AN EMPIRICAL STUDY ON FUZZY C-MEANS CLUSTERING FOR TURKISH BANKING SYSTEM

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

FATİH ALTINEL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF ECONOMICS

SEPTEMBER 2012

Approval of the Graduate School of Social Sciences.

_____
Prof. Dr. Meliha ALTUNIŞIK
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

_____
Prof. Dr. Erdal ÖZMEN
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

_____
Asst. Prof. Dr. Esma GAYGISIZ
Supervisor

Examining Committee Members

Asst. Prof. Dr. Esma GAYGISIZ	(METU, ECON)	_____
Prof. Dr. Erdal ÖZMEN	(METU, ECON)	_____
Dr. Cihan YALÇIN	(TCMB)	_____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Fatih ALTINEL

Signature:

ABSTRACT

AN EMPIRICAL STUDY ON FUZZY C-MEANS CLUSTERING FOR TURKISH BANKING SYSTEM

Altınel, Fatih

M.S., Department of Economics

Supervisor : Asst. Prof. Dr. Esma GAYGISIZ

September 2012, 94 pages

Banking sector is very sensitive to macroeconomic and political instabilities and they are prone to crises. Since banks are integrated with almost all of the economic agents and with other banks, these crises affect entire societies. Therefore, classification or rating of banks with respect to their credibility becomes important. In this study we examine different models for classification of banks. Choosing one of those models, fuzzy c-means clustering, banks are grouped into clusters using 48 different ratios which can be classified under capital, assets quality, liquidity, profitability, income-expenditure structure, share in sector, share in group and branch ratios. To determine the inter-dependency between these variables, covariance and correlation between variables is analyzed. Principal component analysis is used to decrease the number of factors. As a a result, the representation space of data has been reduced from 48 variables to a 2 dimensional space. The observation is that 94.54% of total variance is produced by these

two factors. Empirical results indicate that as the number of clusters is increased, the number of iterations required for minimizing the objective function fluctuates and is not monotonic. Also, as the number of clusters used increases, initial non-optimized maximum objective function values as well as optimized final minimum objective function values monotonically decrease together. Another observation is that the 'difference between initial non-optimized and final optimized values of objective function' starts to diminish as number of clusters increases.

Keywords: Fuzzy C-Means Clustering, Bank Clustering, Bank Rating, Bank Classification, Bank Failures

ÖZ

BANKACILIK SEKTÖRÜNÜN BULANIK KÜMELENMESİ ÜZERİNE DENEYSEL BİR ÇALIŞMA

Altınel, Fatih

Yüksek Lisans, İktisat Bölümü

Tez Yöneticisi : Yrd. Doç. Dr. Esma GAYGISIZ

Eylül 2012, 94 sayfa

Makro iktisadi ve siyasal düzensizliklere karşı duyarlı olan bankacılık sistemi, krizlere maruz kalabilmektedir. Bu yüzden, banka güvenilirliğinin ölçülmesi bu açıdan önem arz etmektedir. Bu çalışmada bankaların gruplanması için bulanık kümeleme yöntemi kullanılmıştır. Yöntem uygulanırken finansal oran olarak sermaye, varlık kalitesi, likidite, karlılık, gelir-gider yapısı, sektör payı, şube yapısı gibi 48 adet oran kullanılmıştır. Finansal oranlar arasındaki bağımlılığın tespiti için kovaryans ve korelasyon matrisleri incelenmiştir. Ardından, ana bileşen analizi (PCA) kullanılarak 48 değişkenle temsil edilen uzay 2 değişkenle temsil edilen başka bir uzaya indirgenmiştir. Seçilen iki değişkenin toplam varyansın %94.54 ünü temsil ettiği görülmektedir. Çalışmalar göstermiştir ki, küme sayısı artırıldıkça hedef fonksiyonunun değerinin minimize edilmesi için gereken işlem sayısı değeri tekdüze olmayıp belirli değerler arasında salınım yapmaktadır. Ayrıca, kullanılan küme sayısı artırıldıkça hedef fonksiyonunun her bir küme için hesaplanan ilk değeri ve son değeri tekdüze bir şekilde azalmaktadır. Ayrıca, kullanılan küme sayısı

artırıldıkça hedef fonksiyonunun her bir küme için hesaplanan ilk değeri ve son değeri arasındaki farkların da tekdüze bir şekilde azaldığı görülmüştür.

Anahtar sözcükler: Bulanık Kümeleme, Banka Kümeleme, Banka Derecelendirme, Banka Sınıflandırma, Banka Başarısızlığı.

ACKNOWLEDGMENTS

The author wishes to express his deepest gratitude to his supervisor Asst. Prof. Dr. Esma GAYGISIZ for her guidance, advice, criticism, encouragements and insight throughout the research.

The author would also like to thank Prof. Dr. Erdal ÖZMEN and Dr. Cihan Yalçın for their suggestions and comments.

The technical assistance of Mr. İsmail GÖKGÖZ and Mustafa Fatih BOYRAZ is gratefully acknowledged.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

BFP	Business Failure Prediction
BPNN	Back Propagation Trained Neural Network
BRSA	Banking Regulation and Supervision Agency
BSH	Best Separating Hyperplane
CAR	Capital Adequacy Ratio
CBR	Case Based Reasoning
DA	Discriminant Analysis
FCM	Fuzzy c-Means Clustering
FSI	Financial Soundness Indicators
$\boldsymbol{k}$**-NN**	k-Nearest Neighbor
LDA	Linear Discriminant Analysis
LP	Linear Programming
Logit	Logistic Regression
MARS	Multivariate Adaptive Regression Splines
MDA	Multivariate Discriminant Analysis
MSD	Minimized Sum of Deviations
PCA	Principal Component Analysis
ROA	Return on Assets
SME	Small and Medium Enterprises
SV	Support Vector
SVM	Support Vector Machine

CHAPTER 1

INTRODUCTION

In financial markets, credit risk is the oldest form of risk (Caouette, Altman, Narayanan, & Nimmo, 2008). It can be defined as "expectation that some amount of money which is owed will not be paid back within a certain limited time". Credit risk is as old as lending itself, which means that it dates back to ancient times. Just like in ancient times, today there is the element of uncertainty that whether a given borrower will repay a particular loan.

In an economy, banking sector holds the role of allocating financial resources and in that sense; they play an important role in economic growth. Nevertheless, banking sector is very sensitive to macroeconomic and political instabilities and they are vulnerable to crises. Since banks are integrated with almost all of the economic agents and with other banks, these crises affect entire societies. The sole cost of recent global crisis according to IMF (2010) is 3283,5 billion dollars and according to FDIC (2012), the number of unsuccessful US banks reached 494 in the recent global crisis.

In this study different models for classification of banks are examined. Choosing one of those models, fuzzy c-means clustering, banks are grouped into clusters using 48 different ratios which can be classified under capital, assets quality, liquidity, profitability, income-expenditure structure, share in sector, share in group and branch ratios. The main research questions are: "How can the banks in Turkish banking system be grouped using fuzzy c-means clustering". "How does the iteration time change as number of clusters is increased?" "How can the clusters of banks be ranked so that their group-by-group credibility compared."

In chapter 2, a brief description of credit risk measurement models have been given under the following headlines (i) credit scoring systems; (ii) intensity based models; and (iii) structural models which are based on Merton's model and its variations. Most of the emphasis has been given to credit scoring models since fuzzy c-means clustering is a

member of credit scoring family. Linear discriminant analysis, logit and probit models, support vector machines, k nearest neighbors algorithm and fuzzy c-means clustering have been studied.

In chapter 3, the theoretical and empirical literature on most widely used credit scoring models have been given in an organized manner. Similarly, literature on linear discriminant analysis, logit and probit models, support vector machines, k nearest neighbors and fuzzy c-means clustering algorithms have been studied.

In chapter 4, the experimental design used in this study is described. Data collection and preparation step is reviewed firstly. Then brief information on analysis is given. In this part, fuzzy clustering algorithm for 23 banks, 48 ratios and 6 clusters has been applied. Also, additional explanation has been given regarding the ranking of these clusters.

In chapter 5, empirical results for fuzzy 6-means clustering are shown. Cluster membership values for banks have been given; banks have been assigned to the most probable cluster using membership values. Further, a heuristic approach has been used to rank these clusters. These results are briefly discussed.

Finally, chapter 6 is the section which summarizes the findings of the study and reviews the results mentioned in chapter 5 with appropriate conclusions.

CHAPTER 2

CREDIT RISK MEASUREMENT MODELS

Rating systems and credit scoring systems are used to evaluate the riskiness of default for the evaluated entity. Calculations are based on the analysis of entity's history and long term economic prospects. The literature on credit risk measurement covers three main types of models: (i) credit scoring systems; (ii) intensity based models; and (iii) structural models which are based on Merton's model and its variations.

In order to be able to segregate "good borrowers" from "bad borrowers" in terms of banks' creditworthiness, mathematical and statistical approaches are required. These models help the banks not only in evaluating customers' riskiness, but also help in competitiveness, as some researchers may argue. (Allen, 1995)

2. 1. Credit Scoring Models[1]

Mathematical and statistical methods are used in credit scoring techniques. In 1932, Fitzpatrick studied the dependence between probability of default and characteristics of corporate credits (Fitzpatrick, 1932). Also Fisher came up with the idea of discriminant analysis to study the relationships between subgroups in a population (Fisher, 1936). In the following period, Durand's study has been published containing discriminant analysis methods for segregating good consumer loans from bad ones (Durand, 1941). Following the WW II, these methods have been used by large companies for different purposes such as marketing (Servigny & Renault, 2004).

According to Stephey (2009), penetration rate of credit cards as a payment method was not very high until 1950s. During the postwar boom in 1950 Diner's Club issued its first card, in 1958 American Express issued cards while Visa issued its first 60000 cards (Stephey, 2009). Increased penetration rate of credit cards resulted in a collection of a large pile of

[1] This section heavily relies on (Servigny & Renault, 2004, pp. 73-88)

consumer data. Also, considering the size of population of card users, automated analysis for lending decisions became a necessity. With the help of computerization in the 1970s and '80s, analyzing such data became comparatively easy. In USA, credit scoring concept has been recognized in The Equal Credit Opportunity Act Of 1974 and 1976 (Smith, 1977). According to this act “A creditor shall not discriminate against any applicant on the basis of sex or marital status in any aspect of a credit transaction.” Moreover, this legislation stated that any discrimination in the granting of credit was outlawed unless it was based on statistical assessments. Thus, legislations had clearly mentioned the usage of credit scoring in evaluation of credits.

In 1960s, several improvements and extensions have been introduced regarding methodology. Credit scoring techniques have been extended from credit card loans to other asset classes. Myers and Forgy compared regression and discriminant analysis for credit scoring applications (Myers & Forgy, 1963). In the following period, Beaver studied bankruptcy prediction models (Beaver, 1966), and Altman applied multiple discriminant credit scoring analysis (MDA) which can be used to classify a firm’s creditworthiness, by assigning a Z-score to it (Altman, 1968). Later, Martin (1977), Ohlson (1980), and Wiginton (1980) dealt with logit analysis and they were the first to apply logit analysis for bankruptcy prediction studies.

In the above mentioned models, failure prediction and the classification of credit quality are both focused. In that sense, distinction between prediction and classification is actually important. The reason is that it may be unclear for users of scores that whether classification or prediction is the most important aspect to focus on. An expert using these scoring models may have difficulties when selecting criteria for performance measures if this distinction is not clear. (Servigny & Renault, 2004)

Credit scoring has been widely used in banks. In USA, The Federal Reserve's November 1996 Senior Loan Officer Opinion Survey of Bank Lending Practices showed that 97 percent of U.S. banks were using internal scoring models for credit card applications. Over the last years, these figures have increased especially with the Basel II focus on probabilities of default. In 1995, Fair, Isaac and Company (FICO) introduced its Small Business Credit Scoring model using data from 17 banks (Servigny & Renault, 2004). Today there are several

providers of credit models and credit information services (i.e. FICO Score, FICO NextGen Score, VantageScore, CE Score) in USA market. (myFICO.com, 2012)

In Turkish legislation, according to Article 9[2] of Bank Cards and Credit Cards Law No: 5464, card issuers are obliged to implement a policy to limit card usage per customer which should be consistent with each customer's credit score (BRSA, 2006). Moreover, Circular Regarding Credit Risk Management (draft document) forces banks to employ policies to limit card usage (BRSA, 2011).

Servigny & Renault (2004) state that the major appeal of scoring models is related to a competitive issue and that scoring models enable growth of productivity by providing a credit assessment with reduced costs in a limited time frame. Researchers report that based on credit scoring, the traditional small-business loan approval process averages about 12 hours per loan, a process that took up to 2 weeks in the past. (Allen, 1995)

In another research, the impact of scoring systems on bank loans given to SME's in the USA has been analyzed for 1995-1997 period. The study has shown that scoring system usage in banks is beneficial since it tends to increase the appetite of banks for that type of risk and that it reduces adverse selection by enabling "marginal borrowers" with higher risk profiles to be financed more easily, but at an appropriate price (Berger, Frame, & Miller, 2002).

Based on an existing data set, selecting the optimal model is not very easy. (Servigny & Renault, 2004). In a comparative study, Galindo and Tamayo (2000) have defined five requisite qualities for the choice of an optimal scoring model: i) Accuracy: Having low error rates, ii) Parsimony: Using small number of explanatory variables, iii) Nontriviality: Outputs and results of the model should be interesting iv) Feasibility: Time and resources required to run the model should be reasonable. v) Transparency and interpretability: A model should provide an abstract level insight into the data relationships and trends.

[2] Article 9: Card issuing organizations are under obligation to determine and apply a limit of use of cards as a result of an assessment to be carried on by them by taking into consideration the prohibitions inflicted on or the legal incompetence of, the economic and social status of, and the monthly or yearly average income of the persons who apply for a credit card, as well as the existing credit card limits previously allocated to these persons by the other card issuing organizations, and the results of a modeling or scoring system, and the "know-your-customer" principles, and the information to be provided pursuant to Article 29 hereof. Card issuing organizations may update the current card limits in accordance with these provisions. However, card issuing organizations may not increase card limits unless otherwise demanded by the card holder.

Overview of the various classes of models is given in *Figure 1*. This classification is taken from Servigny & Renault (2004) and resketched.

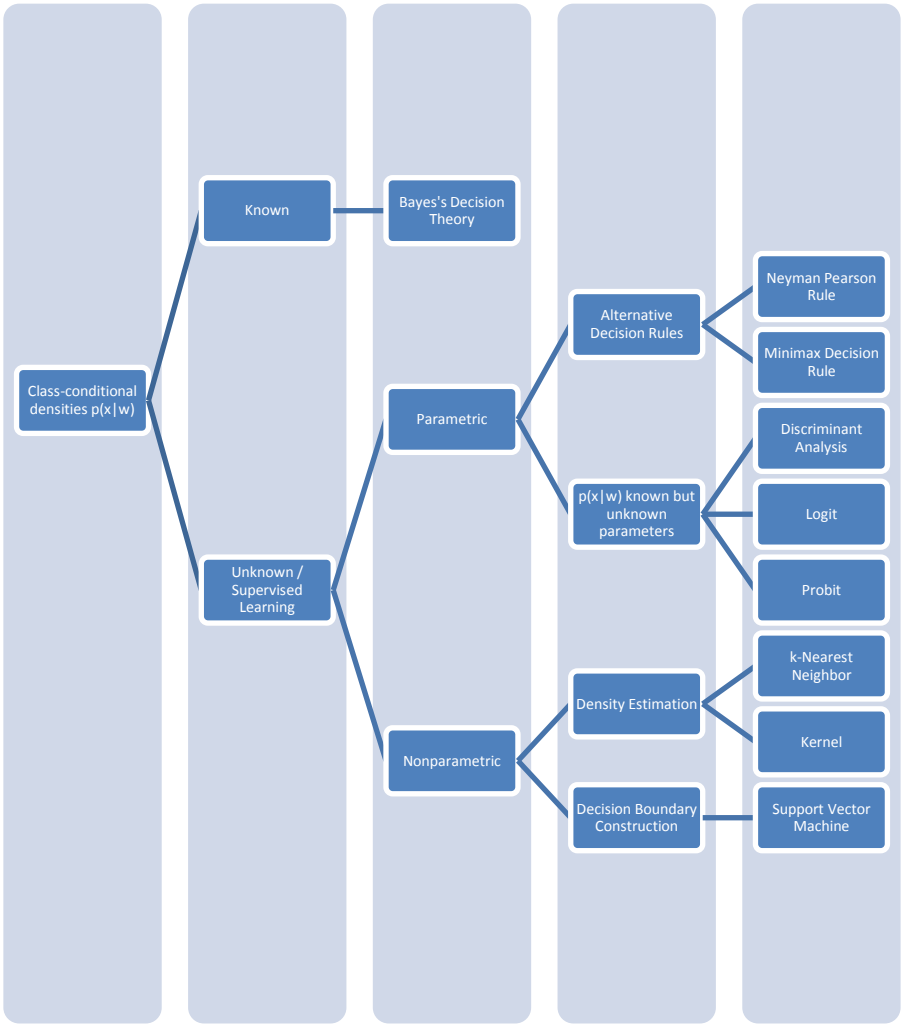


Figure 1: Overview of the various classes of credit scoring models

In statistical pattern recognition literature, there are two classification types, namely supervised and unsupervised classification. (Webb & Copsey, 2011) In supervised classification, the researcher knows according to which criteria he wants to classify firms between groups. In unsupervised classification, the researcher does not know but actually tries to determine the groupings themselves from the model. In the context of this study, only supervised classification case will be covered.

It can be said that, most widely used techniques in the credit scoring field are: i) Fisher linear discriminant analysis, ii) Logistic regression and probit (logit/probit), iii) K-nearest neighbor classifier, iv) Support vector machine classifier. (Servigny & Renault, 2004)

Many other credit scoring techniques exist, namely genetic algorithms, tree induction methods, and neural networks etc. A more detailed and extensive list can be found in statistical pattern recognition books such as Webb & Copsey (2011).

In literature, discriminant and logit/probit classifiers are mostly used as front-end tools[3] since it is easier to understand and to implement these methods; they also give fast and accurate results. On the other hand, nearest neighbor and support vector machines can be seen as back-end tools, since longer computational time is required.

2.1. 1 Linear Discriminant Analysis[4]

Linear discriminant analysis algorithm segregates and classifies a heterogeneous population into different homogeneous subsets. In order determine a useful decision rule, various decision criteria (Perceptron Criterion, Fisher's Criterion etc.) are used.

The procedure works as following: First, number of classes in which data will be segregated is determined. To exemplify, in the case of a credit scoring model, there can be, for example, two classes: default and non-default. Secondly, model searches for the linear combination of explanatory variables which separates the two classes most. That is, optimum linear combination occurs with the optimum weight vector that maximizes the distance between classes, which guarantees maximum separation between groups.

In this study, initial assumption is that the decision boundaries are linear. This analysis may be divided into two parts: the binary classification problem and the multiclass problem. For simplicity, only binary classification algorithm will be described, although the two-class problem is a special case of the multiclass situation.

[3] Front-end users are salespeople who look for a fast and robust model which performs well for new customers.

[4] This sub-section heavily relies on (Webb & Copsey, 2011) and (Servigny & Renault, 2004).

Suppose the set of training patterns is $x_1, x_2, \dots, x_n$ and each x_i is assigned to a class w_1 or w_2. A weight vector w and a threshold value w_o needs to be found such that

$$w^T x + w_o \begin{cases} > 0 \\ < 0 \end{cases} \rightarrow x \in \begin{cases} w_1 \\ w_2 \end{cases} \qquad (2.1)$$

The linear discriminant rule given by Equation (2.1) can be simplified as:

$$v^T z \begin{cases} > 0 \\ < 0 \end{cases} \rightarrow x \in \begin{cases} w_1 \\ w_2 \end{cases}$$

where $z = \left(x_1, x_2, \dots, x_p\right)^T$ is the augmented pattern vector and v is a $(p+1)$ dimensional vector $v = \left(w_0, w_1, \dots, w_p\right)^T$.

A sample in class w_2 is classified correctly if $v^T z < 0$. Let $y = -z = \left(-x_1, -x_2, \dots, -x_p\right)^T$ then $v^T y > 0$.

In an ideal case, a solution for v makes $v^T y$ positive for as many samples as possible, which minimizes the misclassification error. If $v^T y_i > 0$ for all y_i then the data is called linearly separable. Nevertheless, minimizing the number of misclassifications is not very easy. To minimize this number, other criteria may be used.

Perceptron Criterion

Perceptron criterion function is the simplest criterion to be minimized in linear discriminant analysis method. It can be denoted as:

$$J_P(v) = \sum_{y_i \in \mathcal{Y}} -v^T y_i$$

where $\mathcal{Y} = \{y_i | v^T y_i < 0\}$ corresponds to the set of misclassified samples.

It can be seen that the criterion function is continuous which makes it possible to use gradient based optimization procedures. To find the minimum, simply take partial derivative of criterion function to find:

$$\frac{\partial J_P}{\partial v} = \sum_{y_i \in \mathcal{Y}} -y_i$$

Then using the steepest descent method following update rule is found:

$$v_{k+1} = v_k + \rho_k \sum_{y_i \in Y} y_i$$

where ρ_k is a parameter that determines the step size. If samples are separable, then this procedure guarantees to converge to a solution that separates the sets. Then updating pattern is:

$$v_{k+1} = v_k + \rho_k y_i$$

where y_i is a training sample which has been misclassified by v_k. This procedure cycles through the training set, and modifies the weight vector when and if a sample is misclassified.

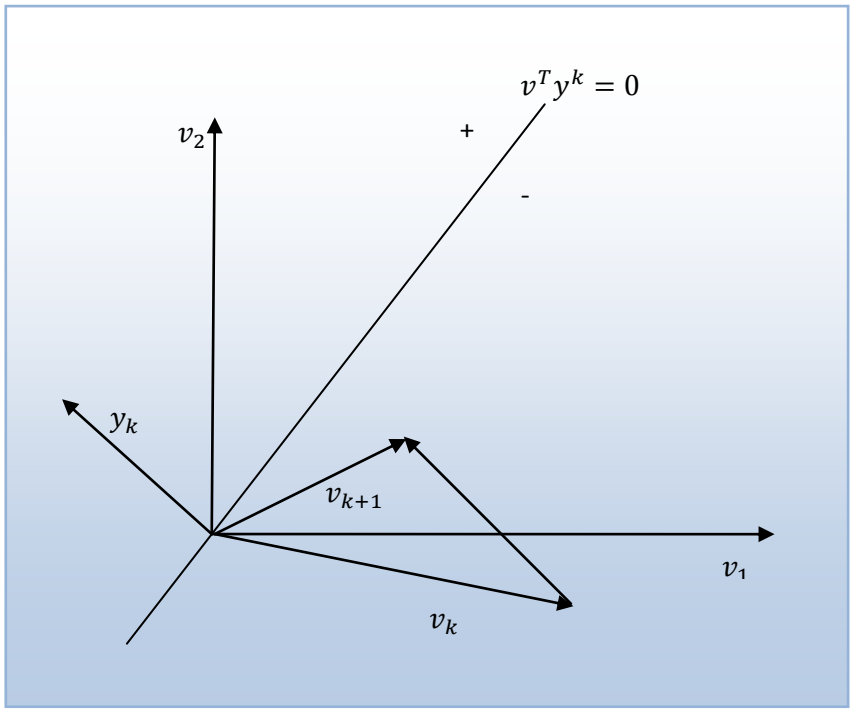


Figure 2: Graph of error correction algorithm

The visual graph of error correction in weight space can be seen in *Figure 2* and *Figure 3*. Let $\rho_k = \rho = 1$. In *Figure 3*, the plane is partitioned by the straight line $v^T y_k = 0$. Since currently $v^T y_k < 0$, the weight vector v_k is updated by addition of the pattern vector y_k.

This updating moves the weight vector v_k towards the region $v^T y_k > 0$. To exemplify, the lines $v^T y_k = 0$ are graphed for four separate training patterns in *Figure 3*.

If the classes are separable, the solution for v lies in the shaded region of *Figure 3*. (solution point is in the region where for which $v^T y_k > 0$ for all y_k). For separable patterns,, eventually a solution with $J_P(v) = 0$ will be obtained.

In this context, for perceptron criterion, only fixed increment rule has been applied and ρ has been assumed constant. Although they will not be explained in this study, other variants may be used for ρ, some of which may be listed as i) absolute correction rule, ii) Fractional correction rule, iii) Introduction of a margin b, iv) Variable increment ρ, v) Relaxation algorithm.

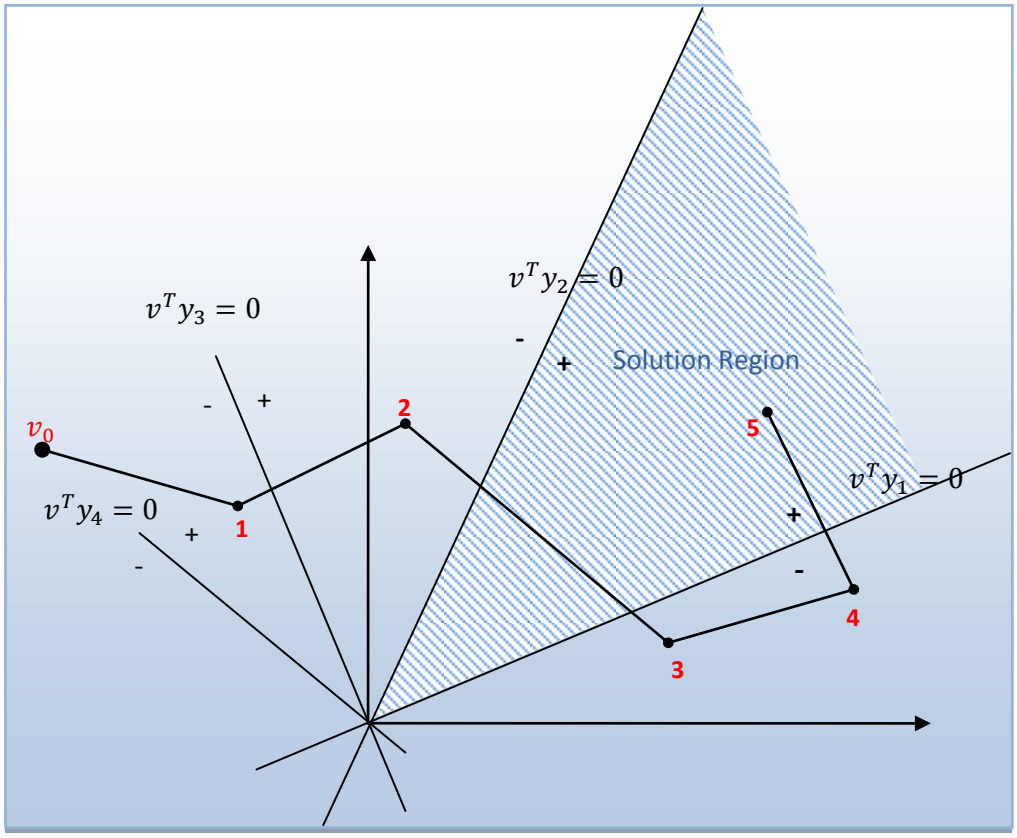


Figure 3: Graph of Error Correction on Weight Space (Inspired from (Servigny & Renault, 2004))

Fisher's Criterion

The idea is to find a linear combination of the variables that separates the two classes as much as possible. In other words, the direction along which the two classes are best separated is searched. The criterion proposed and used by Fisher is the ratio of between-class to within-class variances. Mathematically, a direction w is searched such that:

$$J_F = \frac{|w^T(m_1 - m_2)|^2}{w^T S_W w}$$

is maximized, where m_1 and m_2 are the class means and S_W is the within-class sample covariance matrix, and is given as following:

$$S_W = \frac{1}{n-2}\left(n_1\widehat{\sum_1} + n_2\widehat{\sum_2}\right)$$

where $\widehat{\sum_1}$ and $\widehat{\sum_2}$ are the maximum likelihood estimates of the covariance matrices of classes w_1 and w_1 respectively. n_i corresponds to sample size in class w_i where $n_1 + n_2 = n$.

A solution for the direction w is found by maximizing Fisher's criterion J_F. By differentiating J_F with respect to w and equating it to zero, the following equation is obtained:

$$\frac{2w^T(m_1 - m_2)}{w^T S_W w}\left\{(m_1 - m_2) - \left(\frac{w^T(m_1 - m_2)}{w^T S_W w}\right)S_W w\right\} = 0$$

since $\frac{w^T(m_1 - m_2)}{w^T S_W w} = Q$ is a scalar, the direction w that maximizes Fisher's criterion J_F is calculated as:

$$w = Q^{-1}S_w^{-1}(m_1 - m_2)$$

2.1. 2 Parametric Discrimination Analysis

Fisher's analysis provides a very clear segregation with respect to classes. That is, the classes $w_1$or w_2 are clearly segregated and a bank either falls into class w_1 or w_2. If

$S(x) = w^{\mathrm{T}}x + \propto$ is chosen as the score function and $\propto$ is determined, then it can be said that a clear-cut segregation between classes w_1 and w_2 is established. As it can be seen, that model It does not provide probabilities of being in classes w_1 or w_2.

Most of the times, the information of just being in a class w_1 or w_2 is not enough, in addition to that, probability of being in a certain class may be useful. One approach could be applying linear regression. Letting $S(x) = p(w_2|x)$ so that the probability of being in class w_2 (for example, rate B) conditional on realized values of variables x (for example, financial ratios), is simply equal to the score. This is not a very suitable technique due to differences in structure of probability function and score since the probability is bounded and is between 0 and 1, but we know that generally the score is not bounded.

To get rid of this problem, score value may be transformed so that the probability correspondent falls into the interval [0,1].

Linear Logit and Probit Models

Let the following relationship holds:

$$S(x) = w^{\mathrm{T}}x + \propto = \log\left(\frac{p(x|w_1)}{p(x|w_2)}\right)$$

It can be shown that:

$$p(w_2|x) = \frac{1}{1 + exp(w^{\mathrm{T}}x + \beta)} = \frac{1}{1 + exp\left(\sum_{i=1}^{p} w_i x_i + \beta\right)}$$

Where $\beta = \propto + \log\left(\frac{p(w_1)}{p(w_2)}\right)$

β is a constant and it corresponds to the ratio of banks in two classes.

In the $S(x) = w^{\mathrm{T}}x + \propto = \log\left(\frac{p(x|w_1)}{p(x|w_2)}\right)$ expression, conditional probability of a bank being in class w_2 is shown to be conditional on realized values of its variables (financial ratios). Since the probability probability follows a logistic law, the model is named as *logit* model. Instead of logistic, other transformations may be used. As an example, if normal distribution is used, the model is called as *normit* model, another name for normit model is *probit* model.

In logit or probit models, maximum likelihood estimation is used for estimation of parameters. The normal distribution has thinner tails than the logistic distribution but this difference can be ignored if the sample does not include too many extreme observations.

The decision rule for classifying patterns is similar to the linear discriminant classification scheme. That is, a bank with observed financial ratios x will be assigned to class w_1 if $w^{\mathrm{T}}x+\propto>0$; otherwise it will be assigned to class w_2. The value for w is estimated using maximum likelihood estimation procedures. As the size for training sample increases, w gets closer to its unknown true value.

Nonlinear Logit Models

Linear logit models assume that a linear relationship exists between the score and the variables. More complex relationship types such as non-monotonicity is ignored in linear models.

In their article, Laitinen and Laitinen (2000) suggest that a nonlinear transformation to the factors T(x) may be applied so that:

$$S(x) = w^{\mathrm{T}}T(x)+\propto= \log\left(\frac{p(x|w_1)}{p(x|w_2)}\right)$$

Where where *Box-Cox function* is used as the data transformation function:

$$T_i(x_i) = \frac{{x_i}^{\gamma_i} - 1}{\gamma_i}$$

where γ_i is a convexity parameter such that the transformation is concave whenever $\gamma_i < 1$; and convex otherwise.

Other transformation functions may be used, such as *quadratic logit*. It is similar to the linear logit model may also be used but instead of considering only the first-order terms, second-order terms are also included in the model.

$$F(y) = \frac{1}{1 + exp\left(\beta + \sum_{i=1}^{p}\delta_i x_i + \sum_{i=1}^{p}\sum_{j=1}^{p}\gamma_{ij}x_i x_j\right)}$$

Where δ_i and γ_{ij} are weights and β is a constant. The model includes the nonlinear relationship between the score and each explanatory variable. Moreover, the model includes the interactions between the variables since it includes the product $x_i x_j$.

Generally, it can be said that since these models have more parameters, they will be better predictors and provide a better fit than linear models.

2.1. 3 Non-Parametric Discrimination Analysis[5]: K-Nearest-Neighbour (k-NN)

In the models which have been previously described, class-conditional probability density functions are used. Given these density functions, Bayes rule or the likelihood ratio test may be applied and as a result, a pattern x is assigned to a class. In these models, parameters should be estimated which describe the densities from available data samples.

But many times, density functions do not have a formal structure. Then density cannot be assumed to be described by a set of parameters just like in parametric analysis. In that case, nonparametric methods of density estimation are used.

The k-nearest neighbor (k-NN) classifier is assumed to be one of the easiest approaches to obtaining a nonparametric classification. It has a simple structure, but it is widely used since it performs well when a Bayesian classifier is unknown. Then, classifier is taken as the entire training set. (Servigny & Renault, 2004) An illustrative graph of k-NN is given in *Figure 4*.

The similarities between the input pattern x and a set of reference patterns from the training set are analyzed. A pattern is classified to a class of the majority of its k-nearest neighbors in the training set.

[5] This sub-section heavily relies on (Webb & Copsey, 2011) and (Servigny & Renault, 2004).

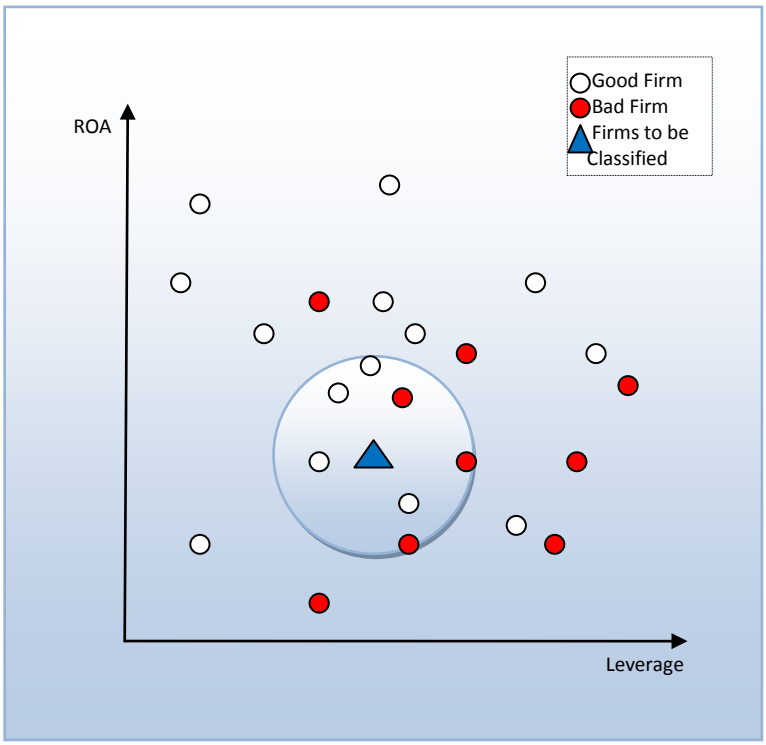


Figure 4: Illustrative graph of k Nearest Neighbor Algorithm (Inspired from (Servigny & Renault, 2004))

The k-nearest neighbor procedure for classifying a measurement x to one of C classes is as follows: i) Determine the k nearest training data vectors to the measurement x, using an appropriate distance metric. ii) Assign x to the class with the most representatives within the set of k nearest vectors.

Before the analysis, the number of neighbors "k", the distance metric, and the training dataset should be determined.

Suppose that the training dataset consists of pairs (x_i, z_i), for $i = 1, \dots, n$, where x_i is the ith training data vector, and x_i is the corresponding class indicator (such that $z_{i=}j$ if the ith training data vector belongs to class w_j).

Let the distance metric between two measurements x and y be denoted by $d(x, y)$. If Euclidean distance metric is to be used then $d(x, y) = |x - y|$.

When applied to a test sample x, the algorithm first calculates $\delta_i = d(x_i, x)$, for $i = 1, \dots, n$. Then, the indices $\{a_1, \dots, a_k\}$ of the k smallest values of δ_i (such that $\delta_{a_i} \leq \delta_j$ for all

$j \notin \{a_1, \dots, a_k\}$, $i = 1, \dots, k$ and $\delta_{a_i} \leq \cdots \leq \delta_{a_k}$). Then k_j is defined to be the number of patterns (data vectors) within the nearest set of k whose class is w_j.

$$k_j = \sum_{i=1}^{k} I\left(z_{a_i} = j\right)$$

where $I(a = b)$ the indicator function which is equal to one if $a = b$ and 0 otherwise. Then x is assigned to w_m if $k_m \geq k_j$ for all j.

This rule may produce more than one winning class, i.e. if $\max_{j=1,\dots C} k_j = k'$. There may be more than one $j \in \{1, \dots, C\}$ for which $k_j = k'$. Various ways for breaking such ties exist: $i)$ Ties may be broken arbitrarily. $ii)$ x may be assigned to the class of the nearest neighbor. $iii)$ x may be assigned to the class, out of the classes with tying values of k', that has nearest mean vector to x (with the mean vector calculated over the k'samples). $iv)$ x may be assigned to the most compact class of the tying classes, i.e. to the one for which the distance to the k'^{th} member is the smallest, which does not require any extra computation.

A distance-weighted rule has been proposed by Dudani in which weights are assigned to the k nearest neighbors, and closest neighbors being weighted more heavily. (Dudani, 1976) A pattern is assigned to the class for which the weights of the representatives among the k-neighbours sum to the greatest value. Then instead of $k_j = \sum_{i=1}^{k} I\left(z_{a_i} = j\right)$, the following equation is used:

$$k'_j = \sum_{i=1}^{k} w_{a_i} I\left(z_{a_i} = j\right)$$

where

$$w_{a_j} = \begin{cases} \dfrac{\delta_{a_k} - \delta_{a_j}}{\delta_{a_k} - \delta_{a_1}}, & if\ \delta_{a_k} \neq \delta_{a_1} \\ 1, & if\ \delta_{a_k} = \delta_{a_1} \end{cases}$$

It can be seen that the weights w_{a_j}take value between 0 for the most distant (k^{th} neighbor) and 1 for the nearest neighbor.

Existence of a realistic split between training and test data is important when using the k-nearest-neighbor method. If there is overlap between the training and test data, the resulting estimates may be over-optimistic.

2.1. 4 Support Vector Machines[6] (SVMs)

SVM which has been introduced by Cortes and Vapnik (1995) aims to separate data into various classes using the best hyperplane. In literature linear support vector machine (LSVM) and nonlinear support vector machine techniques exist.

LSVM method provides a rule for selecting the best separating hyperplane. This hyperplane is the one for which the distance to two parallel hyperplanes on each side is the largest. *Figure 5* shows an example best separating hyperplane (BSH) which separates the data optimally.

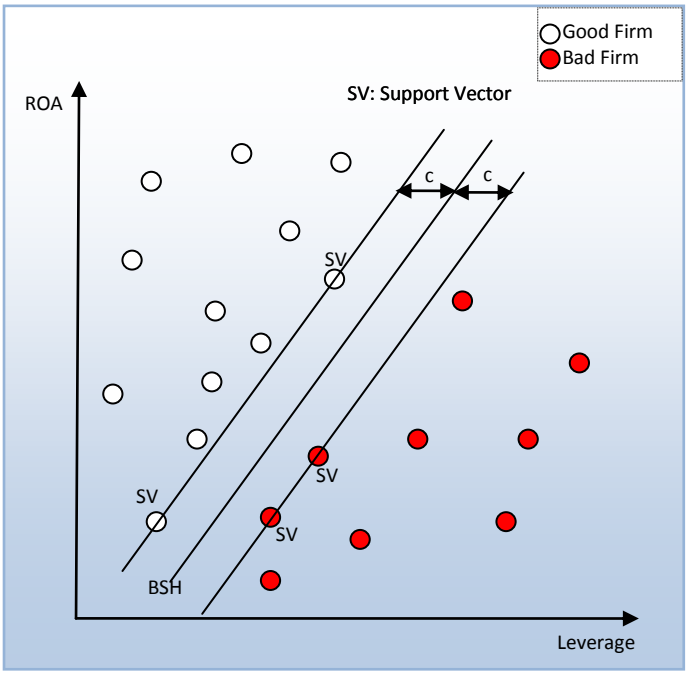


Figure 5: An Illustrative Graph of Support Vector Algorithm

[6] This sub-section heavily relies on (Webb & Copsey, 2011) and (Servigny & Renault, 2004).

As it can be seen, two parallel hyperplanes having same distances to the BSH are tangent to some of the observations. These points are named as support vectors (SV).

Binary classification problem has been studied in a great deal of work on support vector classifiers, and multiclass classifier has been constructed by combining several binary classifiers. (Webb & Copsey, 2011) For simplification binary case will be studied in this thesis.

If Data Is Linearly Separable

Then points x on the best separating hyperplane satisfy the equation $w^{\mathrm{T}}x + w_0 = 0$. Suppose the set of training observations x_i where $i = 1, \dots, n$, then an observation x_i will be assigned to class $w_1$using the rule $w^T x + w_o > 0$ and to class w_2 according to rule $w^T x + w_o < 0$

$$w^T x + w_o \begin{cases} > 0 \\ < 0 \end{cases} \rightarrow x \in \begin{cases} w_1 \\ w_2 \end{cases}$$

Thi rule can be re-expressed as $y_i(w^T x + w_o) > 0$ for all i, where $y_i = 1$ if $x_i \in w_1$ and $y_i = -1$ otherwise.

The best separating hyperplane does not provide the tightest separation rule. Tightest separation can only be achieved when the two hyperplanes which are parallel to the BSH and also tangent to the data.

Then a stricter rule may be imposed so that $y_i(w^T x + w_o) \geq c$ where c represents the distance from the BSH to the tangent hyperplanes.

By minimizing $\|w\|$ subject to the constraint $y_i(w^T x + w_o) \geq c$, the parameters of the hyperplanes may be found.

If Data Is Not Linearly Separable

In most practical cases, a straightforward separation, where a straight line that would separate good firms from bad firms, is not possible. In *Figure 6* the data is not separable by a straight line since one of the members of data set is at the other side of the hyperplanes.

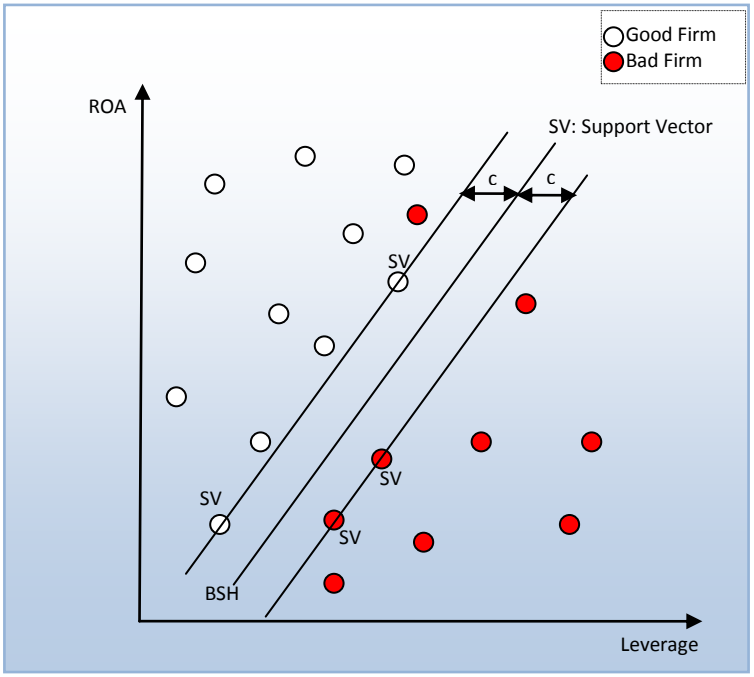


Figure 6: Support Vector for Nonlinarly Seperable data

If the constraints are modified slightly, then the support vector machines may still be used. Instead of using the strict rule $w^T x + w_o > 0$ which actually is the reason for that particular data being assigned to class w_1, another rule $w^T x + w_o > 1 - F_i$ is used where F_i is a positive parameter measuring the degree of fuzziness or slackness of the rule. In a similar fashion, a firm will be classified in class $w_2$if $w^T x + w_o < 1 - F_i$.

If F_i is chosen so that $F_i > 1$ then a firm will be assigned to the wrong category by this method. Therefore, the optimization formulation should include a cost function that penalizes the misclassified observations.

Nonlinear Support Vector Machine

Nonlinearly separable data may also be separated using a nonlinear classifier $h(x)$. Replacing $h(x)$ with x, the following inequality $y_i(w^T h(x_i) + w_o) > 0 \; \forall i$ is found. This inequality defines the best separating surface.

In the determination of optimal parameter vector w, $h(x_i)^T h(x_i)$ term appears. This term can be replaced by a kernel function $K(x_{i,} x_{j,})$.

2.1. 5 Fuzzy C-Means Clustering[7]

The development of the fuzzy c-means algorithm (Dunn, 1973) was the birth of all clustering techniques related to algorithm 1 which is shown in *Figure 7*.

Let a data set $X = \{x_1, x_2, \dots, x_n\}$ be given. Let each cluster can be uniquely characterized by an element of a set K.

Choose the number c of clusters, $2 \leq c < n$
Choose an $m \in \mathbb{R}_{>1}$
Choose a precision for termination β
Initialize $U^{(0)}$, $i := 0$
REPEAT
    Increase i by 1
    Determine $K^{(i)} \in P_c(K)$ such that J is minimized by $K^{(i)}$for $U^{(i-1)}$
    Determine $U^{(i)}$
UNTIL $\left\|U^{(i-1)} - U^{(i)}\right\| < \beta$

Figure 7: Steps of Algorithm 1 (Source: (Höppner, Klawonn, Kruse, & Runkler, 2000))

The first version of this algorithm (Duda & Hart, 1973) performed a hard cluster partition. This way, a data point could be a member of only one cluster. But in order to treat data belonging to several clusters, a fuzzy version of this algorithm was introduced (Dunn, 1973). Later it was generalized (Bezdek, 1973) thus producing the final version with the introduction of the fuzzifier m. This final fuzzy c-means algorithm recognizes spherical clouds of points in a p-dimensional space (Höppner, Klawonn, Kruse, & Runkler, 2000). Each cluster is assumed to have similar sizes and represented by its centre. Euclidean distance between a data point and center is used.

This algorithm uses a predetermined number of clusters, but it does not make an optimization on number of clusters. An illustrative graph for fuzzy c-means clustering is given in *Figure 8*. As it can be seen, nodes are assigned to the geometrically closest cluster with highest probability which takes a value in interval [0,1]. Therefore, membership value

[7] This sub-section heavily relies on (Höppner, Klawonn, Kruse, & Runkler, 2000)

of a certain node takes a positive value for more than one cluster as long as the membership value assigned is different than unity.

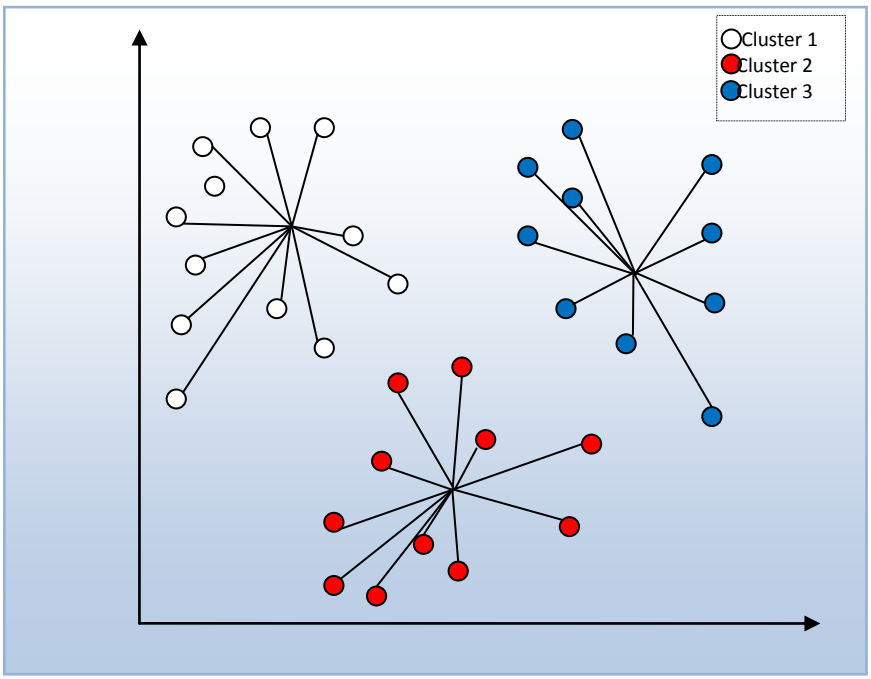


Figure 8: An Illustrative Graph of Fuzzy c-Means Clustering

Theorem 2.1 (Prototypes of FCM)[8] Let $p \in \mathbb{N}$, $D \coloneqq \mathbb{R}^p$, $X = \{x_1, x_2, \dots, x_n\} \subseteq D$, $C \coloneqq \mathbb{R}^p$, $R \coloneqq (C)$, $D \coloneqq \mathbb{R}^p$, $m \in \mathbb{R}_{>1}$,

$$d: DxC \to \mathbb{R}, (x, p) \mapsto \|x - p\|$$

If J is minimized with respect to all probabilistic cluster partitions $X \to F(K)$ with $K = \{k_1, k_2, \dots, k_c\} \in R$ and given memberships $f(x_j)(k_i) = \mu_{ij}$ by $f: X \to F(K)$, then

$$k_i = \frac{\sum_{j=1}^{n} (\mu_{ij})^m \, x_j}{\sum_{j=1}^{n} (\mu_{ij})^m}$$

holds.

[8] Properties of FCM theorem has been taken from the book of Höppner et al. (Höppner, Klawonn, Kruse, & Runkler, 2000)

Proof: The probabilistic cluster partition $f: X \to F(K)$ shall minimize the objective function J. Then, all directional derivatives of J with respect to $k_i \in K, i \in \mathbb{N}_{\leq c}$ must be equal to zero. Then, for all $\xi \in \mathbb{R}^p$ with $t \in \mathbb{R}$.

$$0 = \frac{\partial}{\partial k_i} J = \frac{\partial}{\partial k_i} \sum_{j=1}^{n} \sum_{l=1}^{c} (\mu_{lj})^m \left\| x_j - k_l \right\|^2$$

$$= \sum_{j=1}^{n} (\mu_{ij})^m \frac{\partial}{\partial k_i} \left\| x_j - k_i \right\|^2$$

$$= \sum_{j=1}^{n} (\mu_{ij})^m \lim_{t \to 0} \frac{\left\| x_j - (k_i + t\xi) \right\|^2 - \left\| x_j - k_i \right\|^2}{t}$$

$$= \sum_{j=1}^{n} (\mu_{ij})^m \lim_{t \to 0} \frac{1}{t} \left(\left((x_j - k_i) - t\xi \right)^T \left((x_j - k_i) - t\xi \right) - (x_j - k_i)^T (x_j - k_i) \right)$$

$$= \sum_{j=1}^{n} (\mu_{ij})^m \lim_{t \to 0} \frac{-2t(x_j - k_i)^T \xi + t^2 \xi^T \xi}{t}$$

$$= -2 \sum_{j=1}^{n} (\mu_{ij})^m (x_j - k_i)^T \xi$$

then since,

$$\frac{\partial}{\partial k_i} J = 0$$

$$\Leftrightarrow \sum_{j=1}^{n} (\mu_{ij})^m (x_j - k_i) = 0$$

$$\Leftrightarrow k_i = \frac{\sum_{j=1}^{n} (\mu_{ij})^m x_j}{\sum_{j=1}^{n} (\mu_{ij})^m}$$

■

This algorithm works by assigning membership to each data point corresponding to each cluster center using the distance between the cluster center and the data point. This "membership" takes a value between 0 and 1 and it shows the probability that a particular data point falls into a certain cluster. If the data point is closer to the center of the cluster, the membership value gets a higher value. Since the "membership" shows the probability

that a particular data point falls into a certain cluster, the summation of the membership values for each data is equal to unity. This algorithm is composed of iterations and after each iteration, membership and cluster centers are updated according to the following formula:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} (\frac{d_{ij}}{d_{ik}})^{(2/m-1)}}$$

$$v_j = \frac{\sum_{i=1}^{n} (\mu_{ij})^m x_i}{\sum_{i=1}^{n} (\mu_{ij})^m}, \forall j = 1,2, \dots c$$

where,

n is the number of data points,

v_j represents the j[th] cluster center,

m is the fuzziness index $m \in [1, \infty)$,

c represents the number of cluster center,

μ_{ij} represents the membership of ith data to j[th] cluster center,

d_{ij} represents the Euclidean distance between i[th] data and j[th] cluster center.

Main objective of fuzzy c-means algorithm is to minimize:

$$J(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{c} (\mu_{ij})^m \left\| x_i - v_j \right\|^2$$

Where $\left\| x_i - v_j \right\|$ is the Euclidean distance between i[th] data and j[th] cluster center.

Algorithmic steps for Fuzzy c-means clustering

Let $X = \{x_1, x_2, \dots, x_n\}$ be the set of data points and $V = \{v_1, v_2, \dots, v_c\}$ be the set of cluster centers.

Step 1: c cluster centers are selected randomly

Step 2: Calculate μ_{ij} values (probabilities that a particular data point falls into a certain cluster) using the following formula:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} (\frac{d_{ij}}{d_{ik}})^{(2/m-1)}}$$

Step 3: Compute the fuzzy means using

$$v_j = \frac{\sum_{i=1}^{n} (\mu_{ij})^m x_i}{\sum_{i=1}^{n} (\mu_{ij})^m}, \forall j = 1,2, \dots c$$

Step 4: Repeat steps 2 and 3 until the minimum $J(U,V)$ value is reached or $\|U(k+1) - U(k)\| < \beta$ where,

k represents the iteration step.

β is the termination criterion where $\beta \in [0,1]$,

$U_{n\,x\,c} = (\mu_{ij})$ is the fuzzy membership matrix which is composed of μ_{ij} values,

$J(U,V)$ represents the objective function to be minimized.

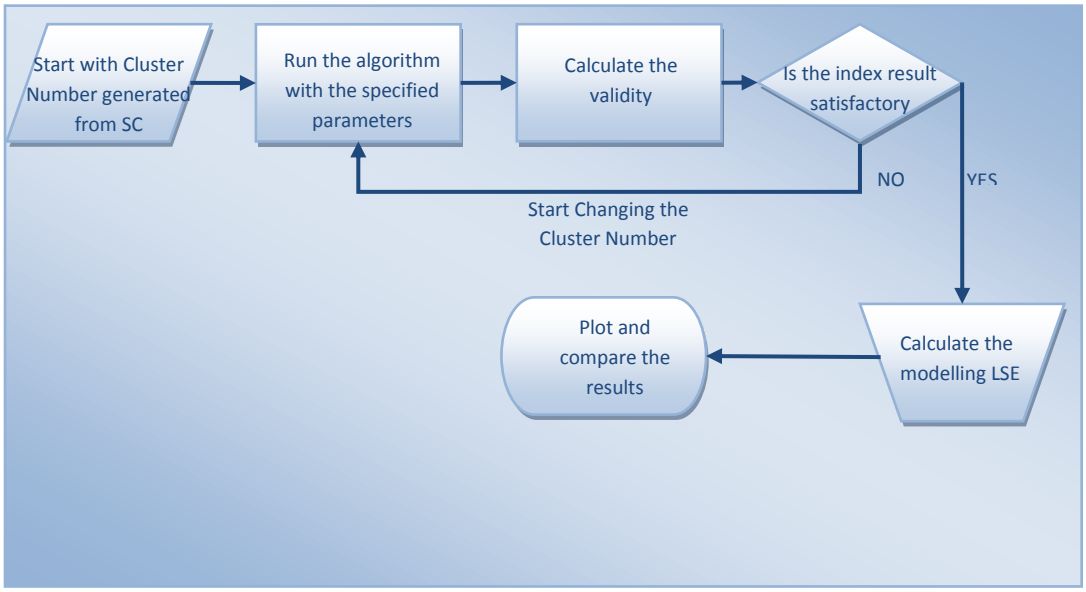


Figure 9: Flow Chart of FCM Clustering Algorithm Procedure
Inspired from (Bataineh, Najia, & Saqera, Aug. 2011))

The flow chart for the algorithm is given in *Figure 9*. The fuzzy c-means clustering algorithm is similar to the k-neighborhood algorithm. It has some advantages together with aspects which may be criticized. Advantages could be listed as: i) Produces comparatively better

results then k-means algorithm. ii) Unlike k-means where data point must exclusively belong to one cluster center in fuzzy c-means, a data point is assigned membership to each cluster. On the other side, disadvantages may be listed as: i) Apriori specification of the number of clusters. ii) With lower value of β we get the better result but at the expense of more number of iteration. iii) Euclidean distance measures can unequally weight underlying factors.

CHAPTER 3

LITERATURE REVIEW

In reality, determining vulnerable banks, determining their relative financial strength, comparing riskiness of banks and bankruptcy studies all have a common ground. They all try to determine which banks are easier to default. In that sense, literatures and theoretical structures for abovementioned study areas have great deal of intersection.

Mathematical and statistical methods are widely used in credit scoring techniques. In 1932, Fitzpatrick studied the dependence between probability of default and characteristics of corporate credits (Fitzpatrick, 1932). Also Fisher came up with the idea of discriminant analysis to study the relationships between subgroups in a population (Fisher, 1936). In the following period, Durand's study has been published containing discriminant analysis methods for segregating good consumer loans from bad ones (Durand, 1941).

In 1960s, several improvements and extensions have been introduced regarding methodology. Credit scoring techniques have been extended from credit card loans to other asset classes, and particularly to the population of SMEs. Myers and Forgy compared regression and discriminant analysis for credit scoring applications (Myers & Forgy, 1963). In the following period, Beaver studied bankruptcy prediction models (Beaver, 1966), and Altman applied multiple discriminant credit scoring analysis (MDA) (Altman, 1968) which is a famous model that can be used to classify a firm's creditworthiness, by assigning a Z-score to it. Next, Martin and Ohlson dealt with logit analysis and they were the first to apply logit analysis for bankruptcy prediction studies. (Martin, 1977) (Ohlson, 1980)

3. 1. Studies on Discriminant Analysis

In 1932, Fitzpatrick studied the dependence between probability of default and characteristics of corporate credits (Fitzpatrick, 1932).

In 1936 Fisher came up with the idea of discriminant analysis to study the relationships between subgroups in a population (Fisher, 1936).

In 1941, Durand's study has been published containing discriminant analysis methods for segregating good consumer loans from bad ones (Durand, 1941).

Myers and Forgy compared regression and discriminant analysis for credit scoring applications (Myers & Forgy, 1963).

In 1966, Beaver studied bankruptcy prediction models using discriminant analysis (Beaver, 1966).

In 1968, Altman used a multivariate discriminant analysis (MDA) model using a data set that consists of 66 firms using the following financial ratios: (i) working capital/total assets; (ii) retained earnings/total assets; (iii) earnings before interest and taxes/total assets; (iv) market value of equity/book value of total debt; (v) sales/total assets. (Altman, 1968) In following times, this pioneering study has been utilized by many researchers.

In 1972, Deakin compared the business failure prediction performance of Beaver's dichotomous classification test and Altman's DA to using data related to 64 firms for 1964-1970 period. The conclusion was that Beaver's dichotomous classification test has been found to be successful in predicting business failure five years in advance while Altman's MDA could have been used to predict business failure from accounting data as far as three years in advance. (Deakin, 1972)

In 1975, Sinkey applied MDA using 10 variables of 110 pairs of banks in health and distress observed in the period 1969-1972. He reported that asset composition, loan characteristics, capital adequacy, sources and uses of revenue, efficiency, and profitability were found to be good discriminators between the groups and the designed model achieved various success rates between 64,09%-75,24%. (Sinkey, 1975)

In 1977, Altman et al. studied and developed a new bankruptcy classification model called Zeta Credit Risk Model using 111 firms with seven variables covering the period 1969–1975 and they reported that the classification accuracy of the model ranged from 96% for one period to 70% for five periods. (Altman, Haldeman, & Narayanan, 1977)

In 1982, Dietrich and Kaplan have developed a loan classification model from ordinary least squares (OLS) and MDA using variables suggested by experts (debt-equity ratio, funds-flow-to-fixed-commitments ratio, and sales trend), and compared it with Zeta model and Wilcox bankruptcy prediction and they found that the simple three variable linear model gave better predictions. (Dietrich & Kaplan, 1982)

In 1987, using a random sample of 50 companies, Karels and Prakash conducted a study in three steps; (i) investigated the normality condition of financial ratios; (ii) when these ratios are non-normal, they constructed ratios which are either multivariate normal or almost normal; (iii) using these ratios to compare the prediction results of DA with other studies and they reported 96% classification rate for non-bankrupt firms and 54.5% for bankrupt firms. (Karels & Prakash, 1987)

In 2001, Grice and Ingram evaluated the generalizability of Altman's (1968) Z-score model using a proportionate sample of distressed and non-distressed companies from time periods, industries, and financial conditions other than those used by Altman to develop his model. (Grice & Ingram, 2001)

In 2011, Li and Sun proposed a new hybrid method for BFP by integrating PCA with MDA and logit. They compared the hybrid method with the two classifiers with features selected by stepwise method of MDA. For the hybrid method's feature selection, PCA was employed in four different means, that is, the use of PCA on all available features, the use of PCA on the features selected by stepwise method of MDA, the use of PCA on the features selected by stepwise method of logit, and the use of PCA on the features selected by independent sample t test. The best way of employing PCA in the two methods of MDA and logit was to use PCA to extract features on the processed results of stepwise method of MDA. The best predictive performance of MDA and logit was not significantly different, though MDA achieved better mean predictive accuracy than logit. They concluded that the employment of PCA with MDA and logit can help them produce significantly better predictive performance in short-term BFP of Chinese listed companies. (Li & Sun, Empirical research of hybridizing principal component analysis with multivariate discriminant analysis and logistic regression for business failure prediction, 2011)

3. 2. Studies on Logit and Probit Models

In 1977, Martin used logit model to predict the probability of failure of banks. (Martin, 1977)

In 1980, Ohlson used logit model to predict firm failure. The author reported that the classification accuracy was 96.12%, 95.55% and 92.84% for prediction within one year, two years and one or two years respectively. Ohlson also criticized studies that used MDA. (Ohlson, 1980)

In 1984, Zmijewski examined two potential biases caused by sample selection/ data collection procedures used in most financial distress studies using probit method with the data covering all firms listed on the American and New York Stock Exchanges for the period 1972- 1978 which have industry (SIC) codes of less than 6000. The biases were; (i) when a researcher first observes the dependent variable and then selects a sample based on that knowledge (ii) when only observations with complete data are used to estimate the model and incomplete data observations occur non-randomly. He reported that for both biases the results were the same, which was, the bias was clearly shown to exist, but, in general, it did not appear to affect the statistical inferences or overall classification rates. (Zmijewski, 1984)

In 1985, West used a combined method of factor analysis and logit to create composite variables to describe banks' financial and operating characteristics, to measure the condition of individual institutions and to assign each of them a probability of being a problem bank and he reported that his method was promising in evaluating bank's condition. (West, 1985)

In 1985, Gentry et al., compared logit with DA and found that logit outperformed DA, and in 1987, they compared probit with DA and found that probit outperformed DA. In both studies they used a dataset consisted of 33 pairs of companies in health and distress observed in the period 1970-1981. (Gentry, Newbold, & Whitford, 1987) (Gentry, Newbold, & Whitford, 1985)

In 2000, Laitinen and Laitinen tested whether Taylor's series expansion can be used to solve the problem associated with the functional form of bankruptcy prediction models. To avoid the problems associated with the normality of variables, logit was applied to describe the insolvency risk. Then Taylor's expansion was used to approximate the exponent of logit. The cash to total assets, cash flow to total assets, and shareholder's equity to total assets ratios was found to be the factors affecting the insolvency risk. They concluded that for one year and two years prior to the bankruptcy, Taylor's expansion was able to increase the classification accuracy but not for three years prior to the bankruptcy. (Laitinen & Laitinen, 2000)

In 2001, Grice and Dugan evaluated whether Zmijewski (1984) and Ohlson (1980) bankruptcy prediction models could be generalized to proportionate samples of distressed and non-distressed companies from time periods, industries, and financial conditions other than those used to develop their models. They concluded that (i) both models were sensitive to time periods, the accuracy of the models declined when applied to time periods different from those used to develop the models (ii) the accuracy of each model continued to decline moving from the 1988–1991 to the 1992–1999 sample period (iii) Ohlson's (Zmijewski's) model was (was not) sensitive to industry classifications. (Grice & Dugan, 2001)

In 2004, Jones and Hensher presented mixed logit model for firm distress prediction and compared it with multinomial logit models by classifying firms into three groups: state 0: non-failed firms; state 1: insolvent firms, state 2: firms which filed for bankruptcy and concluded that mixed logit obtained substantially better predictive accuracy than multinomial logit models. (Jones & Hensher, 2004)

In 2007, Jones and Hensher employed the multinomial nested logit (NL) model to BFP problem using a four-state failure model based on Australian company samples. They concluded that the unordered NL model outperformed the standard logit and unordered multinomial logit models and so, NL's prediction accuracy was satisfactory for BFP problems. (Jones & Hensher, Modelling corporate failure: A multinomial nested logit analysis for unordered outcomes, 2007)

In 2009, Lin examined the bankruptcy prediction ability of MDA, logit, probit and BPNN using a dataset of matched sample of failed and non-failed Taiwan public industrial firms during 1998–2005. They reported that the probit model possessed the best and stable performance; however, if the data did not satisfy the assumptions of the statistical approach, then the BPNN achieved higher prediction accuracy. In addition, the models used in this study achieved higher prediction accuracy and generalization ability than those of Altman's (1968), Ohlson's (1980), and Zmijewski's (1984). (Lin, 2009)

3. 3. Studies on k-NN Models

In 2002, Park and Han proposed an analogical reasoning structure for feature weighting using a new framework called the analytic hierarchy process (AHP)-weighted k-NN algorithm. They compared this AHP – k-NN model with pure k-NN and logit – k-NN and reported that the classification accuracies were 83.0%, 68.3% and 79.2%, respectively. For this comparison they examined several criteria, both quantitative (financial ratios) and qualitative (non-financial variables). (Park & Han, 2002)

In 2004, Yip compared weighted k-NN, pure k-NN and DA to predict Australian firm business failure and she reported that the overall accuracies were 90.9%, 79.5% and 86.4%, respectively. She concluded that weighted k-NN outperformed DA and was an effective and competitive alternative to predict business failure in a comprehensible manner. (Yip, 2004)

In 2009 a new similarity measure mechanism for k-NN algorithm based on outranking relations has been proposed by Li et al. The study included strict difference, weak difference, and indifference between cases for each feature. Accuracy of the CBR prediction method which is based on outranking relations has been determined directly by four parameters. The experimental design used three year's data from 135 pairs of Chinese listed companies, the cross-validation of leave-one-out was utilized to assess models, the method of stepwise DA has been utilized to select features, grid-search technique was used to get optimized model parameters. The study has a conclusion that OR-CBR method has been determined to be superior to MDA, logit, BPNN, SVM, decision tree, Basic CBR, and Grey CBR in financial distress prediction. (Li, Sun, & Sun, 2009)

In 2011, Li and Sun compared a forward ranking-order case-based reasoning (FRCBR) method, with the standalone RCBR, the classical CBR with Euclidean metric at the center, the inductive CBR, logit, MDA, and support vector machines. The conclusion was that FRCBR produced superior performance in short-term business failure prediction of Chinese listed companies. (Li & Sun, Predicting business failure using forward ranking-order case-based reasoning, 2011)

3. 4. Studies on Support Vector Machine Models

In 1995 Cortes and Vapnik introduced SVM to separate data into various classes using the best hyperplane. (Cortes & Vapnik, 1995)

In 2005, bankruptcy prediction performance of SVM has been compared with MDA, Logit and three-layer fully connected BPNN by Min and Lee, using from the Korea's largest credit guarantee organization. The first method to be used was PCA, and t-test for feature selection. Initial number of financial ratios was 38. Among these, 23 financial ratios were selected first. Then by using the stepwise logit, the number of financial variables has been reduced to 11. The data consisted of 1888 firms includes 944 bankruptcy and 944 non-bankruptcy cases placed in random order. The conclusion was that SVM outperforms other methods. (Min & Lee, 2005)

In 2005, BFP performance of SVM and BPNN has been compared. The data was provided by Korea Credit Guarantee Fund and consisted of externally non-audited 2320 medium-size manufacturing firms, which filed for bankruptcy (1160 cases) and non-bankruptcy (1160 cases) from 1996 to 1999. For feature selection they applied a two-staged input variable selection process. At the first stage, they selected 52 variables among more than 250 financial ratios by independent-samples t-test and in the second stage, they selected 10 variables using a MDA stepwise method. About 80% of the data was used for a training set and 20% for a validation set. They concluded that the accuracy and generalization performance of SVM was better than that of BPN as the training set size gets smaller. (Shin, Lee, & Kim, 2005)

In 2006, financial distress prediction performance of performance of SVM with MDA, logit and BPNN has been compared by Hui and Sun. Three-year data of 135 pairs of Chinese listed companies' has been used. Stepwise MDA method has been preferred for feature selection. Cross-validation and grid-search technique has been adopted to find SVM model's good parameters. At the end of the study, they concluded that financial distress early-warning model based on SVM obtained a better balance among fitting ability, generalization ability and model stability than the other models. (Hui & Sun, 2006)

3. 5. Studies on Fuzzy C-Means Clustering Models

In 1973 Duda and Hart introduced a hard cluster partitioning algorithm which assigned data points into only a single cluster. (Duda & Hart, 1973)

In 1973, to be able to treat data belonging to several clusters, Dunn introduced a fuzzy version of Duda and Hart's algorithm (Dunn, 1973).

In 1973, Bezdek introduced the fuzzifier 'm' to the fuzzy clustering algorithm and thus generalized the algorithm. (Bezdek, 1973)

In 1995, Pal and Bezdek studied the role of model parameters since they have effects on validation of clusters. Their study has shown that some validity indexes have unpredictable dependency on elements of the solution. (Pal & Bezdek, 1995)

In 1999, Michael et al. proposed a combined use of a fuzzy rule generation method and a data mining technique for bankruptcy prediction and compared it with LDA, QDA, logit and probit using two samples of data from Greek firms consists of basic sample and holdout sample. They concluded that fuzzy knowledge-based decision aiding method outperformed other classification methods used in this paper. (Michael, Georgios, Nikolaos, & Constantin, 1999)

In 2000, Alam et al. proposed fuzzy clustering algorithm for identifying potentially failed banks and compared it with two SOM networks viz., (i) competitive neural network and (ii) self organizing neural network and concluded that both fuzzy clustering and SOM are good tools in identifying potentially failing banks. They also mentioned that fuzzy clustering

provides an ordinal rating of the data set in terms of failing likelihood possibility. For the experiment they used an unbalanced dataset that consisted of 3% unsuccessful banks. (Alam, Booth, Lee, & Thordarson, 2000)

In 2011 Andrés et al. proposed a hybrid system combining fuzzy clustering and MARS. They have tested the accuracy of their approach in a real setting consisting of a database made up of 59,336 non-bankrupt Spanish companies and 138 distressed firms which went bankrupt during 2007. As benchmarking techniques they have used discriminant analysis, MARS and a feed-forward neural network. Their results show that the hybrid model outperforms the other systems, both in terms of the percentage of correct classifications and in terms of the profit generated by the lending decisions. (Andrés, Lorca, Juez, & Sánchez-Lasheras, 2011)

CHAPTER 4

METHODOLOGY

The aim of fuzzy c-means clustering algorithm is to classify a number of data points (data points represent the banks in this context) into a number of predefined classes. In this formulation, c represents the number of clusters to be formed. Each data point is assigned a membership value which shows the probability that a particular data point falls into a certain cluster and therefore it takes a value between 0 and 1. If the data point is closer to the center of the cluster, the membership value gets a higher value. Since the "membership" shows the probability that a particular data point falls into a certain cluster, the summation of the membership values for each data is equal to unity.

Unlike other algorithms which are hard-type[9], fuzzy c-means clustering algorithm assigns a data point to more than one cluster. Therefore a vector of probabilities is obtained after running the algorithm. This vector has c dimensions since a data point is assigned to c different groups with different probabilities.

Initially, c cluster centers are selected randomly. Then probabilities of data points falling into different certain clusters, together with the cluster center values are calculated. This algorithm is composed of iterations and after each iteration, membership and cluster centers are updated. These iterations are repeated until the objective function $J(U,V)$ is minimized or certain termination criterion is met.

This study is composed of a number of steps, (i) data collection and preparation, (ii) optimization, (iii) conclusion. Data collection and data preparation is a very time-intensive step of this procedure.

As of September 21st of 2012, there are 48 banks in Turkey. These banks are classified under different categories according to their ownership and activity type. With respect to activity type, a bank is either a "deposit bank" or a "development and investment bank".

[9] Hard-type clustering algorithms assign a data point to a single cluster.

On the other side, with respect to ownership, a bank is a state-owned bank, privately-owned or a foreign bank. This structure can be shown in *Figure 10* as following:

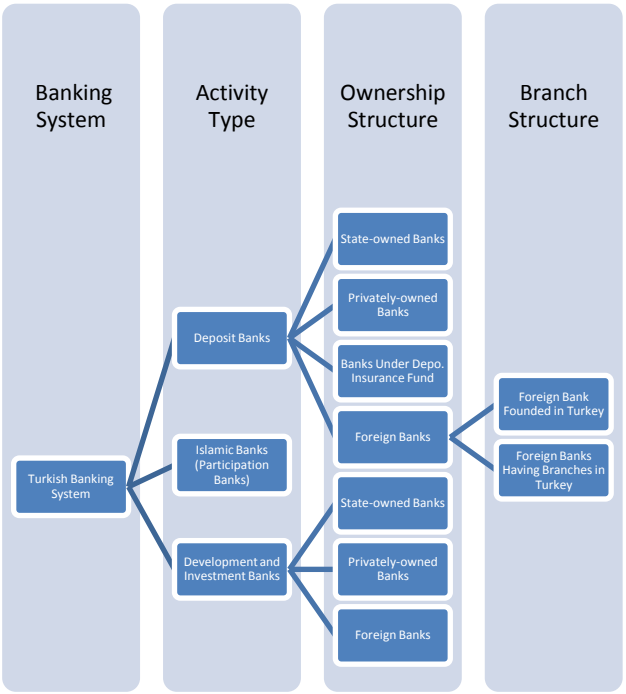


Figure 10: Turkish Banking Sector

Since these banks have different structure, their financial tables and financial ratios differ significantly from each other. Some of the banks have so distinct ratios that their ratios are almost 15 times that of industry average, which may be a good sign for that observation to be accepted as an "outlier". Also some of the banks have no data regarding some ratios, which in turn is the reason why some of the ratios are ignored in certain analysis.

In this study, as a result of these, bank under the control of Deposit Insurance Fund, foreign banks operating as branches in Turkey, development and investment banks and Islamic banks are omitted from the study. As a result, data related to 23 banks which are state-owned deposit banks, privately-owned deposit banks and foreign banks founded in Turkey have been included in the study. Data is analyzed using 48 different ratios which can be classified under capital, assets quality, liquidity, profitability, income-expenditure structure, share in sector, share in group and branch ratios.

4. 1. Data Collection and Preparation

Financial ratios are used very frequently in rating and bankruptcy studies and perhaps they are the most important source for these studies since ratios are better indicators than nominal values. This is particularly true if banks in the sample have very different sizes. As an example, total loans divided by total assets is a better indicator for a bank's financial situation than total loans alone, since the effect of the bank's size is eliminated.

The financial ratios that have been used in this study are financial figures but they have economical implications. To exemplify, capital adequacy ratios measure a bank's capacity to meet the liabilities and other risks such as operational risk, credit risk etc. It can be said that a bank's capital is the buffer or cushion for potential losses, and protects the bank's depositors and other lenders. If this ratio is low, then banks have higher risk of not being able to repay the debt that they owe to deposit holders. Even if "being unable to repay deposit back" occurs a couple of times, customers may rush to their banks and try to draw back their deposits. In that case, if banks do not hold enough capital to repay the deposits, then the whole banking system may be contaminated with the risk of default since probably their customers will rush too. In the end, credit and banking system would stop functioning, credit channels would stop lending. If the real sector that is in need of being financed could not be financed anymore, production activities would stop. Since the lending channels would stop functioning, the value of cash would increase therefore interest rates would reach peak levels too. In most countries, banking regulators monitor and define the capital adequacy ratios (CAR) to protect depositors in order to maintain the confidence in the banking system. Target CAR is 12% for Turkey but the actual level of CAR is 16.5%. (BRSA, June 2012) On the other hand, having a very high CAR is not good since it shows that valuable resources are kept idle instead of being used in financing the economy.

High value for a financial ratio does not always imply a healthier bank. For example, high value for loans under follow up divided by total loans and receivables indicate that some of the loans are not paid back; indicating that the credibility of the customer is not very high. Also it may show that appropriate preventive measures might not have been taken by the bank to control the risk, such as credit scoring. Therefore, it can be said that the lower this

ratio, the healthier the bank. Similarly high value for the ratio interest expenses divided by total assets or total expenses is not desired.

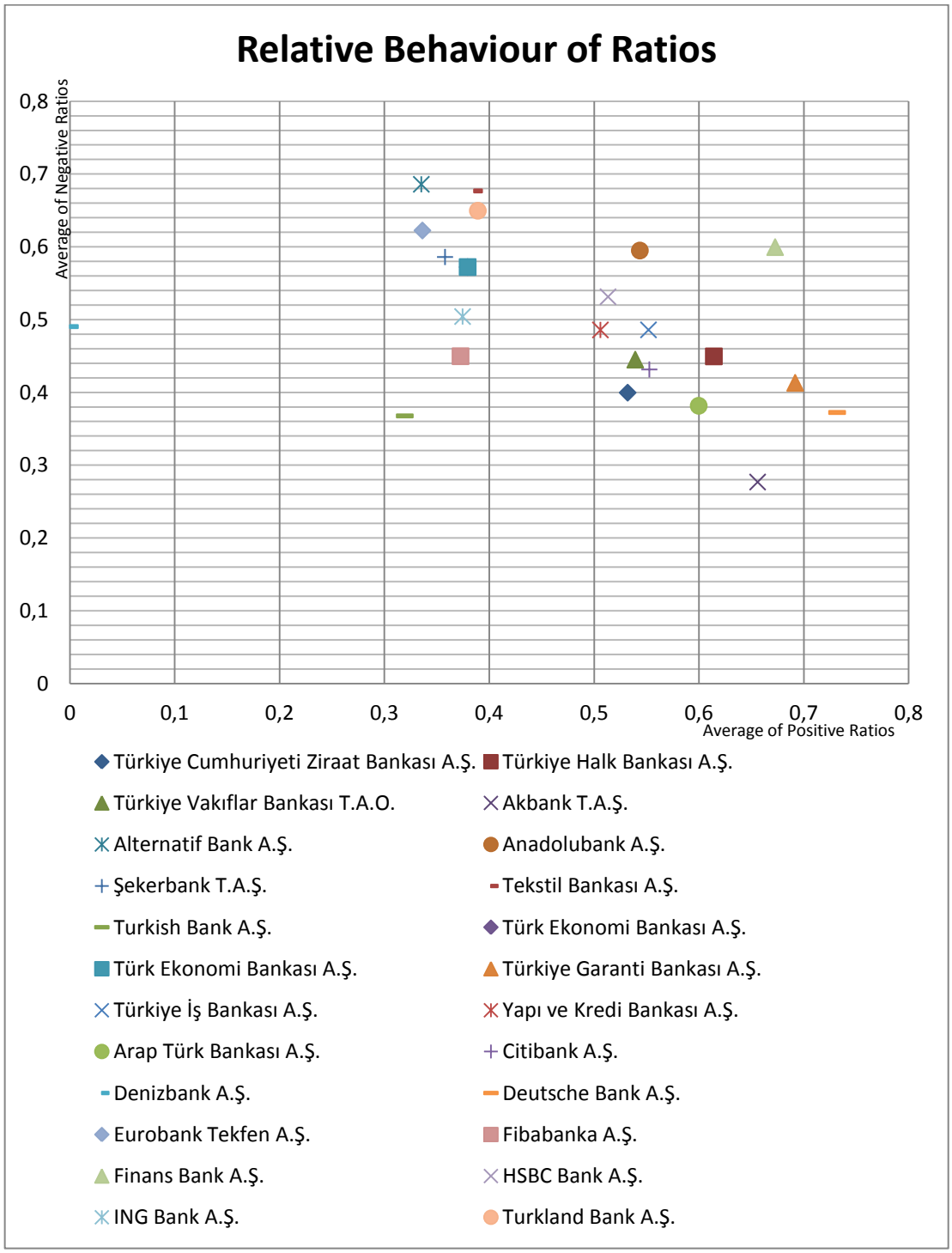


Figure 11: Positive Mannered Ratios vs Negative Mannered ratios

On the other hand, there are financial ratios that are desired to get higher values. Liquidity ratios, profitability ratios, some of the income-expenditure ratios, some of the branch or activity ratios are desired to take high values.

Obviously, some of the ratios are desired to have higher values (positive mannered) where others are desired to have lower (negative mannered) values. The behavior of average positively mannered ratios versus the negatively mannered ratios can be seen in *Figure 11*. For each of the banks, two average values are calculated; one corresponding to the average of the percentile ranks of positively mannered ratios and the other one corresponding to the average of the percentile ranks of negatively mannered ratios. Positive mannered ratios are taken as: C1, C2, C3, C4, A4, L1, L2, L3, P1, P2, P3, P4, I1, I2, I6, I7, I8, I9, I10, I12, S2, S3, B1, B2, B3, B4, B5, B7, Ac6, Ac7 and Positive mannered ratios are taken as: C5, A1, A2,A3, I11, I13, B6, Ac1, Ac2, Ac3 ratios.

It can be seen from Figure 11 that the pattern is negatively sloped which shows that as the average percentile rank of positive-mannered ratios of a bank increases, then most of the time average percentile rank of negative-mannered ratios of a bank decreases.

Following variables are suggested for identifying vulnerable banks in economic theory: net income to total assets, equity capital to total assets, total loans to total assets or liquid assets to total assets, and commercial loans to total assets (Korobow, Stuhr, & Martin, Autumn 1977).

In a similar manner, IMF has worked closely with national agencies and regional and international institutions to develop a set of Financial Soundness Indicators (FSIs) which are statistical measures for monitoring the financial health and soundness of a country's financial sector. (Jose & Georgiou, 2009) (Moorhouse, 2004) These indicators are tabulated in Appendix A. As it can be seen, 25 of these ratios are designed for banking and 14 of them are designed for other financial corporations.

Similarly, in this study, 48 ratios have been used in this study as performance variables which are grouped in various categories as it can be seen from Appendix B. The financial

ratios used in this study have been obtained online from the website[10] of Banks Association of Turkey. (TBB, 2012). Predictive statistics for the ratios used in the study have been given in Appendix C. Explanations regarding the ratios used in the analysis are listed in the Appendix D.

The change in financial ratios has been plotted for some of the representing ratios in Appendix G. It can be seen that some of for some of the ratios plotted, final values tend to converge to a more stable value as time passes since 2001 financial crisis of Turkey.

The dataset originally consisted of 48 financial ratios for each of 48 banks. Some of the banks have been excluded from the analysis since some of their financial ratios are missing. In that sense banks under the control of Deposit Insurance Fund, foreign banks operating as branches in Turkey, development and investment banks and Islamic banks are omitted from the study as well as Adabank. Data related to the remaining 23 banks which are state-owned deposit banks, privately-owned deposit banks and foreign banks founded in Turkey have been analyzed.

The reasons for omission of these banks are obvious. One of the most important reasons is the fact that comparison of private deposit banks with development and investment banks (which do not have permission for collecting deposit) and state-owned commercial banks (which are very different because of its capital and ownership nature) does not produce very meaningful results. Foreign banks operating as branches in Turkey have been excluded too since most of their financial ratios are too deviant from the sector average, which could be a sign that data points for these banks may be seen as outliers in the analysis.

4. 2. Analysis

Matlab version 2008 has been used for analysis of the financial ratios given in Appendix B.

The proposed algorithm, which is based on fuzzy based clustering algorithm, groups banks using the financial ratios which are fed into the system. Banks of similar financial characteristics are grouped together. But this algorithm does not give any information

[10] http://www.tbb.org.tr/eng/Banka_ve_Sektor_Bilgileri/Tum_Raporlar.aspx

regarding which cluster is composed of banks having the best or worst characteristics. To rank the groups, further heuristics will be applied on the results of grouping process.

The algorithm translated into Matlab code executes the following steps:

Step 1: c cluster centers are selected randomly

Step 2: Calculate μ_{ij} values (probabilities that a particular data point falls into a certain cluster) using the following formula:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} (\frac{d_{ij}}{d_{ik}})^{(2/m-1)}}$$

Step 3: Compute the fuzzy means using

$$v_j = \frac{\sum_{i=1}^{n} (\mu_{ij})^m \, x_i}{\sum_{i=1}^{n} (\mu_{ij})^m}, \forall j = 1,2, \dots c$$

Step 4: Repeat steps 2 and 3 until the minimum objective function $J(U,V)$ value is reached or $\|U(k+1) - U(k)\| < \beta$ where represents the iteration step.

4. 3. Ranking the Clusters

As it has been mentioned before, fuzzy clustering algorithm groups the banks having similar financial characteristics. But this algorithm does not give any information regarding which cluster is composed of banks having the best or worst characteristics.

The results of the fuzzy c-means clustering algorithm have been processed further using these ten financial ratios so that the weighted score of the clusters could be used in ranking these clusters.

To rank these groups, further work is required. First a subset of representing ratios has been selected. This set is composed of capital ratios (C1, C2), asset quality ratios (A1, A2, A4), liquidity ratio (L2), profitability ratio (P1), income-expenditure ratio (I6), share ratio (S1) and activity ratio (AC7). Explanatory information is available in *Appendix B* regarding

these ten different financial ratios. Share ratio S1, which is simply total assets[11], has been incorporated into the system for treating large and small banks in a relatively different manner.

First, ten financial ratios corresponding to each bank in the sample have been reevaluated. In the next step, the elements of each column of a certain financial ratio have been ranked. This way, nominal values of these ratios have been replaced by their corresponding rank indices. In the following step, these rank values have been converted into percentile ranks so that the rank values are normalized to fall in interval [0,1].

Next, calculated percentile rank values are linearly combined using an equal weight of 0,1 for each ratio so that a numerical value for each bank has been obtained.

In-cluster average values of these scores are calculated so that each cluster is mapped to a single average numerical score. Finally, these scores have been used for group-by-group comparison of clusters.

4. 4. Principal Component Analysis

The correlation and covariance matrices of the financial ratios are given in *Appendix E* and *Appendix F* respectively. If these matrices are analyzed, it can be seen that there is a great deal of correlation between financial variables. To exemplify, the correlation between I2 and B3; C2 and C3; C2 and C4; B3 and Ac2 are 0,76, 0,89, 0,99 and 0,80 respectively. This co-dependence between variables is plotted on a 2-D Filled Contour Graph in *Figure 11*. As it can be seen from the legend on the right hand side of the plot, darker red corresponds to positive high correlation whereas darker blue regions correspond to negative high correlation. Intuitively, if a method or transformation were used to reduce the density of dark shades of blue and dark shades of red, then some of the correlation between financial ratios could be reduced.

[11] Natural logarithm of total assets has been used by Barniv et al. (1997) and Bell (1997). Also, the variable 'total assets' has been used as an indicator by Dietrich & Kaplan (1982) and Kolari et al. (2002)

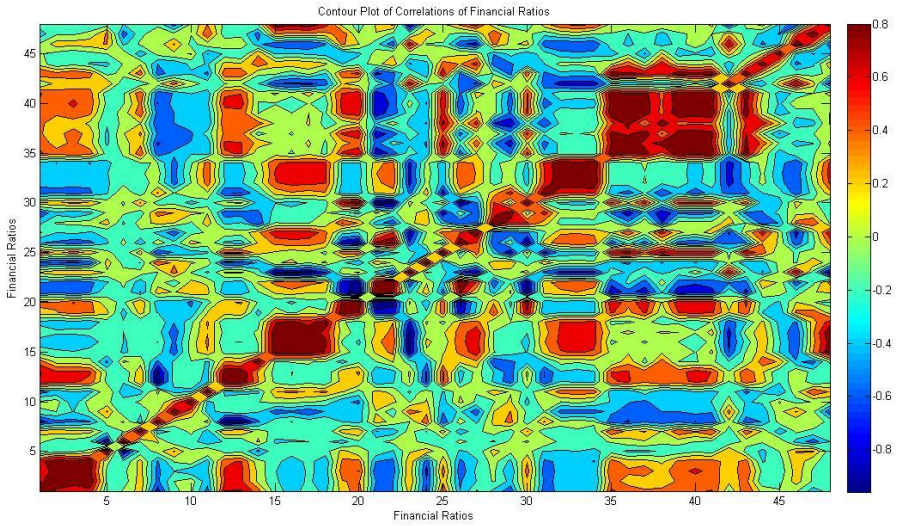


Figure 12: 2-D Contour Plot of Correlations of Financial Ratios

As it can be seen from *Figure 12*, the correlation matrix indicates some strong positive or negative correlations between some of the financial ratios. Since it is clear that there is a great deal of dependence between variables, one wonders whether this problem could be reduced to a problem with less number of dimensions.

Table 1: Eigenvalues and Total Variance Represented by Factors

Factor	Eigenvalues	Cum.Var	Factor	Eigenvalues	Cum.Var	Factor	Eigenvalues	Cum.Var
F1	226176,51	89,30%	**F17**	8,63	100,00%	**F33**	0,00	100,00%
F2	13251,56	94,54%	**F18**	318,25%	100,00%	**F34**	0,00	100,00%
F3	7224,55	97,39%	**F19**	295,22%	100,00%	**F35**	0,00	100,00%
F4	3200,51	98,65%	**F20**	149,56%	100,00%	**F36**	0,00	100,00%
F5	1112,13	99,09%	**F21**	138,22%	100,00%	**F37**	0,00	100,00%
F6	1031,89	99,50%	**F22**	42,52%	100,00%	**F38**	0,00	100,00%
F7	463,87	99,68%	**F23**	0,00%	100,00%	**F39**	0,00	100,00%
F8	289,23	99,80%	**F24**	0,00%	100,00%	**F40**	0,00	100,00%
F9	214,36	99,88%	**F25**	0,00%	100,00%	**F41**	0,00	100,00%
F10	93,69	99,92%	**F26**	0,00%	100,00%	**F42**	0,00	100,00%
F11	56,50	99,94%	**F27**	0,00%	100,00%	**F43**	0,00	100,00%
F12	48,81	99,96%	**F28**	0,00%	100,00%	**F44**	0,00	100,00%
F13	30,56	99,97%	**F29**	0,00%	100,00%	**F45**	0,00	100,00%
F14	22,88	99,98%	**F30**	0,00%	100,00%	**F46**	0,00	100,00%
F15	17,29	99,99%	**F31**	0,00%	100,00%	**F47**	0,00	100,00%
F16	12,88	99,99%	**F32**	0,00%	100,00%	**F48**	0,00	100,00%

For analysis purposes, firstly mean is subtracted from the data variable-wise (row-wise). Using this newly formed, actually rotated, data set, eigenvalues are calculated. These eigenvalues are then ordered by the largest value. Eigenvalues and corresponding variances are given in *Table 1*. It can be seen from *Table 1* that the first two eigenvalues stand for 94,54% of the total variance. Because of this reason, it has been decided to work with the first two principal components F1 and F2 to represent the sample data. It should be emphasized that this result is significant considering that the representation space of data has been reduced from 48 variables to a 2 dimensional space.

CHAPTER 5

EMPIRICAL RESULTS AND DISCUSSION

48 financial ratios for each of 23 banks have been used in this study. Using the fuzzy clustering algorithm for six clusters, the following probability matrix given in *Table 2* has been found:

Table 2: Fuzzy c-Means Clustering Membership Values for Turkish Banking Sector

Bank	p(group1)	p(group2)	p(group3)	p(group4)	p(group5)	p(group6)
T.C. Ziraat Bankası A.Ş.	0,00	0,25	0,01	0,40	0,23	0,11
Türkiye Halk Bankası A.Ş.	0,00	0,16	0,02	0,55	0,15	0,12
T. Vakıflar Bankası T.A.O.	0,00	0,09	0,00	0,83	0,05	0,02
Akbank T.A.Ş.	0,00	0,13	0,01	0,43	0,23	0,20
Alternatif Bank A.Ş.	0,00	0,10	0,02	0,10	0,24	0,54
Anadolubank A.Ş.	0,00	0,09	0,00	0,07	0,67	0,17
Şekerbank T.A.Ş.	0,00	0,30	0,01	0,11	0,49	0,10
Tekstil Bankası A.Ş.	0,00	0,05	0,00	0,03	0,87	0,05
Turkish Bank A.Ş.	0,00	0,27	0,01	0,13	0,41	0,18
T. Ekonomi Bankası A.Ş.	0,00	0,05	0,00	0,04	0,74	0,16
T. Garanti Bankası A.Ş.	0,00	0,07	0,01	0,86	0,05	0,03
T. İş Bankası A.Ş.	0,00	0,05	0,00	0,89	0,04	0,02
Yapı ve Kredi Bankası A.Ş.	0,00	0,14	0,01	0,37	0,28	0,20
Arap Türk Bankası A.Ş.	0,00	0,00	1,00	0,00	0,00	0,00
Citibank A.Ş.	0,00	0,18	0,04	0,25	0,27	0,26
Denizbank A.Ş.	0,00	0,09	0,01	0,11	0,21	0,57
Deutsche Bank A.Ş.	1,00	0,00	0,00	0,00	0,00	0,00
Eurobank Tekfen A.Ş.	0,00	0,87	0,00	0,05	0,06	0,02
Fibabanka A.Ş.	0,01	0,32	0,08	0,25	0,20	0,14
Finans Bank A.Ş.	0,00	0,04	0,00	0,05	0,16	0,74
HSBC Bank A.Ş.	0,00	0,03	0,00	0,03	0,09	0,85
ING Bank A.Ş.	0,00	0,05	0,01	0,05	0,20	0,70
Turkland Bank A.Ş.	0,00	0,13	0,00	0,06	0,72	0,08

After editing the *Table 2*, and adding color codes for easy identification, *Table 3* is obtained, which shows the membership ratios and color codes indicating which bank is most likely to be grouped under which cluster.

Table 3: FCM Clustering Membership Values for Turkish Banking Sector (Color Coded)

Bank (cluster size=6)	p(group1)	p(group2)	p(group3)	p(group4)	p(group5)	p(group6)	group
Deutsche Bank A.Ş.	1,00	0,00	0,00	0,00	0,00	0,00	1
Eurobank Tekfen A.Ş.	0,00	0,87	0,00	0,05	0,06	0,02	2
Fibabanka A.Ş.	0,01	0,32	0,08	0,25	0,20	0,14	2
Arap Türk Bankası A.Ş.	0,00	0,00	1,00	0,00	0,00	0,00	3
T.C. Ziraat Bankası A.Ş.	0,00	0,25	0,01	0,40	0,23	0,11	4
Türkiye Halk Bankası A.Ş.	0,00	0,16	0,02	0,55	0,15	0,12	4
T. Vakıflar Bankası T.A.O.	0,00	0,09	0,00	0,83	0,05	0,02	4
Akbank T.A.Ş.	0,00	0,13	0,01	0,43	0,23	0,20	4
Türkiye Garanti Bankası A.Ş.	0,00	0,07	0,01	0,86	0,05	0,03	4
Türkiye İş Bankası A.Ş.	0,00	0,05	0,00	0,89	0,04	0,02	4
Yapı ve Kredi Bankası A.Ş.	0,00	0,14	0,01	0,37	0,28	0,20	4
Anadolubank A.Ş.	0,00	0,09	0,00	0,07	0,67	0,17	5
Şekerbank T.A.Ş.	0,00	0,30	0,01	0,11	0,49	0,10	5
Tekstil Bankası A.Ş.	0,00	0,05	0,00	0,03	0,87	0,05	5
Turkish Bank A.Ş.	0,00	0,27	0,01	0,13	0,41	0,18	5
Türk Ekonomi Bankası A.Ş.	0,00	0,05	0,00	0,04	0,74	0,16	5
Citibank A.Ş.	0,00	0,18	0,04	0,25	0,27	0,26	5
Turkland Bank A.Ş.	0,00	0,13	0,00	0,06	0,72	0,08	5
Alternatif Bank A.Ş.	0,00	0,10	0,02	0,10	0,24	0,54	6
Denizbank A.Ş.	0,00	0,09	0,01	0,11	0,21	0,57	6
Finans Bank A.Ş.	0,00	0,04	0,00	0,05	0,16	0,74	6
HSBC Bank A.Ş.	0,00	0,03	0,00	0,03	0,09	0,85	6
ING Bank A.Ş.	0,00	0,05	0,01	0,05	0,20	0,70	6

In *Table 3*, red-yellow-green color spectrum has been used to indicate the probability densities. In this table, red circle is an indication of a very low probability that a certain bank is grouped under that cluster, yellow circle indicates that there is medium amount of probability whereas the green circle shows that there is high probability that a certain bank falls under that cluster. Also, the intensity and the length of blue data bar in each cell is positively correlated with the probabilities. This particular grouping is the most probable grouping given the financial ratios as inputs.

As it can be seen from *Table 4* in this setting, Türkiye Cumhuriyeti Ziraat Bankası A.Ş., Türkiye Halk Bankası A.Ş., Türkiye Vakıflar Bankası T.A.O., Akbank T.A.Ş.;Türkiye Garanti Bankası A.Ş., Türkiye İş Bankası A.Ş. and Yapı ve Kredi Bankası A.Ş. have been grouped

together. This group of banks are also the banks which are the largest with respect to their asset size. Another group of banks is Anadolubank A.Ş., Şekerbank T.A.Ş., Tekstil Bankası A.Ş., Turkish Bank A.Ş., Türk Ekonomi Bankası A.Ş., Citibank A.Ş., Turkland Bank A.Ş.. In group 6, Alternatif Bank A.Ş., Denizbank A.Ş., Finans Bank A.Ş., HSBC Bank A.Ş. and ING Bank A.Ş. are clustered together. In group 2, anly Eurobank Tekfen A.Ş., Fibabanka A.Ş. are clustered together whereas group 1 and 3 contain only Deutsche Bank A.Ş. and Arap Türk Bankası A.Ş. respectively.

Table 4: Table of Banks Grouped Under 6 Clusters Using Fuzzy 6-Means Clustering

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Deutsche Bank A.Ş.	Eurobank Tekfen A.Ş.	Arap Türk Bankası A.Ş.	T.C. Ziraat Bankası A.Ş.	Anadolubank A.Ş.	Alternatif Bank A.Ş.
	Fibabanka A.Ş.		Türkiye Halk Bankası A.Ş.	Şekerbank T.A.Ş.	Denizbank A.Ş.
			Türkiye Vakıflar Bankası T.A.O.	Tekstil Bankası A.Ş.	Finans Bank A.Ş.
			Akbank T.A.Ş.	Turkish Bank A.Ş.	HSBC Bank A.Ş.
			Türkiye Garanti Bankası A.Ş.	Türk Ekonomi Bankası A.Ş.	ING Bank A.Ş.
			Türkiye İş Bankası A.Ş.	Citibank A.Ş.	
			Yapı ve Kredi Bankası A.Ş.	Turkland Bank A.Ş.	

The fuzzy c-means clustering minimized the objective function in 24 iterations and found that $\min J(U,V) = 79659$.

To help visualizing the membership probabilities of 23 banks for 6 clusters, a visual 3-D graph is given in *Figure 13*.

At each step of optimization, the value of the objective function is improved and it is monotonically non increasing. The behavior of objection function at each iteration is shown in *Figure 14*.

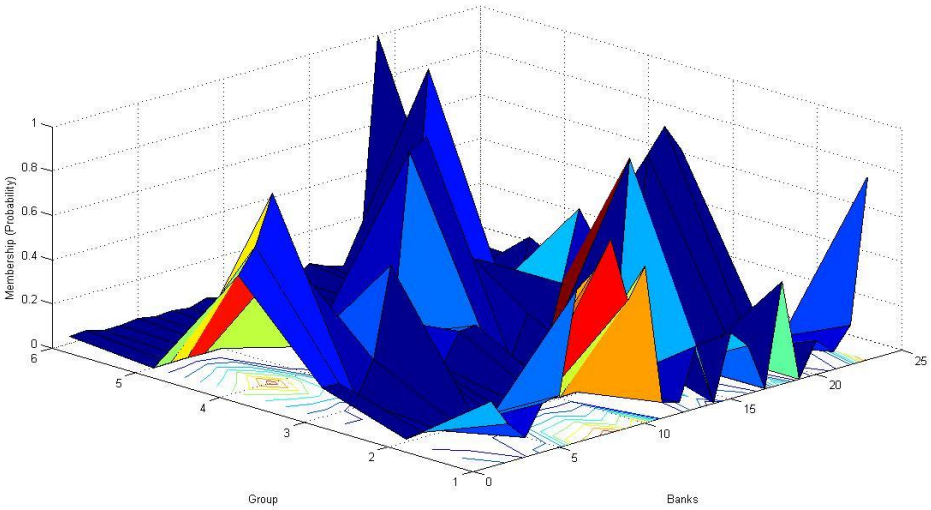


Figure 13: 3-Dimensional Representation of Membership Values

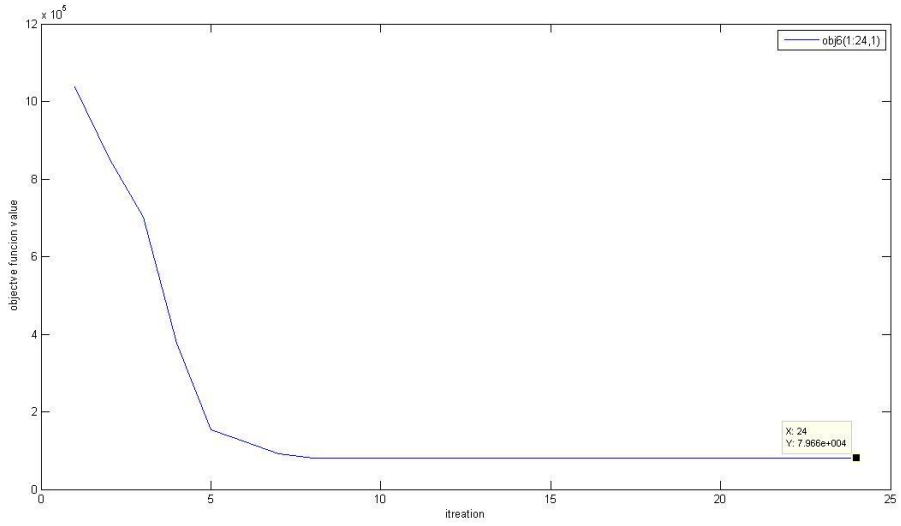


Figure 14: Objective Function Graph for Fuzzy 6-Means Clustering Application

As the number of cluster to be formed changes, the number of iterations required for minimizing the objective function oscillates and is not monotonic. *Figure 15* shows the behavior of number of iterations as the number of clusters is increased.

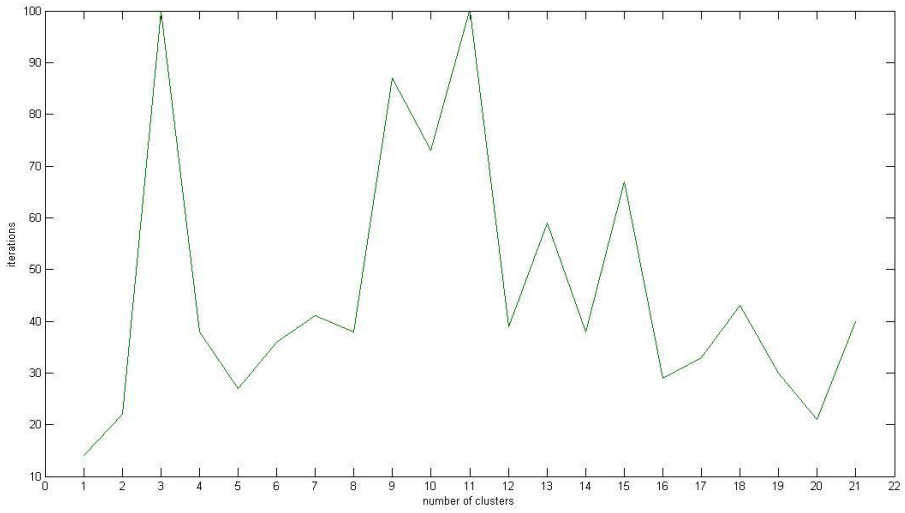


Figure 15: Number of Iterations Required for Optimization vs Number of Clusters

Figure 15 does not help the researcher very much since there is not much useful pattern. What can be inferred from the graph is that the required iteration numbers just fluctuate between very high and low values, as the number of clusters is increased.

Table 5: Iterations, Objective Function Values vs Number of Clusters

clusters	Iteration	max obj	min obj		clusters	Iteration	max obj	min obj
n=2	14	2885816	808383		**n=13**	39	498564	19195
n=3	22	2658495	396355		**n=14**	59	486384	18059
n=4	100	1634187	266235		**n=15**	38	459726	14021
n=5	38	1328352	138249		**n=16**	67	419647	10456
n=6	27	1262191	79659		**n=17**	29	487506	6990
n=7	36	1084974	59466		**n=18**	33	473703	5257
n=8	41	930105	47828		**n=19**	43	382449	4517
n=9	38	853785	38611		**n=20**	30	364289	4225
n=10	87	701401	32427		**n=21**	21	362556	2772
n=11	73	610591	27025		**n=22**	40	345494	691
n=12	100	519079	21040					

Table 5 lists the analysis results of fuzzy c-means clustering study. It summarizes the behavior of number of iterations required for optimization, maximum objective function value and mimimum objective function (optimum) value altogether.

But when analyzed, *Figure 16* shows that as the number of clusters used increases, initial non-optimized maximum objective function values (shown in green) as well as optimized minimum objective function values (shown in blue) monotonically decrease. Another observation is that the 'difference between initial non-optimized and final optimized values of objective function' starts to decrease and these two figures tend to converge to each other.

As it has been mentioned before, fuzzy clustering algorithm groups the banks having similar financial characteristics. But this algorithm does not give any information regarding which cluster is composed of banks having the best or worst characteristics.

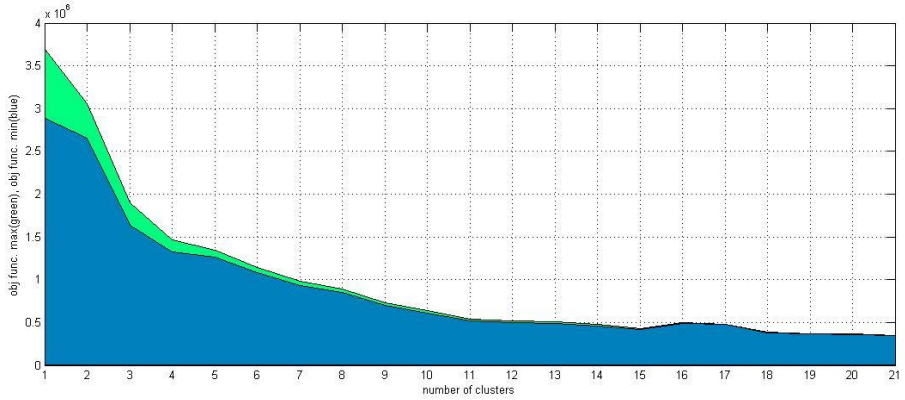


Figure 16: Objective Function Values vs Number of Clusters

The results of the fuzzy c-means clustering algorithm have been processed further using these ten financial ratios so that the weighted score of the clusters could be used in ranking these clusters. The final results are given in Table 6. It can be seen from Table 6 that the cumulative scores are not too much deviant with respect to each other. Therefore these cumulative scores do not imply absolute superiority. But they may be taken as indicators for comparative purposes.

Table 6: Group Scores for Clusters

Bank	Ind. Score	Group	Cum. Gr. Score
Deutsche Bank A.Ş.	0,64	1	0,64
Arap Türk Bankası A.Ş.	0,57	3	0,57
T.C. Ziraat Bankası A.Ş.	0,46	4	0,52
Türkiye Halk Bankası A.Ş.	0,44	4	0,52
T. Vakıflar Bankası T.A.O.	0,48	4	0,52
Akbank T.A.Ş.	0,71	4	0,52
Türkiye Garanti Bankası A.Ş.	0,64	4	0,52
Türkiye İş Bankası A.Ş.	0,48	4	0,52
Yapı ve Kredi Bankası A.Ş.	0,44	4	0,52
Alternatif Bank A.Ş.	0,21	6	0,50
Denizbank A.Ş.	0,58	6	0,50
Finans Bank A.Ş.	0,70	6	0,50
HSBC Bank A.Ş.	0,55	6	0,50
ING Bank A.Ş.	0,46	6	0,50
Anadolubank A.Ş.	0,58	5	0,46
Şekerbank T.A.Ş.	0,28	5	0,46
Tekstil Bankası A.Ş.	0,47	5	0,46
Turkish Bank A.Ş.	0,47	5	0,46
Türk Ekonomi Bankası A.Ş.	0,45	5	0,46
Citibank A.Ş.	0,51	5	0,46
Turkland Bank A.Ş.	0,48	5	0,46
Eurobank Tekfen A.Ş.	0,45	2	0,43
Fibabanka A.Ş.	0,42	2	0,43

Further calculations allowed us to find numerical average scores for clusters. The results show that, according to the selected ratios, Deutsche Bank A.Ş. is the numerically superior bank. Following score belongs to the Arap Türk Bankası A.Ş. This particular rank of Arap Türk Bankası A.Ş may not be very realistic for application purposes. But the financial ratio set selected produces this result. This score mainly comes from the high capital adequacy ratio of Arap Turk A.Ş., which is actually an important ratio but it is only as important as other ratios. Following rank belongs to the group 4 (T.C. Ziraat Bankası A.Ş., Türkiye Halk Bankası A.Ş., T. Vakıflar Bankası T.A.O., Akbank T.A.Ş., Türkiye Garanti Bankası A.Ş., Türkiye İş Bankası A.Ş., Yapı ve Kredi Bankası A.Ş.) which is composed of the largest banks in Turkey. Common sense says that group 4 should have the highest rank because of their financial stability and magnitude. But the selected ratio set does not produce that result. Following

rank belongs to the cluster 6 (Alternatif Bank A.Ş., Denizbank A.Ş., Finans Bank A.Ş., HSBC Bank A.Ş., ING Bank A.Ş.), the group that is composed of the second largest banks of Turkey after group 4. In a way this is expected too, since the largest banks are expected to have better governance structures than the rest of the banks. The last two rank belongs to the group 5 (Anadolubank A.Ş., Şekerbank T.A.Ş., Tekstil Bankası A.Ş., Turkish Bank A.Ş., Türk Ekonomi Bankası A.Ş., Citibank A.Ş., Turkland Bank A.Ş.) and group 2 (Eurobank Tekfen A.Ş., Fibabanka A.Ş.) respectively.

Principal Component Analysis

As it has been mentioned before, the correlation and covariance matrices of the financial ratios are high between some of the financial ratios used. These values are given in *Appendix E* and *Appendix F* respectively and also high correlation between variables is plotted on a 2-D Filled Contour Graph in *Figure 11*. As it can be seen from the legend on the right hand side of the plot, darker red corresponds to positive high correlation whereas darker blue regions correspond to negative high correlation.

To decrease the co-dependence between variables, principal component analysis is applied. Firstly mean is subtracted from the data variable-wise (row-wise). Using this newly formed, actually rotated, data set, eigenvalues are calculated. These eigenvalues are then ordered by the largest value. Eigenvalues and corresponding variances are given in *Table 1*. It can be seen from *Table 1* that the first two eigenvalues stand for 94,54% of the total variance. Because of this reason, it has been decided to work with the first two principal components F1 and F2 to represent the sample data. It should be emphasized that this result is significant considering that the representation space of data has been reduced from 48 variables to a 2 dimensional space. To represent each of the companies from the sample in a unique graph, the values that each of the principal components take for each firm have been calculated and can be seen in *Table 7*.

Table 7: Principal Components

F1	F2	F1	F2
-108,85	-39,03	-100,46	23,71
-97,73	-41,79	359,72	-88,89
-80,60	-79,59	4,48	38,87
-73,29	33,21	-165,23	116,64
-117,47	167,50	2123,61	34,73
-154,67	29,11	-148,11	-86,08
-177,88	-16,93	-80,25	-401,06
-142,30	19,36	-131,89	94,66
-184,53	5,61	-149,76	148,72
-146,54	46,99	-155,42	126,83
-51,60	-72,70	-137,72	-0,47
-83,51	-59,40		

The representation based on the principal components is given in *Figure 17*.

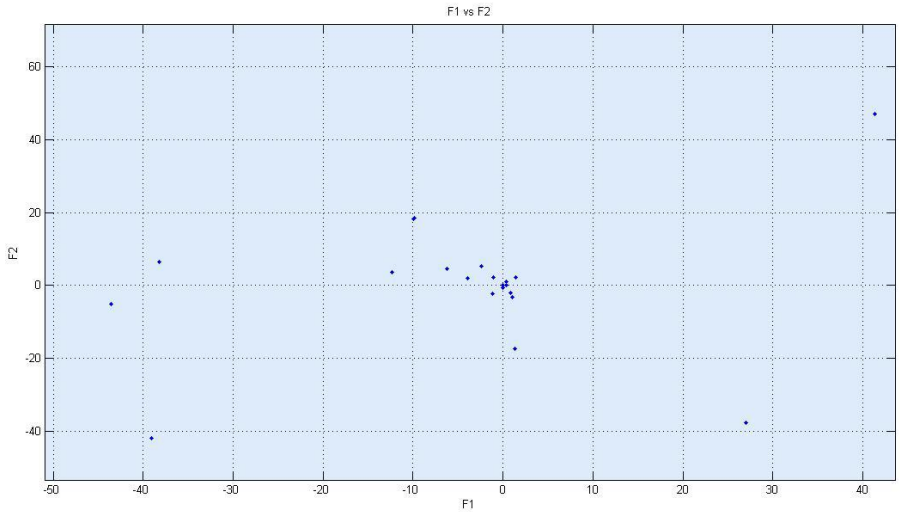


Figure 17: Representation Based on Principal Components

This representation excludes two outliers, i.e. observations with extreme values that deviate from the rest of the sample. It has been decided not to consider these outliers to avoid that the scale of the graph is set in such a way that the rest of the companies can be distinguished.

CHAPTER 6

SUMMARY AND CONCLUSION

Since banking sector holds the role of allocating financial resources, they play important role in economic growth. Nevertheless, banking sector is very sensitive to macroeconomic and political instabilities and they are prone to crises. Since banks are integrated with almost all of the economic agents and with other banks, these crises which may have a very high cost, will probably affect entire society. Since as of September 2012, there are only 48 banks (excluding Odea Bank which is still not a fully authorized bank) in Turkey. Since total number of operating banks is limited, each of the banks may be easily assumed to potentially have a systemic risk triggering affect. In that sense, evaluating banks with respect to their financials and supporting other information could be of help. In this context, fuzzy c-means clustering algorithm has been applied for banks.

The findings of the study are summarized as follows:

Using the financial ratios, banks have been clustered into six clusters. In this setting, Türkiye Cumhuriyeti Ziraat Bankası A.Ş., Türkiye Halk Bankası A.Ş., Türkiye Vakıflar Bankası T.A.O., Akbank T.A.Ş., Türkiye Garanti Bankası A.Ş., Türkiye İş Bankası A.Ş. and Yapı ve Kredi Bankası A.Ş. have been grouped under group 4 and this group contains 7 of the largest banks in Turkey with respect to asset size. Other groups (group 1, 2 ,3, 5, 6) contain 1, 2, 1, 7 and 5 banks respectively.

Number of clusters to be formed and the number of iterations required for minimizing the objective function oscillates and is not monotonic.

As the number of clusters used increases, initial non-optimized maximum objective function values as well as optimized final minimum objective function values monotonically decrease together. Another observation is that the ‘difference between initial non-optimized and final optimized values of objective function’ starts to diminish as number of clusters increases.

Finally, the results of the fuzzy c-means clustering algorithm have been processed further using these ten financial ratios so that the weighted score of the clusters could be used in ranking these clusters. The results show that, according to the selected ratios, Deutsche Bank A.Ş. is the numerically superior bank. Following score belongs to the Arap Türk Bankası A.Ş. Third ranked group is composed of the largest banks in Turkey, whereas fourth ranked group is composed of the second largest banks in Turkish banking system. Some of these results may not be very realistic for application purposes, but the financial ratio set selected produces this result. This score mainly comes from the high capital adequacy ratio of first two banks.

To get rid of the dependency, PCA is applied, thus the representation space of data has been reduced from 48 variables to a 2 dimensional space. These two vectors are found to represent the 94.54% of total variation.

LITERATURE CITED

Alam, P., Booth, D., Lee, K., & Thordarson, T. (2000). The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing bank. *Expert Systems with Applications* , 185-199.

Allen, J. C. (1995). A promise of approvals in minutes, not hours. *American Banker* .

Altman, E. I. (1968, Sep). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance* , pp. 589-609.

Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETATM Analysis A New Model to Identify Bankruptcy Risk of Corporations. *Journal of Banking & Finance* , 29-54.

Andrés, J. D., Lorca, P., Juez, F. J., & Sánchez-Lasheras, F. (2011). Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS). *Expert Systems with Applications* , 1866–1875.

Bataineh, K. M., Najia, M., & Saqera, M. (Aug. 2011). A Comparison Study between Various Fuzzy Clustering Algorithms. *Jordan Journal of Mechanical and Industrial Engineering* , 335 - 343.

Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research, Vol. 4, Empirical Research in Accounting: Selected Studies 1966* , pp. 71-188.

Bell, T. (1997). Neural nets or the logit model? A comparison of each model's ability to predict commercial bank failures. *International Journal of Intelligent Systems in Accounting, Finance and Management* , 249–264.

Berger, A. N., Frame, W. S., & Miller, N. H. (2002). *Credit Scoring and the Availability, Price, and Risk of Small Business Credit.* Atlanta: Federal Reserve Board.

Bezdek, J. (1973). *Fuzzy Mathematics in Pattern Classification.* Ph. D. Thesis, Applied Math. Center, Cornell University: Ithaca.

BRSA. (2006, 2 23). Bank Cards and Credit Cards Law No: 5464. *Article 9* .

BRSA. (2011). Circular Regarding Credit Risk Management. *Kredi Riski Yönetimi Hakkında Tebliğ (Taslak)* . Ankara, Turkey.

BRSA. (June 2012). *Finansal Piyasalar Raporu.* Ankara: BDDK.

Caouette, J. B., Altman, E. I., Narayanan, P., & Nimmo, R. (2008). *The Great Challenge for the Global Financial Markets.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning* , 273-297.

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

Dietrich, J. R., & Kaplan, R. S. (1982). Empirical Analysis of the Commercial Loan Classification Decision. *The Accounting Review* , 18-38.

Duda, R., & Hart, P. (1973). *Pattern Classification and Scene Analysis.* New York : Wiley.

Dudani, S. A. (1976). The Distance-Weighted k-Nearest-Neighbor Rule. *Systems, Man and Cybernetics, IEEE Transactions* , 325-327.

Dunn, J. C. (1973). A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics* , 32 — 5.

Durand, D. (1941). Risk Elements in Consumer Installment Financing. *NBER* .

FDIC. (2012, 9 24). *FDIC: Failed Bank List*. Retrieved 9 24, 2012, from FDIC: http://www.fdic.gov/bank/individual/failed/banklist.html

Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics* , 179-188.

Fitzpatrick, P. J. (1932). A comparison of the ratios of successful industrial enterprises with those of failed companies. *Certified Public Accountant* , 598-605.

Galindo, J., & Tamayo, P. (2000). Credit Risk Assessment Using Statistical and Machine Learning: Basic Methodology and Risk Modeling Applications. *COMPUTATIONAL ECONOMICS* , 107-143.

Gentry, J. A., Newbold, P., & Whitford, D. T. (1985). Classifying Bankrupt Firms with Fund Flow Components. *Journal of Accounting Research* , 146-160.

Gentry, J. A., Newbold, P., & Whitford, D. T. (1987). Funds Flow Components, Financial Ratios, and Bankruptcy. *Journal of Business Finance & Accounting* , 595–606.

Grice, J. S., & Dugan, M. T. (2001). The Limitations of Bankruptcy Prediction Models: Some Cautions for the Researcher. *Review of Quantitative Finance and Accounting* , 151-166.

Grice, J. S., & Ingram, R. W. (2001). Tests of the generalizability of Altman's bankruptcy prediction model. *Journal of Business Research* , 53–61.

Höppner, F., Klawonn, F., Kruse, R., & Runkler, T. (2000). *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition.* West Sussex, England: John Wiley & Sons Ltd.

Hui, X.-F., & Sun, J. (2006). An Application of Support Vector Machine to Companies' Financial Distress Prediction. In V. Torra, Y. Narukawa, A. Valls, & J. Domingo-Ferrer, *Modeling Decisions for Artificial Intelligence* (pp. 274-282). Springer Berlin / Heidelberg.

IMF. (2010). *World Economic Outlook: Recovery, Risk and Rebalancing.* Washington, D.C.: International Monetary Fund.

Jones, S., & Hensher, D. A. (2007). Modelling corporate failure: A multinomial nested logit analysis for unordered outcomes. *The British Accounting Review* , 89–107.

Jones, S., & Hensher, D. A. (2004). Predicting Firm Financial Distress: A Mixed Logit Model. *The Accounting Review* , 1011-1038.

Jose, A. S., & Georgiou, A. (2009). Financial soundness indicators (FSIs): framework and implementation. *Proceedings of the IFC Conference* (pp. 277-282 in Bank for International Settlements). Basel: BIS.

Karels, G. V., & Prakash, A. J. (1987). Multivariate Normality and Forecasting of Business Bankruptcy. *Journal of Business Finance & Accounting* , 573–593.

Kolari, J., Glennon, D., Shin, H., & Caputo, M. (2002). Predicting large US commercial bank failures. *Journal of Economics and Business* , 361–387.

Korobow, L., Stuhr, D. P., & Martin, D. (Autumn 1977). A Nationwide Test of Early Warning Research in Banking. *Federal Reserve Bank of New York Quarterly Review* , 37-52.

Laitinen, E. K., & Laitinen, T. (2000). Bankruptcy prediction: Application of the Taylor's expansion in logistic regression. *International Review of Financial Analysis* , 327-349.

Li, H., & Sun, J. (2011). Empirical research of hybridizing principal component analysis with multivariate discriminant analysis and logistic regression for business failure prediction. *Expert Systems with Applications* , 6244–6253.

Li, H., & Sun, J. (2011). Predicting business failure using forward ranking-order case-based reasoning. *Expert Systems with Applications* , Expert Systems with Applications.

Li, H., Sun, J., & Sun, B.-L. (2009). Financial distress prediction based on OR-CBR in the principle of k-nearest neighbors. *Expert Systems with Applications* , 643–659.

Lin, T.-H. (2009). A cross model study of corporate financial distress prediction in Taiwan: Multiple discriminant analysis, logit, probit and neural networks models. *Neurocomputing* , 3507–3516.

Martin, D. (1977). Early warning of bank failure : A logit regression approach. *Journal of Banking & Finance* , 249-276.

Michael, S., Georgios, D., Nikolaos, M., & Constantin, Z. (1999). A Fuzzy Knowledge-Based Decision: Aiding Method for the Assessment of Financial Risks. *European Symposium on Intelligent Techniques.* Crete.

Min, J. H., & Lee, Y.-C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications* , 603–614.

Moorhouse, A. (2004, February). An introduction to Financial Soundness Indicators. *Monetary & Financial Statistics* .

Myers, J. H., & Forgy, E. W. (1963, 9). The Development of Numerical Credit Evaluation Models. *Journal of the American Statistical Association* , pp. 799-806.

myFICO.com. (2012). Retrieved 8 12, 2012, from myFICO.com: http://www.myfico.com/CreditEducation/CreditScores.aspx

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research* , 109-131.

Pal, N. R., & Bezdek, J. C. (1995). On Cluster Validity for the Fuzzy c-Means Model. *IEEE Transactions on Fuzzy Systems* , 370-379.

Park, C.-S., & Han, I. (2002). A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction. *Expert Systems with Applications* , 255–264.

R. Barniv, A. A. (1997). Predicting the outcome following bankruptcy filing: A three state classification using NN. *International Journal of Intelligent Systems in Accounting, Finance and Management* , 177–194.

Servigny, A. d., & Renault, O. (2004). *Measuring and Managing Credit Risk.* Mc-Graw Hill.

Shin, K.-S., Lee, T. S., & Kim, H.-j. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications* , 127–135.

Sinkey, J. F. (1975). A Multivariate Statistical Analysis of the Characteristics of Problem Banks. *The Journal of Finance* , 21-36.

Smith, J. F. (1977, 5 2). The Equal Credit Opportunity Act Of 1974: A Cost/Benefit Analysis. *The Journal of Finance* , pp. 609-622.

Stephey, M. (2009). A Brief History of: Credit Cards. *Time Magazine* .

TBB. (2012). *TBB*. Retrieved 9 14, 2012, from Turkish Banking Association: http://www.tbb.org.tr/eng/Banka_ve_Sektor_Bilgileri/Tum_Raporlar.aspx

Webb, A. R., & Copsey, K. D. (2011). *Statistical Pattern Recognition.* The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom: John Wiley & Sons, Ltd.

West, R. C. (1985). A factor-analytic approach to bank condition. *Journal of Banking & Finance* , 253–266.

Yip, A. Y. (2004). Predicting Business Failure with a Case-Based Reasoning Approach. In M. Negoita, R. Howlett, & L. Jain, *Knowledge-Based Intelligent Information and Engineering Systems* (pp. 665-671). Springer Berlin / Heidelberg.

Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research* , 59-82.

Appendix A: Financial Soundness Indicators[12]

Core Set	
Deposit-taking institutions (banks)	
Capital adequacy	Regulatory capital to risk-weighted assets
	Regulatory Tier I capital to risk-weighted assets
	Nonperforming loans net of provisions to capital
Asset quality	Nonperforming loans to total gross loans
	Sectoral distribution of loans to total loans
Earnings and profitability	Return on assets
	Return on equity
	Interest margin to gross income
	Noninterest expenses to gross income
Liquidity	Liquid assets to total assets (liquid asset ratio)
	Liquid assets to short-term liabilities
Sensitivity to market risk	Net open position in foreign exchange to capital
Encouraged Set	
Deposit-taking institutions (banks)	Capital to assets
	Large exposures to capital
	Geographical distribution of loans to total loans
	Gross asset position in financial derivatives to capital
	Gross liability position in financial derivatives to capital
	Trading income to total income
	Personnel expenses to noninterest expenses
	Spread between reference lending and deposit rates
	Spread between highest and lowest interbank rate
	Customer deposits to total (non-interbank) loans
	Foreign currency-denominated loans to total loans
	Foreign currency-denominated liabilities to total liabilities
	Net open position in equities to capital
Other financial corporations	Assets to total financial system assets
	Assets to GDP
Nonfinancial corporations sector	Total debt to equity
	Return on equity
	Earnings to interest and principal expenses
	Net foreign exchange exposure to equity
	Number of applications for protection from creditors
Households	Household debt to GDP
	Household debt service and principal payments to income
Market liquidity	Average bid-ask spread in the securities market
	Average daily turnover ratio in the securities market
Real estate markets	Residential real estate prices
	Commercial real estate prices
	Residential real estate loans to total loans
	Commercial real estate loans to total loans

[12] Prepared by IMF

Appendix B: Financial Ratios

Category	Code	Ratio
Capital Ratios	C1	Shareholders' Equity / (Amount Subj. to Credit R.+ Market R. + Op. R.)
	C2	Shareholders' Equity / Total Assets
	C3	(Shareholders' Equity-Permanent Assets) / Total Assets
	C4	Shareholders' Equity / (Deposits + Non-Deposit Funds)
	C5	On Balance-sheet FC Position / Shareholders' Equity
	C6	Net on Balance-sheet Position / Total Shareholders' Equity
	C7	N(on+off) Balance-sheet Position / Total Shareholders' Equity
Assets Quality	A1	Total Loans and Receivables* / Total Assets
	A2	Loans under follow-up (net) / Total Loans and Receivables
	A3	Permanent Assets / Total Assets
	A4	Consumer Loans / Total Loans and Receivables
Liquidity	L1	Liquid Assets / Total Assets
	L2	Liquid Assets / (Deposits + Non-Deposit Funds)
	L3	FC Liquid Assets / FC Liabilities
Profitability	P1	Net Profit (Losses) / Total Assets
	P2	Net Profit (Losses) / Total Shareholders' Equity
	P3	Income Before Taxes / Total Assets
	P4	Net Profit (Losses) / Paid-in Capital
Income-Expenditure	I1	Net Interest Income After Specific Provisions / Total Assets
	I2	Net Interest Income After Specific Provisions / Total Oper.Income
	I3	Non-Interest Income (Net) / Total Assets
	I4	Non-Interest Income (Net) / Other Operating Expenses
	I5	Other Operating Expenses / Total Operating Income
	I6	Provision For Loan or Other Receivables Losses / Total Assets
	I7	Interest Income / Interest Expense
	I8	Non-Interest Income / Non-Interest Expense
	I9	Total Income / Total Expense
	I10	Interest Income / Total Assets
	I11	Interest Expense / Total Assets
	I12	Interest Income / Total Expenses
	I13	Interest Expense / Total Expenses
Share	S1	Total Assets
	S2	Total Loans and Receivables*
	S3	Total deposits
Branch	B1	Total Assets / No. of Branches
	B2	Total Deposits / No. of Branches
	B3	TRY Deposits / No. of Branches
	B4	FX Deposits / No. of Branches
	B5	Total Loans and Receivables* / No. of Branches
	B6	Total Employees / No. of Branches (person)
	B7	Net Income / No. of Branches
Activity	Ac1	(Personnel Exp. + Reserve for Emp. Term. Benefit) / Tot.Assets
	Ac2	(Personnel Exp. + Reserve for Emp. Term. Benefit) / # Personnel
	Ac3	Reserve for Employee Termination Benefit / # Personnel
	Ac4	Personnel Expenses / Other Operating Expenses
	Ac5	Other Operating Expenses / Total Asset
	Ac6	Total Operating Income / Total Assets
	Ac7	Net Operating Income(Loss) / Total Assets

Appendix C: Predictive Statistics for Financial Ratios Used

C1		C2		C3	
Mean	17,31	Mean	11,89	Mean	9,15
Standard Error	1,04	Standard Error	0,56	Standard Error	0,59
Median	16,14	Median	11,39	Median	8,92
Standard Deviation	4,99	Standard Deviation	2,62	Standard Deviation	2,85
Sample Variance	24,86	Sample Variance	6,89	Sample Variance	8,11
Kurtosis	4,83	Kurtosis	1,24	Kurtosis	0,90
Skewness	2,29	Skewness	0,91	Skewness	0,78
Minimum	13,24	Minimum	7,52	Minimum	4,46
Maximum	32,09	Maximum	18,61	Maximum	16,74
Sum	398,04	Sum	261,62	Sum	210,47
Count	23,00	Count	22,00	Count	23,00
Largest(1)	32,09	Largest(1)	18,61	Largest(1)	16,74
Smallest(1)	13,24	Smallest(1)	7,52	Smallest(1)	4,46
Conf. Level(95,0%)	2,16	Conf. Level(95,0%)	1,16	Conf. Level(95,0%)	1,23
C4		**C5**		**C6**	
Mean	14,75	Mean	62,75	Mean	-17,75
Standard Error	0,74	Standard Error	15,69	Standard Error	18,74
Median	14,16	Median	73,50	Median	-24,71
Standard Deviation	3,57	Standard Deviation	75,23	Standard Deviation	89,90
Sample Variance	12,75	Sample Variance	5659,67	Sample Variance	8081,24
Kurtosis	0,66	Kurtosis	2,05	Kurtosis	12,40
Skewness	0,70	Skewness	-0,11	Skewness	3,06
Minimum	9,04	Minimum	-137,34	Minimum	-119,57
Maximum	23,70	Maximum	244,33	Maximum	343,68
Sum	339,16	Sum	1443,22	Sum	-408,15
Count	23,00	Count	23,00	Count	23,00
Largest(1)	23,70	Largest(1)	244,33	Largest(1)	343,68
Smallest(1)	9,04	Smallest(1)	-137,34	Smallest(1)	-119,57
Conf. Level(95,0%)	1,54	Conf. Level(95,0%)	32,53	Conf. Level(95,0%)	38,87
C7		**A1**		**A2**	
Mean	0,69	Mean	56,77	Mean	0,97
Standard Error	1,18	Standard Error	3,18	Standard Error	0,17
Median	0,05	Median	61,69	Median	0,81
Standard Deviation	5,68	Standard Deviation	15,27	Standard Deviation	0,83
Sample Variance	32,26	Sample Variance	233,23	Sample Variance	0,69
Kurtosis	1,93	Kurtosis	0,17	Kurtosis	-0,14
Skewness	1,35	Skewness	-0,75	Skewness	0,83
Minimum	-6,67	Minimum	23,73	Minimum	0,00
Maximum	16,54	Maximum	84,72	Maximum	2,68
Sum	15,92	Sum	1305,77	Sum	22,31
Count	23,00	Count	23,00	Count	23,00
Largest(1)	16,54	Largest(1)	84,72	Largest(1)	2,68
Smallest(1)	-6,67	Smallest(1)	23,73	Smallest(1)	0,00
Conf. Level(95,0%)	2,46	Conf. Level(95,0%)	6,60	Conf. Level(95,0%)	0,36

Appendix C (continued)

A3		*A4*		*L1*	
Mean	2,89	Mean	24,30	Mean	34,22
Standard Error	0,28	Standard Error	3,67	Standard Error	3,35
Median	2,65	Median	27,49	Median	29,27
Standard Deviation	1,33	Standard Deviation	17,61	Standard Deviation	16,07
Sample Variance	1,78	Sample Variance	310,27	Sample Variance	258,37
Kurtosis	-0,87	Kurtosis	-0,87	Kurtosis	0,41
Skewness	0,38	Skewness	0,05	Skewness	1,10
Minimum	0,66	Minimum	0,02	Minimum	13,36
Maximum	5,36	Maximum	61,63	Maximum	72,31
Sum	66,49	Sum	558,96	Sum	787,07
Count	23,00	Count	23,00	Count	23,00
Largest(1)	5,36	Largest(1)	61,63	Largest(1)	72,31
Smallest(1)	0,66	Smallest(1)	0,02	Smallest(1)	13,36
Conf. Level(95,0%)	0,58	Conf. Level(95,0%)	7,62	Conf. Level(95,0%)	6,95
L2		*L3*		*P1*	
Mean	41,78	Mean	31,74	Mean	1,13
Standard Error	4,15	Standard Error	2,75	Standard Error	0,15
Median	37,59	Median	28,81	Median	1,31
Standard Deviation	19,90	Standard Deviation	13,19	Standard Deviation	0,73
Sample Variance	395,93	Sample Variance	174,07	Sample Variance	0,54
Kurtosis	0,79	Kurtosis	2,25	Kurtosis	-1,23
Skewness	1,18	Skewness	1,45	Skewness	0,05
Minimum	15,92	Minimum	15,61	Minimum	0,08
Maximum	92,11	Maximum	67,58	Maximum	2,43
Sum	960,93	Sum	729,92	Sum	26,04
Count	23,00	Count	23,00	Count	23,00
Largest(1)	92,11	Largest(1)	67,58	Largest(1)	2,43
Smallest(1)	15,92	Smallest(1)	15,61	Smallest(1)	0,08
Conf. Level(95,0%)	8,60	Conf. Level(95,0%)	5,71	Conf. Level(95,0%)	0,32
P2		*P3*		*P4*	
Mean	9,93	Mean	1,41	Mean	36,93
Standard Error	1,41	Standard Error	0,18	Standard Error	8,72
Median	8,61	Median	1,73	Median	20,04
Standard Deviation	6,78	Standard Deviation	0,86	Standard Deviation	41,80
Sample Variance	45,97	Sample Variance	0,73	Sample Variance	1747,23
Kurtosis	-0,78	Kurtosis	-1,24	Kurtosis	2,95
Skewness	0,33	Skewness	-0,06	Skewness	1,71
Minimum	0,45	Minimum	0,07	Minimum	0,86
Maximum	23,67	Maximum	2,89	Maximum	163,61
Sum	228,42	Sum	32,33	Sum	849,42
Count	23,00	Count	23,00	Count	23,00
Largest(1)	23,67	Largest(1)	2,89	Largest(1)	163,61
Smallest(1)	0,45	Smallest(1)	0,07	Smallest(1)	0,86
Conf. Level(95,0%)	2,93	Conf. Level(95,0%)	0,37	Conf. Level(95,0%)	18,08

Appendix C (continued)

I1		*I2*		*I3*	
Mean	3,38	Mean	67,12	Mean	1,34
Standard Error	0,24	Standard Error	3,90	Standard Error	0,19
Median	2,95	Median	63,73	Median	1,62
Standard Deviation	1,13	Standard Deviation	18,72	Standard Deviation	0,89
Sample Variance	1,28	Sample Variance	350,53	Sample Variance	0,79
Kurtosis	3,92	Kurtosis	8,17	Kurtosis	7,42
Skewness	1,69	Skewness	2,39	Skewness	-2,23
Minimum	2,04	Minimum	47,93	Minimum	-1,90
Maximum	7,07	Maximum	136,52	Maximum	2,32
Sum	77,73	Sum	1543,83	Sum	30,81
Count	23,00	Count	23,00	Count	23,00
Largest(1)	7,07	Largest(1)	136,52	Largest(1)	2,32
Smallest(1)	2,04	Smallest(1)	47,93	Smallest(1)	-1,90
Conf. Level(95,0%)	0,49	Conf. Level(95,0%)	8,10	Conf. Level(95,0%)	0,38
I4		*I5*		*I6*	
Mean	51,68	Mean	59,07	Mean	0,67
Standard Error	8,05	Standard Error	3,54	Standard Error	0,06
Median	43,62	Median	55,78	Median	0,66
Standard Deviation	38,63	Standard Deviation	16,95	Standard Deviation	0,30
Sample Variance	1491,99	Sample Variance	287,44	Sample Variance	0,09
Kurtosis	2,10	Kurtosis	-0,94	Kurtosis	0,13
Skewness	-0,73	Skewness	0,36	Skewness	0,74
Minimum	-61,46	Minimum	34,15	Minimum	0,25
Maximum	117,00	Maximum	90,81	Maximum	1,33
Sum	1188,69	Sum	1358,62	Sum	15,30
Count	23,00	Count	23,00	Count	23,00
Largest(1)	117,00	Largest(1)	90,81	Largest(1)	1,33
Smallest(1)	-61,46	Smallest(1)	34,15	Smallest(1)	0,25
Conf. Level(95,0%)	16,70	Conf. Level(95,0%)	7,33	Conf. Level(95,0%)	0,13
I7		*I8*		*I9*	
Mean	204,83	Mean	51,68	Mean	133,50
Standard Error	13,06	Standard Error	8,05	Standard Error	4,72
Median	185,14	Median	43,62	Median	132,36
Standard Deviation	62,65	Standard Deviation	38,63	Standard Deviation	22,64
Sample Variance	3924,62	Sample Variance	1491,99	Sample Variance	512,61
Kurtosis	5,61	Kurtosis	2,10	Kurtosis	4,20
Skewness	2,40	Skewness	-0,73	Skewness	1,58
Minimum	134,85	Minimum	-61,46	Minimum	105,29
Maximum	392,75	Maximum	117,00	Maximum	208,08
Sum	4710,99	Sum	1188,69	Sum	3070,55
Count	23,00	Count	23,00	Count	23,00
Largest(1)	392,75	Largest(1)	117,00	Largest(1)	208,08
Smallest(1)	134,85	Smallest(1)	-61,46	Smallest(1)	105,29
Conf. Level(95,0%)	27,09	Conf. Level(95,0%)	16,70	Conf. Level(95,0%)	9,79

Appendix C (continued)

I10		*I11*		*I12*	
Mean	7,72	Mean	4,03	Mean	85,05
Standard Error	0,32	Standard Error	0,23	Standard Error	2,33
Median	7,99	Median	4,04	Median	83,41
Standard Deviation	1,51	Standard Deviation	1,09	Standard Deviation	11,16
Sample Variance	2,29	Sample Variance	1,19	Sample Variance	124,58
Kurtosis	4,10	Kurtosis	3,49	Kurtosis	7,30
Skewness	-1,48	Skewness	-0,84	Skewness	1,77
Minimum	2,79	Minimum	0,71	Minimum	62,35
Maximum	9,81	Maximum	6,35	Maximum	124,75
Sum	177,55	Sum	92,61	Sum	1956,22
Count	23,00	Count	23,00	Count	23,00
Largest(1)	9,81	Largest(1)	6,35	Largest(1)	124,75
Smallest(1)	2,79	Smallest(1)	0,71	Smallest(1)	62,35
Conf. Level(95,0%)	0,65	Conf. Level(95,0%)	0,47	Conf. Level(95,0%)	4,83
I13		*S1*		*S2*	
Mean	57,19	Mean	4,16	Mean	4,16
Standard Error	2,17	Standard Error	1,03	Standard Error	0,98
Median	59,47	Median	1,81	Median	2,08
Standard Deviation	10,40	Standard Deviation	4,93	Standard Deviation	4,72
Sample Variance	108,13	Sample Variance	24,30	Sample Variance	22,24
Kurtosis	0,15	Kurtosis	-0,49	Kurtosis	-0,83
Skewness	-0,57	Skewness	1,03	Skewness	0,87
Minimum	33,03	Minimum	0,08	Minimum	0,04
Maximum	76,35	Maximum	13,93	Maximum	13,79
Sum	1315,44	Sum	95,59	Sum	95,65
Count	23,00	Count	23,00	Count	23,00
Largest(1)	76,35	Largest(1)	13,93	Largest(1)	13,79
Smallest(1)	33,03	Smallest(1)	0,08	Smallest(1)	0,04
Conf. Level(95,0%)	4,50	Conf. Level(95,0%)	2,13	Conf. Level(95,0%)	2,04
S3		*B1*		*B2*	
Mean	4,31	Mean	210,83	Mean	87,32
Standard Error	1,08	Standard Error	94,27	Standard Error	16,96
Median	1,65	Median	102,30	Median	59,10
Standard Deviation	5,19	Standard Deviation	452,12	Standard Deviation	81,34
Sample Variance	26,97	Sample Variance	204411	Sample Variance	6616,20
Kurtosis	-0,27	Kurtosis	20,90	Kurtosis	6,07
Skewness	1,07	Skewness	4,51	Skewness	2,55
Minimum	0,05	Minimum	44,79	Minimum	27,44
Maximum	16,18	Maximum	2242,02	Maximum	334,84
Sum	99,22	Sum	4849,16	Sum	2008,42
Count	23,00	Count	23,00	Count	23,00
Largest(1)	16,18	Largest(1)	2242,02	Largest(1)	334,84
Smallest(1)	0,05	Smallest(1)	44,79	Smallest(1)	27,44
Conf. Level(95,0%)	2,25	Conf. Level(95,0%)	195,51	Conf. Level(95,0%)	35,17

Appendix C (continued)

B3		B4		B5	
Mean	48,46	Mean	38,86	Mean	83,42
Standard Error	7,84	Standard Error	12,39	Standard Error	21,18
Median	38,63	Median	25,01	Median	57,99
Standard Deviation	37,62	Standard Deviation	59,41	Standard Deviation	101,56
Sample Variance	1415,48	Sample Variance	3530,06	Sample Variance	10314
Kurtosis	11,77	Kurtosis	14,23	Kurtosis	19,31
Skewness	3,11	Skewness	3,66	Skewness	4,25
Minimum	12,76	Minimum	8,18	Minimum	13,15
Maximum	198,20	Maximum	283,68	Maximum	531,98
Sum	1114,69	Sum	893,73	Sum	1918,66
Count	23,00	Count	23,00	Count	23,00
Largest(1)	198,20	Largest(1)	283,68	Largest(1)	531,98
Smallest(1)	12,76	Smallest(1)	8,18	Smallest(1)	13,15
Conf. Level(95,0%)	16,27	Conf. Level(95,0%)	25,69	Conf. Level(95,0%)	43,92
B6		**B7**		**Ac1**	
Mean	24,65	Mean	2,79	Mean	1,44
Standard Error	4,27	Standard Error	1,38	Standard Error	0,10
Median	18,37	Median	0,97	Median	1,49
Standard Deviation	20,46	Standard Deviation	6,61	Standard Deviation	0,46
Sample Variance	418,65	Sample Variance	43,67	Sample Variance	0,21
Kurtosis	12,08	Kurtosis	19,61	Kurtosis	-0,97
Skewness	3,38	Skewness	4,33	Skewness	0,15
Minimum	12,98	Minimum	0,03	Minimum	0,72
Maximum	106,00	Maximum	32,07	Maximum	2,34
Sum	566,91	Sum	64,24	Sum	33,11
Count	23,00	Count	23,00	Count	23,00
Largest(1)	106,00	Largest(1)	32,07	Largest(1)	2,34
Smallest(1)	12,98	Smallest(1)	0,03	Smallest(1)	0,72
Conf. Level(95,0%)	8,85	Conf. Level(95,0%)	2,86	Conf. Level(95,0%)	0,20
Ac2		**Ac3**		**Ac4**	
Mean	78,79	Mean	1,60	Mean	48,26
Standard Error	6,81	Standard Error	0,39	Standard Error	1,83
Median	74,12	Median	1,36	Median	46,34
Standard Deviation	32,65	Standard Deviation	1,89	Standard Deviation	8,79
Sample Variance	1065,74	Sample Variance	3,58	Sample Variance	77,22
Kurtosis	14,65	Kurtosis	8,06	Kurtosis	0,18
Skewness	3,63	Skewness	2,54	Skewness	0,70
Minimum	57,14	Minimum	0,00	Minimum	32,60
Maximum	214,51	Maximum	8,58	Maximum	67,22
Sum	1812,15	Sum	36,88	Sum	1110,04
Count	23,00	Count	23,00	Count	23,00
Largest(1)	214,51	Largest(1)	8,58	Largest(1)	67,22
Smallest(1)	57,14	Smallest(1)	0,00	Smallest(1)	32,60
Conf. Level(95,0%)	14,12	Conf. Level(95,0%)	0,82	Conf. Level(95,0%)	3,80

Appendix C (continued)

Ac5		*Ac6*		*Ac7*	
Mean	2,96	Mean	5,03	Mean	1,41
Standard Error	0,20	Standard Error	0,18	Standard Error	0,18
Median	2,98	Median	5,05	Median	1,73
Standard Deviation	0,95	Standard Deviation	0,88	Standard Deviation	0,86
Sample Variance	0,91	Sample Variance	0,78	Sample Variance	0,73
Kurtosis	-0,09	Kurtosis	-1,02	Kurtosis	-1,24
Skewness	0,43	Skewness	0,04	Skewness	-0,06
Minimum	1,44	Minimum	3,66	Minimum	0,07
Maximum	5,24	Maximum	6,49	Maximum	2,89
Sum	68,12	Sum	115,74	Sum	32,33
Count	23,00	Count	23,00	Count	23,00
Largest(1)	5,24	Largest(1)	6,49	Largest(1)	2,89
Smallest(1)	1,44	Smallest(1)	3,66	Smallest(1)	0,07
Conf. Level(95,0%)	0,41	Conf. Level(95,0%)	0,38	Conf. Level(95,0%)	0,37

Appendix D: Explanations for Financial Ratios Used

Liquid Assets = Cash + Due from Banks + Central Bank + Other Financial Institutions + Interbank + Securities + Reserve Requirements

Average Total Assets = (Total Assets (1st Year) + Total Assets (2nd Year)) / 2

Average Shareholders' Equity = (Shareholders' Equity (1st Year)) + Shareholders' Equity (2nd Year)) / 2

Average Share-in Capital = (Share-in Capital (1st Year) + Share-in Capital (2nd Year)) / 2

Non-deposits Funds = Interbank + Central Bank + Other Funds Borrowed + Funds + Securities Issued

Contingencies and Commitments = Total Contingencies and Commitments- Other Contingencies and Commitments

Net Working Capital = Shareholders' Equity + Total Income (Current + Previous) - Permanent Assets except Affiliated Securities

Total Profit = Current Year's Profit + Previous Years' Profit

FX Position = FX Liabilities - FX Assets

Permanent Assets = Non-performing Assets (net) + Equity Participations + Affiliated Securities and Companies + Fixed Assets

Profitable Assets = Loans + Securities Portfolio + Banks + Interbank + Government Bonds Account for Legal Reserves

Non-Profitable Assets = Deposits + Non-deposit Funds

Total Income = Interest Income + Non-Interest Income

Total Expenditures = Interest Expenses + Non-Interest Expenses

Interest Income = Interest on (Loans + Securities Portfolio + Deposits in other Banks + Interbank Funds Sold) + Other Interest Income

Other Interest Income = Income from Reserve Requirements + Other

Interest Expenses = Interest on (Deposits + Non-Deposits Funds Borrowed) + Other Interest Expenses

Other Interest Expenses = Interest on Interbank Funds Borrowed + Interest on Securities Issued + Other

Net Interest Income After Provision for Loan Losses = Interest Income - Interest Expenses - Provisions for Loan Losses

Non-Interest Income = Income from Commissions (net) + Income from FX Transactions (net) + Income from Capital Market Transactions (net) + Other Non-Interest Income

Income from Commissions (net) = Fees and Commissions Received - Fees and Commissions Paid

Income from FX Transactions (net) = Income from FX Transactions - Loss from FX Transactions

Income from Capital Market Transactions (net) = Income from Capital Market Transactions - Loss from Capital Market Transactions

Other Non-Interest Income = Dividends from Equity Participations and Affiliated Companies + Extraordinary Income + Other

Non-Interest Expenses = Salaries & Employee Benefits + Res. for Retire. Pay + Other Provisions + Taxes and Duties + Rental Expenses + Depreciation & Amortization + Other

Other Non-Interest Expenses = Extraordinary Expenses + Other

Operational Expenses = Salaries and Benefits + Reserve for Retirement + Rental Expenses + Depreciation and Amortization

Provisions = Reserves for Retirement Pay + Provision for Loan Losses + Provisions for Taxes + Other Provisions

Income before Tax = Net Interest Income after Provision for Loan Losses + Non-Interest Income - Non-Interest Expenses

Net Income (Loss) = Income before Tax - Provisions for Income Tax

Appendix E: Correlation Matrix of Financial Ratios Used

Ratio	C1	C2	C3	C4	C5	C6	C7	A1	A2	A3	A4	L1	L2	L3
C1	24,9	9,9	9,6	12,9	-69,0	20,3	8,3	-55,6	-0,1	0,3	-41,5	64,1	80,8	33,8
C2	9,9	7,1	6,7	9,4	-4,5	-31,3	4,8	-19,0	0,0	0,4	-16,4	26,0	34,3	4,9
C3	9,6	6,7	8,1	8,8	-7,8	-18,1	4,4	-19,4	-0,5	-1,4	-10,6	28,4	36,9	5,1
C4	12,9	9,4	8,8	12,7	9,0	-53,6	5,7	-23,2	0,0	0,6	-17,9	32,7	43,9	3,5
C5	-69,0	-4,5	-7,8	9,0	5659,7	-5697	-50,6	92,9	26,3	3,3	255,3	-83,8	-62,1	-447
C6	20,3	-31,3	-18,1	-53,6	-5698	8081,2	1,5	310,7	-13,8	-13,1	-543,0	-248,2	-340,3	282,3
C7	8,3	4,8	4,4	5,7	-50,6	1,5	32,3	-27,7	-0,7	0,4	-29,7	35,6	44,2	8,2
A1	-55,6	-19,0	-19,4	-23,2	92,9	310,7	-27,7	233,2	1,7	0,4	68,8	-230,1	-280,2	-88,3
A2	-0,1	0,0	-0,5	0,0	26,3	-13,8	-0,7	1,7	0,7	0,5	-4,2	-1,4	-1,5	-1,9
A3	0,3	0,4	-1,4	0,6	3,3	-13,1	0,4	0,4	0,5	1,8	-5,8	-2,4	-2,6	-0,2
A4	-41,5	-16,4	-10,6	-17,9	255,3	-543,0	-29,7	68,8	-4,2	-5,8	310,3	-73,8	-84,7	-56,3
L1	64,1	26,0	28,4	32,7	-83,8	-248,2	35,6	-230,1	-1,4	-2,4	-73,8	258,4	318,6	87,1
L2	80,8	34,3	36,9	43,9	-62,1	-340,3	44,2	-280,2	-1,5	-2,6	-84,7	318,6	395,9	95,6
L3	33,8	4,9	5,1	3,5	-446,9	282,3	8,2	-88,3	-1,9	-0,2	-56,3	87,1	95,6	174,1
P1	-0,5	-0,3	-0,4	-0,3	-3,2	-15,4	-0,3	-0,8	-0,3	0,1	5,1	-0,9	-1,0	-0,2
P2	-8,9	-6,7	-7,4	-8,3	-35,3	-125,5	-6,6	0,5	-2,5	0,8	52,3	-21,2	-27,3	0,5
P3	-0,5	-0,3	-0,4	-0,4	-7,3	-15,5	-0,3	-1,0	-0,4	0,0	5,9	-1,1	-1,3	0,3
P4	-50,6	-39,2	-35,5	-50,8	-276,5	-642,5	-42,9	-2,2	-13,6	-3,6	314,1	-109,2	-140,7	-29,7
I1	1,7	1,2	1,7	1,8	23,3	-30,6	1,9	-3,5	-0,2	-0,5	4,3	5,1	7,5	-4,2
I2	48,5	21,7	32,5	30,0	57,0	-48,5	42,0	-89,9	-4,4	-10,9	-26,5	114,2	153,7	-6,7
I3	-2,4	-1,0	-1,4	-1,3	-2,5	-3,4	-2,1	4,2	0,1	0,4	4,3	-5,2	-6,8	-0,1
I4	-76,2	-44,0	-57,6	-62,6	-654,8	175,0	-76,2	73,6	-5,1	13,6	106,7	-152,2	-222,2	108,1
I5	21,7	17,9	18,0	23,9	205,9	78,9	8,6	-2,4	7,2	-0,1	-70,0	58,4	80,8	-22,0
I6	-0,9	-0,5	-0,6	-0,6	6,7	-1,8	-0,4	2,6	0,1	0,1	0,8	-3,0	-3,7	-1,5
I7	181,1	59,7	86,5	78,1	276,1	-811,7	108,9	-571,8	-20,0	-26,8	-250,2	648,6	801,2	237,6
I8	-76,2	-44,0	-57,6	-62,6	-654,8	175,0	-76,2	73,6	-5,1	13,6	106,7	-152,2	-222,2	108,1
I9	3,9	-12,0	-10,3	-17,6	-260,7	-134,0	-7,0	-85,3	-10,0	-1,8	8,5	41,4	35,1	107,1
I10	-1,9	0,1	0,2	0,5	34,0	-31,8	1,0	5,8	0,3	-0,1	8,3	-5,9	-5,9	-12,1
I11	-3,0	-0,8	-1,2	-1,0	2,6	4,7	-0,7	7,8	0,4	0,4	3,6	-9,2	-11,3	-6,5
I12	19,4	10,9	15,1	15,6	89,3	-17,3	26,1	-23,8	-0,5	-4,2	-21,7	34,7	52,6	-39,5
I13	-26,3	-12,0	-14,6	-16,1	-175,7	175,9	-9,1	67,5	-0,1	2,6	34,4	-99,5	-125,3	-30,4
S1	-8,3	-4,8	-4,8	-6,5	-83,4	-27,5	-0,8	2,1	-2,0	0,0	44,2	-16,0	-21,5	4,2
S2	-8,7	-4,7	-5,0	-6,1	-74,1	-30,5	-1,0	6,6	-1,9	0,3	43,6	-19,1	-24,7	0,6
S3	-8,9	-5,5	-5,3	-7,5	-97,6	-21,4	-1,1	1,7	-2,1	-0,2	46,5	-17,8	-24,2	6,4
B1	1414,7	613,8	749,3	822,0	-3160	622,2	1496,5	-3792	-127	-135	-2621	4260,2	5489,2	116,3
B2	221,2	63,0	104,8	75,4	-1484	996,5	156,0	-766,2	-34,2	-41,8	-463,4	819,0	990,8	352,8
B3	78,9	31,1	53,3	42,0	-709,3	658,3	91,0	-241,5	-14,3	-22,2	-77,4	267,7	348,3	-37,0
B4	142,3	31,9	51,5	33,4	-774,6	338,2	65,1	-524,7	-19,9	-19,6	-386,0	551,3	642,6	389,8
B5	286,7	121,3	157,6	163,2	-1088	837,9	317,7	-721,3	-33,6	-36,3	-553,0	825,2	1069,9	-9,6
B6	59,1	27,3	37,2	36,7	-70,1	-12,5	50,5	-188,7	-5,7	-9,9	-84,8	218,0	279,3	-2,4
B7	20,3	8,7	10,3	11,6	-51,0	-8,0	21,6	-55,4	-2,1	-1,7	-35,1	60,4	77,7	4,0
Ac1	0,0	0,3	0,3	0,5	13,7	-7,3	-0,4	1,0	0,2	0,0	-0,4	0,0	0,2	-1,2
Ac2	97,2	42,0	52,0	55,4	-182,0	217,7	99,0	-252,3	-8,1	-10,0	-260,6	296,0	376,9	16,3
Ac3	1,2	-0,8	-0,4	-1,5	-34,8	4,4	1,9	-11,4	-0,5	-0,4	-7,0	8,5	8,3	13,4
Ac4	0,7	-1,4	-3,9	-2,8	4,7	121,8	-18,3	11,7	1,1	2,5	-62,8	-23,5	-34,6	43,0
Ac5	0,1	0,8	0,9	1,1	29,6	-22,7	0,3	0,7	0,3	-0,1	2,4	2,3	3,6	-4,6
Ac6	-1,4	-0,1	-0,1	0,2	29,0	-39,9	-0,4	2,2	0,0	0,0	9,0	-1,8	-1,5	-5,7
Ac7	-0,5	-0,3	-0,4	-0,4	-7,3	-15,5	-0,3	-1,0	-0,4	0,0	5,9	-1,1	-1,3	0,3

Appendix E (continued)

Ratio2	P1	P2	P3	P4	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
C1	-0,5	-8,9	-0,5	-50,6	1,7	48,5	-2,4	-76,2	21,7	-0,9	181,1	-76,2	3,9	-1,9
C2	-0,3	-6,7	-0,3	-39,2	1,2	21,7	-1,0	-44,0	17,9	-0,5	59,7	-44,0	-12,0	0,1
C3	-0,4	-7,4	-0,4	-35,5	1,7	32,5	-1,4	-57,6	18,0	-0,6	86,5	-57,6	-10,3	0,2
C4	-0,3	-8,3	-0,4	-50,8	1,8	30,0	-1,3	-62,6	23,9	-0,6	78,1	-62,6	-17,6	0,5
C5	-3,2	-35,3	-7,3	-276,5	23,3	57,0	-2,5	-654,8	205,9	6,7	276,1	-654,8	-260,7	34,0
C6	-15,4	-126	-15,5	-642,5	-30,6	-48,5	-3,4	175,0	78,9	-1,8	-811,7	175,0	-134,0	-31,8
C7	-0,3	-6,6	-0,3	-42,9	1,9	42,0	-2,1	-76,2	8,6	-0,4	108,9	-76,2	-7,0	1,0
A1	-0,8	0,5	-1,0	-2,2	-3,5	-89,9	4,2	73,6	-2,4	2,6	-571,8	73,6	-85,3	5,8
A2	-0,3	-2,5	-0,4	-13,6	-0,2	-4,4	0,1	-5,1	7,2	0,1	-20,0	-5,1	-10,0	0,3
A3	0,1	0,8	0,0	-3,6	-0,5	-10,9	0,4	13,6	-0,1	0,1	-26,8	13,6	-1,8	-0,1
A4	5,1	52,3	5,9	314,1	4,3	-26,5	4,3	106,7	-70,0	0,8	-250,2	106,7	8,5	8,3
L1	-0,9	-21,2	-1,1	-109,2	5,1	114,2	-5,2	-152,2	58,4	-3,0	648,6	-152,2	41,4	-5,9
L2	-1,0	-27,3	-1,3	-140,7	7,5	153,7	-6,8	-222,2	80,8	-3,7	801,2	-222,2	35,1	-5,9
L3	-0,2	0,5	0,3	-29,7	-4,2	-6,7	-0,1	108,1	-22,0	-1,5	237,6	108,1	107,1	-12,1
P1	0,5	4,8	0,6	24,0	0,1	-1,7	0,2	13,0	-10,6	0,0	7,9	13,0	12,1	0,0
P2	4,8	46,0	5,3	249,9	-0,2	-30,0	2,0	142,7	-100,8	0,5	16,9	142,7	108,9	-0,2
P3	0,6	5,3	0,7	25,3	0,1	-1,4	0,2	15,9	-13,1	0,0	9,5	15,9	14,8	0,0
P4	24,0	249,9	25,3	1747,2	1,7	-98,8	8,7	628,7	-471,2	3,0	-157,4	628,7	442,2	3,6
I1	0,1	-0,2	0,1	1,7	1,3	18,1	-0,7	-32,9	2,9	-0,1	36,0	-32,9	-3,9	1,1
I2	-1,7	-30,0	-1,4	-98,8	18,1	350,5	-16,0	-620,9	69,9	-2,2	636,1	-620,9	-81,1	10,8
I3	0,2	2,0	0,2	8,7	-0,7	-16,0	0,8	29,6	-3,7	0,1	-28,7	29,6	4,2	-0,4
I4	13,0	142,7	15,9	628,7	-32,9	-620,9	29,6	1492,0	-382,2	2,1	-642,0	1492,0	523,7	-32,9
I5	-10,6	-101	-13,1	-471,2	2,9	69,9	-3,7	-382,2	287,4	-1,2	-104,6	-382,2	-318,5	4,3
I6	0,0	0,5	0,0	3,0	-0,1	-2,2	0,1	2,1	-1,2	0,1	-6,9	2,1	0,0	0,1
I7	7,9	16,9	9,5	-157,4	36,0	636,1	-28,7	-642,0	-104,6	-6,9	3924,6	-642,0	718,9	-24,4
I8	13,0	142,7	15,9	628,7	-32,9	-620,9	29,6	1492,0	-382,2	2,1	-642,0	1492,0	523,7	-32,9
I9	12,1	108,9	14,8	442,2	-3,9	-81,1	4,2	523,7	-318,5	0,0	718,9	523,7	512,6	-17,7
I10	0,0	-0,2	0,0	3,6	1,1	10,8	-0,4	-32,9	4,3	0,1	-24,4	-32,9	-17,7	2,3
I11	-0,1	-0,2	-0,1	0,7	-0,2	-5,3	0,2	-1,3	1,8	0,1	-55,2	-1,3	-13,3	1,1
I12	-1,7	-22,2	-1,8	-70,1	10,1	191,0	-9,0	-396,2	53,1	-0,5	176,1	-396,2	-107,7	10,6
I13	2,1	27,1	2,9	169,9	-4,1	-58,4	2,4	119,0	-81,3	1,1	-483,9	119,0	-13,9	4,7
S1	2,1	22,2	2,7	127,9	-1,3	-24,2	1,5	101,5	-56,9	0,1	-96,1	101,5	41,4	-0,7
S2	2,1	21,9	2,7	123,9	-1,2	-26,9	1,6	104,4	-55,1	0,2	-93,8	104,4	41,6	-0,7
S3	2,2	23,6	2,8	140,6	-1,3	-23,3	1,4	102,8	-60,3	0,1	-102,8	102,8	42,9	-0,6
B1	42,5	-102	62,9	-1187	334,1	6506,8	-306,0	-9415	-525,6	-48,6	21057,4	-9414,8	1970,0	77,9
B2	10,2	32,6	16,0	-105,5	30,2	734,8	-36,0	-441,5	-328,2	-10,3	4540,2	-441,5	1031,4	-43,5
B3	2,5	-3,7	4,7	88,8	27,3	532,5	-24,4	-742,1	-56,6	-3,1	1333,3	-742,1	81,3	10,6
B4	7,7	36,3	11,2	-194,3	2,9	202,3	-11,6	300,6	-271,6	-7,2	3206,9	300,6	950,0	-54,1
B5	11,6	-1,4	17,5	-167,5	71,7	1417,5	-66,8	-1928	-222,0	-9,7	4585,6	-1928,1	535,0	13,0
B6	-0,2	-22,5	0,3	-116,2	16,0	291,3	-13,4	-445,6	32,9	-2,5	1016,8	-445,6	52,9	2,7
B7	1,2	3,4	1,5	5,3	4,7	89,9	-4,2	-117,4	-19,3	-0,7	309,6	-117,4	42,8	0,9
Ac1	-0,2	-2,0	-0,3	-10,0	0,1	0,6	0,0	-7,8	6,3	0,0	-2,4	-7,8	-6,7	0,2
Ac2	0,7	-29,3	1,9	-287,2	19,8	402,7	-19,6	-584,0	-9,3	-3,4	1614,0	-584,0	173,2	-0,7
Ac3	0,4	4,5	0,6	19,9	-0,4	-3,8	0,2	28,3	-13,5	-0,2	58,4	28,3	30,1	-1,4
Ac4	-1,3	-10,6	-1,8	-106,5	-3,9	-38,9	0,7	55,5	4,4	0,2	19,1	55,5	28,9	-5,2
Ac5	-0,4	-3,7	-0,5	-16,7	0,5	3,8	-0,1	-20,0	13,2	0,0	-0,5	-20,0	-15,0	0,7
Ac6	0,3	2,1	0,3	11,6	0,5	0,2	0,2	-2,0	-1,2	0,1	2,1	-2,0	-0,2	0,8
Ac7	0,6	5,3	0,7	25,3	0,1	-1,4	0,2	15,9	-13,1	0,0	9,5	15,9	14,8	0,0

Appendix E (continued)

Ratio3	I11	I12	I13	S1	S2	S3	B1	B2	B3	B4	B5	B6	B7
C1	-3,0	19,4	-26,3	-8,3	-8,7	-8,9	1414,7	221,2	78,9	142,3	286,7	59,1	20,3
C2	-0,8	10,9	-12,0	-4,8	-4,7	-5,5	613,8	63,0	31,1	31,9	121,3	27,3	8,7
C3	-1,2	15,1	-14,6	-4,8	-5,0	-5,3	749,3	104,8	53,3	51,5	157,6	37,2	10,3
C4	-1,0	15,6	-16,1	-6,5	-6,1	-7,5	822,0	75,4	42,0	33,4	163,2	36,7	11,6
C5	2,6	89,3	-175,7	-83,4	-74,1	-97,6	-3160,3	-1483,8	-709,3	-774,6	-1089,7	-70,1	-51,0
C6	4,7	-17,3	175,9	-27,5	-30,5	-21,4	622,2	996,5	658,3	338,2	837,9	-12,5	-8,0
C7	-0,7	26,1	-9,1	-0,8	-1,0	-1,1	1496,5	156,0	91,0	65,1	317,7	50,5	21,6
A1	7,8	-23,8	67,5	2,1	6,6	1,7	-3792,4	-766,2	-241,5	-524,7	-721,3	-188,7	-55,4
A2	0,4	-0,5	-0,1	-2,0	-1,9	-2,1	-127,4	-34,2	-14,3	-19,9	-33,6	-5,7	-2,1
A3	0,4	-4,2	2,6	0,0	0,3	-0,2	-135,5	-41,8	-22,2	-19,6	-36,3	-9,9	-1,7
A4	3,6	-21,7	34,4	44,2	43,6	46,5	-2621,4	-463,4	-77,4	-386,0	-553,0	-84,8	-35,1
L1	-9,2	34,7	-99,5	-16,0	-19,1	-17,8	4260,2	819,0	267,7	551,3	825,2	218,0	60,4
L2	-11,3	52,6	-125,3	-21,5	-24,7	-24,2	5489,2	990,8	348,3	642,6	1069,9	279,3	77,7
L3	-6,5	-39,5	-30,4	4,2	0,6	6,4	116,3	352,8	-37,0	389,8	-9,6	-2,4	4,0
P1	-0,1	-1,7	2,1	2,1	2,1	2,2	42,5	10,2	2,5	7,7	11,6	-0,2	1,2
P2	-0,2	-22,2	27,1	22,2	21,9	23,6	-102,3	32,6	-3,7	36,3	-1,4	-22,5	3,4
P3	-0,1	-1,8	2,9	2,7	2,7	2,8	62,9	16,0	4,7	11,2	17,5	0,3	1,5
P4	0,7	-70,1	169,9	127,9	123,9	140,6	-1187,1	-105,5	88,8	-194,3	-167,5	-116,2	5,3
I1	-0,2	10,1	-4,1	-1,3	-1,2	-1,3	334,1	30,2	27,3	2,9	71,7	16,0	4,7
I2	-5,3	191,0	-58,4	-24,2	-26,9	-23,3	6506,8	734,8	532,5	202,3	1417,5	291,3	89,9
I3	0,2	-9,0	2,4	1,5	1,6	1,4	-306,0	-36,0	-24,4	-11,6	-66,8	-13,4	-4,2
I4	-1,3	-396,2	119,0	101,5	104,4	102,8	-9414,8	-441,5	-742,1	300,6	-1928,1	-445,6	-117,4
I5	1,8	53,1	-81,3	-56,9	-55,1	-60,3	-525,6	-328,2	-56,6	-271,6	-222,0	32,9	-19,3
I6	0,1	-0,5	1,1	0,1	0,2	0,1	-48,6	-10,3	-3,1	-7,2	-9,7	-2,5	-0,7
I7	-55,2	176,1	-483,9	-96,1	-93,8	-102,8	21057,4	4540,2	1333,3	3206,9	4585,6	1016,8	309,6
I8	-1,3	-396,2	119,0	101,5	104,4	102,8	-9414,8	-441,5	-742,1	300,6	-1928,1	-445,6	-117,4
I9	-13,3	-107,7	-13,9	41,4	41,6	42,9	1970,0	1031,4	81,3	950,0	535,0	52,9	42,8
I10	1,1	10,6	4,7	-0,7	-0,7	-0,6	77,9	-43,5	10,6	-54,1	13,0	2,7	0,9
I11	1,2	1,0	8,2	0,6	0,5	0,8	-218,6	-64,6	-13,5	-51,1	-50,1	-11,5	-3,3
I12	1,0	124,6	-5,0	-14,3	-15,8	-13,8	3404,5	208,9	291,7	-82,8	738,9	142,7	45,9
I13	8,2	-5,0	108,1	31,1	28,8	34,1	-1619,0	-423,1	-66,9	-356,1	-311,8	-112,3	-20,8
S1	0,6	-14,3	31,1	24,3	23,0	25,4	-364,5	-44,0	2,8	-46,9	-56,7	-27,0	-3,0
S2	0,5	-15,8	28,8	23,0	22,2	23,8	-382,3	-50,3	-1,1	-49,2	-57,6	-27,8	-3,2
S3	0,8	-13,8	34,1	25,4	23,8	27,0	-393,9	-45,0	5,5	-50,5	-63,2	-28,4	-3,4
B1	-218,6	3404,5	-1619	-364,5	-382,3	-393,9	204411,6	29393,6	15038,4	14355,2	45425,8	8444,6	2962,9
B2	-64,6	208,9	-423,1	-44,0	-50,3	-45,0	29393,6	6616,2	2250,8	4365,4	6709,1	1383,9	430,3
B3	-13,5	291,7	-66,9	2,8	-1,1	5,5	15038,4	2250,8	1415,5	835,3	3477,7	689,3	212,4
B4	-51,1	-82,8	-356,1	-46,9	-49,2	-50,5	14355,2	4365,4	835,3	3530,1	3231,4	694,6	217,9
B5	-50,1	738,9	-311,8	-56,7	-57,6	-63,2	45425,8	6709,1	3477,7	3231,4	10314,6	1865,3	659,8
B6	-11,5	142,7	-112,3	-27,0	-27,8	-28,4	8444,6	1383,9	689,3	694,6	1865,3	418,6	118,1
B7	-3,3	45,9	-20,8	-3,0	-3,2	-3,4	2962,9	430,3	212,4	217,9	659,8	118,1	43,7
Ac1	0,1	0,6	-2,2	-1,6	-1,5	-1,7	-55,1	-13,3	-5,3	-8,0	-15,5	-0,3	-1,0
Ac2	-17,8	197,7	-154,2	-44,8	-44,1	-49,4	14099,1	2241,0	984,7	1256,3	3156,8	591,0	202,9
Ac3	-0,8	-7,5	-4,1	0,7	0,4	1,1	135,0	83,7	1,7	82,0	28,0	5,0	2,6
Ac4	-1,2	-30,4	-9,7	-14,6	-15,3	-14,5	-1265,6	-16,2	-139,7	123,4	-293,3	-42,6	-18,9
Ac5	0,2	2,9	-5,1	-2,7	-2,5	-2,9	-23,8	-20,4	-1,9	-18,5	-11,5	3,0	-0,9
Ac6	0,2	0,6	-1,1	0,1	0,4	0,0	-9,5	-14,8	-0,3	-14,5	-3,7	0,8	0,0
Ac7	-0,1	-1,8	2,9	2,7	2,7	2,8	62,9	16,0	4,7	11,2	17,5	0,3	1,5

Appendix E (continued)

Ratio	Ac1	Ac2	Ac3	Ac4	Ac5	Ac6	Ac7
C1	0,0	97,2	1,2	0,7	0,1	-1,4	-0,5
C2	0,3	42,0	-0,8	-1,4	0,8	-0,1	-0,3
C3	0,3	52,0	-0,4	-3,9	0,9	-0,1	-0,4
C4	0,5	55,4	-1,5	-2,8	1,1	0,2	-0,4
C5	13,7	-182,0	-34,8	4,7	29,6	29,0	-7,3
C6	-7,3	217,7	4,4	121,8	-22,7	-39,9	-15,5
C7	-0,4	99,0	1,9	-18,3	0,3	-0,4	-0,3
A1	1,0	-252,3	-11,4	11,7	0,7	2,2	-1,0
A2	0,2	-8,1	-0,5	1,1	0,3	0,0	-0,4
A3	0,0	-10,0	-0,4	2,5	-0,1	0,0	0,0
A4	-0,4	-260,6	-7,0	-62,8	2,4	9,0	5,9
L1	0,0	296,0	8,5	-23,5	2,3	-1,8	-1,1
L2	0,2	376,9	8,3	-34,6	3,6	-1,5	-1,3
L3	-1,2	16,3	13,4	43,0	-4,6	-5,7	0,3
P1	-0,2	0,7	0,4	-1,3	-0,4	0,3	0,6
P2	-2,0	-29,3	4,5	-10,6	-3,7	2,1	5,3
P3	-0,3	1,9	0,6	-1,8	-0,5	0,3	0,7
P4	-10,0	-287,2	19,9	-106,5	-16,7	11,6	25,3
I1	0,1	19,8	-0,4	-3,9	0,5	0,5	0,1
I2	0,6	402,7	-3,8	-38,9	3,8	0,2	-1,4
I3	0,0	-19,6	0,2	0,7	-0,1	0,2	0,2
I4	-7,8	-584,0	28,3	55,5	-20,0	-2,0	15,9
I5	6,3	-9,3	-13,5	4,4	13,2	-1,2	-13,1
I6	0,0	-3,4	-0,2	0,2	0,0	0,1	0,0
I7	-2,4	1614,0	58,4	19,1	-0,5	2,1	9,5
I8	-7,8	-584,0	28,3	55,5	-20,0	-2,0	15,9
I9	-6,7	173,2	30,1	28,9	-15,0	-0,2	14,8
I10	0,2	-0,7	-1,4	-5,2	0,7	0,8	0,0
I11	0,1	-17,8	-0,8	-1,2	0,2	0,2	-0,1
I12	0,6	197,7	-7,5	-30,4	2,9	0,6	-1,8
I13	-2,2	-154,2	-4,1	-9,7	-5,1	-1,1	2,9
S1	-1,6	-44,8	0,7	-14,6	-2,7	0,1	2,7
S2	-1,5	-44,1	0,4	-15,3	-2,5	0,4	2,7
S3	-1,7	-49,4	1,1	-14,5	-2,9	0,0	2,8
B1	-55,1	14099,1	135,0	-1265,6	-23,8	-9,5	62,9
B2	-13,3	2241,0	83,7	-16,2	-20,4	-14,8	16,0
B3	-5,3	984,7	1,7	-139,7	-1,9	-0,3	4,7
B4	-8,0	1256,3	82,0	123,4	-18,5	-14,5	11,2
B5	-15,5	3156,8	28,0	-293,3	-11,5	-3,7	17,5
B6	-0,3	591,0	5,0	-42,6	3,0	0,8	0,3
B7	-1,0	202,9	2,6	-18,9	-0,9	0,0	1,5
Ac1	0,2	-2,9	-0,3	1,2	0,4	0,1	-0,3
Ac2	-2,9	1065,7	14,4	-52,8	-1,0	-2,5	1,9
Ac3	-0,3	14,4	3,6	3,8	-0,8	-0,4	0,6
Ac4	1,2	-52,8	3,8	77,2	-1,7	-3,3	-1,8
Ac5	0,4	-1,0	-0,8	-1,7	0,9	0,4	-0,5
Ac6	0,1	-2,5	-0,4	-3,3	0,4	0,8	0,3
Ac7	-0,3	1,9	0,6	-1,8	-0,5	0,3	0,7

Appendix F: Covariance Matrix of Financial Ratios Used

Ratio	C1	C2	C3	C4	C5	C6	C7	A1	A2	A3	A4	L1	L2	L3
C1	1	0,75	0,68	0,72	-0,18	0,05	0,29	-0,73	-0,02	0,04	-0,47	0,80	0,81	0,51
C2	0,75	1	0,89	0,99	-0,02	-0,13	0,31	-0,47	0,00	0,11	-0,35	0,61	0,65	0,14
C3	0,68	0,89	1	0,87	-0,04	-0,07	0,27	-0,45	-0,21	-0,37	-0,21	0,62	0,65	0,14
C4	0,72	0,99	0,87	1	0,03	-0,17	0,28	-0,43	0,01	0,13	-0,28	0,57	0,62	0,07
C5	-0,18	-0,02	-0,04	0,03	1	-0,84	-0,12	0,08	0,42	0,03	0,19	-0,07	-0,04	-0,45
C6	0,05	-0,13	-0,07	-0,17	-0,84	1	0,00	0,23	-0,19	-0,11	-0,34	-0,17	-0,19	0,24
C7	0,29	0,31	0,27	0,28	-0,12	0,00	1	-0,32	-0,16	0,05	-0,30	0,39	0,39	0,11
A1	-0,73	-0,47	-0,45	-0,43	0,08	0,23	-0,32	1	0,14	0,02	0,26	-0,94	-0,92	-0,44
A2	-0,02	0,00	-0,21	0,01	0,42	-0,19	-0,16	0,14	1	0,44	-0,29	-0,10	-0,09	-0,17
A3	0,04	0,11	-0,37	0,13	0,03	-0,11	0,05	0,02	0,44	1	-0,25	-0,11	-0,10	-0,01
A4	-0,47	-0,35	-0,21	-0,28	0,19	-0,34	-0,30	0,26	-0,29	-0,25	1	-0,26	-0,24	-0,24
L1	0,80	0,61	0,62	0,57	-0,07	-0,17	0,39	-0,94	-0,10	-0,11	-0,26	1	1,00	0,41
L2	0,81	0,65	0,65	0,62	-0,04	-0,19	0,39	-0,92	-0,09	-0,10	-0,24	1,00	1	0,36
L3	0,51	0,14	0,14	0,07	-0,45	0,24	0,11	-0,44	-0,17	-0,01	-0,24	0,41	0,36	1
P1	-0,12	-0,15	-0,18	-0,11	-0,06	-0,23	-0,08	-0,07	-0,51	0,09	0,39	-0,07	-0,07	-0,02
P2	-0,26	-0,37	-0,39	-0,34	-0,07	-0,21	-0,17	0,01	-0,44	0,08	0,44	-0,19	-0,20	0,01
P3	-0,12	-0,15	-0,15	-0,13	-0,11	-0,20	-0,05	-0,08	-0,55	0,03	0,39	-0,08	-0,08	0,03
P4	-0,24	-0,35	-0,30	-0,34	-0,09	-0,17	-0,18	0,00	-0,39	-0,06	0,43	-0,16	-0,17	-0,05
I1	0,30	0,39	0,53	0,44	0,27	-0,30	0,29	-0,20	-0,23	-0,36	0,21	0,28	0,33	-0,28
I2	0,52	0,43	0,61	0,45	0,04	-0,03	0,40	-0,31	-0,28	-0,44	-0,08	0,38	0,41	-0,03
I3	-0,54	-0,41	-0,56	-0,41	-0,04	-0,04	-0,41	0,31	0,18	0,38	0,27	-0,36	-0,39	-0,01
I4	-0,40	-0,43	-0,52	-0,45	-0,23	0,05	-0,35	0,12	-0,16	0,26	0,16	-0,25	-0,29	0,21
I5	0,26	0,40	0,37	0,39	0,16	0,05	0,09	-0,01	0,51	0,00	-0,23	0,21	0,24	-0,10
I6	-0,62	-0,59	-0,65	-0,55	0,30	-0,07	-0,21	0,56	0,35	0,22	0,14	-0,63	-0,63	-0,38
I7	0,58	0,36	0,49	0,35	0,06	-0,14	0,31	-0,60	-0,39	-0,32	-0,23	0,64	0,64	0,29
I8	-0,40	-0,43	-0,52	-0,45	-0,23	0,05	-0,35	0,12	-0,16	0,26	0,16	-0,25	-0,29	0,21
I9	0,03	-0,20	-0,16	-0,22	-0,15	-0,07	-0,05	-0,25	-0,53	-0,06	0,02	0,11	0,08	0,36
I10	-0,26	0,03	0,04	0,09	0,30	-0,23	0,11	0,25	0,20	-0,04	0,31	-0,24	-0,20	-0,60
I11	-0,54	-0,28	-0,37	-0,26	0,03	0,05	-0,12	0,47	0,40	0,25	0,19	-0,53	-0,52	-0,45
I12	0,35	0,37	0,47	0,39	0,11	-0,02	0,41	-0,14	-0,05	-0,28	-0,11	0,19	0,24	-0,27
I13	-0,51	-0,43	-0,49	-0,43	-0,22	0,19	-0,15	0,43	-0,01	0,19	0,19	-0,60	-0,61	-0,22
S1	-0,34	-0,37	-0,34	-0,37	-0,22	-0,06	-0,03	0,03	-0,49	0,00	0,51	-0,20	-0,22	0,07
S2	-0,37	-0,37	-0,37	-0,36	-0,21	-0,07	-0,04	0,09	-0,49	0,05	0,52	-0,25	-0,26	0,01
S3	-0,34	-0,40	-0,36	-0,41	-0,25	-0,05	-0,04	0,02	-0,49	-0,03	0,51	-0,21	-0,23	0,09
B1	0,63	0,51	0,58	0,51	-0,09	0,02	0,58	-0,55	-0,34	-0,22	-0,33	0,59	0,61	0,02
B2	0,55	0,29	0,45	0,26	-0,24	0,14	0,34	-0,62	-0,51	-0,39	-0,32	0,63	0,61	0,33
B3	0,42	0,31	0,50	0,31	-0,25	0,19	0,43	-0,42	-0,46	-0,44	-0,12	0,44	0,47	-0,07
B4	0,48	0,20	0,30	0,16	-0,17	0,06	0,19	-0,58	-0,40	-0,25	-0,37	0,58	0,54	0,50
B5	0,57	0,45	0,54	0,45	-0,14	0,09	0,55	-0,47	-0,40	-0,27	-0,31	0,51	0,53	-0,01
B6	0,58	0,50	0,64	0,50	-0,05	-0,01	0,43	-0,60	-0,34	-0,36	-0,24	0,66	0,69	-0,01
B7	0,62	0,49	0,55	0,49	-0,10	-0,01	0,58	-0,55	-0,38	-0,19	-0,30	0,57	0,59	0,05
Ac1	-0,01	0,27	0,23	0,30	0,39	-0,18	-0,16	0,14	0,50	0,06	-0,05	0,00	0,02	-0,20
Ac2	0,60	0,48	0,56	0,47	-0,07	0,07	0,53	-0,51	-0,30	-0,23	-0,45	0,56	0,58	0,04
Ac3	0,13	-0,15	-0,07	-0,22	-0,24	0,03	0,17	-0,39	-0,35	-0,15	-0,21	0,28	0,22	0,54
Ac4	0,02	-0,06	-0,16	-0,09	0,01	0,15	-0,37	0,09	0,15	0,21	-0,41	-0,17	-0,20	0,37
Ac5	0,02	0,30	0,32	0,34	0,41	-0,26	0,05	0,05	0,42	-0,08	0,14	0,15	0,19	-0,36
Ac6	-0,31	-0,02	-0,03	0,06	0,44	-0,50	-0,07	0,16	0,04	0,01	0,58	-0,13	-0,08	-0,49
Ac7	-0,12	-0,15	-0,15	-0,13	-0,11	-0,20	-0,05	-0,08	-0,55	0,03	0,39	-0,08	-0,08	0,03

Appendix F (continued)

Ratio	P1	P2	P3	P4	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
C1	-0,12	-0,26	-0,12	-0,24	0,30	0,52	-0,54	-0,40	0,26	-0,62	0,58	-0,40	0,03	-0,26
C2	-0,15	-0,37	-0,15	-0,35	0,39	0,43	-0,41	-0,43	0,40	-0,59	0,36	-0,43	-0,20	0,03
C3	-0,18	-0,39	-0,15	-0,30	0,53	0,61	-0,56	-0,52	0,37	-0,65	0,49	-0,52	-0,16	0,04
C4	-0,11	-0,34	-0,13	-0,34	0,44	0,45	-0,41	-0,45	0,39	-0,55	0,35	-0,45	-0,22	0,09
C5	-0,06	-0,07	-0,11	-0,09	0,27	0,04	-0,04	-0,23	0,16	0,30	0,06	-0,23	-0,15	0,30
C6	-0,23	-0,21	-0,20	-0,17	-0,30	-0,03	-0,04	0,05	0,05	-0,07	-0,14	0,05	-0,07	-0,23
C7	-0,08	-0,17	-0,05	-0,18	0,29	0,40	-0,41	-0,35	0,09	-0,21	0,31	-0,35	-0,05	0,11
A1	-0,07	0,01	-0,08	0,00	-0,20	-0,31	0,31	0,12	-0,01	0,56	-0,60	0,12	-0,25	0,25
A2	-0,51	-0,44	-0,55	-0,39	-0,23	-0,28	0,18	-0,16	0,51	0,35	-0,39	-0,16	-0,53	0,20
A3	0,09	0,08	0,03	-0,06	-0,36	-0,44	0,38	0,26	0,00	0,22	-0,32	0,26	-0,06	-0,04
A4	0,39	0,44	0,39	0,43	0,21	-0,08	0,27	0,16	-0,23	0,14	-0,23	0,16	0,02	0,31
L1	-0,07	-0,19	-0,08	-0,16	0,28	0,38	-0,36	-0,25	0,21	-0,63	0,64	-0,25	0,11	-0,24
L2	-0,07	-0,20	-0,08	-0,17	0,33	0,41	-0,39	-0,29	0,24	-0,63	0,64	-0,29	0,08	-0,20
L3	-0,02	0,01	0,03	-0,05	-0,28	-0,03	-0,01	0,21	-0,10	-0,38	0,29	0,21	0,36	-0,60
P1	1	0,96	0,95	0,78	0,10	-0,12	0,24	0,46	-0,86	0,15	0,17	0,46	0,73	0,01
P2	0,96	1	0,91	0,88	-0,03	-0,24	0,34	0,54	-0,88	0,27	0,04	0,54	0,71	-0,02
P3	0,95	0,91	1	0,71	0,11	-0,09	0,22	0,48	-0,90	0,06	0,18	0,48	0,76	-0,01
P4	0,78	0,88	0,71	1	0,04	-0,13	0,23	0,39	-0,66	0,24	-0,06	0,39	0,47	0,06
I1	0,10	-0,03	0,11	0,04	1	0,85	-0,71	-0,75	0,15	-0,16	0,51	-0,75	-0,15	0,61
I2	-0,12	-0,24	-0,09	-0,13	0,85	1	-0,96	-0,86	0,22	-0,39	0,54	-0,86	-0,19	0,38
I3	0,24	0,34	0,22	0,23	-0,71	-0,96	1	0,86	-0,24	0,33	-0,51	0,86	0,21	-0,29
I4	0,46	0,54	0,48	0,39	-0,75	-0,86	0,86	1	-0,58	0,18	-0,27	1,00	0,60	-0,56
I5	-0,86	-0,88	-0,90	-0,66	0,15	0,22	-0,24	-0,58	1	-0,24	-0,10	-0,58	-0,83	0,17
I6	0,15	0,27	0,06	0,24	-0,16	-0,39	0,33	0,18	-0,24	1	-0,37	0,18	0,00	0,29
I7	0,17	0,04	0,18	-0,06	0,51	0,54	-0,51	-0,27	-0,10	-0,37	1	-0,27	0,51	-0,26
I8	0,46	0,54	0,48	0,39	-0,75	-0,86	0,86	1,00	-0,58	0,18	-0,27	1	0,60	-0,56
I9	0,73	0,71	0,76	0,47	-0,15	-0,19	0,21	0,60	-0,83	0,00	0,51	0,60	1	-0,52
I10	0,01	-0,02	-0,01	0,06	0,61	0,38	-0,29	-0,56	0,17	0,29	-0,26	-0,56	-0,52	1
I11	-0,12	-0,03	-0,11	0,01	-0,15	-0,26	0,25	-0,03	0,10	0,36	-0,81	-0,03	-0,54	0,67
I12	-0,20	-0,29	-0,19	-0,15	0,80	0,91	-0,91	-0,92	0,28	-0,15	0,25	-0,92	-0,43	0,63
I13	0,28	0,38	0,32	0,39	-0,35	-0,30	0,26	0,30	-0,46	0,35	-0,74	0,30	-0,06	0,30
S1	0,59	0,66	0,64	0,62	-0,23	-0,26	0,34	0,53	-0,68	0,07	-0,31	0,53	0,37	-0,10
S2	0,62	0,69	0,67	0,63	-0,23	-0,31	0,39	0,57	-0,69	0,11	-0,32	0,57	0,39	-0,10
S3	0,57	0,67	0,63	0,65	-0,23	-0,24	0,31	0,51	-0,69	0,07	-0,32	0,51	0,36	-0,08
B1	0,13	-0,03	0,16	-0,06	0,65	0,77	-0,76	-0,54	-0,07	-0,36	0,74	-0,54	0,19	0,11
B2	0,17	0,06	0,23	-0,03	0,33	0,48	-0,50	-0,14	-0,24	-0,43	0,89	-0,14	0,56	-0,35
B3	0,09	-0,01	0,15	0,06	0,64	0,76	-0,73	-0,51	-0,09	-0,28	0,57	-0,51	0,10	0,19
B4	0,18	0,09	0,22	-0,08	0,04	0,18	-0,22	0,13	-0,27	-0,41	0,86	0,13	0,71	-0,60
B5	0,16	0,00	0,20	-0,04	0,62	0,75	-0,74	-0,49	-0,13	-0,32	0,72	-0,49	0,23	0,08
B6	-0,01	-0,16	0,02	-0,14	0,69	0,76	-0,74	-0,56	0,09	-0,41	0,79	-0,56	0,11	0,09
B7	0,24	0,08	0,27	0,02	0,63	0,73	-0,71	-0,46	-0,17	-0,35	0,75	-0,46	0,29	0,09
Ac1	-0,60	-0,63	-0,68	-0,52	0,22	0,07	-0,05	-0,44	0,80	0,04	-0,08	-0,44	-0,65	0,30
Ac2	0,03	-0,13	0,07	-0,21	0,54	0,66	-0,68	-0,46	-0,02	-0,35	0,79	-0,46	0,23	-0,01
Ac3	0,30	0,35	0,35	0,25	-0,21	-0,11	0,10	0,39	-0,42	-0,29	0,49	0,39	0,70	-0,50
Ac4	-0,21	-0,18	-0,24	-0,29	-0,39	-0,24	0,09	0,16	0,03	0,07	0,03	0,16	0,15	-0,39
Ac5	-0,53	-0,58	-0,60	-0,42	0,44	0,21	-0,12	-0,54	0,82	-0,01	-0,01	-0,54	-0,70	0,46
Ac6	0,40	0,35	0,35	0,31	0,52	0,01	0,20	-0,06	-0,08	0,38	0,04	-0,06	-0,01	0,59
Ac7	0,95	0,91	1,00	0,71	0,11	-0,09	0,22	0,48	-0,90	0,06	0,18	0,48	0,76	-0,01

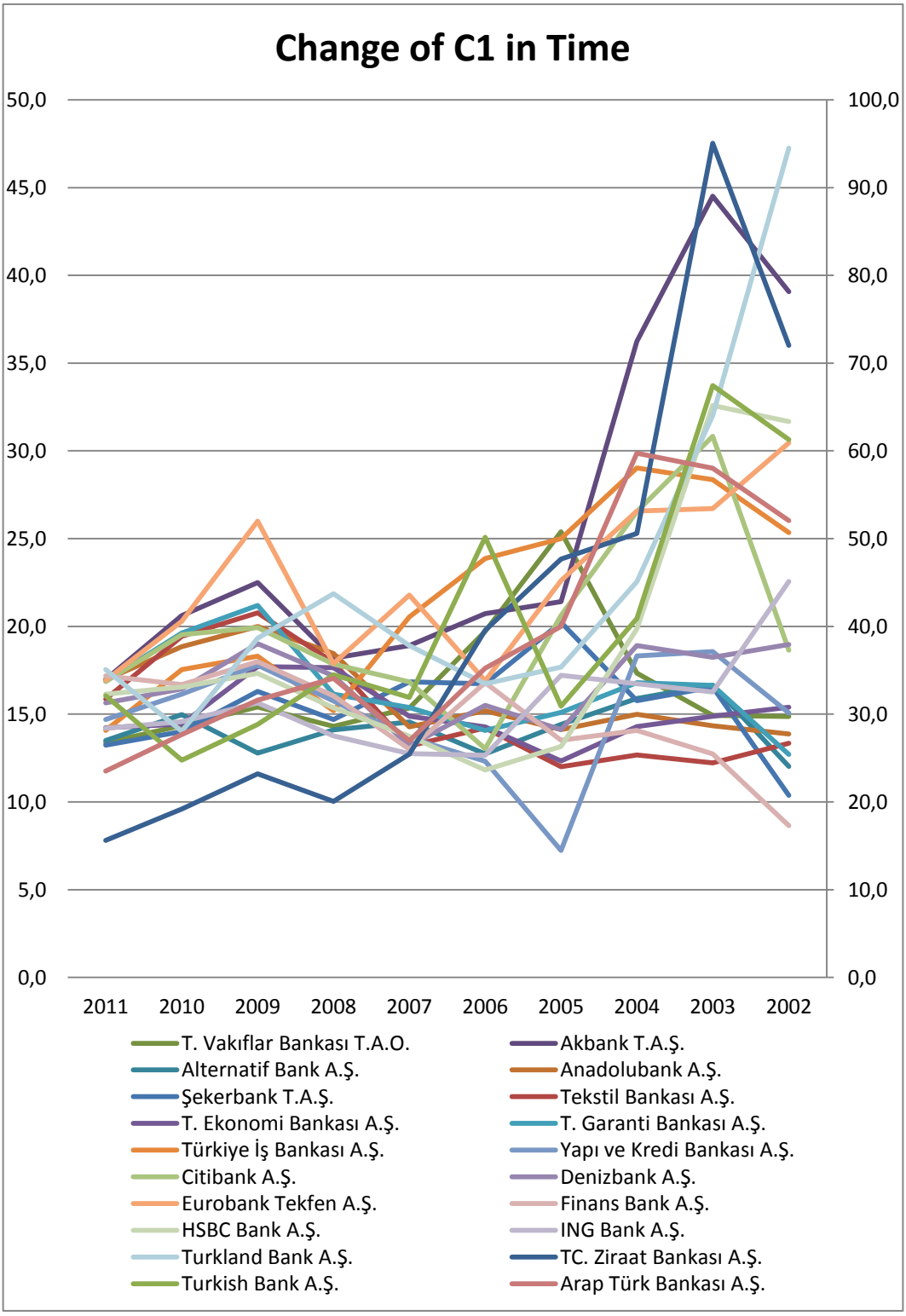
Appendix F (continued)

Ratio	I11	I12	I13	S1	S2	S3	B1	B2	B3	B4	B5	B6	B7
C1	-0,54	0,35	-0,51	-0,34	-0,37	-0,34	0,63	0,55	0,42	0,48	0,57	0,58	0,62
C2	-0,28	0,37	-0,43	-0,37	-0,37	-0,40	0,51	0,29	0,31	0,20	0,45	0,50	0,49
C3	-0,37	0,47	-0,49	-0,34	-0,37	-0,36	0,58	0,45	0,50	0,30	0,54	0,64	0,55
C4	-0,26	0,39	-0,43	-0,37	-0,36	-0,41	0,51	0,26	0,31	0,16	0,45	0,50	0,49
C5	0,03	0,11	-0,22	-0,22	-0,21	-0,25	-0,09	-0,24	-0,25	-0,17	-0,14	-0,05	-0,10
C6	0,05	-0,02	0,19	-0,06	-0,07	-0,05	0,02	0,14	0,19	0,06	0,09	-0,01	-0,01
C7	-0,12	0,41	-0,15	-0,03	-0,04	-0,04	0,58	0,34	0,43	0,19	0,55	0,43	0,58
A1	0,47	-0,14	0,43	0,03	0,09	0,02	-0,55	-0,62	-0,42	-0,58	-0,47	-0,60	-0,55
A2	0,40	-0,05	-0,01	-0,49	-0,49	-0,49	-0,34	-0,51	-0,46	-0,40	-0,40	-0,34	-0,38
A3	0,25	-0,28	0,19	0,00	0,05	-0,03	-0,22	-0,39	-0,44	-0,25	-0,27	-0,36	-0,19
A4	0,19	-0,11	0,19	0,51	0,52	0,51	-0,33	-0,32	-0,12	-0,37	-0,31	-0,24	-0,30
L1	-0,53	0,19	-0,60	-0,20	-0,25	-0,21	0,59	0,63	0,44	0,58	0,51	0,66	0,57
L2	-0,52	0,24	-0,61	-0,22	-0,26	-0,23	0,61	0,61	0,47	0,54	0,53	0,69	0,59
L3	-0,45	-0,27	-0,22	0,07	0,01	0,09	0,02	0,33	-0,07	0,50	-0,01	-0,01	0,05
P1	-0,12	-0,20	0,28	0,59	0,62	0,57	0,13	0,17	0,09	0,18	0,16	-0,01	0,24
P2	-0,03	-0,29	0,38	0,66	0,69	0,67	-0,03	0,06	-0,01	0,09	0,00	-0,16	0,08
P3	-0,11	-0,19	0,32	0,64	0,67	0,63	0,16	0,23	0,15	0,22	0,20	0,02	0,27
P4	0,01	-0,15	0,39	0,62	0,63	0,65	-0,06	-0,03	0,06	-0,08	-0,04	-0,14	0,02
I1	-0,15	0,80	-0,35	-0,23	-0,23	-0,23	0,65	0,33	0,64	0,04	0,62	0,69	0,63
I2	-0,26	0,91	-0,30	-0,26	-0,31	-0,24	0,77	0,48	0,76	0,18	0,75	0,76	0,73
I3	0,25	-0,91	0,26	0,34	0,39	0,31	-0,76	-0,50	-0,73	-0,22	-0,74	-0,74	-0,71
I4	-0,03	-0,92	0,30	0,53	0,57	0,51	-0,54	-0,14	-0,51	0,13	-0,49	-0,56	-0,46
I5	0,10	0,28	-0,46	-0,68	-0,69	-0,69	-0,07	-0,24	-0,09	-0,27	-0,13	0,09	-0,17
I6	0,36	-0,15	0,35	0,07	0,11	0,07	-0,36	-0,43	-0,28	-0,41	-0,32	-0,41	-0,35
I7	-0,81	0,25	-0,74	-0,31	-0,32	-0,32	0,74	0,89	0,57	0,86	0,72	0,79	0,75
I8	-0,03	-0,92	0,30	0,53	0,57	0,51	-0,54	-0,14	-0,51	0,13	-0,49	-0,56	-0,46
I9	-0,54	-0,43	-0,06	0,37	0,39	0,36	0,19	0,56	0,10	0,71	0,23	0,11	0,29
I10	0,67	0,63	0,30	-0,10	-0,10	-0,08	0,11	-0,35	0,19	-0,60	0,08	0,09	0,09
I11	1	0,08	0,72	0,12	0,10	0,14	-0,44	-0,73	-0,33	-0,79	-0,45	-0,51	-0,45
I12	0,08	1	-0,04	-0,26	-0,30	-0,24	0,67	0,23	0,69	-0,12	0,65	0,62	0,62
I13	0,72	-0,04	1	0,61	0,59	0,63	-0,34	-0,50	-0,17	-0,58	-0,30	-0,53	-0,30
S1	0,12	-0,26	0,61	1	0,99	0,99	-0,16	-0,11	0,02	-0,16	-0,11	-0,27	-0,09
S2	0,10	-0,30	0,59	0,99	1	0,97	-0,18	-0,13	-0,01	-0,18	-0,12	-0,29	-0,10
S3	0,14	-0,24	0,63	0,99	0,97	1	-0,17	-0,11	0,03	-0,16	-0,12	-0,27	-0,10
B1	-0,44	0,67	-0,34	-0,16	-0,18	-0,17	1	0,80	0,88	0,53	0,99	0,91	0,99
B2	-0,73	0,23	-0,50	-0,11	-0,13	-0,11	0,80	1	0,74	0,90	0,81	0,83	0,80
B3	-0,33	0,69	-0,17	0,02	-0,01	0,03	0,88	0,74	1	0,37	0,91	0,90	0,85
B4	-0,79	-0,12	-0,58	-0,16	-0,18	-0,16	0,53	0,90	0,37	1	0,54	0,57	0,55
B5	-0,45	0,65	-0,30	-0,11	-0,12	-0,12	0,99	0,81	0,91	0,54	1	0,90	0,98
B6	-0,51	0,62	-0,53	-0,27	-0,29	-0,27	0,91	0,83	0,90	0,57	0,90	1	0,87
B7	-0,45	0,62	-0,30	-0,09	-0,10	-0,10	0,99	0,80	0,85	0,55	0,98	0,87	1
Ac1	0,15	0,12	-0,47	-0,70	-0,69	-0,70	-0,26	-0,36	-0,30	-0,29	-0,33	-0,03	-0,34
Ac2	-0,50	0,54	-0,45	-0,28	-0,29	-0,29	0,96	0,84	0,80	0,65	0,95	0,88	0,94
Ac3	-0,41	-0,36	-0,21	0,07	0,04	0,11	0,16	0,54	0,02	0,73	0,15	0,13	0,21
Ac4	-0,13	-0,31	-0,11	-0,34	-0,37	-0,32	-0,32	-0,02	-0,42	0,24	-0,33	-0,24	-0,33
Ac5	0,14	0,27	-0,51	-0,57	-0,56	-0,59	-0,06	-0,26	-0,05	-0,33	-0,12	0,15	-0,14
Ac6	0,17	0,06	-0,12	0,03	0,09	0,01	-0,02	-0,21	-0,01	-0,28	-0,04	0,05	0,00
Ac7	-0,11	-0,19	0,32	0,64	0,67	0,63	0,16	0,23	0,15	0,22	0,20	0,02	0,27

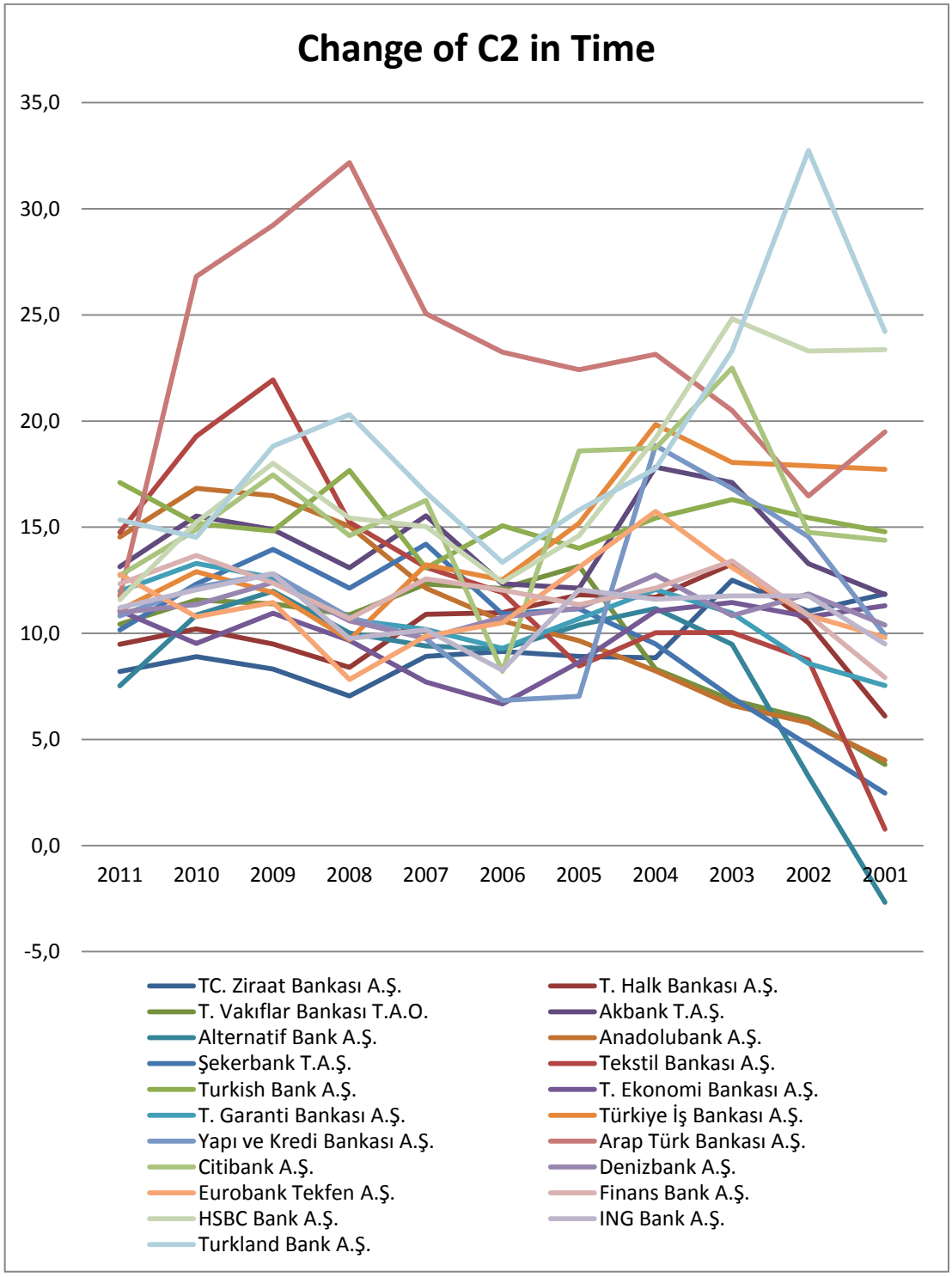
Appendix F (continued)

Ratio	Ac1	Ac2	Ac3	Ac4	Ac5	Ac6	Ac7
C1	-0,01	0,60	0,13	0,02	0,02	-0,31	-0,12
C2	0,27	0,48	-0,15	-0,06	0,30	-0,02	-0,15
C3	0,23	0,56	-0,07	-0,16	0,32	-0,03	-0,15
C4	0,30	0,47	-0,22	-0,09	0,34	0,06	-0,13
C5	0,39	-0,07	-0,24	0,01	0,41	0,44	-0,11
C6	-0,18	0,07	0,03	0,15	-0,26	-0,50	-0,20
C7	-0,16	0,53	0,17	-0,37	0,05	-0,07	-0,05
A1	0,14	-0,51	-0,39	0,09	0,05	0,16	-0,08
A2	0,50	-0,30	-0,35	0,15	0,42	0,04	-0,55
A3	0,06	-0,23	-0,15	0,21	-0,08	0,01	0,03
A4	-0,05	-0,45	-0,21	-0,41	0,14	0,58	0,39
L1	0,00	0,56	0,28	-0,17	0,15	-0,13	-0,08
L2	0,02	0,58	0,22	-0,20	0,19	-0,08	-0,08
L3	-0,20	0,04	0,54	0,37	-0,36	-0,49	0,03
P1	-0,60	0,03	0,30	-0,21	-0,53	0,40	0,95
P2	-0,63	-0,13	0,35	-0,18	-0,58	0,35	0,91
P3	-0,68	0,07	0,35	-0,24	-0,60	0,35	1,00
P4	-0,52	-0,21	0,25	-0,29	-0,42	0,31	0,71
I1	0,22	0,54	-0,21	-0,39	0,44	0,52	0,11
I2	0,07	0,66	-0,11	-0,24	0,21	0,01	-0,09
I3	-0,05	-0,68	0,10	0,09	-0,12	0,20	0,22
I4	-0,44	-0,46	0,39	0,16	-0,54	-0,06	0,48
I5	0,80	-0,02	-0,42	0,03	0,82	-0,08	-0,90
I6	0,04	-0,35	-0,29	0,07	-0,01	0,38	0,06
I7	-0,08	0,79	0,49	0,03	-0,01	0,04	0,18
I8	-0,44	-0,46	0,39	0,16	-0,54	-0,06	0,48
I9	-0,65	0,23	0,70	0,15	-0,70	-0,01	0,76
I10	0,30	-0,01	-0,50	-0,39	0,46	0,59	-0,01
I11	0,15	-0,50	-0,41	-0,13	0,14	0,17	-0,11
I12	0,12	0,54	-0,36	-0,31	0,27	0,06	-0,19
I13	-0,47	-0,45	-0,21	-0,11	-0,51	-0,12	0,32
S1	-0,70	-0,28	0,07	-0,34	-0,57	0,03	0,64
S2	-0,69	-0,29	0,04	-0,37	-0,56	0,09	0,67
S3	-0,70	-0,29	0,11	-0,32	-0,59	0,01	0,63
B1	-0,26	0,96	0,16	-0,32	-0,06	-0,02	0,16
B2	-0,36	0,84	0,54	-0,02	-0,26	-0,21	0,23
B3	-0,30	0,80	0,02	-0,42	-0,05	-0,01	0,15
B4	-0,29	0,65	0,73	0,24	-0,33	-0,28	0,22
B5	-0,33	0,95	0,15	-0,33	-0,12	-0,04	0,20
B6	-0,03	0,88	0,13	-0,24	0,15	0,05	0,02
B7	-0,34	0,94	0,21	-0,33	-0,14	0,00	0,27
Ac1	1	-0,19	-0,38	0,31	0,86	0,28	-0,68
Ac2	-0,19	1	0,23	-0,18	-0,03	-0,09	0,07
Ac3	-0,38	0,23	1	0,23	-0,46	-0,25	0,35
Ac4	0,31	-0,18	0,23	1	-0,20	-0,42	-0,24
Ac5	0,86	-0,03	-0,46	-0,20	1	0,50	-0,60
Ac6	0,28	-0,09	-0,25	-0,42	0,50	1	0,35
Ac7	-0,68	0,07	0,35	-0,24	-0,60	0,35	1

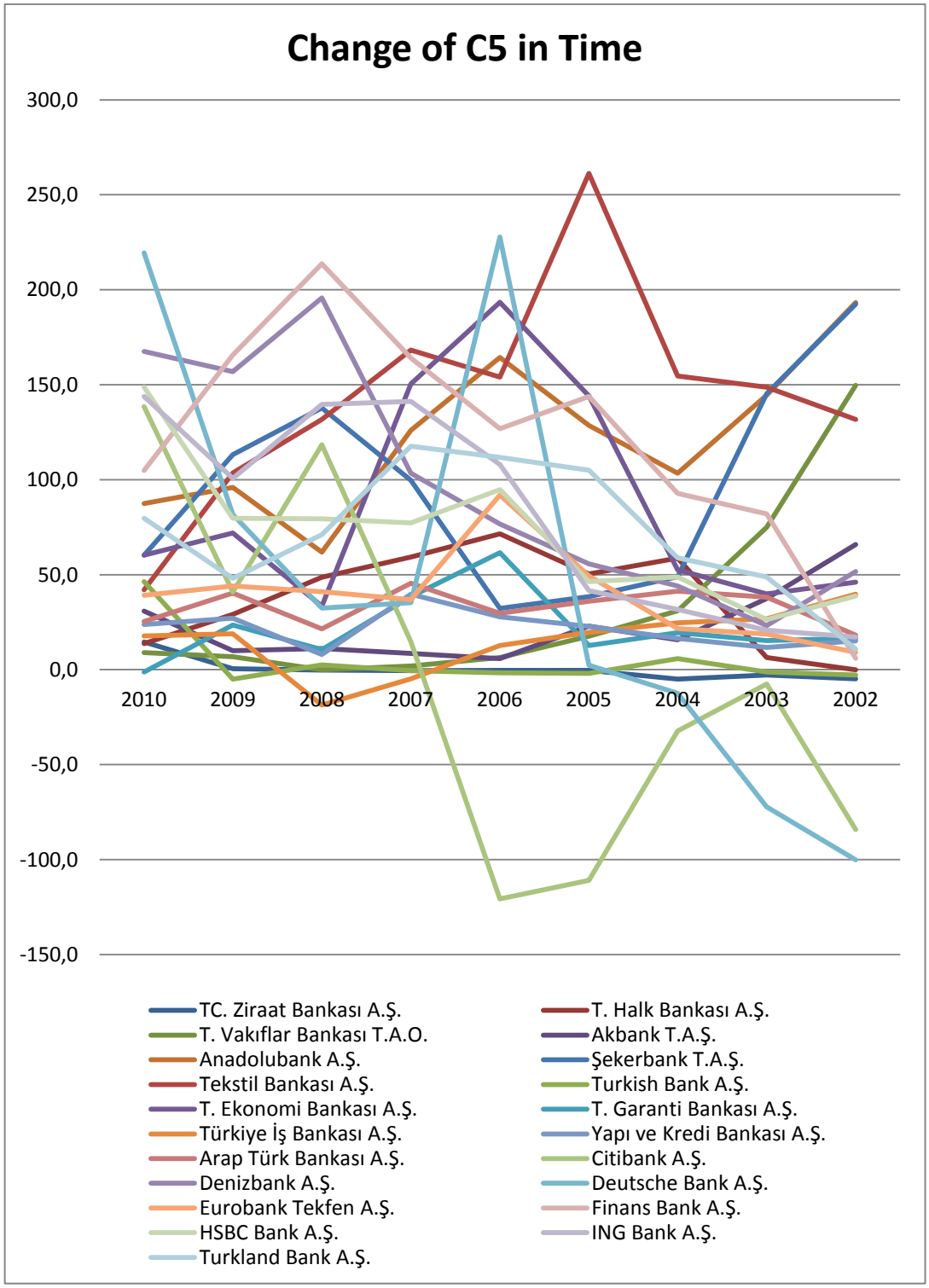
Appendix G: Financial Ratios vs Time



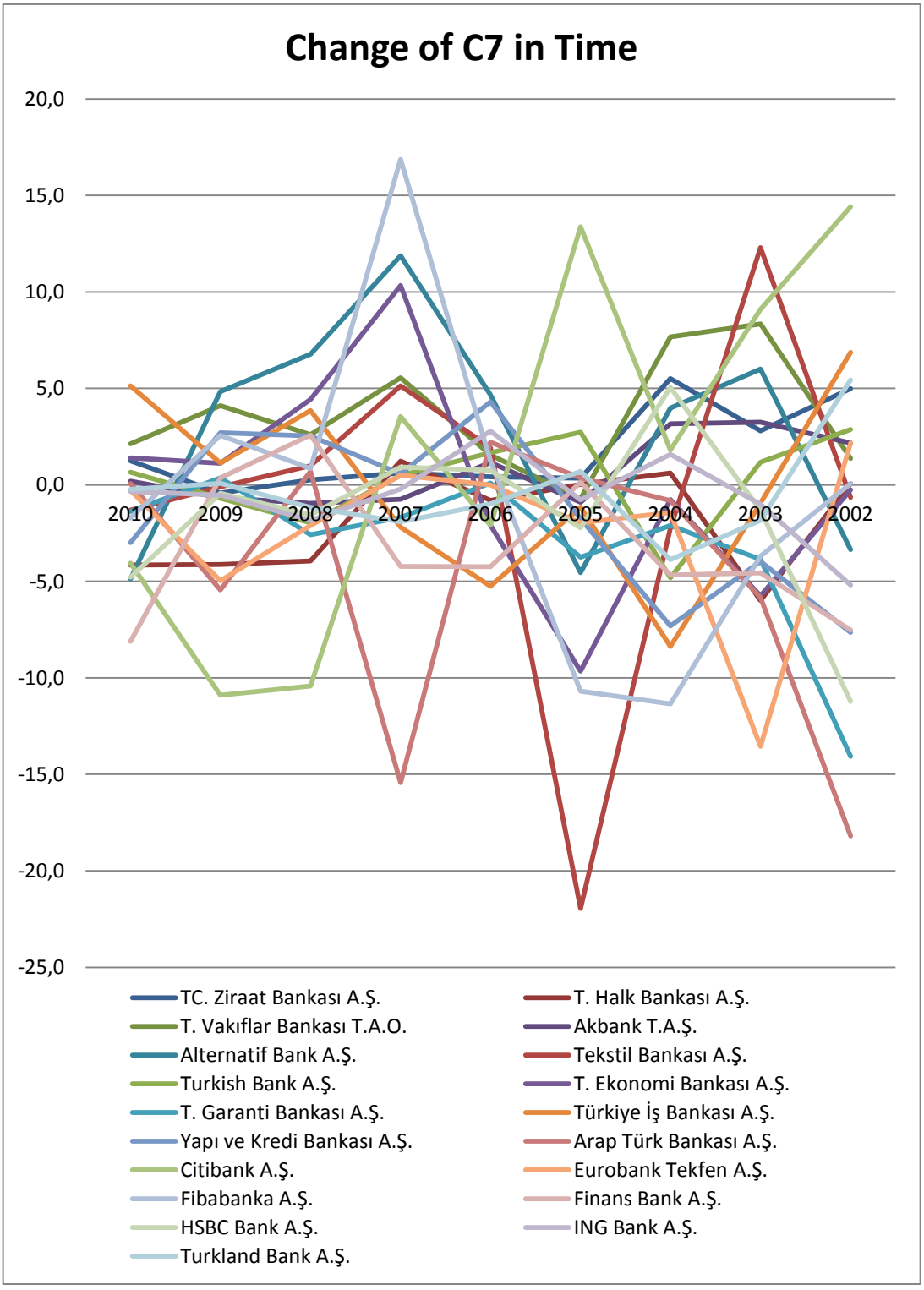
Appendix G (continued)



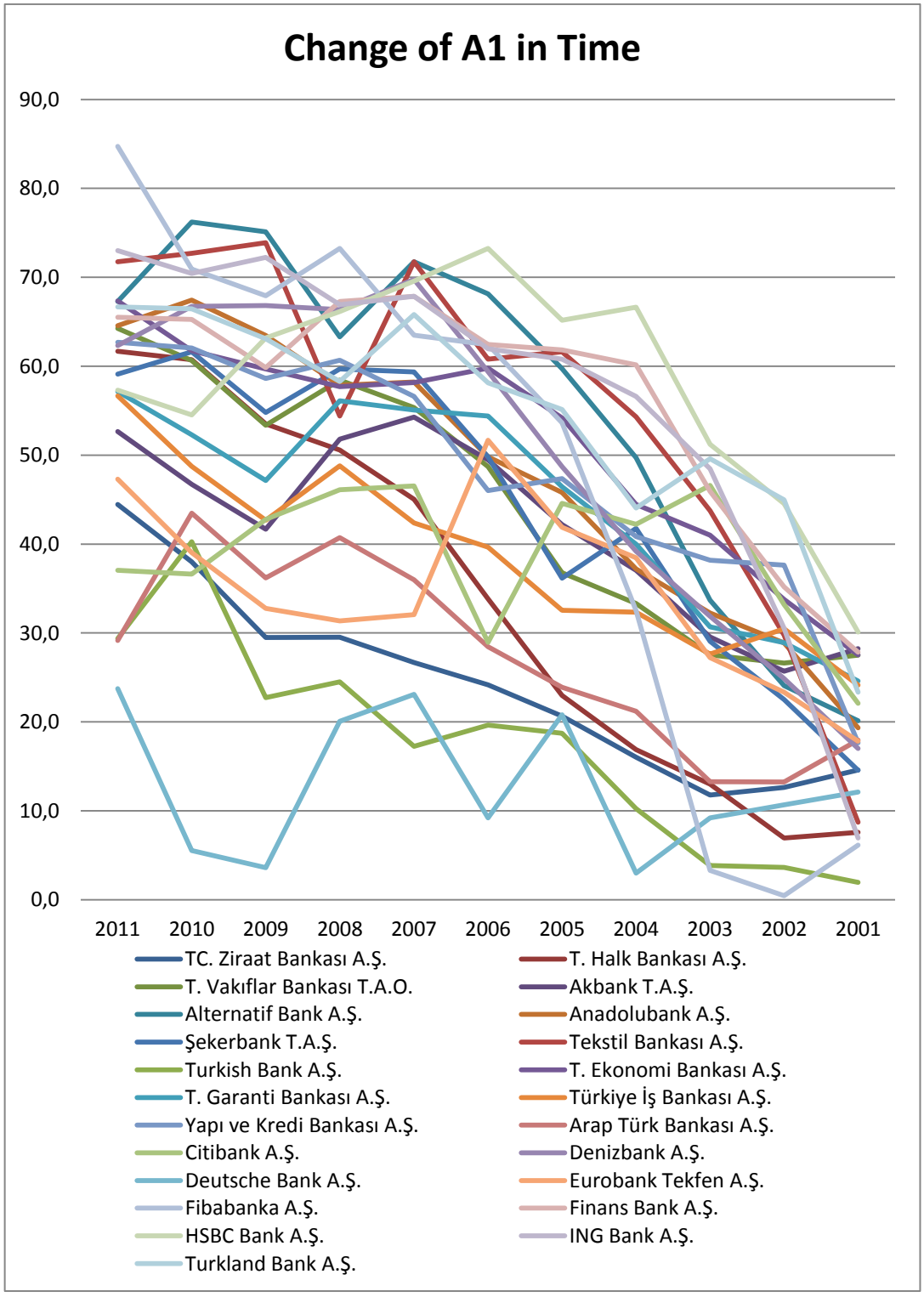
Appendix G (continued)



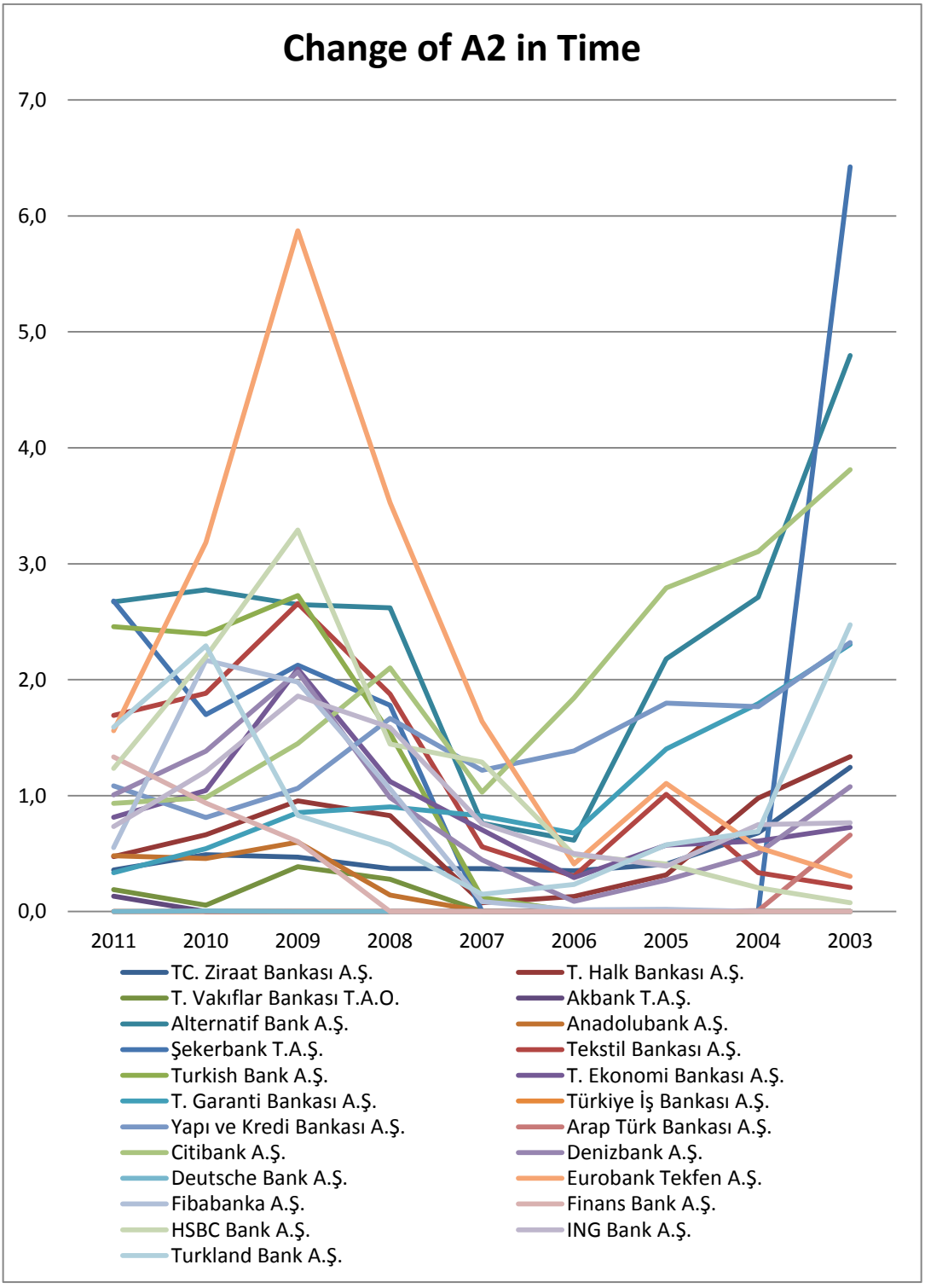
Appendix G (continued)



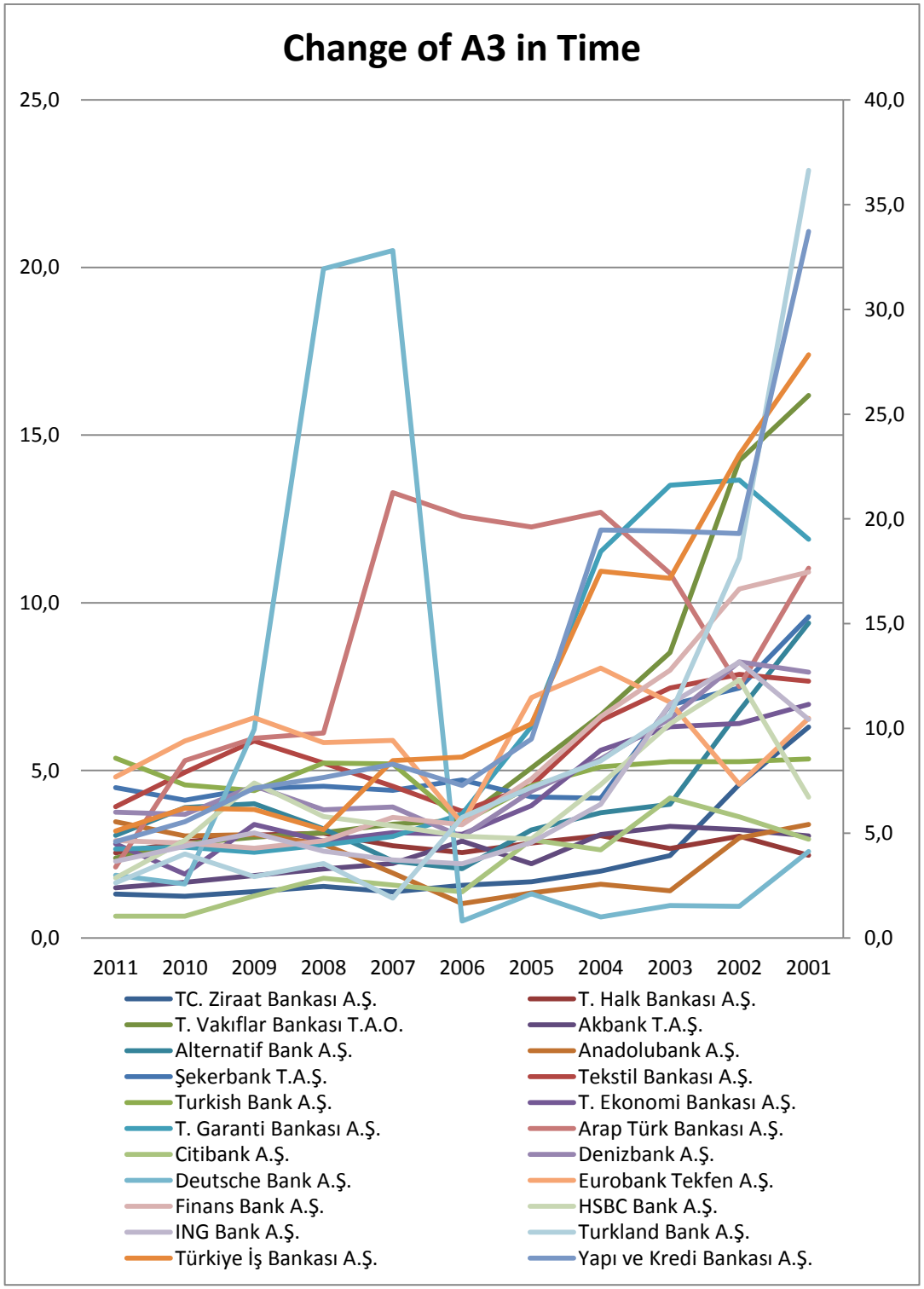
Appendix G (continued)



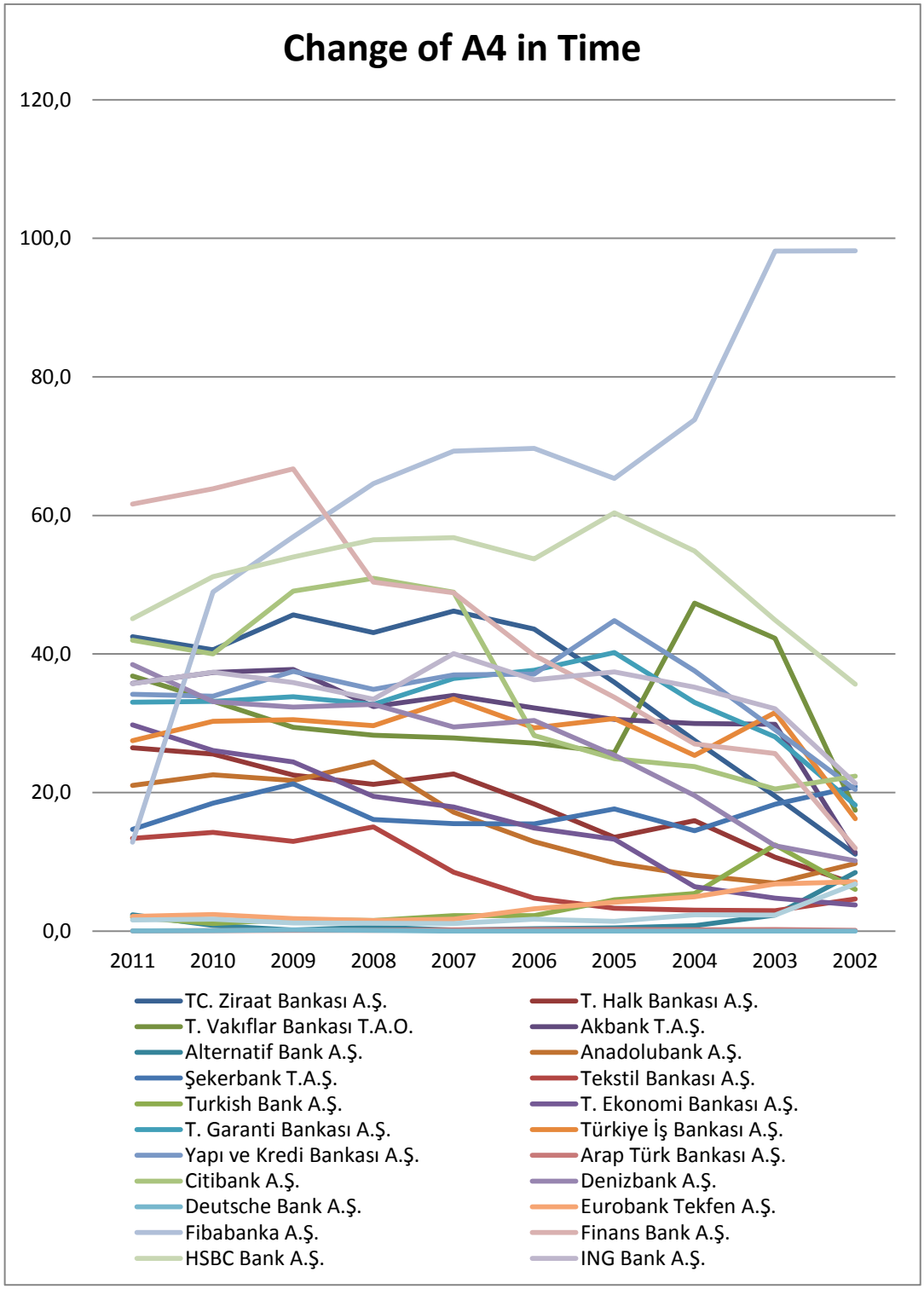
Appendix G (continued)



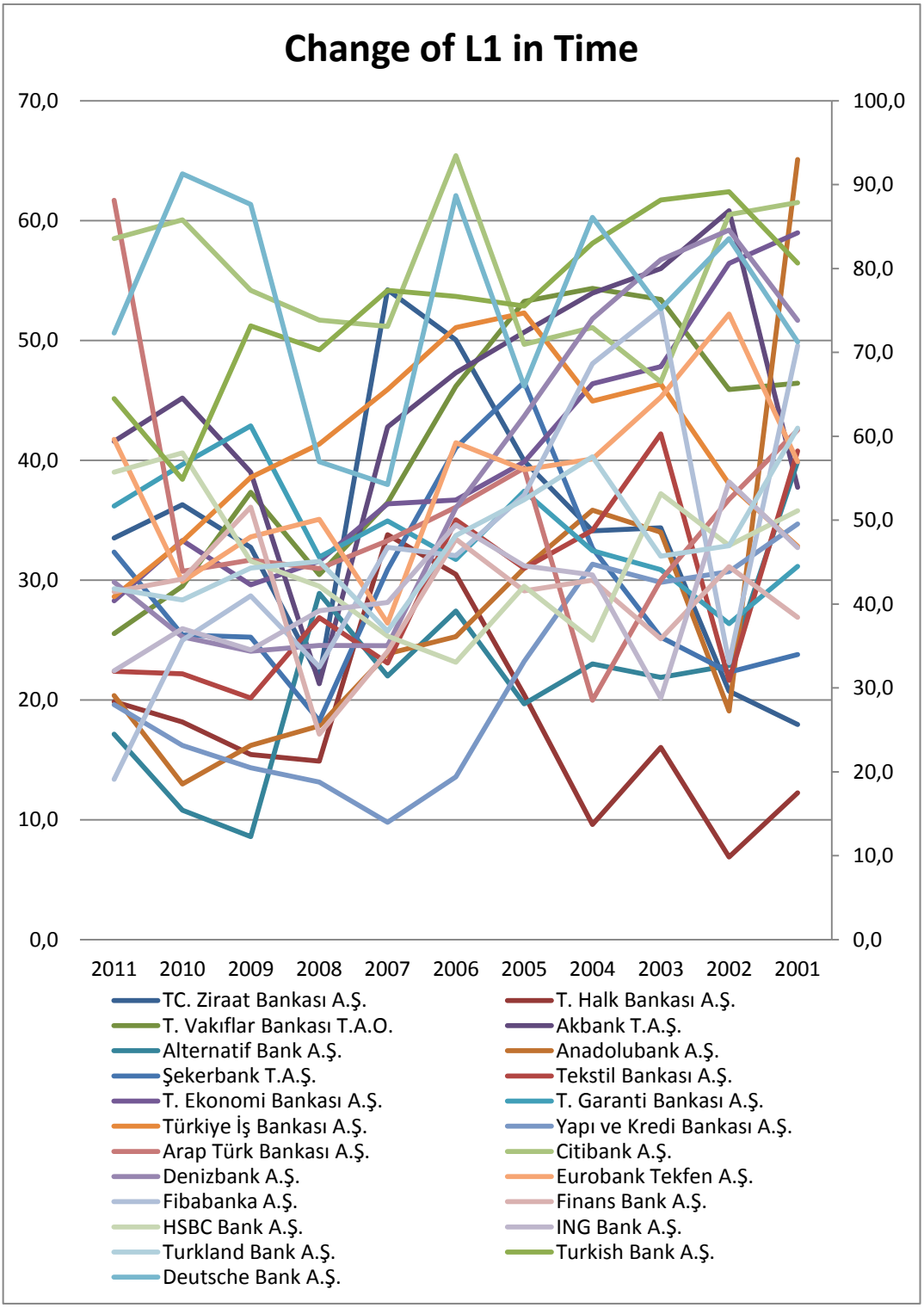
Appendix G (continued)



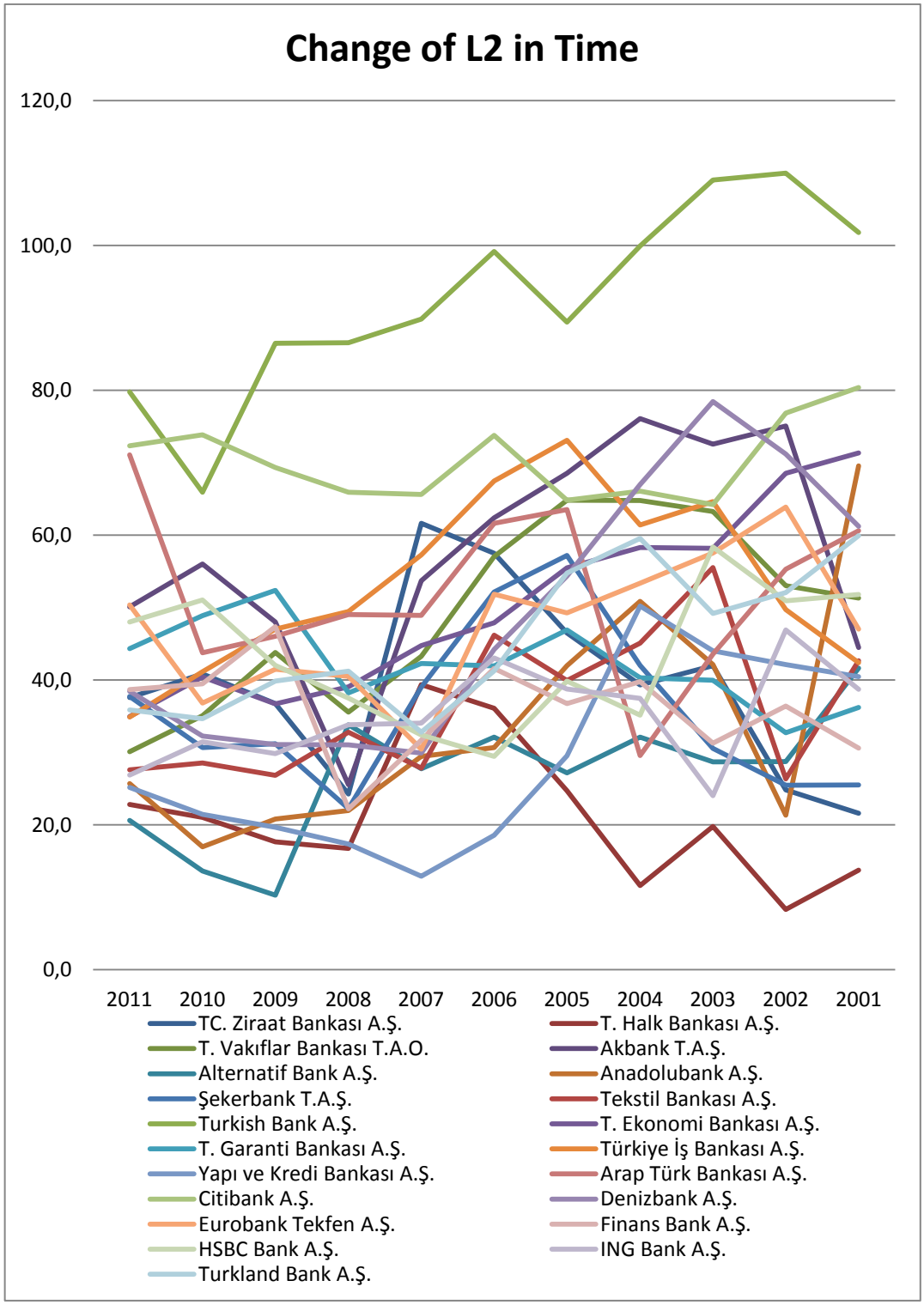
Appendix G (continued)



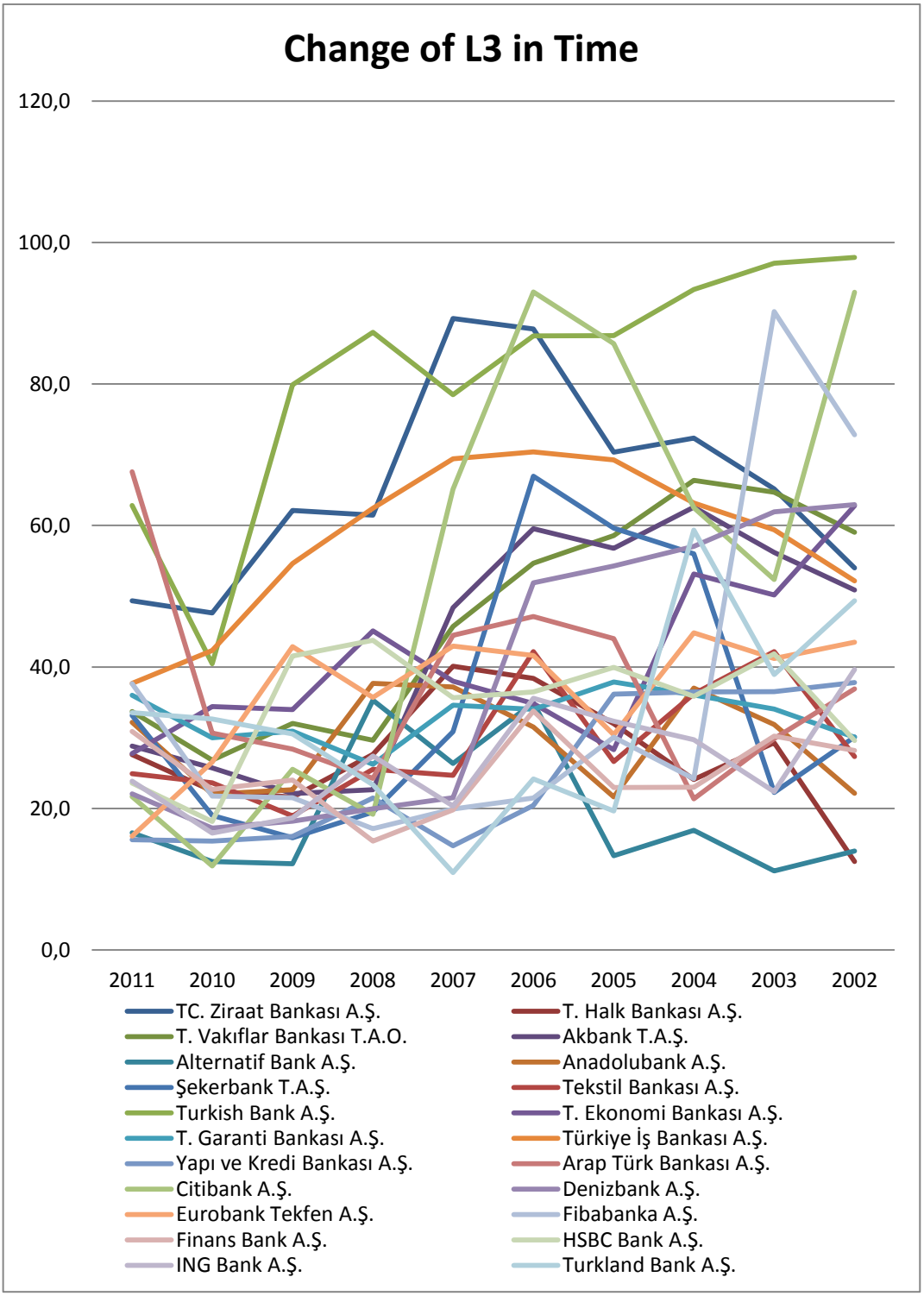
Appendix G (continued)



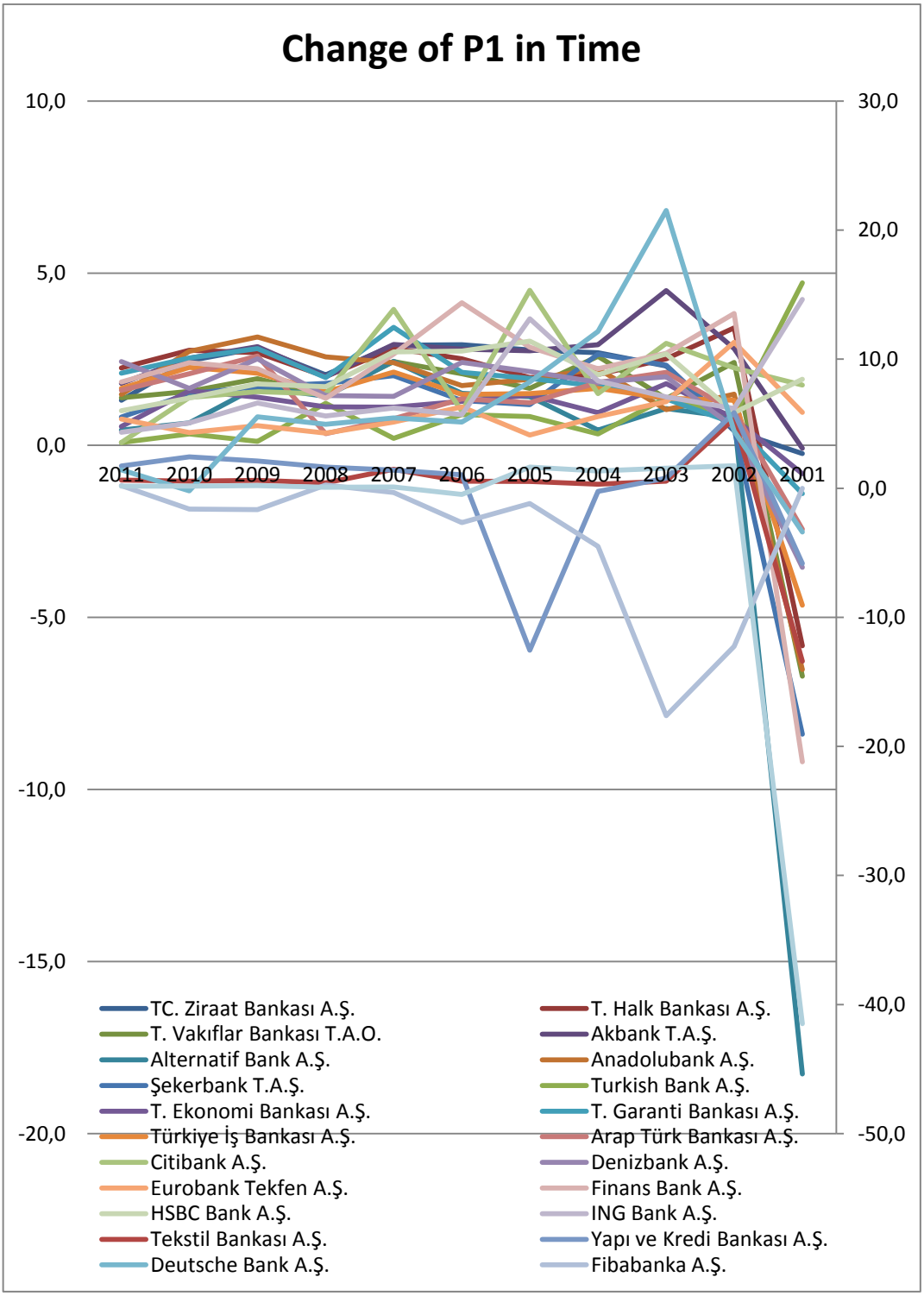
Appendix G (continued)



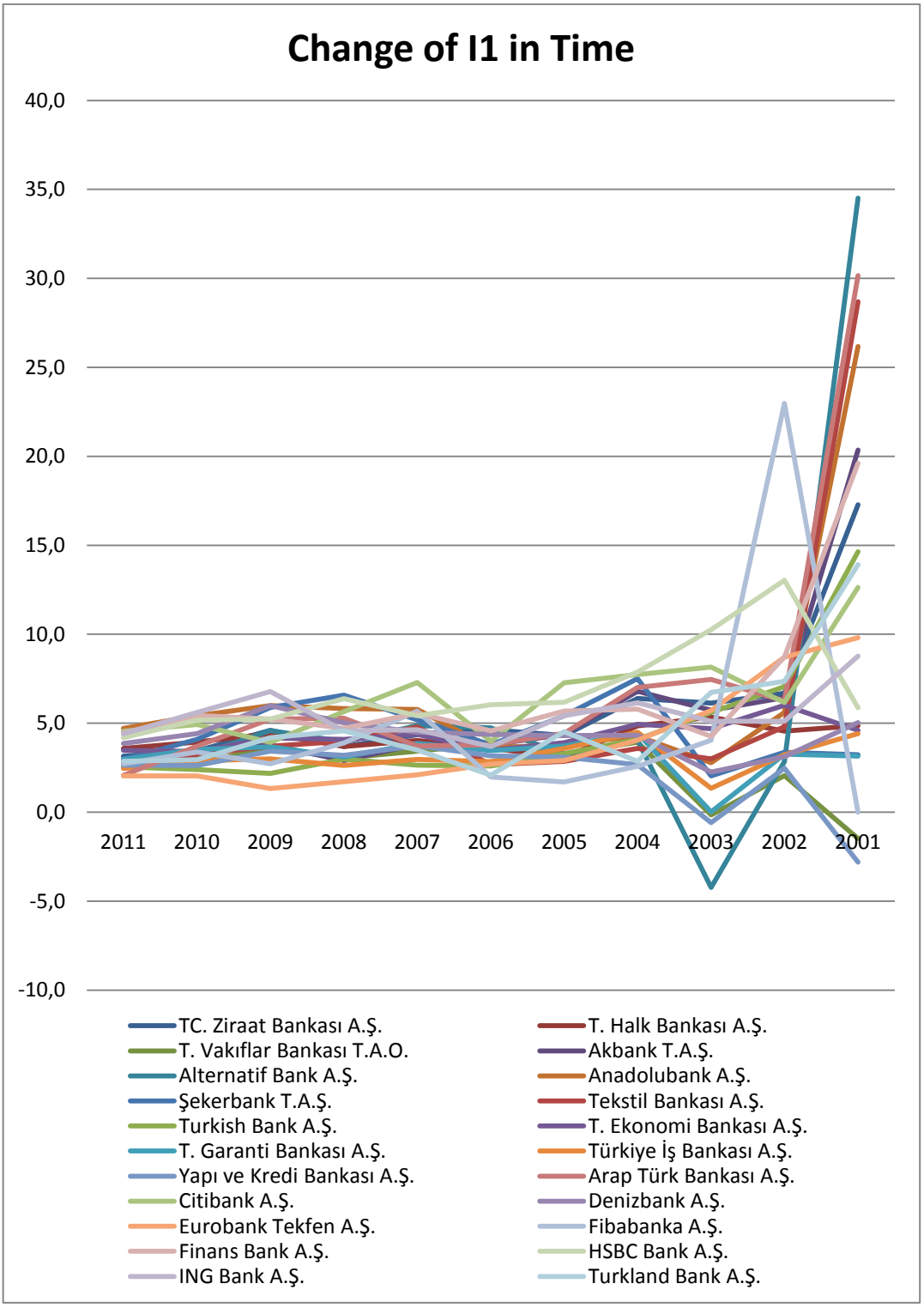
Appendix G (continued)



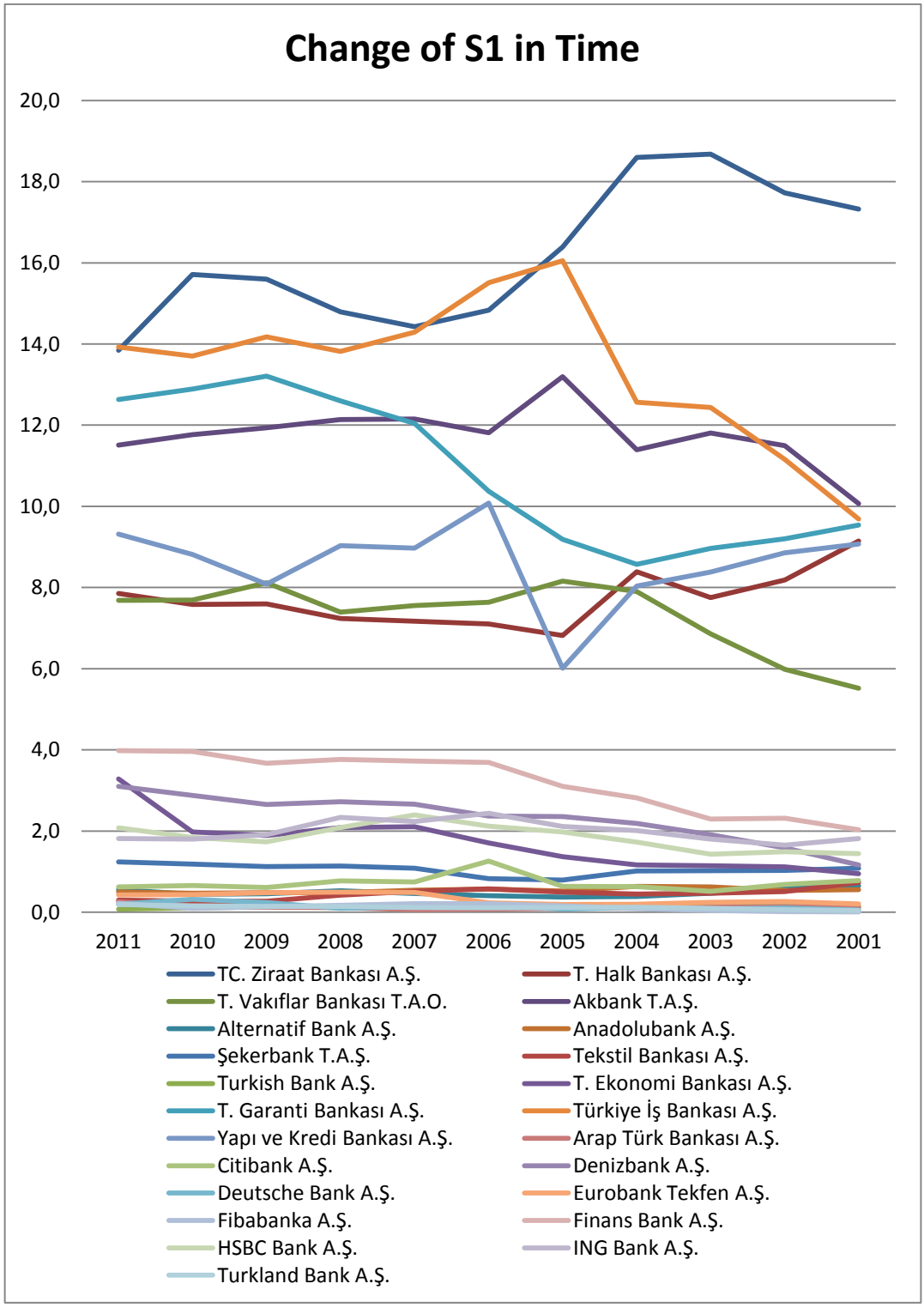
Appendix G (continued)



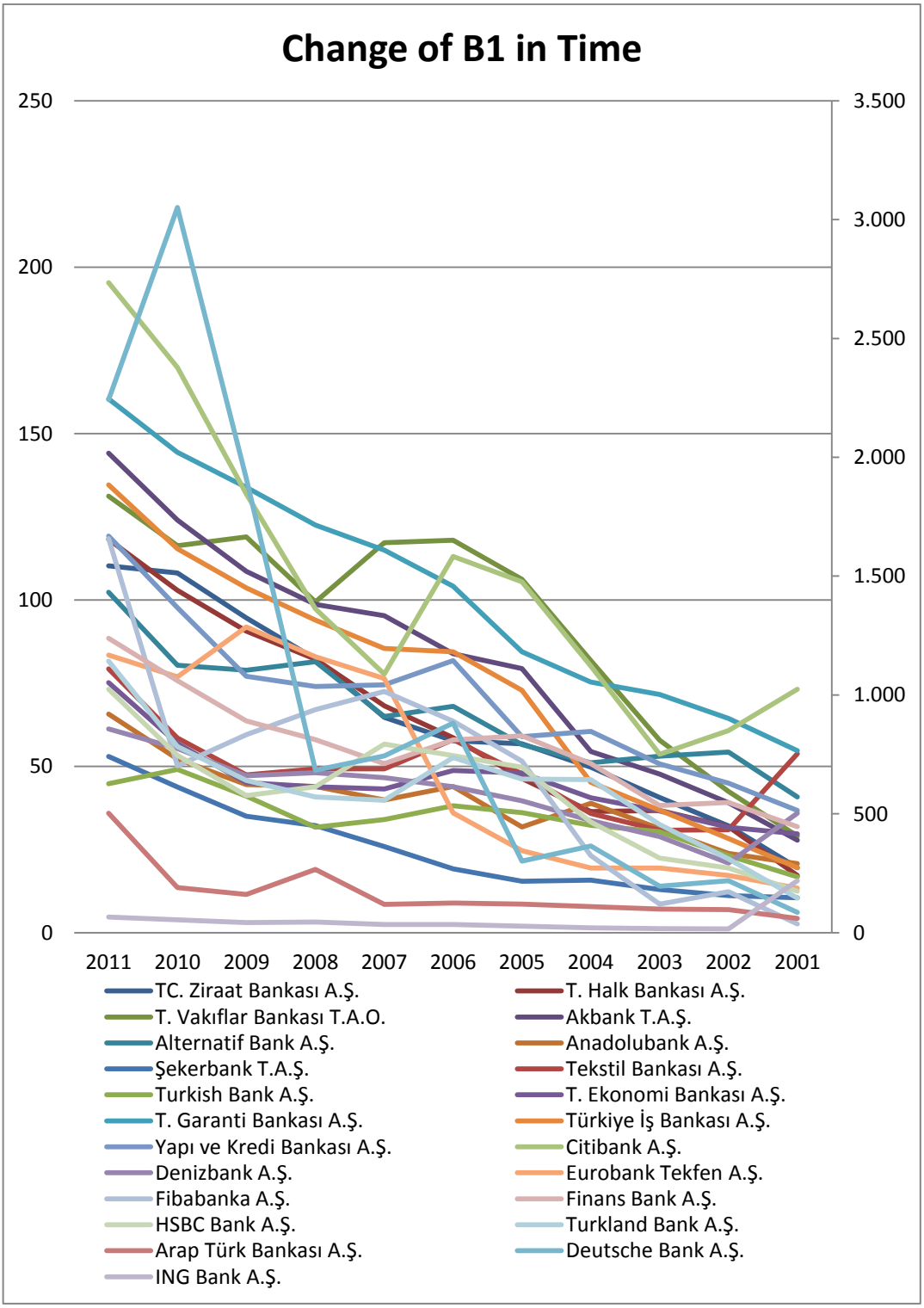
Appendix G (continued)



Appendix G (continued)



Appendix G (continued)



METU LIBRARY

TEZ FOTOKOPİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü ☐

Sosyal Bilimler Enstitüsü ☒

Uygulamalı Matematik Enstitüsü ☐

Enformatik Enstitüsü ☐

Deniz Bilimleri Enstitüsü ☐

YAZARIN

Soyadı : ALTINEL
Adı : Fatih
Bölümü : İktisat

TEZİN ADI (İngilizce) : An Empirical Study On Fuzzy C-Means Clustering For Turkish Banking System

TEZİN TÜRÜ : Yüksek Lisans ☒ Doktora ☐

1. Tezimin tamamı dünya çapında erişime açılsın ve kaynak gösterilmek şartıyla tezimin bir kısmı veya tamamının fotokopisi alınsın. ☐
2. Tezimin tamamı yalnızca Orta Doğu Teknik Üniversitesi kullanıcılarının erişimine açılsın. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.) ☐
3. Tezim bir (1) yıl süreyle erişime kapalı olsun. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.) ☒

Yazarın imzası ..... Tarih .....