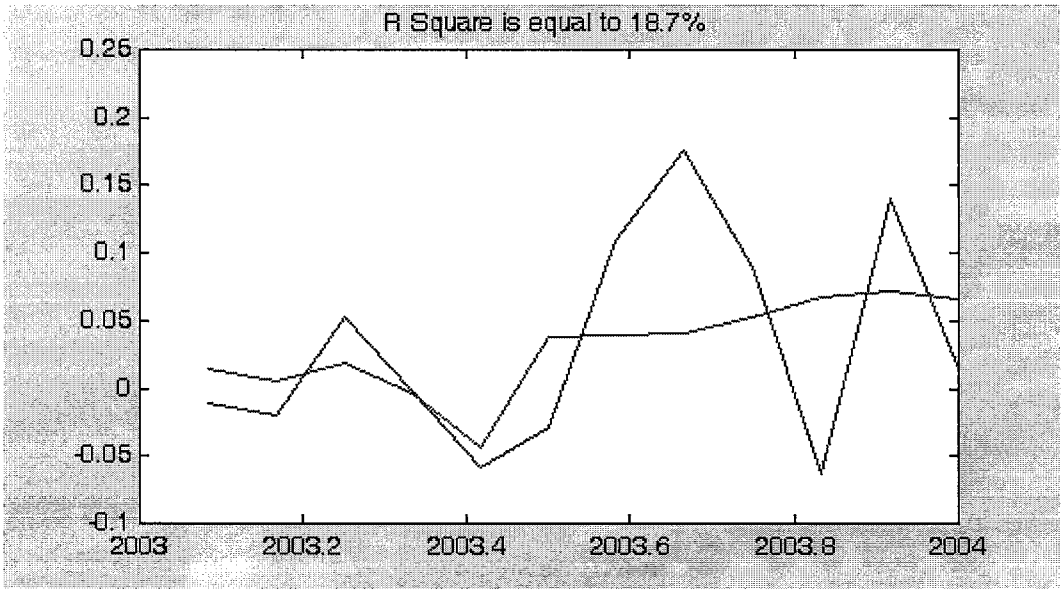


Figure 4.14. Actual and Estimated Output (Monthly Frequency)

Finally, there are around 20-21 working days in a month and 20 different results can be calculated as in the figure above. By using bootstrap re-sampling method, these 20 different outputs can be used to analyze the results of the model statistically. This topic will be analyzed in the next section.

4.7 PERFORMANCE EXPECTATIONS

Performance expectations for a financial forecasting application depend on one's viewpoint of the underlying market dynamics of the target market. Since the fuzzy logic model is designed to forecast a completely random time series, the ISE100 Index, then it should not be surprising if large prediction errors occur, since, by definition, such a time series is unpredictable. While both stochastic and deterministic forces are driving this kind of markets, only some portion of the market is predictable. It can be said that the maximum achievable accuracy of financial forecasting is unknown. It is unrealistic to expect to achieve zero error, since this would require a model that accounts for events that are today deemed random. Nevertheless, events are currently unpredictable does not mean that they are necessarily random. For the time being, it is unknown where the "ceiling" on performance accuracy is located. Successful fuzzy logic applications for nontrivial problems require considerable talent and expertise in a number of disciplines, including market analysis, computer science, and applied mathematics.

By the early 1970's a consensus had emerged among financial economists suggesting that stock prices could be well approximated by a random walk model and that changes stock returns were basically unpredictable starting with the analysis of

Fama (1970). Recently, a large number of studies in the finance literature have confirmed that stock returns can be predicted to some degree by means of interest rates, dividend yields and a variety of macroeconomic variables exhibiting business cycle variations, while the vast majority of these literature, which includes Campbell (1987), Fama and French (1988, 1989), Balvers, Cosimano and MacDonald (1990), Breen, Glosten and Jagannathan (1990), Ferson and Harvey (1993), Pesaran and Timmermann (1994,1995) and Campbell, Chan and Viceira (2003) have looked at the US stock market. Granger (1992) made a survey of the methods and results in the literature. Predicting stock returns is difficult and the R square results of the models are changed between 1% to 11% R square value, which depends on the frequency of the data. Since the estimation of the daily data is harder than the estimation of weekly, quarterly or yearly data.

As a good example for the modelling and testing the goodness of fit, Keim and Stambaugh (1986) constructs a linear model for the return prediction of a chosen stock portfolio in the New York Stock Exchange. Their sample consists of monthly data during the period January 1928 through November 1978. In their studies, the R-square values are ranging from 0.1% to 1.6%, depending on their equations. Campbell (1987) also tests another model for the return prediction of a chosen stock portfolio with 11.2% R-square value during the period May 1959 through August 1979 in the New York Stock Exchange. In similar linear models, Harvey (1989) and Ferson and Harley (1991) predict the returns of a chosen stock portfolio in the New York Stock Exchange. Finally, Campbell, Chan and Viceira (2003) develops a multivariate model by using VAR estimation. The R-square results of the models for predicting excess stock return during the second quarter of 1952 through the fourth quarter of 1999 in both the quarterly and the annual sample are 8.6% and 5.0%, respectively.

As also explained in the previous sections, the in-sample R square result of the fuzzy logic model is 11.8% and the out-sample R square result of the model is 17.6%. The results can be comparable with the findings in the literature. Besides, the Istanbul Stock Exchange one of the stock markets with the highest degree of volatility in the world. Because of this volatility, the modelling accuracy can be distorted.

Furthermore, the data used and tested for the modelling the ISE100 Index covers three main crises and other political and financial crisis during the period. After the devaluation in Russia on August 17th, the Turkish economy was also hit due to its strong financial relations with Russia. The bankruptcy of Demirbank in November 2000 and the currency crisis of the Turkish Economy in April 2001 can be considered as the other two major crises during the period. To provide examples for other political and financial crises, the Brazilian devaluation and Korkmaz Yigit crises are worth noting. As a result, not only the stock market movements but also the macroeconomic fundamentals are very volatile in Turkey. Being one of the emerging markets in the world, the forecast of the stock market may be possible since the market is not efficient. While comparing the accuracy of the models, it should be kept in mind that it is harder to design models for emerging markets than for developed markets due to the significant volatility difference.

As for the first step, the model can be compared by the simple linear regression in the next section. Then, the model will be tested by the bootstrap re-sampling method. The original data is overlapping and by using these data, the statistical distribution of the results can be calculated, whereafter the results will be statistically analyzed.

4.7.1. Comparing the Results with Ordinary Least Square Method

In the fuzzy logic case, the model is constructed by using the data between 1996 and 2003. After the development of the model in the learning section, the data in the period between 2003 and 2004 are used to test how well the model is working. To make a comparative analysis, the same method can be used to produce the results of the ordinary least square method.

As a result, the data for the period between 1996 and 2003 is regressed. The coefficients, which produced the model by using OLS method, are used to estimate the returns of the ISE100 Index in the period between 2003 and 2004. In the analysis, the residual sum of squares, the square of the difference of the actual and estimated output is greater than total sum of squares, the square of the difference of the real output and the average of the real output. Consequently, the alternative model in the linear regression case fails because the prediction of the stock market with linear models has disadvantages in explaining the chaotic variables.

4.7.2. Testing the Results with Bootstrap Method

To test the performance of the model, it will be particularly useful in helping to determine entry and exit times for periodic trading or position holding by using the results of the output (which the monthly returns of the ISE100 Index). If the forecasted information suggests that the next period will be an up in the market, investors enter a long position. The reverse would involve to stay in TL position or to sell the stocks. The profitability of those trades that are executed should be substantial. Besides, there should be four different types of benchmarks to compare the results of the model; TL benchmark, USD benchmark, Repo benchmark and Stock Market Index benchmark. Thus, the corresponding benchmarks are based on the returns of the USD, based on the

returns of the repo and also on the returns of the ISE100 Index. If the benchmarks are achieved successfully, it can be said that the model works good enough. To justify how well the model is working statistically, the distribution of the data should be known. By using bootstrap re-sampling method, statistical tables can be constructed and used for testing the model's suggestion.

Consequently, if the model estimates a positive return for the month ahead, the investor will prefer to use the model to fully invest in the ISE100 Index. Conversely, if the model estimates a negative return for the month ahead, the investor will prefer to fully invest in repo. After the end of the one-year period (from the beginning of 2003 to the beginning of 2004) the investor will have reached his new portfolio amount. Thus, the three hypotheses to be tested will be:

- Hypothesis 1: Does the difference between investing in the model and fully investing in the ISE100 Index yield a positive return?
- Hypothesis 2: Does the difference between investing in the model and fully investing in USD yield a positive return?
- Hypothesis 3: Does the difference between investing in the model and fully investing in Repo yield a positive return?
- Hypothesis 4: Does investing in the model yield a positive return?

These hypotheses are tested with the results of the data in the period between 2003 and 2004. Since the model is built in the period between 1996 and 2003, the model and the data for the testing period are independent from each other. Furthermore, the data is overlapping. Since there are 21 working days in a month, there exists a set of 21 series to test for the four hypotheses above.

As a result, we construct a consecutive 12-month series (estimated returns are in monthly frequency) and compute the difference of the estimated monthly returns from the benchmark returns (0 for TL benchmark, monthly USD return for the USD benchmark, monthly repo rates for repo benchmark and monthly ISE100 Index return for the ISE100 benchmark). The average and standard deviation of the returns for hypothesis i are then calculated as μ_i and σ_i, respectively. We aim to test whether the true mean $\hat{\mu}_i$ of returns in hypothesis i is zero. The test cannot be immediately performed since the actual test statistic $Z_{\mu_i} = (\mu_i - \hat{\mu}_i)/\sigma_i$ does not have a known distribution with tables. So, we try to approximate the distribution of Z_{μ_i}. By bootstrap re-sampling method, we simulate 20,000 values of pivotal test statistic $Z_{\mu_i^B} = (\mu_i^B - \hat{\mu}_i)/\sigma_i^B$ where μ_i^B and σ_i^B are the average and the standard deviation of bootstrap samples of returns, respectively. By using the definition of quantiles, we then construct $1-\alpha$ confidence interval $C_{i,1-\alpha}$ for the pivotal test statistic. The null hypothesis of zero average return in class i is accepted with $1-\alpha$ confidence if the actual test statistic Z_{μ_i} lies within the interval $C_{i,1-\alpha}$.

The algorithm of the bootstrap re-sampling method is described below: (All computations are performed with the software Matlab 5.3.)

1. For $i = 1,2,3,4$, the data for each hypothesis is recorded. Since monthly returns are tested and there are 21 working days in a month, the total number of data for each hypothesis is denoted by 21. The data (different of the returns according to their benchmarks) for hypothesis i is indexed by $j = 1,2,...,21$ and denoted by $X_{i,j}$.

2. For $i=1,2,3,4$, $j=1,2,...,21$ and observation $X_{i,j}$ for hypothesis i, the consecutive 12-month series $S_{i,j}=(X_{i,j},X_{i,j}+1,...,X_{i,j}+11)$ is constructed (there is one year investment period).
3. For $i=1,2,3,4$, and $j=1,2,...,21$, the mean return $\mu_{i,j}$ of the series $S_{i,j}$ is calculated. The set $M_i=\{\mu_{i,1},\mu_{i,2},...,\mu_{i,21}\}$ is recorded.
4. For $i=1,2,3,4$, the average μ_i and the standard deviation σ_i of the set of returns M_i are calculated.
5. For $i=1,2,3,4$, the actual test statistic $Z_{\mu_i}=(\mu_i-\hat{\mu}_i)/\sigma_i$ for the postulated true average return $\hat{\mu}_i=0$ is calculated.
6. For $i=1,2,3,4$, some 20,000 samples denoted by $(B_{i,1},B_{i,2},...,B_{i,20000})$ of size *ceiling*[21/4] (i.e. the highest integer less than or equal to *ceiling*[21/4]=6) are sequentially drawn from the set M_i with replacement.
7. For $i=1,2,3,4$, and $k=1,2,...,20000$, the average $\mu_{i,k}^B$ and the standard deviation $\sigma_{i,k}^B$ of the returns in the bootstrap sample $B_{i,k}$ are calculated.
8. For $i=1,2,3,4$, and $k=1,2,...,20000$, the pivotal statistic $Z_{\mu_{i,k}^B}=(\mu_{i,k}^B-\mu_i)/\sigma_{i,k}^B$ is calculated.
9. For $i=1,2,3,4$, $(1-\alpha)$ confidence interval $C_{i,1-\alpha}=[z_{i,(\alpha/2)},z_{i,(1-\alpha/2)}]$ for $Z_{\mu_{i,k}^B}$'s is obtained noting that Prob($z_{i,(\alpha/2)}<Z_{\mu_{i,k}^B}<z_{i,(1-\alpha/2)})=1-\alpha$.
10. For $i=1,2,3,4$, the hypothesis $H_0:\hat{\mu}_i=0$ is accepted with $(1-\alpha)$ confidence if (and only if) $Z_{\mu_i}\in\mu_{i,1-\alpha}$.

Table 4.8. Statistical Results of the Hypothesis

Hyp.	μ_i	σ_{μ_i}	Z_{μ_i}	$C_{i,0.98}$	$C_{i,0.95}$	$C_{i,0.90}$	$C_{i,0.80}$	p	H_0
1	0.1322	0.094	1.41	[-1.51,1.32]	[-0.91,0.92]	[-0.62,0.67]	[-0.35,0.42]	0.98	NO
2	1.4132	0.156	9.04	[-1.68,1.12]	[-1.09,0.82]	[-0.75,0.61]	[-0.41,0.40]	1.00	NO
3	0.5227	0.127	4.10	[-1.74,1.09]	[-1.17,0.81]	[-0.77,0.60]	[-0.40,0.40]	0.99	NO
4	1.0046	0.149	6.76	[-1.71,1.09]	[-1.15,0.81]	[-0.78,0.60]	[-0.41,0.40]	0.99	NO

Consequently, the answers to the four hypotheses are positive and in favour of the fuzzy logic model. It is statistically proved that the suggestion of the model has positive returns in USD terms and higher returns than the ISE100 Index and the return of the repo by using bootstrap re-sampling method. However, we also state that the testing period of one year is not enough to feel confident on this out of sample results. In other words, the model is built by using the past data and there is no guarantee for the future performances of the model. Additionally, there is an inevitable risk of underestimation and overestimation in all forecasting models. The main goal of this study is to try a new forecast technique and then to compare with other mathematical methods. Other than this academic motivation, we have no intention to obtain practical and strategic results for investors in the Istanbul Stock Exchange.

CHAPTER 5. CONCLUSION & FUTURE WORK

To guess what a man will do may not be possible without significant deviations, but to estimate what a lot of men will do is quite possible and is the area of science. In order to forecast an index in a stock market, data set should be enlarged as much as possible. The return of the index represents different types of data such as: macroeconomy, balance sheets and income statements of the companies quoted in the index, technical analysis of the index etc. Meanwhile, modelling dynamical environments needs improvement in learning and dynamic ability of the algorithms.

This study implements the rule based fuzzy logic modelling in the Istanbul Stock Exchange, our nonlinear dynamical environment. To represent the Istanbul Stock Exchange, the most extensive index ISE100 is used. Because of the methodology of the fuzzy logic, lots of rules are needed to construct the dependent variables of the model. History has shown that numerous economic crises in the world have identifiable effects on the behaviour of financial markets. Indeed, it has been long highlighted that Turkey is a textbook example of a very volatile economy; in other words, unanticipated economic and political events leading to financial crises are frequently observed in Turkey. Building some rules for the nonlinear models in these kinds of volatile financial markets should be more indicative to gauge where the market is heading when compared with linear models. Before developing the rules, the model is divided into three main parts: the technical analysis, the fundamental analysis and the macroeconomy to forecast the monthly return of the ISE100 Index.

Recent studies show that active traders may underperform those who trade less. Such a result is partly due to the persistent high annual net returns earned by the well-

diversified portfolios in the examined stock markets. For the technical analysis part, first an index level observation by using classifier systems is found and then used for the rule based modelling of the fuzzy logic model. The inverse correlation of the level of the ISE100 Index in USD terms and the daily returns are expressed as an observation in this context. This result shows that under the prescriptions of a classifier system buy-and-hold strategy may yield huge net returns even in an extremely volatile and horizontal market such as the Istanbul Stock Exchange. This observation is statistically tested and found to be significant by using bootstrap re-sampling method. Besides, a basic technical analysis rule using moving average is also modified as a short period input for the fuzzy logic model.

For the fundamental analysis, a factor analysis is done with 81 financial ratios of the stocks and then by using Altman's famous bankruptcy prediction method, a model for predicting next period's stock return is developed with the discriminant analysis and the logistic regression. The results are statistically significant so that "good" companies perform better than "bad" companies in the ANOVA analysis. Then, the weighted average of the z score or the logit score of the stock exchange companies forecast the next period return of the index for the stock market. In case of restrictive properties of the discriminant analysis for the distributions of the variables, the result of the logistic regression analysis of the model is used as a financial data input for the final rule based modelling.

The other input of the model, macroeconomic data in Turkey is gathered in three main groups: real economy, FX market and TL market. The data is ruled and modelled by using fuzzy logic. Furthermore, the final model by using technical analysis, stocks' financial analysis and macroeconomic analysis is constructed within the period from the beginning of 1996 to the beginning of 2003 and optimized to

perform better with steepest descent learning algorithm. The R-Square of the model in other words, the ratio of the explained sum of squares to the total sum of squares, to estimate the monthly return of the ISE100 Index is calculated as 11.8%, which is comparable in literature. The reasons for better results are: a) the ISE is not a developed and an efficient market; thus, predicting future prices is easier than is the case with other developed markets; b) rule based fuzzy logic modelling with a large enough data set improves the explanation of the future movements.

To test the model, extra data is needed in addition to the data used for the model building. Then, the model is tested by the data in the period from the beginning of 2003 to the beginning of 2004. The R-Square of the model or the ratio of the explained sum of squares to the total sum of squares to estimate the monthly return of the ISE100 Index is calculated as 17.6%, which is above the results of the data used for the development and optimization of the model. The data is overlapping, in other words the model uses the daily frequency of the data but forecasts the monthly returns of the ISE100 Index. If the first days of every month in 2003, the testing period is taken as the starting point of the one-month period and the estimation results are used to calculate how successful the model is. The ratio of the explained sum of squares to the total sum of squares for the model will be calculated as 18.7% in monthly frequency, which is also above the results of the data.

To make a comparative analysis for the fuzzy logic analysis, the same method can be used for reaching the results of the ordinary least square method. Then, the data for the period from the beginning of 1996 to the beginning of 2003 is regressed. The coefficients, which produced the model by using OLS method, are used to estimate the returns of the ISE100 Index in the period from the beginning of 2003 to the beginning of 2004. The residual sum of squares is greater than the total sum of squares. Roughly

speaking, the absolute difference of the actual and estimated output is greater than the absolute value of the real output, which means the alternative model with the linear regression case fails.

Consequently, the model is used to achieve an optimal investment decision in the period from the beginning of 2003 to the beginning of 2004. The testing algorithm of the model for the investor is as follows: If the next month's predicted return of the index is positive, the customer will decide to invest all of the money in the stock market on the first day of the every one-month period in 2003. Otherwise, the investor will decide to wait or to invest in repo. Since, there are 21 working days in a month and 12 months in year, there are 21 12-monthly distinct series for the suggestion of the model. By using these series, the performance of the model can be tested statistically by using bootstrap re-sampling method. As a result, four different hypotheses are produced:

- Hypothesis 1: Does the difference between investing in the model and fully investing in the ISE100 Index yield a positive return?
- Hypothesis 2: Does the difference between investing in the model and fully investing in USD yield a positive return?
- Hypothesis 3: Does the difference between investing in the model and fully investing in Repo yield a positive return?
- Hypothesis 4: Does investing in the model yield a positive return?

The benchmarks: the return of the ISE100 Index, the return of the USD and the return of the repo market are calculated according to the model suggestion series format. The return of the suggested model is tested under the hypotheses written above by using bootstrap re-sampling method. The result is hopefully good and the suggestion of the model performs statistically better than the benchmarks for the one-year period in

2003. However, we also state that the testing period of one year is not enough to feel confident on this out of sample results. The model is built by using the past data and there is no guarantee for the future performances of the model. Inevitably, in all prediction models there is a risk of underestimation and overestimation. In order to minimize forecast error, all available information as to the predicted variables should be exploited with best mathematical tools and/or econometric models. Thus, fuzzy logic model is a useful tool in this aspect. We use this technique in order to forecast the monthly return of the ISE100 Index. Definitely, our aim is not to give a clue or a hint for the future behaviour of the stock market index. The aim in this modelling technique is solely to try this new methodology and to compare with other mathematical methods.

Since all the stock markets in the world are very volatile and the Istanbul Stock Exchange has one of the highest fluctuations, modelling these extreme increases or decreases is hardly possible. In spite of this difficulty, the fuzzy logic models can be developed and new rules can be added to the previous models. As for the future work, the rules for the fuzzy models may not be enough to explain the real outputs. There may be some other rules to explain the behaviours more explicitly and the models can be developed and can be redesigned better with new rules. Similarly, there may be some new variables to better explain the characteristics of the behaviours. In modelling of this study, there are no lagged variables of the inputs. There may be one or more period lagged variables to extend the models' modelling perspectives. Besides, the data set can be enlarged by the introduction of some new variables such as population growth or GNP per capita. Additionally, incorporating "combining forecast" method with fuzzy logic may yield profitable results in the framework of chaotic time series forecasting. As in Fiordaliso (1998), the different forecasting tools can be combined to perform better results with fuzzy logic.

REFERENCES

Akgiray, V. (1989). Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts. The Journal of Business. Vol: 62/1. pp: 55-80.

Aharony, J., Jones, C.P., and Swary, I. (1980). An Analysis of Risk and Return Characteristics of Corporate Bankruptcy Using Capital Market Data. The Journal of Finance. Vol: 35. September. pp: 1001-1016.

Aksoy, H., and Sağlam, I. (2001). Sınıflayıcı (Classifier) Sistem ile İMKB'de Yeni Bir Anomali Gözlemi. Working Paper. Department of Economics. Boğaziçi University. No: 15.

Aksoy, H. (2003). A Logistic Regression Analysis in the Istanbul Stock Exchange. Unpublished Working Paper.

Altman, E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. The Journal of Finance. Vol: 23. September. pp: 589-609.

Altman, E.I., and Eisenbeis, R. (1976). Financial Applications of Discriminant Analysis: A Clarification. Working Paper. No: 79. New York University.

Altman, E.I., Haldeman, R., and Narayanan. P. (1977). ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations. Journal of Banking and Finance. Vol: 1. June. pp: 29-54.

Altman, E.I. (1984). A Further Empirical Investigation of the Bankruptcy Cost Question. Journal of Finance. Vol: 39. September. pp: 1067-89.

Altman, E. (2000). Predicting financial distress of companies: Revisiting the Z-score and Zeta Models. Working Paper. Stern School of Business. New York University.

Arthur, W.B., Holand, J.H., LeBaron, B., Palmer, R., and Tayler, P. (1997). Asset Pricing under Endogenous Expectations in an Artificial Stock Market. The Economy as an Evolving Complex System II. Edited By W.B. Arthur, S.N. Durlauf and D.A. Lane. Addison-Wesley. pp: 15-40.

Aydoğan, K., and Güney, A. (1997). P/E Ratio and Dividend Yield as Forecasting Tools in the ISE. Istanbul Stock Exchange Review. Vol: 1/1. pp: 83-96.

Back, B., Laitinen, S.T., and Wezel, M. (1996). Choosing Bankruptcy Predictors Using Discriminant Analysis, Logit Analysis and Genetic Algorithms. Tech. Report. Turcu Center for Computer Science. No: 40.

Bak, P., Paczuski, M., and Shubik, M. (1997). Price Variations in a Stock Market with Many Agents. Physica A. Vol: 246. pp: 430-453.

Balvers, R.J., Cosimono, T.F., and MacDonald, B. (1990). Predicting Stock Returns in an Efficient Market. Journal of Finance. Vol: 45. pp: 1109-1128.

Basu, S. (1977). Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of Efficient Market Hypothesis. The Journal of Finance. Vol: 32. No: 2. pp: 663-682.

Beaver, W. (1966). Financial Ratios as Predictors of Failure. Empirical Research in Accounting, Journal of Accounting Research. Vol: 4. pp: 71-111.

Black, F. (1988). A Equilibrium Model of the Crash. National Bureau of Economic Research Macroeconomics Annual. pp: 269-275.

Blum, M. (1974). Failing Company Discriminant Analysis. Journal of Accounting Research. Spring. pp: 1-21.

Breen, W., Glosten, L.R., and Jagannathan, R. (1990). Predictable Variations in Stock Index Returns. Journal of Finance. Vol: 44. pp: 1177-1189.

Brischetto, A., and Voss, G. (1999). A Structural Vector Autoregression Model of Monetary Policy in Australia. Research Discussion Paper. No: 9811, Economic Research Department, Reserve Bank of Australia.

Brown, M., and Harris, C. (1995). Neurofuzzy adaptive modelling and control. Prentice Hall.

Brownstone, D. (1995). Using Percentage Accuracy to measure Neural Network Predictions in Stock Market Movements. Neurocomputing. Vol: 10. pp: 237-250.

Campbell, J.Y. (1987). Stock Returns and the Term Structure. Journal of Financial Economics. Vol: 18. pp: 373-379.

Campbell, J.Y., and Shiller, R.J. (1998). Valuation Ratios and the Long-Run Stock Market Outlook. Journal of Portfolio Management. Vol: 24. pp: 11-26.

Campbell, J.Y., Chan, Y.L., and Viceira, L.M. (2003). A Multivariate Model of Strategic Asset Allocation. Journal of Financial Economics. Vol: 67. pp. 41-80.

Çetiner, I. (1993). Test of Overreaction in the Istanbul Stock Exchange. Unpublished MBA Thesis. Department of Management. Bilkent University. Ankara.

Chen, S.M. (1996). Forecating Enrollments Based on Fuzzy Time Series. Fuzzy Sets and Systems. Vol: 81. pp: 311-319.

Chen, S., and Yeh, C. (2001). Evolving Traders and the Business School with Genetic Programming: A New Architecture of the Agent-based Artificial Stock Market. Journal of Economic Dynamics and Control. Vol: 25. pp: 363-393.

Chenoweth, T., and Obradovic, Z. (1996). A Multi-component Nonlinear Prediction System for the S&P 500 Index. Neurocomputing. Vol: 10. pp: 275-290.

Cremers, K.J.M. (2002). Stock Return Predictability: A Bayesian Model Selection Perspective. The Review of Financial Studies. Vol: 15. No: 4.

Conrad, J., and Kaul, G. (1988). Time-Variation in Expected Returns. Journal of Business. Vol: 61. pp: 409-425.

Deakin, E. (1972). A Discriminant Analysis of Predictors of Business Failure. Journal of Accounting Research. Spring. pp: 167-179.

De Silva, C.W. (1995). Intelligent control: Fuzzy Logic Applications. CRC.

Economakos, E. (1979). Application of Fuzzy Concepts to Power Demand Forecasting. IEEE Transactions on Systems, Man and Cybernetics. Vol: 9. pp: 651-657.

Edey, M., and Romalis, J. (1996). Issues in Modelling Monetary Policy, Economic Research Discussion Paper. No: 9604. Reserve Bank of Australia.

Fama, E.F. (1965). The Behaviour of Stock Market Prices. Journal of Business. Vol: 38. pp: 34-105.

Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. Journal of Finance. Vol: 25. pp: 338-417.

Fama, E.F. (1981). Stock Returns, Real Activity, Inflation and Money. American Economic Review. Vol: 71. pp: 545-565.

Fama, E.F. (1991). Efficient Capital Markets II. Journal of Finance. Vol: 46. pp: 1575-1617.

Fama, E.F., and French, K.R. (1988). Permanent and Temporary Components of Stock Prices. Journal of Political Economy. Vol: 96. pp: 246-273.

Fama, E.F., and French, K.R. (1989). Business Conditions and Expected Returns on Stocks and Bonds. Journal of Financial Economics. Vol: 25. pp: 23-49.

Fama, E.F., and French, K.R. (1995). Size and Book-to-Market Factors in Earnings and Returns. The Journal of Finance. Vol: L. No: 1. pp: 131-155.

Fama, E.F., and French, K.R. (1996). Multifactor Explanations of Asset Pricing Anomalies. The Journal of Finance. Vol: LI. No: 1. pp: 55-84.

Fama, E.F., and French, K.R. (1998). Value versus Growth: The International Evidence. The Journal of Finance. Vol: LIII. No: 6. pp:1975-1999.

Ferson, W.E., and Harvey, C.R. (1991). Sources of Predictability in Portfolio Returns. Financial Analysts Journal. Vol: 47. pp: 49-56.

Ferson, W.E., and Harvey, C.R. (1993). The Risk and Predictability of International Equity Returns. Review of Financial Studies. Vol: 6. pp: 527-566.

Firth, M. (1979). The Relationship Between Stock Market Returns and Rates of Inflation. The Journal of Finance. Vol: 34. No: 3. pp: 743-749.

Fiordaliso, A. (1998). A Nonlinear Forecasts Combination Method Based on Takagi-Sugeno Fuzzy Systems. International Journal of Forecasting. Vol: 14. pp: 367-379.

Fitzpatrick, P. (1932). A Comparison of the Ratios of Successful Industrial Enterprises with Those of Failed Companies. The Accountants Publishing Company.

French, K.R., and Roll, R. (1986). Stock Return Variances: The Arrival of Information and the Reaction of Traders. Journal of Financial Economics. Vol: 17. pp: 5-26.

Foster, D.F., Smith, T., and Whaley, R.E. (1997) Assessing Goodness-of-fit of Asset Pricing Models: The Distribution of the Maximal R-square. The Journal of Finance. Vol: 52. No: 2. pp: 591-607.

Fukuda, T., and Shibata, T. (1994). Fuzzy-neuro-GA Based Intelligent Robotics. Computational Intelligence Imitating Life. pp: 352-362. New York. IEEE Press.

Gencay, R. (1998). The Predictability of Security Returns with Simple Technical Trading Rules. Journal of Empirical Finance. Vol: 5. pp: 347-359.

Gilbert, A.C. (1999). Marketing Research Methodological Foundations. 7th Edition. Dryden Publishing Company.

Granger, C.W.J. (1992). Forecasting Stock Market Prices: Lessons for Forecasters. International Journal of Forecasting. Vol: 8. pp: 3-13.

Granger, C.W.J., and Terasvirta, T. (1993). Modelling Nonlinear Economic Relationships. Oxford University Press.

Guo, H. (2002). On the Out-of-Sample Predictability of Stock Market Returns. Working Paper. No: 008A. Federal Reserve Bank of St. Louis.

Hair, J., Anderson, R., Tatham, R., and Black, W. (1998). Multivariate Data Analysis. Macmillan Publishing Company.

Hall, R. (2000). The stock market and capital accumulation. Working Paper. No: 7180. National Bureau of Economic Research.

Harvey, C.R. (1989). Time Varying Conditional Covariances in Tests of Asset Pricing Models. Journal of Financial Economics. Vol: 24. pp: 289-317.

Hasbrouck, J. (1984). Stock Returns, Inflation, and Economic Activity: The Survey Evidence. The Journal of Finance. Vol: XXXIX. No: 5.

Hellström, T., and Holmström, K. (1998). Predicting the Stock Market. Technical Report. No: IMa-TOM-1997-07.

Hosmer, D.W., and Lemeshow, S. (1989). Applied Logistic Regression. New York. Wiley.

Hsieh, D.A. (1991). Chaos and Nonlinear Dynamics: Application to Financial Markets. The Journal of Finance. Vol: XLVI. No: 5.

Jegadeesh, N. (1990). Evidence of Predictable Behaviour of Security Returns. Journal of Finance. Vol: 45. pp: 881-898.

Jegadeesh, N., and Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. Journal of Finance. Vol: 48. pp:65-91.

Joy, O.M., and Tollefson, J.O. (1975). On the Financial Applications of Discriminant Analysis. Journal of Financial and Quantitative Analysis. Vol: 10. December. pp: 723-739.

Kandel, S., and Stambaugh, R.F. (1996). On the Predictability of Stock Returns: An Asset Allocation Perspective. The Journal of Finance. Vol: LI. No: 2.

Keim, D.B., and Stambaugh, R.F. (1986). Predicting Returns in the Stock and Bond Markets. Journal of Financial Economics. Vol: 17. pp: 357-390.

Keynes, J.M. (1936). The General Theory of Employment, Interest, and Money. Harcourt. Brace. New York. Chapter 12.

Kimoto, T., and Asakawa, K. (1990). Stock Market Prediction System with Modular Neural Network Approach. Proc. IJCNN. Vol: 1. pp: 215-221.

Komo, D., Chang, C.I., and Ko, H. (1994). Stock Market Index Prediction Using Neural Networks. Applications of Artificial Neural Networks. Vol: V. pp: 516-526.

Kooths, S. (1999). Modelling Rule and Experience Based Expectations Using Neuro Fuzzy Systems. Paper in the 5th International Conference of the Society for Computational Economics. Boston. USA.

LeBaron, B., Arthur, W.B., and Palmer, R. (1999). Time Series Properties of an Artificial Stock Market. Journal of Economic Dynamics and Control. Vol: 23. pp: 1487-1516.

Lederman, J., and Klein, R.A. (1995). Virtual Trading: How Any Trader with a PC can Use the Power of Neural Nets and Expert Systems to Boost Trading Profits. Probus.

Lee, C.-Y., and Antonsson, E.K. (2000). A Pareto Optimal Approach to Takagi-Sugeno Fuzzy System Synthesis. Unpublished Working Paper. Caltech. USA.

Leung, M.T., Daouk, H., and Chen, A. (2000). Forecasting Stock Indices: A Comparison of Classification and Level Estimation Models. International Journal of Forecasting. Vol: 16. pp: 173-190.

Levy, H. (1978). Equilibrium in an Imperfect Market: A Constraint on the Number of Securities in the Portfolio. American Economic Review. Vol: 68. September.

Lin, C.T. (1994). Neural Fuzzy Control Systems with Structure and Parameter Learning. World Scientific.

Lin, C.T., and Lee, C.S.G. (1996). Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems. Prentice Hall.

Linn, S.C., and Tay, N.S.P. (2001). Fuzzy Inductive Reasoning, Expectation Formation, and the Behaviour of Security Prices. Journal of Economic Dynamic and Control. Vol: 25. pp: 321-361.

Lo, A.W., and MacKinlay, A.C. (1988). Stock Market Prices do not Follow Random Walks: Evidence from A Simple Specification Test. Review of Financial Studies. Vol: 1. pp: 41-66.

Lo, A.W., Mamaysky, H., and Wang, J. (2000). Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. The Journal of Finance. Vol: LV. No: 4. pp: 1705-1765.

Lo, A.W., and Wang, J. (2001). Trading Volume: Implications of An Intertemporal Capital Asset Pricing Model. Working Paper. No: 8565. National Bureau of Economic Research.

Lo, A.W., Mamaysky, H., and Wang, J. (2001). Asset Prices and Trading Volume Under Fixed Transaction Costs. Unpublished Working Paper. MIT.

Malkiel, B.G. (1996). A Random Walk down Wall Street: Including a Life-Cycle Guide to Personal Investing. W.W. Norton, New York-London.

Malkiel, B.G. (2003). The Efficient Market Hypothesis and its Critics. Journal of Economic Perspectives. Vol: 17. No: 1. pp: 59-82.

Mamdani, E.H. (1974). Application of Fuzzy Algorithms for Simple Dynamic Plant. Proc. IEE. Vol: 121. pp: 1585-1588.

Mamdani, E.H. (1977). Applications of Fuzzy Set Theory to Control Systems: A Survey in Fuzzy Automata and Decision Process. North Holland. New York. pp: 1-13.

McQueen, G., and Thorley, S., (1991). Are Stock Returns Predictable? A Test Using Markov Chains. Journal of Finance. Vol: 46. pp: 239-264.

Metin, K., Muradoğlu, G., and Yazıcı, B. (1997). An Analysis of Day of the Week Effect on the ISE. Istanbul Stock Exchange Review. Vol: 1. pp: 15-27.

Mizumoto, M. (1989). Improvement Methods of Fuzzy Controls. Proceedings of the 3rd IFSA Congress. pp: 60-62.

Modis, T. (1999). Technological Forecasting at the Stock Market. Technological Forecasting and Social Change. Vol: 62. pp: 173-202.

Muradoğlu, G., and Oktay, T. (1993). Türk Hisse Senedi Piyasasında Zayıf Etkinlik: Takvim Anomalileri. Hacettepe Üniversitesi İktisadi İdari Bilimler Fakültesi Dergisi. No: 11. pp: 51-62.

Muradoğlu, G., and Ünal, M. (1994). Week Form Efficiency in the Thinly Traded Istanbul Stock Exchange. Middle East Business and Economic Review. Vol: 6. pp: 37-44.

Nomura, H., Hayashi, I., and Wakami, N. (1992). A Learning Method of Fuzzy Inference by Descent Method. Proceedings of the IEEE International Conference on Fuzzy Systems. pp: 203-210.

Obstfeld, M., and Rogoff, K. (2000). Perspective on OECD Economic Integration: Implications for U.S. Current Account Adjustment. Working Paper. Harvard University. http://www.economics.harvard.edu/~krogoff/jackson_hole.pdf.

Ohlson, J.S. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research. Vol: 19. pp: 109-131.

Özmen, T. (1997). Dünya Borsalarında Gözlemlenen Anomaliler ve İstanbul Menkul Kıymetler Borsası Üzerine Bir Deneme. Sermaye Piyasası Kurulu Yayınları. No: 61. Ankara.

Patel, A.V. (2004). Transformation Functions for Trapezoidal Membership Functions. International Journal of Computational Cognition. Vol: 2. No: 3. September. pp: 115-135.

Pesaran, M.H., and Timmermann, A. (1994). Forecasting Stock Returns: An Examination of Stock Market Trading in the Presence of Transaction Costs. Journal of Forecasting. Vol: 13. pp: 330-365.

Pesaran, M.H., and Timmermann, A. (1995). Predictability of Stock Returns: Robustness and Economic Significance. The Journal of Finance. Vol: L. No: 4.

Pesaran, M.H., and Timmermann, A. (2002). Market Timing and Return Prediction under Model Instability. Journal of Empirical Finance. Vol: 9. pp: 495.

Poterba, J., and Summers, L. (1988). Mean Reversion in Stock Prices: Evidence and Implications. Journal of Financial Economics. Vol: 22. pp: 27-59.

Ramser, J., and Foster, L. (1931). A Demonstration of Ratio Analysis. Bureau of Business Research Bulletin 40. University of Illinois.

Rapach, D.E., and Wohar, M.E. (2002). In-Sample and Out-of-Sample Tests of Stock Return Predictability in the Context of Data Mining. Working Paper. University of Nebraska-Omaha.

Rigobon, R., and Sack, B. (2001). Measuring the Reaction of Monetary Policy to the Stock Market. Working Paper. No: 8350. National Bureau of Economic Research.

Sayın, G. (1993). An Investigation of Anomalies at the Istabul Stock Exchange: Winner-Loser Effect. Unpublished MBA Thesis. Department of Management. Bilkent University. Ankara.

Sharma, S. (1996). Applied Multivariate Techniques. Wiley Publishing Company.

Shen, P. (2000). The P/E Ratio and Stock Market Performance. Federal Reserve Bank of Kansas City Economic Review. Vol: 85. pp: 23-36.

Shiller R.J. (1984). Stock Prices and Social Dynamics. Brookings Papers on Economic Activity. Vol: 2. pp: 457-510.

Shnaider, E., and Kandel, A. (1989). The Use of Fuzzy Set Theory for Forecasting Corporate Tax Revenues. Fuzzy Sets and Systems. Vol: 31. pp: 187-204.

Song, Q., and Chissom, B.S. (1993). Forecasting Enrollments with Fuzzy Time Series – part I. Fuzzy Sets and Systems. Vol: 54. pp: 1-9.

Song, Q., and Chissom, B.S. (1994). Forecasting Enrollments with Fuzzy Time Series – part II. Fuzzy Sets and Systems. Vol: 62. pp: 1-8.

Sugeno, M., and Kang, G.T. (1988). Structure Identification of Fuzzy Model. Fuzzy Sets and Systems. Vol: 28. pp: 15-33.

Summers, L.H. (1986). Does the Stock Market Rationally Reflect Fundamental Values? Journal of Finance. Vol: 41. pp: 591-601.

Sun, C.-T. (1994). Rule-base Structure Identification in an Adaptive-network-based Fuzzy Inference System. IEEE Transactions on Fuzzy Systems. Vol: 2. pp: 64-73.

Taffler, R.J. (1983). The Z-Score Approach to Measuring Company Solvency. The Accountant's Magazine. Vol: 87. No: 921. pp: 91-96.

Takagi, T., and Sugeno, M. (1985). Fuzzy Identification System and Its Applications to Modelling and Control. IEEE Transactions on Systems, Man, and Cybernetics. Vol: 15. pp: 116-132.

Tanaka, H., Uejima, S., and Asai, K. (1982). Linear Regression Analysis with Fuzzy Model. IEEE Transactions on Systems, Man and Cybernetics. Vol: 12. pp: 903-907.

Tay, N.S.P., and Linn, S.C. (2001). Fuzzy Inductive Reasoning, Expectation Formation and the Behaviour of Security Prices. Journal of Economic Dynamics and Control. Vol: 25. pp: 321-361.

Taylor, J.B. (1998). A Historical Analysis of Monetary Policy Rules. Working Paper. No: 6768. National Bureau of Economic Research.

Uğurlu, M. (2000). Agency Costs and Corporate Control Devices in the Turkish Manufacturing Industry. Journal of Economic Studies. Vol: 27. pp: 566-600.

Winakor, A., and Smith, R. (1935). Changes in the Financial Structure of Unsuccessful Industrial Corporations. Bureau of Business Research Bulletin 51. University of Illinois.

Zadeh, L.A. (1965). Fuzzy Sets. Information and Control. Vol: 8. pp: 338-353.

Zadeh, L.A. (1973). Outline of a New Approach to the Analysis of Complex Systems. IEEE Transactions on Systems, Man and Cybernetics. Vol: 3.

Zadeh, L.A. (1975). The Calculus of Fuzzy Restrictions. Fuzzy Sets and Applications to Cognitive and Decision Making Processes. Edited by L.A. Zadeh, K.-S. Fu, K. Tanaka, and M. Shimura. Academic Press. New York. pp 1-39.

Zavgren, C.V. (1985). Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis. Journal of Business Finance and Accounting. Vol: 12. pp: 19-45.

APPENDIX – A Program Codes of The Classifier System (Prepared with Gauss 3.1.1.)

```
"INDEX LEVEL";
load yi[] =ISE100.txt;                /* Data Loading */
ni=rows(yi)/2;
xi=reshape(yi,ni,2);
xi[.,2]=xi[.,2]./100;
psubsamp=0.20;
ntrial=20000;
"OUTPUT FILE NAME:";
foname=cons;
cls;
output file=^foname reset;
print"Number of data used";
ni;
print"Proportion of number of subsample to number of sample";
psubsamp;
print"Number of trial";
ntrial;
horiz=0;
k=0;
do while k<3;
        k=k+1;
        if      k==1;   horiz=21;
                elseif  k==4;   horiz=62;
                elseif  k==5;   horiz=125;
                elseif  k==6;   horiz=187;          /* Mean-Stdev Ratio =*/
                elseif  k==2;   horiz=251;
                elseif  k==7;   horiz=313;
                elseif  k==8;   horiz=376;
                elseif  k==9;   horiz=438;
                elseif  k==10;  horiz=502;
                elseif  k==3;   horiz=753;
```

```
endif;
num18=0;
num16=0;
num14=0;            /* Number of Obsevations */
num12=0;
num10=0;
num8=0;
num6=0;
num4=0;
num1=0;
ii=0;
do  while ii<(ni-horiz+1);
            ii=ii+1;     /* Classification */
            if xi[ii,2]>=0.018;                                    num18=num18+1;
            elseif (0.018>xi[ii,2]) and (xi[ii,2]>=0.016);         num16=num16+1;
            elseif (0.016>xi[ii,2]) and (xi[ii,2]>=0.014);         num14=num14+1;
            elseif (0.014>xi[ii,2]) and (xi[ii,2]>=0.012);         num12=num12+1;
            elseif (0.012>xi[ii,2]) and (xi[ii,2]>=0.010);         num10=num10+1;
            elseif (0.010>xi[ii,2]) and (xi[ii,2]>=0.008);         num8=num8+1;
            elseif (0.008>xi[ii,2]) and (xi[ii,2]>=0.006);         num6=num6+1;
            elseif (0.006>xi[ii,2]) and (xi[ii,2]>=0.004);         num4=num4+1;
            elseif (0.004>xi[ii,2]) and (xi[ii,2]>=0.001);         num1=num1+1;
            endif;
endo;
num18ms=zeros(num18,2);
num16ms=zeros(num16,2);
num14ms=zeros(num14,2);          /* Means */
num12ms=zeros(num12,2);          /* and Standard Deviations */
num10ms=zeros(num10,2);
num8ms=zeros(num8,2);
num6ms=zeros(num6,2);
num4ms=zeros(num4,2);
num1ms=zeros(num1,2);
mm=zeros(horiz,2);
xii=zeros(ni-horiz+1,3);
in18=0;
in16=0;
in14=0;                    /* Index Numbers */
```

```
in12=0;
in10=0;
in8=0;
in6=0;
in4=0;
in1=0;
i=0;
do  while i<(ni-horiz);
        i=i+1;
        j=0;
        do while j<horiz;               /* Mean Calculations */
                j=j+1;
                if  (xi[i+j,1]-xi[i+j-1,1])==1;
                mm[j,1]=(xi[i+j,2]/xi[i+j-1,2])-1;
                mm[j,2]=1;
                else;
                mm[j,1]=((xi[i+j,2]/xi[i+j-1,2])^(1/(xi[i+j,1]-xi[i+j-1,1])))-1;
                mm[j,2]=(xi[i+j,1]-xi[i+j-1,1]);
                endif;
        endo;
        xii[i,1]=xi[i,2];
        ss=sumc(mm[.,2]);
        xii[i,2]=(mm[.,1]'*mm[.,2])/ss;
        xii[i,3]=sqrt(((ss*((mm[.,1].*mm[.,1])'*mm[.,2]))-((ss*xii[i,2])^2))/(ss*(ss-1)));
        if xii[i,1]>=0.018;             /* Classification */
                in18=in18+1;
                num18ms[in18,1]=(xii[i,2]);
                num18ms[in18,2]=(xii[i,3]);
        elseif (0.018>xii[i,1]) and (xii[i,1]>=0.016);
                in16=in16+1;
                num16ms[in16,1]=(xii[i,2]);
                num16ms[in16,2]=(xii[i,3]);
        elseif (0.016>xii[i,1]) and (xii[i,1]>=0.014);
                in14=in14+1;
                num14ms[in14,1]=(xii[i,2]);
                num14ms[in14,2]=(xii[i,3]);
        elseif (0.014>xii[i,1]) and (xii[i,1]>=0.012);
                in12=in12+1;
```

```
                num12ms[in12,1]=(xii[i,2]);
                num12ms[in12,2]=(xii[i,3]);
        elseif (0.012>xii[i,1]) and (xii[i,1]>=0.010);
                in10=in10+1;
                num10ms[in10,1]=(xii[i,2]);
                num10ms[in10,2]=(xii[i,3]);
        elseif (0.010>xii[i,1]) and (xii[i,1]>=0.008);
                in8=in8+1;
                num8ms[in8,1]=(xii[i,2]);
                num8ms[in8,2]=(xii[i,3]);
        elseif (0.008>xii[i,1]) and (xii[i,1]>=0.006);
                in6=in6+1;
                num6ms[in6,1]=(xii[i,2]);
                num6ms[in6,2]=(xii[i,3]);
        elseif (0.006>xii[i,1]) and (xii[i,1]>=0.004);
                in4=in4+1;
                num4ms[in4,1]=(xii[i,2]);
                num4ms[in4,2]=(xii[i,3]);
        elseif (0.004>xii[i,1]) and (xii[i,1]>=0.001);
                in1=in1+1;
                num1ms[in1,1]=(xii[i,2]);
                num1ms[in1,2]=(xii[i,3]);
        endif;
endo;
conf=zeros(ntrial,9);
confs=zeros(ntrial,9);
iii=0;
do while iii<ntrial;
        iii=iii+1;
        num18su=ceil(num18*rndu(ceil(num18*psubsamp),1));
        iv=0;
        asam=zeros(ceil(num18*psubsamp),2);
        do while iv<ceil(num18*psubsamp);
                iv=iv+1;
                asam[iv,1]=num18ms[num18su[iv,1],1];
                asam[iv,2]=num18ms[num18su[iv,1],2];
        endo;
        conf[iii,9]=(meanc(asam[.,1])-meanc(num18ms[.,1]))/stdc(asam[.,1]);
```

```
confs[iii,9]=(meanc(asam[.,1])-meanc(num18ms[.,1]))/(meanc(asam[.,2])/sqrt(ceil(num18*psubsamp)));
num16su=ceil(num16*rndu(ceil(num16*psubsamp),1));
iv=0;
asam=zeros(ceil(num16*psubsamp),2);
do while iv<ceil(num16*psubsamp);
        iv=iv+1;
        asam[iv,1]=num16ms[num16su[iv,1],1];
        asam[iv,2]=num16ms[num16su[iv,1],2];
endo;
conf[iii,8]=(meanc(asam[.,1])-meanc(num16ms[.,1]))/stdc(asam[.,1]);
confs[iii,8]=(meanc(asam[.,1])-meanc(num16ms[.,1]))/(meanc(asam[.,2])/sqrt(ceil(num16*psubsamp)));
num14su=ceil(num14*rndu(ceil(num14*psubsamp),1));
iv=0;
asam=zeros(ceil(num14*psubsamp),2);
do while iv<ceil(num14*psubsamp);
        iv=iv+1;
        asam[iv,1]=num14ms[num14su[iv,1],1];
        asam[iv,2]=num14ms[num14su[iv,1],2];
endo;
conf[iii,7]=(meanc(asam[.,1])-meanc(num14ms[.,1]))/stdc(asam[.,1]);
confs[iii,7]=(meanc(asam[.,1])-meanc(num14ms[.,1]))/(meanc(asam[.,2])/sqrt(ceil(num14*psubsamp)));
num12su=ceil(num12*rndu(ceil(num12*psubsamp),1));
iv=0;
asam=zeros(ceil(num12*psubsamp),2);
do while iv<ceil(num12*psubsamp);
        iv=iv+1;
        asam[iv,1]=num12ms[num12su[iv,1],1];
        asam[iv,2]=num12ms[num12su[iv,1],2];
endo;
conf[iii,6]=(meanc(asam[.,1])-meanc(num12ms[.,1]))/stdc(asam[.,1]);
confs[iii,6]=(meanc(asam[.,1])-meanc(num12ms[.,1]))/(meanc(asam[.,2])/sqrt(ceil(num12*psubsamp)));
m10su=ceil(num10*rndu(ceil(num10*psubsamp),1));
iv=0;
asam=zeros(ceil(num10*psubsamp),2);
do while iv<ceil(num10*psubsamp);
        iv=iv+1;
        asam[iv,1]=num10ms[num10su[iv,1],1];
        asam[iv,2]=num10ms[num10su[iv,1],2];
```

```
endo;
conf[iii,5]=(meanc(asam[.,1])-meanc(num10ms[.,1]))/stdc(asam[.,1]);
confs[iii,5]=(meanc(asam[.,1])-meanc(num10ms[.,1]))/(meanc(asam[.,2])/sqrt(ceil(num10*psubsamp)));
num8su=ceil(num8*rndu(ceil(num8*psubsamp),1));
iv=0;
asam=zeros(ceil(num8*psubsamp),2);
do while iv<ceil(num8*psubsamp);
        iv=iv+1;
        asam[iv,1]=num8ms[num8su[iv,1],1];
        asam[iv,2]=num8ms[num8su[iv,1],2];
endo;
conf[iii,4]=(meanc(asam[.,1])-meanc(num8ms[.,1]) )/stdc(asam[.,1]);
confs[iii,4]=(meanc(asam[.,1])-meanc(num8ms[.,1]))/(meanc(asam[.,2])/sqrt(ceil(num8*psubsamp)));
num6su=ceil(num6*rndu(ceil(num6*psubsamp),1));
iv=0;
asam=zeros(ceil(num6*psubsamp),2);
do while iv<ceil(num6*psubsamp);
        iv=iv+1;
        asam[iv,1]=num6ms[num6su[iv,1],1];
        asam[iv,2]=num6ms[num6su[iv,1],2];
endo;
conf[iii,3]=(meanc(asam[.,1])-meanc(num6ms[.,1]) )/stdc(asam[.,1]);
confs[iii,3]=(meanc(asam[.,1])-meanc(num6ms[.,1]))/(meanc(asam[.,2])/sqrt(ceil(num6*psubsamp)));
num4su=ceil(num4*rndu(ceil(num4*psubsamp),1));
iv=0;
asam=zeros(ceil(num4*psubsamp),2);
do while iv<ceil(num4*psubsamp);
        iv=iv+1;
        asam[iv,1]=num4ms[num4su[iv,1],1];
        asam[iv,2]=num4ms[num4su[iv,1],2];
endo;
conf[iii,2]=(meanc(asam[.,1])-meanc(num4ms[.,1]) )/stdc(asam[.,1]);
confs[iii,2]=(meanc(asam[.,1])-meanc(num4ms[.,1]))/(meanc(asam[.,2])/sqrt(ceil(num4*psubsamp)));
num1su=ceil(num1*rndu(ceil(num1*psubsamp),1));
iv=0;
asam=zeros(ceil(num1*psubsamp),2);
do while iv<ceil(num1*psubsamp);
        iv=iv+1;
```

```
                    asam[iv,1]=num1ms[num1su[iv,1],1];
                    asam[iv,2]=num1ms[num1su[iv,1],2];
            endo;
            conf[iii,1]=(meanc(asam[.,1])-meanc(num1ms[.,1]) )/stdc(asam[.,1]);
            confs[iii,1]=(meanc(asam[.,1])-meanc(num1ms[.,1]))/(meanc(asam[.,2])/sqrt(ceil(num1*psubsamp)));
endo;
horiz;
"   region";;                    /* Output */
"   mean        deviation     mean/deviation   num.of obs.";
"   conf1%      conf99%       conf2.5%      conf97.5%  ";
"   conf5%      conf95%       conf10%       conf90%    ";
"   num.of subsamp.  p-value                          ";
">0.018";
meanc(num18ms[.,1])~stdc(num18ms[.,1])~meanc(num18ms[.,1])/
stdc(num18ms[.,1])~num18;
ind=sortc(conf,9);
ind[round(0.01*ntrial),9]~ind[round(0.99*ntrial),9]~ind[round(0.025*ntrial),9]~ind[round(0.975*ntrial),9];
ind[round(0.05*ntrial),9]~ind[round(0.95*ntrial),9]~ind[round(0.100*ntrial),9]~ind[round(0.900*ntrial),9];
v=1;
do while (ind[v,9]<(meanc(num18ms[.,1])/stdc(num18ms[.,1]))) and (v<ntrial-1);
            v=v+1;
endo;
ceil(num18*psubsamp)~v/ntrial;
"0.018-0.016";
meanc(num16ms[.,1])~stdc(num16ms[.,1])~meanc(num16ms[.,1])/stdc(num16ms[.,1])~num16;
ind=sortc(conf,8);
ind[round(0.01*ntrial),8]~ind[round(0.99*ntrial),8]~ind[round(0.025*ntrial),8]~ind[round(0.975*ntrial),8];
ind[round(0.05*ntrial),8]~ind[round(0.95*ntrial),8]~ind[round(0.100*ntrial),8]~ind[round(0.900*ntrial),8];
v=1;
do while (ind[v,8]<(meanc(num16ms[.,1])/stdc(num16ms[.,1]))) and (v<ntrial-1);
            v=v+1;
endo;
ceil(num16*psubsamp)~v/ntrial;
"0.016-0.014";
meanc(num14ms[.,1])~stdc(num14ms[.,1])~meanc(num14ms[.,1])/stdc(num14ms[.,1])~num14;
ind=sortc(conf,7);
ind[round(0.01*ntrial),7]~ind[round(0.99*ntrial),7]~ind[round(0.025*ntrial),7]~ind[round(0.975*ntrial),7];
ind[round(0.05*ntrial),7]~ind[round(0.95*ntrial),7]~ind[round(0.100*ntrial),7]~ind[round(0.900*ntrial),7];
```

```
v=1;
do while (ind[v,7]<(meanc(num14ms[.,1])/stdc(num14ms[.,1]))) and (v<ntrial-1);
        v=v+1;
endo;
ceil(num14*psubsamp)~v/ntrial;
"0.014-0.012";
meanc(num12ms[.,1])~stdc(num12ms[.,1])~meanc(num12ms[.,1])/stdc(num12ms[.,1])~num12;
ind=sortc(conf,6);
ind[round(0.01*ntrial),6]~ind[round(0.99*ntrial),6]~ind[round(0.025*ntrial),6]~ind[round(0.975*ntrial),6];
ind[round(0.05*ntrial),6]~ind[round(0.95*ntrial),6]~ind[round(0.100*ntrial),6]~ind[round(0.900*ntrial),6];
v=1;
do while (ind[v,6]<(meanc(num12ms[.,1])/stdc(num12ms[.,1]))) and (v<ntrial-1);
        v=v+1;
endo;
ceil(num12*psubsamp)~v/ntrial;
"0.012-0.010";
meanc(num10ms[.,1])~stdc(num10ms[.,1])~meanc(num10ms[.,1])/stdc(num10ms[.,1])~num10;
ind=sortc(conf,5);
ind[round(0.01*ntrial),5]~ind[round(0.99*ntrial),5]~ind[round(0.025*ntrial),5]~ind[round(0.975*ntrial),5];
ind[round(0.05*ntrial),5]~ind[round(0.95*ntrial),5]~ind[round(0.100*ntrial),5]~ind[round(0.900*ntrial),5];
v=1;
do while (ind[v,5]<(meanc(num10ms[.,1])/stdc(num10ms[.,1]))) and (v<ntrial-1);
        v=v+1;
endo;
ceil(num10*psubsamp)~v/ntrial;
"0.010-0.008";
meanc(num8ms[.,1])~stdc(num8ms[.,1])~meanc(num8ms[.,1])/stdc(num8ms[.,1])~num8;
ind=sortc(conf,4);
ind[round(0.01*ntrial),4]~ind[round(0.99*ntrial),4]~ind[round(0.025*ntrial),4]~ind[round(0.975*ntrial),4];
ind[round(0.05*ntrial),4]~ind[round(0.95*ntrial),4]~ind[round(0.100*ntrial),4]~ind[round(0.900*ntrial),4];
v=1;
do while (ind[v,4]<(meanc(num8ms[.,1])/stdc(num8ms[.,1]))) and (v<ntrial-1);
        v=v+1;
endo;
ceil(num8*psubsamp)~v/ntrial;
"0.008-0.006";
meanc(num6ms[.,1])~stdc(num6ms[.,1])~meanc(num6ms[.,1])/stdc(num6ms[.,1])~num6;
ind=sortc(conf,3);
```

```
        ind[round(0.01*ntrial),3]~ind[round(0.99*ntrial),3]~ind[round(0.025*ntrial),3]~ind[round(0.975*ntrial),3];
        ind[round(0.05*ntrial),3]~ind[round(0.95*ntrial),3]~ind[round(0.100*ntrial),3]~ind[round(0.900*ntrial),3];
        v=1;
        do while (ind[v,3]<(meanc(num6ms[.,1])/stdc(num6ms[.,1]))) and (v<ntrial-1);
                v=v+1;
        endo;
        ceil(num6*psubsamp)~v/ntrial;
        "0.006-0.004";
        meanc(num4ms[.,1])~stdc(num4ms[.,1])~meanc(num4ms[.,1])/stdc(num4ms[.,1])~num4;
        ind=sortc(conf,2);
        ind[round(0.01*ntrial),2]~ind[round(0.99*ntrial),2]~ind[round(0.025*ntrial),2]~ind[round(0.975*ntrial),2];
        ind[round(0.05*ntrial),2]~ind[round(0.95*ntrial),2]~ind[round(0.100*ntrial),2]~ind[round(0.900*ntrial),2];
        v=1;
        do while (ind[v,2]<(meanc(num4ms[.,1])/stdc(num4ms[.,1]))) and (v<ntrial-1);
                v=v+1;
        endo;
        ceil(num4*psubsamp)~v/ntrial;
        "0.004-0.002";
        meanc(num1ms[.,1])~stdc(num1ms[.,1])~meanc(num1ms[.,1])/stdc(num1ms[.,1])~num1;
        ind=sortc(conf,1);
        ind[round(0.01*ntrial),1]~ind[round(0.99*ntrial),1]~ind[round(0.025*ntrial),1]~ind[round(0.975*ntrial),1];
        ind[round(0.05*ntrial),1]~ind[round(0.95*ntrial),1]~ind[round(0.100*ntrial),1]~ind[round(0.900*ntrial),1];
        v=1;
        do while (ind[v,1]<(meanc(num1ms[.,1])/stdc(num1ms[.,1]))) and (v<ntrial-1);
                v=v+1;
        endo;
        ceil(num1*psubsamp)~v/ntrial;
endo;
output off;
```

APPENDIX – B Financial Ratios

1. Company Name
2. Year
3. Quarter
4. Sector
5. Segment
6. A Priori Rule for Failled and Unfailled Companies (0 or 1)
7. **TOAS** Total Assets/ Wholesale Price Index
8. **SALE** Sales/ Wholesale Price Index
9. **TAFAS** Tangible Fixed Assets/ Wholesale Price Index
10. **LTD** Long Term Debt/ Wholesale Price Index
11. **CATA** Current Assets/ Total Assets
12. **QATA** Quick Assets/ Total Assets
13. **LAMSTA** (Liquid Assets+Mark. Securities)/ Total Assets
14. **STARTA** Short Term Accounts Receivables/ Total Assets
15. **WCTA** Working Capital/ Total Assets
16. **DR** Debt Ratio = Total Debt/ Total Assets
17. **LTDTA** Long Term Debt/ Total Assets
18. **STDTA** Short Term Debt/ Total Assets
19. **FDTA** Financial Debt/ Total Assets
20. **STFDTA** Short Term Financial Debt/ Total Assets
21. **SETA** Shareholders' Equity/ Total Assets
22. **APNPTA** (Account Payables+Notes Payables)/ Total Assets

23. **ROA** Return on Assets = Net Income/ Total Assets
24. **RETA** Retained Earnings/ Total Assets
25. **TAT** Total Asset Turnover = Sales/ Total Assets
26. **SCA** Sales/ Current Assets
27. **SLA** Sales/ Liquid Assets
28. **ART** Accounts Receivable Turnover = Sales/ Accounts Receivables
29. **SWC** Sales/ Working Capital
30. **FAT** Fixed Asset Turnover = Sales/ Fixed Assets
31. **STFA** Sales/ Tangible Fixed Assets
32. **SSE** Sales/ Shareholders' Equity
33. **CAS** Current Assets/ Sales
34. **QAS** Quick Assets/ Sales
35. **WCS** Net Working Capital/ Sales
36. **LAMSS** (Liquid Assets+Mark. Securities)/ Sales
37. **EXS** Exports/ Sales
38. **FES** Financial Expenses/ Sales
39. **RDS** R&D Expenses/ Sales
40. **SES** Shareholders' Equity/ Sales
41. **GPM** Gross Profit Margin = (Sales – Cost of Goods Sold)/ Sales
42. **OPS** Operating Profit/ Sales
43. **NIS** Net Income/ Sales
44. **APT** Accounts Payable Turnover=Cost of Goods Sold/ Short Term Accounts Payables
45. **INVT** Inventory Turnover = Cost of Goods Sold/ Inventories
46. **QAI** Quick Assets/ Inventories

47. **LTDTD** Long Term Debt/ Total Debt
48. **LTFDTD** Long Term Financial Debt/ Total Debt
49. **IETD** Interest Expenses/ Total Debt
50. **MVETD** Market Value of Equity/ Total Debt
51. **FDLTD** Financial Debt/ Long Term Debt
52. **WCLTD** Net Working Capital/ Long Term Debt
53. **WCSTD** Net Working Capital/ Short Term Debt
54. **STFDSTD** Short Term Financial Debt/ Short Term Debt
55. **MVETL** Market Value of Equity/ Total Liabilities
56. **PCTL** Paid Capital/ Total Liabilities
57. **APTL** Accounts Payables/ Total Liabilities
58. **CR** Current Ratio = Current Assets/ Current Liabilities
59. **QR** Quick Ratio = (Current Assets – Inventories)/ Current Liabilities
60. **CASH** Cash Ratio = (Cash + Marketable Securities)/ Current Liabilities
61. **CCL** Cash/ Current Liabilities
62. **APCL** Accounts Payables/ Current Liabilities
63. **CLCA** Current Liabilities/ Current Assets
64. **FEFL** Financial Expenses/ Financial Liabilities
65. **SFEFE** Short Term Financial Expenses/ Financial Expenses
66. **FASE** Fixed Assets/ Shareholders' Equity
67. **SEFA** Shareholders' Equity/ Fixed Assets
68. **FASELTD** Fixed Assets/ (Shareholders' Equity+Long Term Debt)
69. **TFASE** Tangible Fixed Assets/ Shareholders' Equity
70. **TDSE** Total Debt/ Shareholders' Equity
71. **LTDSE** Long Term Debt/ Shareholders' Equity

72. **ROE** Return on Equity = Net Income/ Shareholders' Equity
73. **RSSE** Revaluation Surplus/ Shareholders' Equity
74. **EBITTA** EBIT/ Total Assets
75. **EBITSE** EBIT/ Shareholders' Equity
76. **EBITPC** EBIT/ Paid Capital
77. **EBITS** EBIT/ Sales
78. **EBTTA** EBT/ Total Assets
79. **EBITDT** EBITDA/ Total Assets
80. **EBITDS** EBITDA/ Sales
81. **TIE** Times Interest Earned = EBIT/ Interest Expenses
82. **NIPC** Net Income/ Paid Capital
83. **OIBOIA** Other Income Before/ Other Income After
84. **IDIOOI** Interest and Dividend Income/ Other Operating Income
85. **PROFIT** (Price_end – Price_beg)/ Price_beg
86. **VOL** Annual Volatility in Quarterly Period
87. **RET_3** Prev. 3 months return with repect to ISE100 Index
88. **RET3** Next 3 months return with repect to ISE100 Index
89. **RET6** Next 6 months return with repect to ISE100 Index

APPENDIX – C Outlier Companies for Disc. Anal. And Logistic Reg.

OUTLIERS

ABANA	1996	2	CARSI	2000	1	FVORI	1999	3	METAS	2001	4	SANKO	2000	1
ABANA	1996	3	CARSI	2000	2	FVORI	2002	2	METAS	2002	2	SANKO	2002	1
ABANA	1996	4	CARSI	2000	3	FVORI	2002	3	METAS	2000	2	SERVE	1997	2
AKENR	2001	1	CARSI	2000	4	GIMA	1996	1	METAS	2000	4	SERVE	1996	4
AKENR	2001	2	CARSI	2001	1	GIMA	2002	1	MIGRS	1997	1	SERVE	1997	3
AKENR	2001	3	CARSI	2001	2	GIMA	2002	3	MIGRS	1997	2	SERVE	1997	4
AKENR	2001	4	CARSI	2001	3	GIMA	1996	3	MIGRS	1997	3	SIFAS	2001	3
AKSUE	2000	1	CARSI	2002	1	GIMA	1996	4	MIGRS	1996	1	SIFAS	2001	1
AKSUE	2000	4	CBSBO	2001	3	GIMA	1997	1	MIGRS	1996	2	SIFAS	2001	2
AKTAS	2000	1	CBSBO	2001	4	GIMA	1997	2	MIGRS	1996	3	SIFAS	2001	4
ALKIM	1997	4	CBSBO	2001	2	GIMA	1997	4	MIGRS	1996	4	SNPAM	1998	1
ANACM	1997	3	CBSBO	2002	1	GIMA	1998	4	MIGRS	1997	4	SNPAM	1998	2
ARAT	2001	4	CBSBO	2002	3	GIMA	2001	4	NERGS	2001	4	SNPAM	1998	3
ARENA	1998	4	CBSBO	2002	2	GSRAY	2001	4	NERGS	2001	3	SNPAM	1998	4
ARENA	1999	2	CBSBO	2001	1	GSRAY	2000	3	NERGS	2001	2	SOKSA	2001	4
ARENA	1999	3	CLEBI	2001	3	GSRAY	2000	4	NERGS	2001	1	SOKSA	2001	2
ARENA	1997	4	CMLOJ	2001	2	GSRAY	2001	3	NERGS	1999	1	SOKSA	2001	1
ARENA	1999	4	CMLOJ	2001	1	GSRAY	2001	2	NERGS	1999	3	SOKSA	1999	3
ARENA	2000	1	CUKEL	2002	2	GUBRF	2000	3	NERGS	1999	4	SOKSA	2002	1
ARENA	2000	2	CUKEL	2002	3	GUBRF	2001	3	NERGS	2000	1	TCELL	2001	1
ARENA	2000	3	CUKEL	2001	1	GUBRF	1997	1	NERGS	2002	3	TCELL	1997	4
ARENA	2000	4	CUKEL	2001	2	GUBRF	1998	3	NETAS	1999	4	TNSAS	1998	1
ARENA	2001	4	CUKEL	2001	3	GUBRF	1999	3	PKENT	2000	4	TNSAS	1998	2
ARENA	2002	1	CUKEL	1998	2	GUBRF	2001	4	PKENT	2001	1	TNSAS	2000	4
ARENA	2002	2	CUKEL	2000	3	GUBRF	2002	1	PKENT	2001	2	TNSAS	1996	3
ARENA	2002	3	CUKEL	2000	4	GUBRF	2002	3	PKENT	2000	2	TNSAS	1997	1
ASLAN	2000	1	CUKEL	2001	4	GUMUS	2002	2	PKENT	2000	3	TNSAS	1997	3
ASUZU	1999	4	DARDL	2001	4	GUMUS	2001	2	PKENT	2001	3	TNSAS	1997	4
AYEN	2000	1	DARDL	2002	1	GUMUS	1999	2	PKENT	2001	4	TNSAS	1998	3
AYEN	2000	2	DARDL	2002	2	GUMUS	1999	3	PKENT	2002	2	TNSAS	1998	4
AYEN	1997	4	DARDL	2001	3	GUMUS	1999	4	PKENT	2002	3	TNSAS	1999	1
AYEN	1998	4	DARDL	2001	1	GUMUS	2000	1	PKENT	1999	4	TNSAS	1999	2
AYEN	1999	3	DARDL	2001	2	GUMUS	2001	1	PKENT	1999	1	TNSAS	1999	3
AYEN	1999	4	DARDL	2002	3	IHEVA	1999	2	PKENT	1999	2	TNSAS	1999	4
AYEN	2001	1	DARDL	2000	4	IHEVA	2002	1	PKENT	2002	1	TNSAS	2000	1
AYEN	2001	3	DOGUB	2001	2	IHEVA	1996	1	PKENT	2000	1	TNSAS	2000	2
AYEN	2001	4	DOGUB	1996	2	IHEVA	1999	4	PKENT	1998	4	TNSAS	2000	3
AYEN	2002	1	DUROF	1999	3	IHEVA	2000	3	PKENT	1999	3	TRCAS	1999	2
AYEN	2002	2	EDIP	2000	4	INTEM	1998	3	POLYL	1998	1	TUMTK	2001	1
AYEN	2002	3	EDIP	2001	1	INTEM	1999	3	POLYL	1998	2	TUMTK	2001	2
BFREN	1999	2	EDIP	2001	2	INTEM	2000	2	POLYL	1999	2	TUMTK	1999	4
BFREN	1999	3	EDIP	2001	4	INTEM	2000	3	POLYL	1999	4	TUMTK	2000	1
BFREN	1999	4	EDIP	2000	2	INTEM	2000	4	POLYL	2000	4	TUMTK	2000	3
BFREN	2000	1	EDIP	2000	3	INTEM	2001	1	POLYL	2001	2	TUMTK	2001	4
BFREN	2000	2	EDIP	2001	3	INTEM	2001	3	POLYL	2001	3	TUMTK	2002	1
BFREN	2000	3	EDIP	2002	2	INTEM	2002	2	POLYL	2001	4	TUMTK	1999	3
BFREN	2001	1	EDIP	2002	3	INTEM	2002	3	PRKTE	2000	3	TUMTK	2000	2
BFREN	2001	2	EDIP	1999	4	IPMAT	2002	3	RAKSE	2001	2	TUMTK	2000	4
BFREN	2001	3	EDIP	2002	1	KENT	2001	3	RAKSE	2001	4	TUMTK	2001	3
BISAS	1996	1	EDIP	1999	1	KEPEZ	2001	4	RAKSE	2001	3	TUMTK	2002	2
BISAS	1999	3	EDIP	1999	2	KEPEZ	2002	1	RAKSE	2001	1	TUMTK	2002	3
BJKAS	2001	1	EDIP	2000	1	KEPEZ	2002	3	RAKSE	2000	2	TUMTK	1999	2
BJKAS	2000	4	EDIP	1998	4	KEPEZ	1999	4	RAKSE	2000	3	UCAK	1997	3
BROVA	1998	3	EDIP	1999	3	KEPEZ	2001	3	RAKSE	2002	1	UCAK	1997	4
BROVA	1998	4	EGGUB	1996	4	KEPEZ	2002	2	RAKSE	2002	2	UCAK	1998	1
BROVA	1999	1	EGGUB	2001	3	KERVT	2001	2	RAKSE	2002	3	UCAK	1997	2
BROVA	1999	2	EGIYM	1998	4	KERVT	2001	3	RAKSE	1999	4	UCAK	1998	2
BROVA	1999	3	EGIYM	1999	4	KERVT	2001	4	RAKSE	2000	4	UCAK	1998	3
BYSAN	1999	4	EGIYM	1999	1	KERVT	2002	3	RKSEV	2001	3	UCAK	1998	4
BYSAN	2000	3	EGIYM	1999	2	KERVT	2002	2	RKSEV	2001	1	UCAK	1999	1
BYSAN	2001	3	EGIYM	1999	3	KIPA	1999	3	RKSEV	2001	2	UCAK	1999	2
BYSAN	2000	4	EGIYM	2000	1	KOYTS	1999	4	RKSEV	2001	4	UCAK	1999	3
BYSAN	2001	2	EGIYM	2000	2	KOYTS	2001	4	RKSEV	2000	1	UCAK	1999	4
BYSAN	2001	4	EGIYM	2001	1	KOYTS	2002	2	RKSEV	2000	2	UCAK	2000	1
BYSAN	2000	1	EGIYM	2001	2	KOYTS	2000	4	RKSEV	2000	3	UCAK	2000	2
BYSAN	2001	1	EGIYM	2001	3	KOYTS	2001	2	RKSEV	2002	1	UCAK	2000	3
BYSAN	2002	1	EGIYM	2001	4	LINK	2001	2	RKSEV	2002	2	UCAK	2000	4
BYSAN	2002	2	EGSER	2001	3	LINK	2001	3	RKSEV	2002	3	UCAK	2001	1
BYSAN	2002	3	EMKEL	1997	1	LINK	2001	4	RKSEV	1999	4	UCAK	2001	2
BYSAN	1999	2	EMKEL	1997	2	LINK	2002	1	SABAH	2001	3	UCAK	2001	3
BYSAN	1999	3	ENKAI	1999	4	LINK	2002	2	SABAH	2002	1	UCAK	2001	4
BYSAN	2000	2	ENKAI	2000	1	LINK	2002	3	SABAH	2001	4	UNYEC	1999	4
CARSI	1999	3	ENKAI	2001	2	LUKSK	2001	2	SABAH	2001	2	UNYEC	2000	2
CARSI	1999	4	EPLAS	2002	1	LUKSK	2001	4	SABAH	2002	3	UNYEC	2000	3
CARSI	2001	4	ERSU	2002	2	MDRNU	2001	2	SABAH	2002	2	VKING	1997	1
CARSI	1996	1	ERSU	2000	2	MDRNU	2001	3	SABAH	1999	4	VKING	1997	2
CARSI	1996	2	ERSU	2001	1	MDRNU	2001	4	SABAH	2000	3	ZOREN	2000	1
CARSI	1996	3	ERSU	2001	2	MDRNU	2001	1	SABAH	2001	1	ZOREN	2000	2
CARSI	1996	4	ESCOM	1997	4	MDRNU	2000	4	SANKO	1997	4	ZOREN	1998	4
CARSI	1997	1	ESCOM	2001	1	METAS	2002	3	SANKO	1998	4	ZOREN	1999	4
CARSI	1997	2	FENIS	1998	2	METAS	2001	2	SANKO	2000	2	ZOREN	2000	3
CARSI	1997	4	FENIS	2001	2	METAS	2002	1	SANKO	2001	2	ZOREN	2001	3
CARSI	1998	2	FENIS	2001	3	METAS	2000	3	SANKO	2002	2	ZOREN	2001	4
CARSI	1998	4	FVORI	1997	4	METAS	2001	1	SANKO	2002	3	ZOREN	2002	1
CARSI	1999	2	FVORI	1999	2	METAS	2001	3	SANKO	1999	4	ZOREN	2002	2
												ZOREN	2002	3

APPENDIX – D Program Codes for Developing The Fuzzy Logic Model (Prepared with Matlab 5.3.)

```
clc;
clear;
disp('Fuzzy Logic Model Programming for the Istanbul Stock Exchange 100 Index');

phddvardata;                                    %Variable data
datax=12;                                       %Number of variables
datay=1729;                                     %Number of data, between 1997 and 2000
totalrulenum=6;                                 %there are 48 months
maxcomprule=3;                                  %Number of variables produced
output=1;                                       %Output variable
datev=1;                                        %Date variable
loopnum=100;                                    %Loop number
bloopnum=3;                                     %Big loop number
phdvardata=zeros(datax+output,datay);
phdvardata(1:datax+output,1:datay)=phddvardata(1+datev:datev+datax+output,1:datay);
data=zeros(datax+totalrulenum+output,datay);
data(1:datax,1:datay)=phdvardata(1:datax,1:datay);
data(datax+totalrulenum+output,1:datay)=phdvardata(datax+output,1:datay);
rule=zeros(totalrulenum,(maxcomprule+output));
rule(1,1:(maxcomprule+output)) =[ 1  2  0 16];    %Design of the model and the rules.
rule(2,1:(maxcomprule+output)) =[ 6  5  4 13];    %As an example [1 2 0 16] means:
rule(3,1:(maxcomprule+output)) =[ 7  8  9 14];    %1st and 2nd var. produces 16th var.
rule(4,1:(maxcomprule+output)) =[12 11 10 15];    %If we have 3 var. entering such as:
rule(5,1:(maxcomprule+output)) =[15 13 14 17];    %[7 8 9 14] 7th,8th and 9th variable
rule(6,1:(maxcomprule+output)) =[16 17  3 18];    %produces 14th variable.
rulemat2=zeros(5,5);
rulemat3=zeros(5,5,5);
i=0;
while i<5
   j=0;
   i=i+1;
   while j<5                              %Rule Matrices 2by2 for 2 rules, 3by3 for 3 rules:
```

```
      j=j+1;
      rulemat2(i,j)=round((i+j)/2);
      k=0;
      while k<5
         k=k+1;
         rulemat3(i,j,k)=round((i+j+k)/3);
      end;
   end;
end;
phdrangedata;
range=zeros(datax+totalrulenum+output,7);
datacor=data(1:datax+output,1:datay);
datacor(datax+output,1:datay)=data(datax+totalrulenum+output,1:datay);
cc=corrcoef(datacor');
aa=zeros(datax+output,1);
aa=cc(1:datax+output,datax+output);
aa=[1 -1; 2 1; 3 -1; 4 -1; 5 1; 6 1; 7 1; 8 -1; 9 1; 10 1; 11 1; 12 1 ];
for i=1:datax,
   if aa(i,2)>0
        range(i,1)=phdrangedata(i,7);
        range(i,2)=phdrangedata(i,6);
        range(i,3)=phdrangedata(i,5);
        range(i,4)=phdrangedata(i,4);
        range(i,5)=phdrangedata(i,3);
        range(i,6)=phdrangedata(i,2);
        range(i,7)=phdrangedata(i,1);
   else
       range(i,1)=-1*phdrangedata(i,1);
        range(i,2)=-1*phdrangedata(i,2);
        range(i,3)=-1*phdrangedata(i,3);
        range(i,4)=-1*phdrangedata(i,4);
        range(i,5)=-1*phdrangedata(i,5);
        range(i,6)=-1*phdrangedata(i,6);
        range(i,7)=-1*phdrangedata(i,7);
      data(i,1:datay)=-1*data(i,1:datay);
   end;
end;
i=0;
while i<totalrulenum
```

```
    i=i+1;
    range(datax+i,1)=-2.5;
    range(datax+i,2)=-2;
    range(datax+i,3)=-1;
    range(datax+i,4)=0;
    range(datax+i,5)=1;
    range(datax+i,6)=2;
    range(datax+i,7)=2.5;
end;
range(datax+totalrulenum+output,1)=phdrangedata(datax+output,7);   %Initialization
range(datax+totalrulenum+output,2)=phdrangedata(datax+output,6);
range(datax+totalrulenum+output,3)=phdrangedata(datax+output,5);
range(datax+totalrulenum+output,4)=phdrangedata(datax+output,4);
range(datax+totalrulenum+output,5)=phdrangedata(datax+output,3);
range(datax+totalrulenum+output,6)=phdrangedata(datax+output,2);
range(datax+totalrulenum+output,7)=phdrangedata(datax+output,1);
data=model(data,range,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
    %Normalization of the output
    yreg=data(datax+totalrulenum+output,1:datay)';
    xreg=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
    [B,BINT,R,RINT,STATS] =regress( yreg, xreg,.05);
    out=xreg*B;
    exp=(1-((yreg-out)'*(yreg-out))/((yreg-mean(yreg))'*(yreg-mean(yreg))));
    t=[1996+(1/datay):((2003-1996)/datay):2003];
    figure(1);
    plot(t,yreg,t,out);
    title(exp);
con=0.00000018;
abdif=con+1;
alpham=1;
opvec=[3 16 17 18];
n1=max(size(opvec));
n2=7-1;
iii=0;
while and ((abdif>con),(iii<loopnum));
    iii=iii+1;
    disp('iii1');
    disp(iii);
```

```
grad=zeros(n1,n2-1);
epsmax=zeros(n1,n2-1);
eps=0.000001;
rangeeps=zeros(datax+totalrulenum+output,7);
check=zeros(1,6);
docheck=0;
rangeeps=range;                                          %Initialization
data=model(data,range,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';          %Normalization of the output
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
exp=(1-((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo))));
ii=0;
while ii<n1
      ii=ii+1;
      jj=1;
      disp([opvec(ii) exp]);
      while jj<6;
         jj=jj+1;
         epsp=.4*(range(opvec(ii),jj+1)-range(opvec(ii),jj));
         epsm=.4*(range(opvec(ii),jj)-range(opvec(ii),jj-1));
         epsmax(ii,jj-1)=min([epsp epsm]);
         if eps>epsmax(ii,jj-1)
            eps=epsmax(ii,jj-1);
         end;
         rangeeps(opvec(ii),jj)=range(opvec(ii),jj)+eps;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
         yo=data(datax+totalrulenum+output,1:datay)';    %Normalization of the output
         xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
         xob=xo*B;
         expo=(1-((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo))));
         grad(ii,jj-1)=(expo-exp)/eps;
      end;
end;
alphamax=alpham;
rangeeps(opvec,2:n2)=range(opvec,2:n2);
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
```

```
xob=xo*B;
f0=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
h=0;
l=0;
if f1>f0;
 while and ((h<100),(f1>f0));
      h=h+1;
      rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
 data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
      yo=data(datax+totalrulenum+output,1:datay)';
      xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
      xob=xo*B;
      f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
      alphamax=alphamax/2;
 end;
   alphamax=2*alphamax;
   rangeeps(opvec,2:n2)=range(opvec,2:n2);
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f0=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
   rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
```

```
    f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
    rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
    data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
    yo=data(datax+totalrulenum+output,1:datay)';
    xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
    xob=xo*B;
    f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
elseif f1>f2;
 while and((l<100),(f1>f2))
       l=l+1;
       rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
 data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
       yo=data(datax+totalrulenum+output,1:datay)';
       xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
       xob=xo*B;
       f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
       alphamax=2*alphamax;
    end;
    alphamax=alphamax/2;
    rangeeps(opvec,2:n2)=range(opvec,2:n2);
    data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
    yo=data(datax+totalrulenum+output,1:datay)';
    xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
    xob=xo*B;
    f0=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
    rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
    data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
    yo=data(datax+totalrulenum+output,1:datay)';
    xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
    xob=xo*B;
    f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
    rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
    data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
    yo=data(datax+totalrulenum+output,1:datay)';
    xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
    xob=xo*B;
    f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
end;
if (min([f0 f1 f2])==f1);
```

```
            sol=(inv([0              0           1;
                (alphamax/2)^2  (alphamax/2) 1;
                (alphamax)^2    (alphamax)   1]))*([ f0 ; f1 ; f2 ]);
            alphamax=-sol(2)/(2*sol(1));
    end;
    disp('f0 f1 f2 alphamax h l');
    disp([f0 f1 f2 alphamax h l]);
    if f1 > f0   'there is something going on wrong'; end;
    rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
    while docheck==0;
       check(1)=min(rangeeps(opvec,1)<rangeeps(opvec,2));
       check(2)=min(rangeeps(opvec,2)<rangeeps(opvec,3));
       check(3)=min(rangeeps(opvec,3)<rangeeps(opvec,4));
       check(4)=min(rangeeps(opvec,4)<rangeeps(opvec,5));
       check(5)=min(rangeeps(opvec,5)<rangeeps(opvec,6));
       check(6)=min(rangeeps(opvec,6)<rangeeps(opvec,7));
       docheck=min(check);
       if docheck==0;
           alphamax=alphamax/2;
           rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
        else;
           range=rangeeps;
        end;
    end;
    abdif=sum(sum((alphamax*grad).*(alphamax*grad)));
    disp(abdif);
end;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
bloop=0;
while bloop<bloopnum;
bloop=bloop+1;
con=0.00000018;
abdif=con+1;
alpham=1;
opvec=[1 2 3 4 5 6 7 8 9 10 11 12];
n1=max(size(opvec)); n2=7-1;
iii=0;
while and ((abdif>con),(iii<loopnum));
   iii=iii+1;
```

```
disp('bloop iii2');
disp([bloop iii]);
grad=zeros(n1,n2-1);
epsmax=zeros(n1,n2-1);
eps=0.000001;
rangeeps=zeros(datax+totalrulenum+output,7);
check=zeros(1,6);
docheck=0;
rangeeps=range;                                          %Initialization
data=model(data,range,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';          %Normalization of the output
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
exp=(1-((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo))));
ii=0;
while ii<n1
      ii=ii+1;
      jj=1;
      disp([opvec(ii) exp]);
      while jj<6;
         jj=jj+1;
         epsp=.4*(range(opvec(ii),jj+1)-range(opvec(ii),jj));
         epsm=.4*(range(opvec(ii),jj)-range(opvec(ii),jj-1));
         epsmax(ii,jj-1)=min([epsp epsm]);
         if eps>epsmax(ii,jj-1)
            eps=epsmax(ii,jj-1);
         end;
         rangeeps(opvec(ii),jj)=range(opvec(ii),jj)+eps;
 data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
         yo=data(datax+totalrulenum+output,1:datay)';    %Normalization of the output
         xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
         xob=xo*B;
         expo=(1-((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo))));
         grad(ii,jj-1)=(expo-exp)/eps;
      end;
end;
alphamax=alpham;
rangeeps(opvec,2:n2)=range(opvec,2:n2);
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
```

```
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
f0=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
h=0;
l=0;
if f1>f0;
 while and ((h<100),(f1>f0));
      h=h+1;
      rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
 data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
      yo=data(datax+totalrulenum+output,1:datay)';
      xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
      xob=xo*B;
      f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
      alphamax=alphamax/2;
 end;
   alphamax=2*alphamax;
   rangeeps(opvec,2:n2)=range(opvec,2:n2);
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f0=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
   rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
```

```
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
   rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
elseif f1>f2;
 while and((l<100),(f1>f2))
      l=l+1;
      rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
 data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
      yo=data(datax+totalrulenum+output,1:datay)';
      xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
      xob=xo*B;
      f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
      alphamax=2*alphamax;
   end;
   alphamax=alphamax/2;
   rangeeps(opvec,2:n2)=range(opvec,2:n2);
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f0=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
   rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
   rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
```

```
    end;
    if (min([f0 f1 f2])==f1);
            sol=(inv([0              0           1;
                      (alphamax/2)^2  (alphamax/2) 1;
                      (alphamax)^2    (alphamax)   1]))*([ f0 ; f1 ; f2 ]);
            alphamax=-sol(2)/(2*sol(1));
    end;
    disp('f0 f1 f2 alphamax h l');
    disp([f0 f1 f2 alphamax h l]);
    if f1 > f0   'there is something going on wrong'; end;
    rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
    while docheck==0;
       check(1)=min(rangeeps(opvec,1)<rangeeps(opvec,2));
       check(2)=min(rangeeps(opvec,2)<rangeeps(opvec,3));
       check(3)=min(rangeeps(opvec,3)<rangeeps(opvec,4));
       check(4)=min(rangeeps(opvec,4)<rangeeps(opvec,5));
       check(5)=min(rangeeps(opvec,5)<rangeeps(opvec,6));
       check(6)=min(rangeeps(opvec,6)<rangeeps(opvec,7));
       docheck=min(check);
       if docheck==0;
           alphamax=alphamax/2;
           rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
        else;
           range=rangeeps;
        end;
    end;
    abdif=sum(sum((alphamax*grad).*(alphamax*grad)));
    disp(abdif);
end;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
con=0.00000018;
abdif=con+1;
alpham=1;
opvec=[3 16 17 18];
n1=max(size(opvec));
n2=7-1;
iii=0;
while and ((abdif>con),(iii<loopnum));
    iii=iii+1;
```

```
disp('bloop iii3');
disp([bloop iii]);
grad=zeros(n1,n2-1);
epsmax=zeros(n1,n2-1);
eps=0.000001;
rangeeps=zeros(datax+totalrulenum+output,7);
check=zeros(1,6);
docheck=0;
rangeeps=range;                                          %Initialization
data=model(data,range,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';          %Normalization of the output
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
exp=(1-((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo))));
ii=0;
while ii<n1
      ii=ii+1;
      jj=1;
      disp([opvec(ii) exp]);
      while jj<6;
         jj=jj+1;
         epsp=.4*(range(opvec(ii),jj+1)-range(opvec(ii),jj));
         epsm=.4*(range(opvec(ii),jj)-range(opvec(ii),jj-1));
         epsmax(ii,jj-1)=min([epsp epsm]);
         if eps>epsmax(ii,jj-1)
            eps=epsmax(ii,jj-1);
         end;
         rangeeps(opvec(ii),jj)=range(opvec(ii),jj)+eps;
 data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
         yo=data(datax+totalrulenum+output,1:datay)';    %Normalization of the output
         xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
         xob=xo*B;
         expo=(1-((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo))));
         grad(ii,jj-1)=(expo-exp)/eps;
      end;
end;
alphamax=alpham;
rangeeps(opvec,2:n2)=range(opvec,2:n2);
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
```

```
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
f0=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
xob=xo*B;
f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
h=0;
l=0;
if f1>f0;
 while and ((h<100),(f1>f0));
      h=h+1;
      rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
 data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
      yo=data(datax+totalrulenum+output,1:datay)';
      xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
      xob=xo*B;
      f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
      alphamax=alphamax/2;
 end;
   alphamax=2*alphamax;
   rangeeps(opvec,2:n2)=range(opvec,2:n2);
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f0=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
   rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
```

```
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
   rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
elseif f1>f2;
 while and((l<100),(f1>f2))
      l=l+1;
      rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
 data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
      yo=data(datax+totalrulenum+output,1:datay)';
      xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
      xob=xo*B;
      f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
      alphamax=2*alphamax;
   end;
   alphamax=alphamax/2;
   rangeeps(opvec,2:n2)=range(opvec,2:n2);
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f0=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
   rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax/2)*grad;
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f1=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
   rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
   data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
   yo=data(datax+totalrulenum+output,1:datay)';
   xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
   xob=xo*B;
   f2=((yo-xob)'*(yo-xob))/((yo-mean(yo))'*(yo-mean(yo)));
```

```
    end;
    if (min([f0 f1 f2])==f1);
            sol=(inv([0               0            1;
                      (alphamax/2)^2  (alphamax/2) 1;
                      (alphamax)^2    (alphamax)   1]))*([ f0 ; f1 ; f2 ]);
            alphamax=-sol(2)/(2*sol(1));
    end;
    disp('f0 f1 f2 alphamax h l');
    disp([f0 f1 f2 alphamax h l]);
    if f1 > f0   'there is something going on wrong'; end;
    rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
    while docheck==0;
       check(1)=min(rangeeps(opvec,1)<rangeeps(opvec,2));
       check(2)=min(rangeeps(opvec,2)<rangeeps(opvec,3));
       check(3)=min(rangeeps(opvec,3)<rangeeps(opvec,4));
       check(4)=min(rangeeps(opvec,4)<rangeeps(opvec,5));
       check(5)=min(rangeeps(opvec,5)<rangeeps(opvec,6));
       check(6)=min(rangeeps(opvec,6)<rangeeps(opvec,7));
       docheck=min(check);
       if docheck==0;
           alphamax=alphamax/2;
           rangeeps(opvec,2:n2)=range(opvec,2:n2)+(alphamax)*grad;
        else;
           range=rangeeps;
        end;
    end;
    abdif=sum(sum((alphamax*grad).*(alphamax*grad)));
    disp(abdif);
end;
data=model(data,rangeeps,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output);
end;                                                                    %bloop
yo=data(datax+totalrulenum+output,1:datay)';
xo=[ones(1,datay) ; data(datax+totalrulenum,1:datay) ]';
[Bo,BINT,R,RINT,STATS] =regress( yo, xo,.05);
xout=xo*Bo;
explast=(1-((yo-xout)'*(yo-xout))/((yo-mean(yo))'*(yo-mean(yo))));
t=[1996+(1/datay):((2003-1996)/datay):2003];
figure(2);
plot(t,yo,t,xout);
```

```
title(explast);
sampleyo=zeros((ceil(datay/20)-1),1);
samplexout=zeros((ceil(datay/20)-1),1);
for i=1:(ceil(datay/20)-1),
   sampleyo(i,1)=yo(1+20*i,1);
   samplexout(i,1)=xout(1+20*i,1);
end;
samplexplast=(1-((sampleyo-samplexout)'*(sampleyo-samplexout))/((sampleyo-
mean(sampleyo))'*(sampleyo-mean(sampleyo))));
t2=[(2003+1/(ceil(datay/20)-1)):(2004-2003)/(ceil(datay/20)-1):2004];
figure(22);
plot(t2,sampleyo,t2,samplexout);
title(samplexplast);
a1=[ 1 1 0 0 0 0 0];
a2=[ 0 0 1 0 0 0 0];
a3=[ 0 0 0 1 0 0 0];
a4=[ 0 0 0 0 1 0 0];
a5=[ 0 0 0 0 0 1 1];
figure(3);
plot(range(3,1:7),a1,range(3,1:7),a2,range(3,1:7),a3,range(3,1:7),a4,range(3,1:7),a5);
title(3);
figure(4);
plot(range(16,1:7),a1,range(16,1:7),a2,range(16,1:7),a3,range(16,1:7),a4,range(16,1:7),a5);
title(16);
figure(5);
plot(range(17,1:7),a1,range(17,1:7),a2,range(17,1:7),a3,range(17,1:7),a4,range(17,1:7),a5);
title(17);
figure(6);
plot(range(18,1:7),a1,range(18,1:7),a2,range(18,1:7),a3,range(18,1:7),a4,range(18,1:7),a5);
title(18);
figure(7);
y3=data(3,1:datay);
plot(t,yo,t,y3);
y3cor=corrcoef(yo,y3);
title(y3cor(1,2));
figure(8);
y16=data(16,1:datay);
plot(t,yo,t,y16);
y16cor=corrcoef(yo,y16);
```

```
title(y16cor(1,2));
figure(9);
y17=data(17,1:datay);
plot(t,yo,t,y17);
y17cor=corrcoef(yo,y17);
title(y17cor(1,2));
figure(10);
y18=data(18,1:datay);
plot(t,yo,t,y18);
y18cor=corrcoef(yo,y18);
title(y18cor(1,2));
cc2=corrcoef(data');
aa2=zeros(datax+totalrulenum+output,1);
aa2=cc2(1:datax+totalrulenum+output,datax+totalrulenum+output);
```

```
function data=model(data,range,rule,rulemat2,rulemat3,datax,datay,totalrulenum,output)
   member=zeros((datax+totalrulenum+output),datay,3);
    i=0;
   while i<(datax+output)
     i=i+1;
     if i==(datax+output)
         i=datax+output+totalrulenum;
     end;
     j=0;
     while j<datay
      j=j+1;
      if data(i,j)<range(i,2)
         member(i,j,1)=1;
         member(i,j,2)=1;
         member(i,j,3)=1;
      elseif data(i,j)<range(i,3)
```

```
        member(i,j,1)=1;
        member(i,j,2)=2;
        member(i,j,3)=(data(i,j)-range(i,3))/(range(i,2)-range(i,3));
    elseif data(i,j)<range(i,4)
        member(i,j,1)=2;
        member(i,j,2)=3;
        member(i,j,3)=(data(i,j)-range(i,4))/(range(i,3)-range(i,4));
    elseif data(i,j)<range(i,5)
        member(i,j,1)=3;
        member(i,j,2)=4;
        member(i,j,3)=(data(i,j)-range(i,5))/(range(i,4)-range(i,5));
    elseif data(i,j)<range(i,6)
        member(i,j,1)=4;
        member(i,j,2)=5;
        member(i,j,3)=(data(i,j)-range(i,6))/(range(i,5)-range(i,6));
    else
        member(i,j,1)=5;
        member(i,j,2)=5;
        member(i,j,3)=1;
     end;
   end;
end;
i=0;
numfuz=zeros(totalrulenum,datay);
while i<totalrulenum
  i=i+1;
  j=0;
  while j<datay
   j=j+1;
   if rule(i,3)==0;
      fend=4;
      numt=zeros(4,1);
      numt(1)=rulemat2(member(rule(i,1),j,1),member(rule(i,2),j,1));
      numt(2)=rulemat2(member(rule(i,1),j,2),member(rule(i,2),j,2));
      numt(3)=rulemat2(member(rule(i,1),j,1),member(rule(i,2),j,2));
      numt(4)=rulemat2(member(rule(i,1),j,2),member(rule(i,2),j,1));
      nump=zeros(4,1);
      nump(1)=min([member(rule(i,1),j,3),member(rule(i,2),j,3)]);
      nump(2)=min([(1-member(rule(i,1),j,3)),(1-member(rule(i,2),j,3))]);
```

```
            nump(3)=min([member(rule(i,1),j,3),(1-member(rule(i,2),j,3))]);
            nump(4)=min([(1-member(rule(i,1),j,3)),member(rule(i,2),j,3)]);
            numa=zeros(4,4);
            nummp=zeros(4,2);
        else
            fend=8;
            numt=zeros(8,1);
    numt(1)=rulemat3(member(rule(i,1),j,1),member(rule(i,2),j,1),member(rule(i,3),j,1));
    numt(2)=rulemat3(member(rule(i,1),j,1),member(rule(i,2),j,1),member(rule(i,3),j,2));
    numt(3)=rulemat3(member(rule(i,1),j,1),member(rule(i,2),j,2),member(rule(i,3),j,1));
    numt(4)=rulemat3(member(rule(i,1),j,1),member(rule(i,2),j,2),member(rule(i,3),j,2));
    numt(5)=rulemat3(member(rule(i,1),j,2),member(rule(i,2),j,1),member(rule(i,3),j,1));
    numt(6)=rulemat3(member(rule(i,1),j,2),member(rule(i,2),j,1),member(rule(i,3),j,2));
    numt(7)=rulemat3(member(rule(i,1),j,2),member(rule(i,2),j,2),member(rule(i,3),j,1));
    numt(8)=rulemat3(member(rule(i,1),j,2),member(rule(i,2),j,2),member(rule(i,3),j,2));
    nump=zeros(8,1);
    nump(1)=min([member(rule(i,1),j,3),member(rule(i,2),j,3),member(rule(i,3),j,3)]);
    nump(2)=min([member(rule(i,1),j,3),member(rule(i,2),j,3),(1-member(rule(i,3),j,3))]);
    nump(3)=min([member(rule(i,1),j,3),(1-member(rule(i,2),j,3)),member(rule(i,3),j,3)]);
    nump(4)=min([member(rule(i,1),j,3),(1-member(rule(i,2),j,3)),(1-member(rule(i,3),j,3))]);
    nump(5)=min([(1-member(rule(i,1),j,3)),member(rule(i,2),j,3),member(rule(i,3),j,3)]);
    nump(6)=min([(1-member(rule(i,1),j,3)),member(rule(i,2),j,3),(1-member(rule(i,3),j,3))]);
    nump(7)=min([(1-member(rule(i,1),j,3)),(1-member(rule(i,2),j,3)),member(rule(i,3),j,3)]);
nump(8)=min([(1-member(rule(i,1),j,3)),(1-member(rule(i,2),j,3)),(1-member(rule(i,3),j,3))]);
            numa=zeros(8,4);
            nummp=zeros(8,2);
        end;
        f=0;
        while f<fend;
              f=f+1;
              numa(f,1)=range(datax+i,numt(f));
              if (numt(f)==1)
                 numa(f,2)=range(datax+i,numt(f));
    numa(f,3)=range(datax+i,numt(f)+2)-((range(datax+i,numt(f)+2)-range(datax+i,numt(f)+1))*nump(f));
              elseif (numt(f)==5)
    numa(f,2)=range(datax+i,numt(f))-((range(datax+i,numt(f))-range(datax+i,numt(f)+1))*nump(f));
                 numa(f,3)=range(datax+i,numt(f)+2);
            else
    numa(f,2)=range(datax+i,numt(f))-((range(datax+i,numt(f))-range(datax+i,numt(f)+1))*nump(f));
```

```
numa(f,3)=range(datax+i,numt(f)+2)-((range(datax+i,numt(f)+2)-range(datax+i,numt(f)+1))*nump(f));
          end;
          numa(f,4)=range(datax+i,numt(f)+2);
          nummp(f,1)=(numa(f,1)+numa(f,2)+numa(f,3)+numa(f,4))/4;
          nummp(f,2)=(numa(f,4)-numa(f,1)+numa(f,3)-numa(f,2))*nump(f)/2;
    end;
    numfuz(i,j)=sum((nummp(1:fend,1).*nummp(1:fend,2)))/sum(nummp(1:fend,2));
    data(i+datax,j)=numfuz(i,j);
    if numfuz(i,j)<range(i+datax,2)
       member(i+datax,j,1)=1;
       member(i+datax,j,2)=1;
       member(i+datax,j,3)=1;
    elseif numfuz(i,j)<range(i+datax,3)
       member(i+datax,j,1)=1;
       member(i+datax,j,2)=2;
member(i+datax,j,3)=(numfuz(i,j)-range(i+datax,3))/(range(i+datax,2)-range(i+datax,3));
    elseif numfuz(i,j)<range(i+datax,4)
       member(i+datax,j,1)=2;
       member(i+datax,j,2)=3;
member(i+datax,j,3)=(numfuz(i,j)-range(i+datax,4))/(range(i+datax,3)-range(i+datax,4));
    elseif numfuz(i,j)<range(i+datax,5)
       member(i+datax,j,1)=3;
       member(i+datax,j,2)=4;
member(i+datax,j,3)=(numfuz(i,j)-range(i+datax,5))/(range(i+datax,4)-range(i+datax,5));
    elseif numfuz(i,j)<range(i+datax,6)
       member(i+datax,j,1)=4;
       member(i+datax,j,2)=5;
 member(i+datax,j,3)=(numfuz(i,j)-range(i+datax,6))/(range(i+datax,5)-range(i+datax,6));
    else
       member(i+datax,j,1)=5;
       member(i+datax,j,2)=5;
       member(i+datax,j,3)=1;
    end;
   end;
end;
```

APPENDIX – E Program Codes for Testing The Fuzzy Logic Model

(Prepared with Matlab 5.3.)

```
disp('Fuzzy Logic Model Prediction for the Istanbul Stock Exchange 100 Index');
phdddvpredict;                              %Variable data
datayp=245;                                 %Number of data in 2003
phddvpredict=zeros(datax+output,datayp);
phddvpredict(1:datax+output,1:datayp)=phdddvpredict(1+datev:datev+datax+output,1:datayp);
datap=zeros(datax+totalrulenum+output,datayp);
datap(1:datax,1:datayp)=phddvpredict(1:datax,1:datayp);
datap(datax+totalrulenum+output,1:datayp)=phddvpredict(datax+output,1:datayp);
datapcor=datap(1:datax+output,1:datayp);
datapcor(datax+output,1:datayp)=datap(datax+totalrulenum+output,1:datayp);
ccp=corrcoef(datapcor');
aap=zeros(datax+output,1);
aap=ccp(1:datax+output,datax+output);
aap=aa;
for i=1:datax,
   if aap(i,2)>0              datap(i,1:datayp)=datap(i,1:datayp);
   elseif aap(i,2)<0          datap(i,1:datayp)=-1*datap(i,1:datayp);
   end;
end;
datap=model(datap,range,rule,rulemat2,rulemat3,datax,datayp,totalrulenum,output);
yop=datap(datax+totalrulenum+output,1:datayp)';
xop=[ones(1,datayp) ; datap(datax+totalrulenum,1:datayp) ]';
xoutp=xop*Bo;
explastp=(1-((yop-xoutp)'*(yop-xoutp))/((yop-mean(yop))'*(yop-mean(yop))));
t=[(2003+1/datayp):(2004-2003)/datayp:2004];
figure(11);
plot(t,yop,t,xoutp);
title(explastp);
sampleyop=zeros((ceil(datayp/20)-1),1);
samplexoutp=zeros((ceil(datayp/20)-1),1);
for i=1:(ceil(datayp/20)-1),          sampleyop(i,1)=yop(1+20*i,1);
                                      samplexoutp(i,1)=xoutp(1+20*i,1);
```

```
end;
samplexplastp=(1-((sampleyop-samplexoutp)'*(sampleyop-samplexoutp))/((sampleyop-
mean(sampleyop))'*(sampleyop-mean(sampleyop))));
t2=[(2003+1/(ceil(datayp/20)-1)):(2004-2003)/(ceil(datayp/20)-1):2004];
figure(112);
plot(t2,sampleyop,t2,samplexoutp);
title(samplexplastp);
[Bop,BINT,R,RINT,STATS] =regress( yop, xop,.05);
xoutp2=xop*Bop;
explastp2=(1-((yop-xoutp2)'*(yop-xoutp2))/((yop-mean(yop))'*(yop-mean(yop))));
figure(12);
plot(t,yop,t,xoutp2);
title(explastp2);
figure(13);
y3p=datap(3,1:datayp);
plot(t,yop,t,y3p);
y3corp=corrcoef(yop,y3p);
title(y3corp(1,2));
figure(14);
y16p=datap(16,1:datayp);
plot(t,yop,t,y16p);
y16corp=corrcoef(yop,y16p);
title(y16corp(1,2));
figure(15);
y17p=datap(17,1:datayp);
plot(t,yop,t,y17p);
y17corp=corrcoef(yop,y17p);
title(y17corp(1,2));
figure(16);
y18p=datap(18,1:datayp);
plot(t,yop,t,y18p);
y18corp=corrcoef(yop,y18p);
title(y18corp(1,2));
ccp2=corrcoef(datap');
aap2=zeros(datax+totalrulenum+output,1);
aap2=ccp2(1:datax+totalrulenum+output,datax+totalrulenum+output);
```

APPENDIX – F Program Codes for Testing the Model's Output With Bootstrap Re-Sampling Method (Prepared with Matlab 5.3.)

```
disp('Testing Fuzzy Logic Model for the ISE100 Index With Bootstrap Method');
phdddboot;                                          %Variable data
ndm=21;                                             %Number of working days in a month
nmy=12;                                             %Number of months in a year
datayb=252+ndm;                                     %Number of data
dataxb=6;                                           %Number of columns
comm=2.1/1000;                                      %Transaction Commision
bootsub=20000;                                      %Number of bootstrap sample
phdbs=zeros(dataxb+1,datayb);
phdbs(1,1:datayb)=phdddboot(1,1:datayb);
phdbs(2,1:datayp)=xoutp(1:datayp)';
phdbs(3:dataxb+1,1:datayb)=phdddboot(2:dataxb,1:datayb);
serymodel=zeros(ndm,nmy);
seryise=zeros(ndm,nmy);
seryusd=zeros(ndm,nmy);
seryrepo=zeros(ndm,nmy);
serytl=zeros(ndm,nmy);
serysug=zeros(ndm,nmy);
for i=1:ndm,
   for j=1:nmy,
      if phdbs(2,(i+(ndm*(j-1))))>=0,
         serysug(i,j)=1;
         serymodel(i,j)=(phdbs(4,(i+(ndm*(j-1))+21))/phdbs(4,(i+(ndm*(j-1)))))-1;
         if (i+(ndm*(j-1)))>21,                     %Transaction Commision Paid for Buying
            if phdbs(2,(i+(ndm*(j-1))-21))<=0, serymodel(i,j)=((1+serymodel(i,j))/(1+comm))-1;
            end;                                    %Transaction Commision Paid for Buying
         else;     serymodel(i,j)=((1+serymodel(i,j))/(1+comm))-1;
         end;
      else;
         serysug(i,j)=-1;
         serymodel(i,j)=(phdbs(6,(i+(ndm*(j-1))+21))/phdbs(6,(i+(ndm*(j-1)))))-1;
         if (i+(ndm*(j-1)))>21,
```

```
            if phdbs(2,(i+(ndm*(j-1))-21))>0, %Transaction Commision Paid for Selling
               serymodel(i,j)=((1+serymodel(i,j))/(1+comm))-1;
            end;
         end;
      end;
      seryise(i,j)=(phdbs(4,(i+(ndm*(j-1))+21))/phdbs(4,(i+(ndm*(j-1)))))-1;
      seryusd(i,j)=(phdbs(5,(i+(ndm*(j-1))+21))/phdbs(5,(i+(ndm*(j-1)))))-1;
      seryrepo(i,j)=(phdbs(6,(i+(ndm*(j-1))+21))/phdbs(6,(i+(ndm*(j-1)))))-1;
      serytl(i,j)=(phdbs(7,(i+(ndm*(j-1))+21))/phdbs(7,(i+(ndm*(j-1)))))-1;
   end;
end;
modise=zeros(ndm,nmy);
modusd=zeros(ndm,nmy);
modrepo=zeros(ndm,nmy);
modtl=zeros(ndm,nmy);
modise=serymodel-seryise;
modusd=serymodel-seryusd;
modrepo=serymodel-seryrepo;
modtl=serymodel-serytl;
cmodise=zeros(ndm,nmy);
cmodusd=zeros(ndm,nmy);
cmodrepo=zeros(ndm,nmy);
cmodtl=zeros(ndm,nmy);
for i=1:ndm,
   j=1;
   cmodise(i,j)=(1+modise(i,j));
   cmodusd(i,j)=(1+modusd(i,j));
   cmodrepo(i,j)=(1+modrepo(i,j));
   cmodtl(i,j)=(1+modtl(i,j));
end;
for i=1:ndm,
   for j=2:nmy,
      cmodise(i,j)=((1+modise(i,j))*(cmodise(i,j-1)));
      cmodusd(i,j)=((1+modusd(i,j))*(cmodusd(i,j-1)));
      cmodrepo(i,j)=((1+modrepo(i,j))*(cmodrepo(i,j-1)));
      cmodtl(i,j)=((1+modtl(i,j))*(cmodtl(i,j-1)));
   end;
end;
benisem=mean(cmodise(1:ndm,nmy)-1);
```

```
benises=std(mean(cmodise'));
beniset=benisem/benises;
benusdm=mean(cmodusd(1:ndm,nmy)-1);
benusds=std(mean(cmodusd'));
benusdt=benusdm/benusds;
benrepom=mean(cmodrepo(1:ndm,nmy)-1);
benrepos=std(mean(cmodrepo'));
benrepot=benrepom/benrepos;
bentlm=mean(cmodtl(1:ndm,nmy)-1);
bentls=std(mean(cmodtl'));
bentlt=bentlm/bentls;
disp('benisem   benusdm   benrepom   bentlm');
disp('benises   benusds   benrepos   bentls');
disp('beniset   benusdt   benrepot   bentlt');
disp([benisem benusdm benrepom bentlm]);
disp([benises benusds benrepos bentls]);
disp([beniset benusdt benrepot bentlt]);
subsampmat=zeros(bootsub,ceil(ndm/4));
subsampmat=round(rand(bootsub,ceil(ndm/4))*20)+1;
zise=zeros(bootsub,1);
zusd=zeros(bootsub,1);
zrepo=zeros(bootsub,1);
ztl=zeros(bootsub,1);
for i=1:bootsub,
   testise=zeros(ceil(ndm/4),1);
   for j=1:ceil(ndm/4),   testise(j,1)=cmodise(subsampmat(i,j),nmy)-1;   end;
   zise(i)=(mean(testise(1:ceil(ndm/4)),1)-benisem)/std(testise(1:ceil(ndm/4)),1);
   testusd=zeros(ceil(ndm/4),1);
   for j=1:ceil(ndm/4),   testusd(j,1)=cmodusd(subsampmat(i,j),nmy)-1;   end;
   zusd(i)=(mean(testusd(1:ceil(ndm/4)),1)-benusdm)/std(testusd(1:ceil(ndm/4)),1);
   testrepo=zeros(ceil(ndm/4),1);
   for j=1:ceil(ndm/4),   testrepo(j,1)=cmodrepo(subsampmat(i,j),nmy)-1; end;
   zrepo(i)=(mean(testrepo(1:ceil(ndm/4)),1)-benrepom)/std(testrepo(1:ceil(ndm/4)),1);
   testtl=zeros(ceil(ndm/4),1);
   for j=1:ceil(ndm/4),   testtl(j,1)=cmodtl(subsampmat(i,j),nmy)-1;      end;
   ztl(i)=(mean(testtl(1:ceil(ndm/4)),1)-bentlm)/std(testtl(1:ceil(ndm/4)),1);
end;
zise=sort(zise);                              %Pivotal statistic is calculated
zusd=sort(zusd);
```

```
zrepo=sort(zrepo);
ztl=sort(ztl);
disp('2 percent   5 percent   10 percent   20 percent');              %outputs
disp('98 percent   95 percent   90 percent   80 percent');
disp([zise(bootsub*0.02,1) zise(bootsub*0.05,1) zise(bootsub*0.10,1)
zise(bootsub*0.20,1)]);
disp([zise(bootsub*0.98,1) zise(bootsub*0.95,1) zise(bootsub*0.90,1)
zise(bootsub*0.80,1)]);
disp('2 percent   5 percent   10 percent   20 percent');
disp('98 percent   95 percent   90 percent   80 percent');
disp([zusd(bootsub*0.02,1) zusd(bootsub*0.05,1) zusd(bootsub*0.10,1)
zusd(bootsub*0.20,1)]);
disp([zusd(bootsub*0.98,1) zusd(bootsub*0.95,1) zusd(bootsub*0.90,1)
zusd(bootsub*0.80,1)]);
disp('2 percent   5 percent   10 percent   20 percent');
disp('98 percent   95 percent   90 percent   80 percent');
disp([zrepo(bootsub*0.02,1) zrepo(bootsub*0.05,1) zrepo(bootsub*0.10,1)
zrepo(bootsub*0.20,1)]);
disp([zrepo(bootsub*0.98,1) zrepo(bootsub*0.95,1) zrepo(bootsub*0.90,1)
zrepo(bootsub*0.80,1)]);
disp('2 percent   5 percent   10 percent   20 percent');
disp('98 percent   95 percent   90 percent   80 percent');
disp([ztl(bootsub*0.02,1) ztl(bootsub*0.05,1) ztl(bootsub*0.10,1) ztl(bootsub*0.20,1)]);
disp([ztl(bootsub*0.98,1) ztl(bootsub*0.95,1) ztl(bootsub*0.90,1) ztl(bootsub*0.80,1)]);
pise=0;                                                  %p value is calculated
pusd=0;
prepo=0;
ptl=0;
for i=1:bootsub,
   if beniset>zise(i,1),          pise=pise+1;       end;
   if benusdt>zusd(i,1),          pusd=pusd+1;       end;
   if benrepot>zrepo(i,1),        prepo=prepo+1;     end;
   if bentlt>ztl(i,1),            ptl=ptl+1;         end;
end;
disp('p_ise   p_usd   p_repo   p_tl 20');
disp([pise/bootsub pusd/bootsub prepo/bootsub ptl/bootsub]);
```