THE ACCURACY OF FINANCIAL DISTRESS PREDICTION MODELS IN TURKEY:
A COMPARATIVE INVESTIGATION WITH SIMPLE MODEL PROPOSALS

Öğr. Grv. Erol MUZIR* Yrd.Doç.Dr. Nazan ÇAĞLAR**

ABSTRACT

This study aims to test eight well-known and widely used financial distress prediction models and to compare their performance in Turkey for the first year prior to failure. The comparison is enriched with the details of four new and simple model proposals, namely Failure Score (F-Score) Models, developed using four statistical techniques. The results show that none of the existing models could achieve satisfactorily high correct-classification rates over 90 %. Ohlson's O-Score Model seems to be superior to other existing models and has the highest rate of correct classification, 81,6 %. However, our new model proposal based on logistic regression outperforms the O-Score model in terms of the overall accuracy t-value and may be viewed as an equally worthy model for predicting bankruptcy.

Keywords: *Corporate failure, financial distress prediction, failure risk assessment, discriminant analysis, binary logistic regression, probit analysis, one-zero linear regression.*

FİNANSAL BAŞARISIZLIK TAHMİN MODELLERİNİN TÜRKİYE'DE GEÇERLİLİĞİ: BASİT MODEL ÖNERİLERİYLE KARŞILAŞTIRMALI BİR ARAŞTIRMA

ÖZ

Bu çalışma, uygulamada yaygın olarak kullanılan sekiz adet finansal başarısızlık modelinin Türkiye'de test edilmesi ve başarısızlık öncesi ilk yıl için tahmin performanslarının karşılaştırılması amacını taşımaktadır. Karşılaştırma çalışmamız, F-Skor Modelleri adını taşıyan ve dört farklı istatistik tekniğin kullanılması neticesinde ortaya konulan dört yeni ve basit model önerisinin detayları ile zenginleştirilmiştir. Çalışmamızın sonuçları doğrultusunda, uygulamada yer bulan mevcut model önerilerinin hiçbirinin 90 % düzeyinde veya daha yüksek bir doğru sınıflandırma oranına sahip olamadığı görülmüştür. Ohlson tarafından önerilen O-Skor modelinin, 81,6 % doğru sınıflandırma oranı ile diğer mevcut modellere kıyasla daha başarılı olduğu anlaşılmaktadır. Buna karşın, ikili lojistik regresyon tekniğine dayalı yeni model önerimiz, genel doğruluk t-değeri açısından O-Skor modelinden bile daha iyi bir tahmin performansı sergilemiştir. Bu bağlamda, lojistik regresyon model önerimiz, O-Skor modeliyle eşdeğer bir tahmin modeli olarak değerlendirilebilir.

Anahtar Kelimeler: *Kurumsal başarısızlık, finansal başarısızlık tahmini,, başarısızlık riskinin tespiti, diskriminant analizi, ikili lojistik regresyon, probit analizi 0-1 doğrusal regresyon*

---

* Yıldız Teknik Üniversitesi Meslek Yüksekokulu Muhasebe Programı, e-mail:emuzir@yildiz.edu.tr.

** Kültür Üniversitesi İ.İ.B.F. İktisat Bölümü, e-mail:ncaglar@iku.edu.tr.

1. INTRODUCTION

Failure of a firm yields harmful effects on its stakeholders. For this reason, a priori assessment of corporate failure risk can help the stakeholders take proactive actions against possible damages of such a fruitless situation. As the prediction of financial distress has informational value and creates substantial contributions for firm-related parties, many recent studies and research were carried out to develop properly working prediction models so as to foresee an impending corporate failure.

Continuous interest in searching for better prediction models has led to the discovery of numerous models. The basic assumption on which many of these models were based is the belief that it is possible to observe the manifests of a probable business failure case in the current accounts and financial ratios of a firm. This belief has encouraged the researchers of this field to study with financial statement data, especially financial ratios as the predictor variables. On the other hand, it is possible to talk about exceptional studies in which both different predictor variables other than financial statement data and more sophisticated techniques apart from basic econometric models were used.

The models using financial statement data can be regarded as of four main types: a) The prediction models based on financial ratios, b) the models based on cash flows, c) the models based on stock returns and return variation, and d) the models based on industry averages (Mossman et al, 1998, p.36). Among the famous advocates of the paradigm suggesting the use of financial statement data in modeling are Altman, Beaver, Deakin, Casey, and many others. The variables of the models in which non-financial data were incorporated include macroeconomic indicators, price changes for particular commodities, and certain qualitative variables such as management structure, quality of accounting information system, and so on. (Argenti, 1991, p.3)

Each of the studies aimed at constructing better models to help predict financial distress more accurately obtained a certain satisfactory level of accuracy for its own sample. Although the performance of each model was considered to be sufficient by its creator, it is inevitable to take into account the reality that each model was designed using a separate sample from a different country. It is therefore a question mark whether these models would generate the same levels of accuracy if they were tested in other countries with different macroeconomic, social, political, and financial structures and conditions. Furthermore, the somewhat unrealistic assumptions underlying the mathematical techniques used in modeling may act as challenging circumstances reducing model accuracy. These problematic issues bring up a serious dilemma, which is *Transferability Problem of Models across Countries*. (Keasey et al, 1991, p.89)

Since Turkey is a country with both a history composed of dramatic economic crises and their dangerous impacts, and a long experience of inconsistent and unstable trends of socio-economic and political developments, Turkish firms have very frequently faced higher levels of failure risk when compared to their counterparts in the countries with more stable economic conditions. It is another unfavorable factor potentially reducing the predictability of economic circumstances that Turkish economy is a dependent economy, which makes Turkish national economy more sensitive to the changes in international conditions and be affected much more by global fluctuations and crises.

It becomes more important to be able to accurately predict corporate failure risk for companies which operate in the countries with extreme levels of volatility in such economic conditions like Turkey. A this point, the existing financial distress prediction models may serve as mechanisms to assess failure risk, but their transferability to Turkey and the accuracy of predictions constitute the main concerns.

Our study aims to test and compare eight reputable and widely respected financial distress prediction models within one year prior to failure in Turkey and is enriched with an essential attempt to create a simple model proposal that could be used to produce better prediction results.

In the scope of the first part of the study, the following models are tested and compared to each other in terms of their prediction performance:

- Beaver's Univariate Model (Beaver, 1966, pp.71 - 111)

- Altman's Z-Score and Revised Z-Score Models (Altman, 1968, pp.589 - 609)

- ZETA Model (Altman et al, 1977, pp.29 - 54)

- Deakin's Business Failure Model (Deakin, 1972, pp.167 - 179)

- Ohlson's O-Score Model (Ohlson, 1980, pp.109 - 131)

- Zavgren's BPR Model (Zavgren, 1982, pp.19 - 45)

- Zmijewski's Financial Distress Prediction Model (Zmijewski, 1984, pp.59 - 82)

Our main attempt to construct simple model proposals is aimed at creating new and easily applicable model equations using such major econometric techniques as multiple discriminant, logit, probit, and multiple regression. Then, the results of these two parts are integrated to compare all the models to one another with respect to their correct-classification rates and some specific measures.

2. PURPOSE OF THE STUDY

The main purposes of our study are to investigate which prediction model can produce the best results in Turkey and to see the extent to which it is possible to develop a new model that potentially produces superior results. The problems that we are concerned with can be summarized as the following:

a)   Are the existing financial distress prediction models applicable in and transferable to Turkey?

b)   Is it possible to design a new model with a superior potential of accurately predicting financial distress for our country?

c)   Are the statistical techniques adequate in modeling?

To test and compare both the existing models and our new model designs, a separate equation is constructed for each model and then its significance is analyzed statistically. The correct-classification rates of all the models are determined using a specific cut off point and then compared to each other.

The univariate analysis is the main tool used to assess the accuracy of linear models and to determine the explanatory strengths of the variables included. We will base our conclusion about which statistical technique is the most adequate one in predicting failure on R^2 parameter (coefficient of determination), which is referred to as an appropriate measure for the power of any model.

3. HYPOTHESES

It is more convenient to separate and present the hypotheses of the study under two main titles: a) Hypothesis for the tests on the existing models, and b) hypotheses regarding our new model proposals.

The general hypothesis to be tested through our investigation on the existing models is derived from a report about the original-sample performance levels of these models (see Table VI), which was presented in Zmijewski, 1984.

The Hypotheses of the Tests on the Existing Models:

H$_1$: ……… Model is significantly accurate in predicting financial distress at the 5 % significance level.

H$_2$: The most accurate model among Altman's models is ZETA model. (claimed by Scott, 1981, p.317)

Hypotheses Relating New Model Proposals:

H$_3$: The predictor variables (financial ratios) have a significantly high explanatory power at 5%.

H$_4$: Probit and logit techniques perform better than linear probability (zero-one regression) and linear discriminant techniques.

H$_5$: Our new model proposal that has proved to be the best one has a performance level better than those of the existing models.

4. METHODOLOGY

Our study is a typical example of the studies involving numerical and continuous independent variables. However, the dependent variable is a dichotomous variable taking the value of 0, or 1 for two complementary situations; failure, and success. The statistical purpose is to construct linear equations which can appropriately represent the relationship between a certain set of independent variables (i.e. financial ratios) and failure status.

4.1. Model Variables

The independent variables used in testing the existing models are the same as the variables that were used in the original studies in which the models were proposed. Differently, the variables included in our new model equations are selected from the first ten financial ratios that were proved to be the best predictors of failure by the past researchers of this field. (See Table I)

4.2. Statistical Techniques and Assumptions

The statistical techniques which will be used have been decided with respect to the basic methodologies followed in the original studies. Therefore, we have used mainly the following four statistical techniques with different characteristics and assumptions: Linear discriminant, logit, multiple linear regression, and probit techniques.

The *linear discriminant technique* is a statistical method that basically tries to compute relevant scores to be used in evaluating sample units as being a member of any of the two complementary (binary) groups or cases; for example, failing or non-failing. It is useful for situations where we need to build a predictive model of group membership based on observed characteristics of each case. The procedure generates a discriminant function (or, for more than two groups, a set of discriminant functions) based on linear combinations of the predictor variables that provide the best discrimination between the groups. Dependent variable takes the value of 1 or 0 (may be more than 2 discrete values provided that it is needed to separate sample units into more than 2 groups) according to the actual status of each sample unit. The final score that the technique produces is compared to a certain cut-off point to conclude which group each unit falls into. This cut-off point is defined as the middle point between the means of the units of two case. It can also be computed determining the extreme scores of each group that are generated through a normal probability function. (Tatlidil, 1996, pp.72-74)

The linear discriminant technique has some assumptions to simplify the real situation. It matters for the accuracy and reliability of discriminant models whether or not these assumptions are met in real life.

The use of *linear multiple regression technique* to predict corporate financial distress follows an approximately same procedure as in its use for general forecasting purposes. However, the dependent variable takes two discrete values (more than 2 for multinomial cases) such as 0, or 1 with respect the observed status of each sample unit. The technique used in this case thus resembles a multivariate linear regression method. A linear equation is the final output of this analysis to be used in determining failure scores of firms. (Tatlidil, 1996, p.74)

Although the final equation is the same as in linear discriminant technique, the cut-off point used in classification differs. For most cases, the value of 0,5 is taken as the classification limit. On the other hand, it is not obvious whether this value is the exact one that helps assess the true classification performance. Normal probability density function and mathematical optimization techniques can be applied so as to determine the true cut-off point.

The assumptions on which the multiple regression technique is based are so restrictive and unfortunately irrational in most cases as the assumptions of linear discriminant analysis such as, normal distribution, linear relationship, absence of multi-collinearity, and so on.

The linear discriminant technique significantly deviates from the multiple linear regression analysis especially at the point that it uses final discriminant score to compute squared Mahalanobis distance to centroid that is considered in computing group membership probability. However, linear regression models are designed to directly compare regression scores with the predetermined cut-off point.

The *binary logistic regression* is a method that partly resembles multiple regression technique, but deviates from it due to the distinguishing formula used to compute failure probability scores. A typical linear regression equation is a semi-final product of the analysis. This equation is used to estimate failure probabilities assuming that the probability distribution is logistic. Like in the other techniques, two or more discrete values are assigned to the dependent variable which represents the identity number of any group. (Ozdamar, 1996, p.364)

The Binary Logistic Regression technique has fewer assumptions which are less restrictive and more rational when compared to the assumptions of the techniques presuming normality such as the Multiple Discriminant and Regression techniques.

Failure probability scores are the key indicators to determine to what extent sample units can be considered a member of a particular group. For example, if the score approaches 0 and is below the cut-off point, it means that the sample unit can be regarded as an affiliate of Group 0 and the chance that it takes place in this group rises proportionately to an increasing degree of closeness to 0. The probability of fifty per cent is usually taken as the classification limit in evaluating individual scores. But, there may be more convenient cut-off values that could yield better results. Several methods like the Gini Coefficient, or mathematical programming solutions aimed at minimizing the costs of misclassifying can be used to determine the most appropriate cut-off figure for a prediction model that would lead to a better performance.

One of the techniques that partially differ from the linear forecasting methods is the ***Probit Technique***. This statistical technique does not necessitate the normality rule of variable distributions. Like the Binary Logistic Regression Technique, it also has fewer assumptions relative to other linear methods. (Tatlidil, 1996, p.74)

As for the Binary Logistic Regression, the procedure of computing failure probability scores consists of two stages. The first stage includes the derivation of a linear equation to be used in generating semi-final scores that are processed in the second stage. Following the computation of semi-final scores for all firms, these scores are taken as input to the process to be applied in the second stage in order to compute failure probabilities according to the cumulative normal distribution function. (Maddala, 2004, p.322)

Since the final score is the probability of failure, it is a practical custom to take the classification cut-off score as 0,5. On the other hand, the cut-off point may be determined to be a value other than 0,5.

Our study has some other assumptions in addition to those of the statistical techniques applied. These basic assumptions valid for the whole study can be itemized in the following way.

- Prior probability of failure for the entire population is 50 %.

- The critical significance level is 5 %.

- All sample units are selected on a random basis.

- The cut-off point is assumed to be 0,5 for our analyses employing the Regression, Probit and Logit techniques whereas it is determined as the mid-point between group mean scores in the case of applying the multiple discriminant technique.

To assess how significant the predictor variables are we have applied the univariate and stepwise filtering techniques.

4.3. Sampling Procedure and Data Collection

The target population in our study is publicly held corporations with shares traded on the Istanbul Stock Exchange (ISE). We have detected exactly 55 failed firms satisfying at least one of the necessary conditions for business failure. These firms were identified after a careful examination of the failure statistics for the period 1998-2003 using our predetermined definitions of corporate failure in accordance with the legal framework regulating trade relations and corporate activities in Turkey,

namely Turkish Commercial Law (TTK) (for the composition of the samples, see Table X). Under the relevant act in this law, the necessary conditions for a corporation to be considered a failed or bankrupt firm are composed of three main circumstances, which are (Berk, 2000, p.488):

- To have been sued due to inability to repay matured debts,

- To have a negative net (equity) value,

- Voluntary petition by the firm itself for bankruptcy.

Thirty five of these failed firms have been included in the original sample used to construct model equations. The selected failed firms have been paired with appropriate non-failed firms of similar asset size in the same industries. The remaining twenty failed firms have been held to construct a secondary sample to be used in measuring real correct-classification rates more objectively. In the end, we have studied 70 firms, 35 failed and 35 non-failed, in the original sample and 56 firms, 20 failed and 36 non-failed, in the secondary sample. Banking institutions were excluded from the sample because of the different bankruptcy environment confronting them.

Stratified sampling procedure has been applied in setting the original sample. On the other hand, the secondary sample units have been selected on a random basis regardless of size and industry differences. Consequently, the number of failed firms does not equal the number of non-failed firms, different from our original sample.

Timing of the financial statement data has been arranged according to the dates on which the failed firms fell in financial distress. The first year prior to failure of a failed firm has been taken as the year of analysis for both that firm and the non-failed firm paired with it when constructing model equations. This alignment has not been made for the secondary sample firms for the purpose of determining the prediction sensitivities of the models to changing circumstances.

As a requirement of our model construction attempts and our further investigation in which we have tried to determine the correct classification rates for the second, third and fourth years before failure, the four-year financial ratios were computed using the year-end balance sheets and income statements as of these years. Non-financial statement data and industry averages were collected from the official web sites of the Istanbul Stock Exchange and the Turkish Statistics Institution.

4.4. Problems and Restrictions

It is a challenging factor for our study that the capital market in Turkey is a new and emerging one. Inaccessibility to full financial statement data was one of the major problems and restrictions that we faced since it did not allow us to study equally-sized samples in testing all the models.

The nature of financial ratios data has challenged us in modelling in some cases because the circumstances when the ratios took negative values made it impossible for us to transform these values by using the natural logarithm as suggested by some of the models. Due to this restriction, we have been confronted with the problem of having to study on smaller samples.

Another problematic issue in our study has resulted from the disharmony of Turkish financial reporting practices to the international standards. In other words, it challenged us that there was no appropriate account called 'Retained Earnings' in the accounting statements. Therefore, we had to perform more detailed calculations so as to reach the balance of retained earnings. To do that, the balance of reserves

account for the first year for which the financial statements are completely available has been aggregated with the net earnings (losses) of each of the years prior to the reporting date, which are the residual net income (loss) in excess of paid dividends.

Differences between the firms' choices on the major reporting and valuation methods regarding inventory, depreciation, and so on, can be considered to be a critical factor that could substantially, maybe negatively, affect the model prediction performances, and we, unfortunately, could do nothing in the context of adjustments to eliminate possible consequences of this unpleasant situation. Moreover, the restrictive assumptions of the statistical techniques applied here may be sought to be the weak sides of our analysis.

5. IMPLEMENTATION & RESULTS

A separate model equation has been created for each model and the models' correct-classification rates have been examined separately for both the original and secondary samples, and also for the entire sample including the former two. The task of examining correct classification rates has been carried out by comparing the fitted scores of the firms to a certain cut off point pre-identified according to the nature of the statistical technique used. The prediction performance comparison between the models has been performed in terms of the Type I, Type II and Overall Performances (See Table V) taking into account the individual accuracy t-score of each in a way that a higher score corresponds to a better model.

- *Type I Error (Misclassification Rate)*, which refers to the percentage of the failed firms that are classified by the model as non-failed. (Type I Performance)

- *Type II Error (Misclassification Rate)*, which refers to the percentage of the non-failed firms that are classified by the model as failed. (Type II Performance)

Predicted Status	Actual	
	Failed	Non-Failed
Failed	a	b
Non-Failed	c	d

b= Type I Error

c= Type II Error

- *Overall-accuracy t-values.*

Therefore, the final comparison of the prediction performances of the new and existing models to one another is based especially on their overall accuracy t-scores or joint prediction performance (Count R^2), which is the overall percentage of correct classification: (Maddala, 2004, p.329)

Number of correct predictions

Count R^2 = ----------------------------------

Total number of observations

An accuracy t-score is calculated using the following formula: (Altman, 1993, p.193)

$$t = \frac{r - p}{\sqrt{\dfrac{p(1-p)}{n}}}$$

where,

t: Accuracy t-score

r: Correct classification rate (Count R^2)

p: Assumed priori probability

n: Sample size

Our study has also been enriched through the inclusion of some other performance measures in the discussion. So, the statistical findings resulting from the evaluation focusing on correct-classification rates and   have been re-tested, even supported in almost all the cases, by use of the following test statistics of goodness of fit: (Maddala, 2004, pp.327 - 328)

A- Simple R-Square: Square of correlation between actual and predicted.

B- Effron's R^2 :

$$R^2 = 1 - \frac{n}{n_1 n_2} \sum (y_i - \hat{y}_i)^2$$

where,

n: Total sample size, n_1: Size of group 1, n_2: Size of group 2, y_i: Actual value (status), $\hat{y}_i$: Predicted value

C- Relative measures based on residual sum of squares and number of independent variables: (Maddala, 2004, p.485)

- Theil's R^2 = RSS / (n-k)

- Hocking's S_P = RSS / (n-k-1)

- Amemiya's PC = RSS (n+k) / (n-k)

-            Akaike's AIC = $RSS.e^{2(k+1)/n}$

where,

       RSS : Residual sum of squares of prediction errors

         n: Sample size

         k: Number of predictor variables

Minimization of the above measures is regarded as evidence for the superior success of a linear model in any comparison.

D- Posterior Odds Ratio

Another measure we have used to assess the superiority of a model to the others is the Posterior Odds Ratio, which can be summarized as: (Maddala, 2004, p. 492)

$$P = \left[\frac{ESS_0}{ESS_1} \right]^{\frac{n}{2}} .n^{\frac{k_0}{k_1}}$$

where;

         P: Posterior odds ratio

        ESS: Sum of squares of prediction errors

         n: Sample size

         k_i: Number of independent variables included in model *i*

As long as the ratio exceeds 1 and approaches to infinity, it can be concluded that the Model 1 is much better than the Model 0. This measure has been used for examining the performance of our new models on a comparative basis. (For results, see Table VIII)

E-   LR - R^2 Test Statistic

The final measure we have considered when comparing in-sample prediction performances is the R^2 statistic based on likelihood ratios, which is summarized as: (Maddala, 2004, p.328)

$$R^2 = 1 - \left(\frac{L_R}{L_{UR}} \right)^{2/n}$$

where,

         L_R: The maximum when maximized with the restriction $\beta_i = 0$ for i = 1,2,3, …., k.

         L_{UR}: The maximum when maximized with respect to all the parameters.

F- Theil's U Test Statistic

Finally, we have added Theil's U Statistic into our comparison analysis so as to comparatively evaluate secondary-sample(out-of-sample) prediction performances. This measure is simply computed using the accompanying formula:

$$U = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{\sum (y_t - y_{t-1})^2}}$$

where;

y_t: Actual status in the year of analysis

y_{t-1}: Actual status in the first year prior to the time point of analysis

$\hat{y}_t$: Predicted status (0 or 1)

Theil's U statistic is a measure used to compare any linear model with a naïve (basic) model. If U is smaller than 1, it can be claimed that the model produce better forecasts than those by the naïve model. While the reverse situation suggests the superiority of the naïve model, a U value of 1 leads the analyst to be indecisive about the adequacy of the model being tested. In the case of a U statistic score equal to zero, it is possible to say that the model forecasts perfectly.

Since it is a common case that any model likely reveals subjectively better performance on its own original (modeling) sample, the basic comparison on the models has been conducted regarding the correct-classification rates, Type I and Type II performance levels for the entire sample to neutrally evaluate the model performances by putting away the biases that the original sample results could lead to. But, the correct-classification percentages for both the original and secondary samples are also reported in this part. SPSS and Excel programs have been used in constructing model equations and computing function scores. The details of our analyses are given in the following parts separately for each model. (See all the detailed statistical results in Tables II and III)

5.1. Results of the Analysis on Altman's Z-Score Model and Concluding Remarks

<u>**Sample:**</u> The sample used in modeling consists of 59 firms, 29 failed and 30 non-failed. The secondary sample contains only 46 firms, 13 failed and 33 non-failed firms.

<u>**Model Equation:**</u>

$Z = 0{,}339 + 2{,}298X_1 + 1{,}963X_2 - 1{,}962X_3 + 0{,}008X_4 - 0{,}182X_5$

<u>**Classification Limits:**</u>

$Z < -0{,}01 \Rightarrow$ Failed,

Otherwise $\Rightarrow$ Non-Failed

Concluding Remarks:

- The discriminant function seems to be able to classify the firms as failed and non-failed. The canonical discriminant function can be used. (Significance Level for Wilks' Lambda Value of 0,696 is 0,001, which is smaller than 0,05.)

- The covariance matrices of the failed and non-failed firms are not equal to each other. This may have decreased the accuracy of the model. A quadratic model can serve better. (F: 14,152 => 0,000 = Significance F < 0,05)

- 73,3 % of the firms are correctly classified by the model. The correct prediction percentages in classifying the failed and non-failed firms are respectively 71,4 % and 74,6 %.

- Among the financial ratios, the best two ratios are NWC/TA and RE/TA, having approximately the same discriminative power over failure. The least powerful ratio is SALES / TA. These results are in tune with the suggestions of Altman.

- According to the stepwise method, only the NWC/TA ratio is worth being kept in a linear discriminant model at the 0,05 significance. This conclusion is contrary to the results declared by Altman. He concluded that RE/TA was the most powerful ratio. Respecting the results of the univariate analysis, three ratios, which are NWC/TA, RE/TA, and EBIT/TA, seem to be significant enough in a linear model at the 0,05 significance.

- The model can be accepted to be accurate enough in classifying all the firms at the 0,05 significance level (Accuracy t = 4,76 > 1,99[1]). The model proves to be accurate enough in classifying both the original sample firms. (t = 4,01 > 2,01[2]) and the firms in the secondary sample

(t = 2,67 > 2,021[3])

- The Z-Score changes in the same direction as the changes in NWC/TA, RE/TA, and MVE/TA. On the other hand, this relationship is reverse for EBIT/TA and SALES/TA

5.2. Results of the Analysis on Altman's Revised Z-Score Model and Concluding Remarks

Sample: The sample used in modeling consists of 70 firms, 35 failed and 35 non-failed. The secondary sample used for re-testing the model contains 55 firms, 20 failed and 35 non-failed firms.

Model Equation:

$$Z = -0,160 + 1,246X_1 + 2,081X_2 - 1,725X_3 + 0,361X_4 - 0,048X_5$$

---

1 t-table value = 1,99 for n=105 at 0,05 significance level. (df= n-k-1)

2 t-table value = 2,01 for n=59 at 0,05 significance level

3 t-table value = 2,021 for n=46 at 0,05 significance level

Classification Limits:

Z < 0 => Failed,

Otherwise => Non-failed

Concluding Remarks:

- The discriminant function seems to be able to classify the firms as failed and non-failed. The canonical discriminant function can be used. (Significance Level for Wilks' Lambda Value of 0,639 is 0,000, a probability smaller than 0,05)

- The variance matrices of the failed and non-failed firms are not equal to each other. This likely decreases the accuracy of the model. Quadratic Model can serve better. (F: 12,993 => 0,000 Significance F < 0,05)

- 76 % of the firms are correctly classified by the model. The correct prediction percentages in classifying the failed and non-failed firms are respectively 80,0 % and 72,9 %.

- Among the financial ratios, the best performing ratios are RE/TA and BVE/TL. The least powerful ratio is SALES / TA. These results are the same as the conclusions of Altman.

- Our stepwise results have concluded that NWC/TA and BE/TL are worth including in a linear model. On the other hand, the univariate analysis concludes that the ratios RE/TA, EBIT/TA, and BE/TL can be regarded as being significant at the 0,05 significance level for a linear model.

- The model is accurate enough in correctly classifying all the firms at the 0,05 significance(Accuracy t=5,82> 1,98[4]). It can also be treated as an accurate model in classifying the firms in the original and secondary samples correctly (t=5,48>1,99[5]   and t=2,58 > 2,01[6])

- The Z-Score changes are in the same direction as the changes in all the ratios except EBIT/TA.

5.3. Results of the Analysis on Altman's ZETA Model and Concluding Remarks

Sample: The original sample consists of 50 firms with complete data, 23 failed and 27 non-failed firms. There are 36 firms in the secondary sample, 9 failed and 27 non-failed firms.

Model Equation:

$$Z = -2,532 - 2,190X_1 - 4,802X_2 - 0,016X_3 + 2,415X_4 + 0,855X_5 - 0,489X_6 + 0,481X_7$$

Classification Limits:

Z < -0,05 => Failed,

Otherwise => Non-Failed

---

4 t-table value = 1,98 for n=125 at 0,05 significance

5 t-table value = 1,99 for n=70 at 0,05 significance

6 t-table value = 2,01 for n=55 at 0,05 significance

Concluding Remarks:

- The discriminant function can be considered to be able to classify the firms as failed and non-failed. The canonical discriminant function can be used to discriminate firms. (Significance Level for Wilks' Lambda Value of 0,691 is 0,021, smaller than 0,05).

- The covariance matrices of the failed and non-failed groups are not equal to each other. Quadratic Discriminant Model could serve better.(F: 3,398 => Significance F = 0,000 < 0,05)

- 72,1 % of the firms are correctly classified by the model. The correct prediction percentages in classifying the failed and non-failed firms are respectively 68,8 % and 74,1 %.

- Only the ratio of CA/CL is found to be a significant predictor for a linear relationship at 0,05 significance level by the univariate technique. As to the Stepwise Analysis, only the RE/TA ratio can be considered significant enough at 0,05 significance level.

- Among the variables, the most contributing ratios are RE/TA and CA/CL. The least contributing one seems to be SMVE / TC. The superior performance of RE/TA is in parallel with Altman's past conclusions.

- The model can be accepted as accurate enough in classifying all the firms at the 0,05 significance. (Accuracy t = 4,10 > 1,99[7]).

- The model can also be accepted as accurate enough in classifying the original sample, but not accurate for the secondary sample firms. (t=3,68 > 2,01[8] and t=2,01<2,05[9])

- The ZETA Score changes in the same direction as the changes in the variables; RE/TA, CA/CL, and LOG (TA), while this change is in the reverse direction to those in the variables; EBIT/TA, STE, LOGEBINT, and SMEV/TC.

5.4. Results of the Analysis on Zmijewski's Probit Model and Concluding Remarks

Sample: The original sample consists of 70 firms with complete data, 35 failed and 35 non-failed firms. There are 55 firms in the secondary sample, 20 failed and 35 non-failed firms.

Model Equation:

P (Z') = Normal Standardized Probability for Z' standardized value.

Subject to:

$\quad\quad\quad$ Z' = Z/Standard Deviation of Z Values

Z = - 2,065 - 2,117X_1 + 3,618X_2 − 0,337X_3

---

7 t-table value = 1,99 for n=86 at 0,05 significance

8 t-table value = 2,01 for n=50 at 0,05 significance

9 t-table value = 2,05 for n=36 at 0,05 significance

The Accuracy of Financial Distress Prediction Models in Turkey: A Comparative Investigation with Simple Model Proposals

29

Classification Limits:

P (Z') > 0,50 => Failed

Otherwise => Non-Failed

Concluding Remarks:

- The probit model can be considered to be capable of classifying the firms as failed and non-failed. (The Chi-Square of the model is 47,440 and bigger than the critical chi-square value, 5,991[10] for the significance level of 0,05). The groups seem to be homogeneous (47,440 < 85,96 -Critical Chi-square for 0,05 and df=66).

- The model classifies 76,0 % of the firms correctly. The correct classification percentages for the failed and non-failed firms are respectively 78,1 % and 74,3 %.

- Only two ratios, TL / TA and CA / CL are considered good predictors and have a significant impact on the Z value on a linear basis.

- The model can be thought as an accurate model in classifying all the firms at the 0,05 significance level for the original sample and for the entire sample, but not for the secondary sample:

Original Sample t-value= 5,00 > 1,99 (t-table value for n=70 at the 0,05 significance level)

Secondary Sample t-value=3,33 < 2,01 (t-table value for n=55 at the 0,05 significance level)

The entire sample t-value= 5,82 > 1,96 (t-table value for n=125 at the 0,05 significance level)

- The Z value changes in the reverse direction as the changes in the variable, NPAT.TA. For the other variables, the Z value changes in the same direction (positive relationship).

5.5. Results of the Analyses on Beaver's Univariate and Deakin's Discriminant Models, and Concluding Remarks

Sample: The original sample and the secondary sample consist of respectively 70 firms - 35 failed and 35 non-failed - and 55 firms - 20 failed and 35 non-failed.

Since Beaver used each ratio separately as the predictor of financial distress and favored   CF / TA most, his system is a typical example of a univariate models.

Classification Limits for Beaver's Model:

CF / TL < 0,13[11] => Failed

CF / TL ≥ 0,13 => Non-failed

---

10 Chi-square table value for 0,05 significance and 2 degree of freedom = (number of groups - 1) x (number of variables - 1)

11 Although the cut-off point was determined to be 0,10 by Beaver, the results of our study have showed that the cut-off point of 13 percent is the best one that yields relatively better results. The normal distribution 100 % - confidence rule has been used in determining that cut-off point.

Model Equation for Deakin's Model:

$$Z = -0,875X_1 + 0,807X_2 + 1,773X_3 - 1,649X_4 + 0,862X_5 - 0,032X_6 - 0,959X_7 - 0,227X_8 + 0,906X_9 + 0,668X_{10} + 7,778X_{11} - 8,340X_{12} + 0,221X_{13}$$

The CASH / SALES ratio is not included in the model, because it was observed in SPSS that the ratio was out of tolerance limit.

Classification Limits for Deakin's Model:

Z < 0 => Non-Failed,

Otherwise => Failed

Concluding Remarks:

- Beaver's univariate model classifies 76,8 % of the firms correctly. The proportions of correct classification of the model for the failed and non-failed groups are respectively 80,0 % and 74,3 %.

- The most powerful ratio in discriminating the failed and non-failed firms according to the results of both univariate and stepwise analyses is CF / TL. This result has proved Beaver claims. The second best variable is TL / TA. The variable called NPAT/TA which was claimed to be the most significant ratio by Beaver just after CF / TL is the third ratio in our analysis. The least significant two ratios are CASH / CL and QA / TA.

- Beaver's univariate model is accurate in predicting distress for the original and secondary samples and also for the whole sample.

Original Sample t-value= 5,00 > 1,98 (t-table value for n=70 at 0,05 significance level)

Secondary Sample t-value= 3,50 > 2,00 (t-table value for n=55 at 0,05 significance level)

The entire sample t-value= 6,00 > 1,96 (t-table value for n=125 at 0,05 significance level).

- While the ratios found significant enough to be included in a linear model according to the stepwise analysis are CF / TL and TL / TA, for the univariate analysis the favored ratios at the 0,05 significance level are CF / CL, NPAT / TA, TL / TA, QA / TA, CASH / TA, CASH / SALES, QA / SALES, and CA / TA.

- The discriminant function developed for Deakin's Model can classify the firms as failed and non-failed firms. The canonical discriminant function can be used. (Significance Level for Wilk's Lambda Value of 0,481 is 0,000 < 0,05)

- Deakin's Model classifies 77,6 % of the firms correctly. The correct classification percentages for the failed and non-failed firms are respectively 76,4 % and 78,6 %.

- The Deakin's re-estimated model can be thought to be accurate enough in classifying the firms at the 0,05 significance level for the both samples and also for the entire sample including the former both.

Original Sample t-value= 5,78 > 2,01   (t-table value for n=70 at 0,05 significance level)

Secondary Sample t-value= 2,82 > 2,02 (t-table value for n=55 at 0,05 significance level)

The entire sample t-value= 6,18 > 1,97 (t-table value for n=125 at 0,05 significance level)

    - The variance matrices of the groups are not equal to each other. (F: 6,885 => Significance Level = 0,000 < 0,05). This situation may affect the accuracy of the model, so a quadratic discriminant function should be used instead.

    - The changes in the values of the ratios, CF/TL, CA/TA, NWC/TA, CASH/TA, CA/ CL, and QA / SALES, create reverse changes in the Z-Score. The changes in the other ratios change it in the same direction.

5.6. Results of the Analysis on Zavgren's BPR Model and Concluding Remarks

Sample: The original sample consists of 69 firms with complete data, 35 failed and 34 non-failed firms. There are 53 firms in the secondary sample, 19 failed and 34 non-failed firms.

Model Equation:

$$I \text{ (Failure Index)} = (1 / (1 + e^{-Z})$$

where,

$$Z = -0,756 + 3,162X_1 + 0,119X_2 + 0,925X_3 - 1,593X_4 - 0,186X_5 + 0,864X_6 - 0,151X_7$$

Classification Limits:

    I > 0,50 => Failed

    Otherwise => Non-Failed

Concluding Remarks:

    - The logistic regression model can be considered to be capable of classifying the firms as failed and non-failed. (The Chi-Square of the model is 53,283 and bigger than the critical chi-square, 12,59[12] for the significance level of 0,05. p-value is 0, 584, a value bigger than 5 per cent, which means that the distribution can be explained by a logistic approach). The groups are homogeneous. (53,283<80,23 (Chi-Square Value for 0,05 sign. and df = 61). A linear function can be used to estimate failure probabilities.

    - The model classifies 74,1 % of the firms correctly. The percentages of correct classification for the failed and non-failed firms are respectively 68,5 % and 79,3 %.

    - None of the ratios is a good predictor and has a significant impact on the Z-value on a linear basis.

    - The model can be considered to be accurate enough in classifying all the firms at the 0,05 significance level for the original sample and for the entire sample, but not for the secondary sample:

---

12   Chi-square table value for 0,05 significance and 6 degree of freedom (# of groups - 1) x (# of ind. Variables - 1).

Original Sample t-value= 4,41 > 2,00 (t-table value for n=69 at 0,05 significance level)

Secondary Sample t-value= 0,95<2,01 (t-table value for n=53 at 0,05 significance level)

The entire sample t-value= 3,93 > 1,97 (t-table value for n=125 at 0,05 significance level)

- The Z-value changes in the same direction as the changes in the variables AINVSAL, ARECINV, CASHTA, and LTLFC. For the remaining variables, the Z-value changes in the reverse direction (negative relationship).

5.7. Results of the Analysis Ohlson's O-Score Model and Concluding Remarks

Sample: The original sample consists of 70 firms with complete data, 35 failed and 35 non-failed firms. There are 55 firms in the secondary sample, 20 failed and 35 non-failed firms.

Model Equation:

$$\text{O-Score } (P(Z)) = (1 / (1 + e^{-Z})$$

where:

$$Z = -4{,}582 - 0{,}228X_1 + 7{,}186X_2 - 0{,}073X_3 + 0{,}613X_4 - 1{,}714X_5 + 3{,}264X_6 - 4{,}187X_7 + 0{,}438X_8 - 0{,}154X_9$$

Classification Limits:

O-Score > 0,50 => Failed

Otherwise => Non-Failed

Concluding Remarks:

- A binary logistic regression model can be considered to be able to classify the firms as failed and non-failed. (The Chi-Square of the model is 38,861 and bigger than the critical chi-square, 15,55[13] for the significance level of 0,05. p-value is 0, 984, a value bigger than 5 per cent, which means that the distribution can be explained by a logistic approach). The groups are homogeneous (38,861 < 79,08 (Chi-square for the 0,05 significance and df=60)

- The model classifies 81,6 % of all the firms correctly. The percentages of correct classification for the failed and non-failed firms are respectively 89,1 % and 75,7 %.

- Only two ratios, TL/TA and CF/TA are considered a good predictor and have a significant impact on the Z value on a linear basis.

- The model can be accepted to be accurate enough in classifying all the firms at the 0,05 significance level for the both sample and secondary samples and for the entire sample.

Original Sample t-value= 6,20 > 2,00 (t-table value for n=70 at 0,05 significance level)

Secondary Sample t-value=3,98 > 2,00 (t-table value for n=55 at 0,05 significance level)

The entire sample t-value= 7,07 > 1,98 (t-table value for n=125 at 0,05 significance level)

---

13 Chi-square table value for 0,05 significance and 8 degree of freedom. (# of gropus - 1) x (# of ind. variables - 1)

- The Z value changes in the same direction as the changes in the variables TL/TA, CL/CA, NPAT/TA, and DUMNPATC. For the others, the Z-value changes in the reverse direction (negative correlation).

5.8. Simple Model Proposals as an Alternative to the Existing Models

Sample: For all of the studies presented below, an original sample of 70 firms, 35 failed and 35 non-failed firms has been used. The secondary sample contains 54 firms, 20 failed and 34 non-failed, with a set of complete data.

Methodology: Four separate models based on the ten financial ratios presented in Table I as the predictors of failure have been constructed using four different statistical techniques that we previously mentioned. Furthermore, the prediction performances of these models have been compared to each other and to those of the existing models that we have tested and detailed in the previous parts of our empirical study.

Model Equations:

a) MDA MODEL:

$$F = 0{,}780 + 2{,}095X_1 + 1{,}146X_2 - 0{,}627X_3 - 0{,}382X_4 - 0{,}813X_5 - 0{,}967X_6 + 1{,}126X_7 - 0{,}166X_8 + 0{,}093X_9 + 0{,}382 \cdot X_{10}$$

b) MRA MODEL:

$$F = 0{,}695 + 0{,}525X_1 + 0{,}287X_2 - 0{,}157X_3 - 0{,}096X_4 - 0{,}204X_5 - 0{,}242X_6 + 0{,}282X_7 - 0{,}042 X_8 + 0{,}023X_9 + 0{,}096 X_{10}$$

c) PROBIT MODEL:

F = P (Z') = Standardized normal probability for Z' standard value

where,

Z' = Z / Standard Deviation of Z Values

$$Z = 2{,}408 + 3{,}436X_1 + 1{,}389X_2 - 1{,}650X_3 - 0{,}920X_4 - 2{,}783X_5 - 1{,}776X_6 + 2{,}542X_7 + 0{,}151X_8 - 0{,}317X_9 + 0{,}806X_{10}$$

d) LOGREG MODEL:

$$F = (1 / (1 + e^{-Z}))$$

where,

$$Z = 4{,}981 + 5{,}006X_1 + 2{,}178X_2 - 4{,}481X_3 - 0{,}924X_4 - 7{,}076X_5 - 3{,}165X_6 + 6{,}093X_7 + 0{,}097X_8 - 0{,}804X_9 + 2{,}492X_{10}$$

Classification Limits:

MDA MODEL:

   F > 0 => Non-failed

   Otherwise => Failed

MRA MODEL:

   F > 0,5 => Non-failed

   Otherwise => Failed

PROBIT and LOG-REG MODELS:

   F > 0,5 => Non-failed

   Otherwise => Failed

Concluding Remarks:

   - According to the stepwise analysis, the ratios CF / TL and TL /TA are the most significant variables at the 0,05 significance level. But, respecting the results of the univariate analysis, one of the ratio, CF/TL, can be considered to be powerful enough in predicting failure in a linear model with the 95 % confidence. . It is also concluded by the univariate analysis that these two variables are the most favored two with respect to their discriminating powers. QA / SALES and CASH / SALES are the least significant variables with respect to the results of the both analyses.

a) MDA Model:

   - The discriminant function can classify the firms as failed and non-failed. The canonical linear discriminant function can be used (The significance level for Wilks' Lambda Value of 0,578 is 0,000 < 0,05). But, since the variance matrices of the groups are not equal to each other (F: 9,094 => 0,000 significance < 0,05) , a quadratic discriminant model can perform better.

   - The model classifies 78,2 of the firms correctly. 81,2 % of the failed firms and 75,7 % of the non-failed firms are truly classified by the model.

b) MRA Model:

   - The regression function can be considered to be properly working at the 0,05 significance level, because the significance level of the model is 0,000 (< 0,05). The explanatory power of the model is moderate (R^2= 42,2 %). None of the variables in the model is significant for the 0,05 significance level.

   - The model classifies 79,0 % of the firms correctly. The correct classification percentages for the failed and non-failed firms are respectively 81,8 % and 77,1 %.

    - We cannot conclude that no autocorrelation exists among the variables in the model. There is an indecisive position. (Durbin-Watson Statistic (d) = 1,821 < 1,86 and > 1,34)[14]

---

14   d_L=1,27 and d_U=1,82 for 0,025 significance level and k=10

c) PROBIT MODEL

- The probit model can classify firms accurately at the 0,05 significance level (Chi-square = 38,513 > 15,51). The groups seem to be homogeneous (38,513 < 77,93). A linear function can be used to estimate failure probabilities.

- The model classifies 80,6 % of the firms correctly. This percentage is 83,6 % for the failed firms and is 78,3 % for the non-failed firms.

- None of the variables are significant enough in the model because the absolute t-values of all are smaller than 2,02 (Critical t-value for df: 59)

d) LOG-REG MODEL

- The log-reg model can classify the firms accurately at the 0,05 significance level because the model's significance level corresponding to its chi-square value of 57,301 is 0,000, a value smaller than 0,05. The R-square value of the model is 0,75. None of the variables has been found significant enough at the 0,05 significance level.

- The model classifies 82,3 % of the firms correctly. (83,6 % of the failed firms and 81,2 % of the non-failed firms.)

6. COMPARISON OF THE OVERALL PREDICTION PERFORMANCES

The comparison of the models is carried out theoretically according to their performances in terms of the quantitative criteria we have mentioned in the fourth part. The hypotheses are tested with respect to the models' accuracy t-scores:

6.1. Testing the Hypotheses related to the Existing Models

H_1: ……… Model is significantly accurate in predicting financial distress at the 5 % significance level.

Ohlson's O-Score Model proves to be the best model in our study and can be treated as significantly accurate at 5 %. Its accuracy t-value is 7,07. H_1 is accepted.

The second best performing model among the models we have tested according to the overall prediction performance is Deakin's Discriminant Model with an accuracy t-value of 6,18. H_1 is accepted. Beaver's one-ratio model can also be considered accurate enough at 5 %.

Zmijewski's model can be considered to be significantly accurate at the significance level of 5 %.(Accuracy t-value = 5,82). H_1 is accepted.

The Z-Score Model is the third of the worst models in our study with the accuracy t-value of 4,76. On the other hand, the model has yielded an accuracy level significant at 5 %. H_1 is accepted.

Altman's Revised Z-Score Model has an accuracy t-value of 5.48. It means that the model has yielded a degree of accuracy which is significant at 5 %. H_1 is accepted.

The hypothesis is accepted since Altman's ZETA Model has an overall accuracy t-value of 4,10 corresponding a significance level smaller than 0,05.

The worst model of our analysis is Zavgren's BPR Model. It has the lowest accuracy t-value, which is 3,93. But, this value is enough for the model to be regarded as significantly accurate at 5 %. H_1 is accepted. (The results of significance tests on the correct-classification performances are summarized in Table IV)

H_2: The most accurate and valid one of Altman's models is ZETA model.

Our results show that this hypothesis should be rejected, because the model with the highest accuracy t-value among Altman's models is the Revised Z-Score Model (5,82). It is obvious if we consider the results of other performance measures (see Table VIII) that the Revised Z-Score Model outperforms the other two and is the best one.

Since none of those models is satisfactory in predicting financial distress, it becomes important to search for a more appropriate prediction model of financial distress for our country using different and best functioning financial ratios and statistical methods.

6.2. Testing the Hypotheses related to the F-Score Models

H_3: The independent variables have collectively a considerably high degree of explanatory power on a linear basis at the 0,05 significance level.

Since the p-values of the financial ratios excluding CF/TL all are larger than 0,025 (0,05 / 2), it can be said that none of the financial ratios has a significant impact on the dependent variable. This conclusion is supported by the finding that the explanatory power of the variables are moderate ($R^2 = 0,422 < 0,70$ – See Table II). H_3 is rejected. But, taking into consideration the stepwise statistics, only the variables CF / TL and TL / TA have high discriminative powers.

H_4: The new models developed by using Probit and Logistic Regression methods perform better than the models derived as based on the Multiple Regression and Multiple Discriminant Analyses.

Accuracy t-Value

F-SCORE PROBIT MODEL	**6,82**
F-SCORE LOG-REG MODEL	**7,27**
Mean t-Value	**7,05**

Accuracy t-Value

F-SCORE MDA MODEL	**6,31**
F-SCORE MRA MODEL	**6,49**
Mean t-Value	**6,40**

Posterior Odds Ratios against the LOG-REG and PROBIT

	LOG-REG	PROBIT
F-SCORE MDA	0,000194	0,001540
F-SCORE MRA	0,000003	0,000024

H_4 is accepted, because 7,05 > 6,40. It means that the normality-based models are less accurate than the models not assuming normal variable distributions. In addition, the posterior odds ratios support our judgment, because the ratios calculated for the MDA and MRA F-Score models against the LOG-REG and PROBIT F-Score models exist much below one. (See Table IX)

H_5: The new model design that has proved to be the best has a superior prediction performance to those the existing models have.

The model that has proved to be the best among our new models is the LOG-REG. Its overall accuracy t-value is 7,27, which is a number higher than all of the overall-accuracy t-values that the other new and existing models could generate. So, it can be said that our most favored F-Score Model is better than all the other models. Besides accuracy t-scores, other performance measures also have produced obvious values supporting that decision. H_5 should be accepted.

According to the results of our attempt to use some additional model performance measures, we also conclude and suggest that Ohlson's O-Score Model and our F-Score Log-Reg Model are two competent and competing models for our country. Their accuracy measures are approximately the same and close to one another (See Table VIII). Moreover, Ohlson's Model seems to be slightly better than the F-Score Log-Reg model (its R^2 is 74,6 % whereas the R^2 for our Log-Reg Model is 74,5%) . The statistical findings that we have received using both accuracy t-scores and other measures produce the identical conclusions about the performance of the models. Logistic regression and probit models have performed better than multiple regression and discriminant models (See Table IX)

7.A FURTHER INVESTIGATION

We have also tried to carry out a multi-term performance evaluation study as a further investigation out of the empirical study of which the details have been given before in order to assess the prediction performance of these eight existing and four newly developed models for two to four years before failure on the initial (original) and secondary samples that we have also used in deriving the first-year model estimations. For this purpose, using the previously derived functions of the models in assessing the failure probabilities of the firms the Type I, Type II, and overall accuracy levels of the models have been investigated for the corresponding years. (For the details on the results of this further effort, see Table VII)

Consequently, the results of our further empirical investigation on the post-modeling correct classification rates of the models for two to four years prior to failure yield some findings and conclusions which are exactly in parallel with the results of our first-year performance analysis. The F-Score Log-Reg Model again has proved to be the best one among all the existing and new prediction models, but with the highest variation of accuracy. Ohlson's O-Score Model is, on average, the best of

the existing models included here. Similarly, the poorest performance has been achieved by Zavgren's BPR Model again. As different from the first-year results, Altman's Z-Score Model has produced a higher percentage of prediction accuracy when the time horizon of data is enlarged, and becomes the best among Altman's models.

8. CONCLUSION & RECOMMENDATIONS

The most important conclusion that the results of this study point out is that it is possible to predict financial distress and bankruptcy more accurately using the domestic data in Turkey through the foundation of new models based on different combinations of financial ratios, when compared to the old and more prestigious model proposals. It is so interesting and amazing to have produced a new model proposal with a slightly better prediction performance using the binary logistic regression technique against the existing model proposals of the past researchers

Our findings show that none of the existing prediction models could achieve a satisfactory level of prediction performance. The correct-classification percentages of the existing models are all below 82 %. None could generate accuracy level obtained in its own original sample (For the original performances, see Table VI).

The best model among the existing financial distress prediction models is Ohlson's O-Score Model. Surprisingly, Altman's Models have produced lower levels of accuracy. The best of his three models is the Revised Z-Score Model. His Z-Score Model seems to be the worst model in Turkey.

Although all of the existing models have relatively higher prediction performance on our modeling sample, they could not repeat the same success for the secondary sample. Ohlson's O-Score Model is again at the top in terms of prediction performance on the secondary sample (75 %). The least accurate model is Zavgren's BPR Model with a correct classification percentage of 68 %.

The best of our new F-Score Models is the Binary Logistic Regression (LOG-REG) model. It has revealed a superior performance when compared to the performance levels of all the existing models, even including Ohlson's O-Score Model in terms of accuracy t-scores for the entire sample. But, if regarding their other performance measures (see Table IX), and Type I and Type II performances for each sample set (see Table VI), they seem to be equally worthy for predicting bankruptcy. Although the accuracy t-scores lead us to a conclusion that our Log-Reg F-Score Model is superior to the Ohlson's O-Score Model, the additional R^2 performance measures like Effron's R^2 and LR-R^2, relative measures based on RSS, and Theil's U statistic, presented in Table VIII, are slightly in favor of the O-Score model against our model. In the light of all these findings we have mentioned, we are sure that our Log-Reg F-Score model and O-Score model both could reveal apparently better performances, but inconclusive about which one of these two models outperforms the other.

It is one of our descriptive findings that the statistical techniques that postulate the normality of variable distributions have been less successful than the probit and log-reg techniques not postulating that circumstance. Posterior odds ratios favor the probit and log-reg models against the MDA and MRA models. (See Table IX)

Unfortunately, neither the existing, nor any of the new models could reveal any favorable success over 90 %. The correct-classification percentages of all have remained much less than the high levels that they achieved in the past original studies for one-year period before failure. With this study, it is targeted

mainly to give the basic fundamentals of and insights into financial distress prediction modeling process. We hope this study can challenge and encourage the further studies and is viewed as a leading theoretical and empirical source to prospective researchers.   But, the existing problems and restrictions due to our arbitrary choices on such methodological issues as the cut-off points used, statistical techniques employed, and predictor variables included in each model, must be handled more carefully and comprehensively. Some appropriate quantitative techniques like the Gini Coefficient should be undertaken to determine which cut-off value could better serve for each model. The performances of the models must be re-tested under all possible circumstances.

It can be viewed as an unfavorable characteristic of our empirical study that only the manufacturing and service providing firms were included in the samples, which may have negatively impacted and biased our research results. Therefore, in further studies the industry coverage of sampling should be kept large as possible to be able to increase the generalization and validity of the findings. In addition, new combinations of financial ratios can be added and tested in order to search for better model designs and reach more satisfactory results.

REFERENCES

Altman, Edward I. (1968). *Financial Ratios, Discriminant Analysis, and Prediction of Corporate Bankruptcy*, Journal of Finance, Vol. 4, 1968, pp. 589 - 609

Altman, Edward I. (1993). *Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy*, 1st Edition, New York: John Wiley and Sons.

Altman, Edward, R. Haldeman, and P. Narayanan. (1977). *ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations,* Journal of Banking and Finance, March, pp. 29 - 54

Argenti, J. (1991). *Predicting Corporate Failure*, Accountants Digest, Vol. 138, p. 3

Beaver, William H. (1966). *Financial Ratios as Predictors of Failures*, Journal of Accounting Research pp.71 – 111

Berk, Niyazi, (2000). *Finansal Yonetim*, 5th Edition, Turkmen Kitabevi, Istanbul

Deakin, Edward B. (1972). *A Discriminant Analysis of Predictrors of Business Failure*, Journal of Accounting Research, Vol. 10, No. 1, pp. 167 – 179

Keasey, K. and Robert Watson (1991). *Financial Distress Prediction Models: A Review of Their Usefulness*, British Journal of Management, Vol 2, pp. 89 - 102

Maddala, G.S. (2004). *Introduction to Econometrics*, 3rd Edition, Wiley, New York,

Mossman, Charles E., Geoffrey G. Bell, L. Mick Swartz, and Harry Turtle (1998). *An Empirical Comparison of Bankruptcy Models*, The Financial Review, Eastern Finance Association, Vol. 33, pp. 35 - 54

Ohlson, James A. (1980). *Financial Ratios and the Probabilistic Prediction of Bankruptcy*, Journal of Accounting Research, Vol. 18, No.1, pp. 109 - 131

Ozdamar, Kazım. (2002). *Paket Programlar ile Istatistiksel Veri Analizi 2 – Çok Degiskenli Analizler*, 4th Edition, Eskisehir, Kaan Press, 2002

Scott, J. (1981). *The Probability of Bankruptcy: A Comparison of Empirical Predictions and Theoretical Models*, Journal of Banking and Finance, Summer, pp. 317 - 344

Tatlidil, Huseyin (1996). *Uygulamali Cok Degiskenli Istatistiksel Analiz*, 1st Edition, Ankara: Cem Web Ofset Ltd. Co.

Zavgren, Christine (1985). *Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis*, Journal of Business Finance & Accounting, Vol. 12, No. 1, pp. 19 - 45

Zmijewski, Mark E. (1984). *Methodological Issues Related to the Estimation of Financial Distress Prediction Models*, Journal of Accounting Research, Vol. 22 Supplement, pp. 59 - 82

Table 1. Model Variables

RATIO	RATIO DESCRIPTION	Z-Score	Revised Z-Score	ZETA	Beaver	Deakin	Zavgren	Zmijewski	Ohlson	F-Score
ainv.sal	AVERAGE INVENTORY / SALES						(X_1)			
arec.inv	AVERAGE RECEIVABLES / SALES						(X_2)			
bve.tl	BOOK VALUE OF EQUITY / TOTAL LIABILITIES		(X_4)							(X_4)
ca.cl	CURRENT ASSETS / CURRENT LIABILITIES			(X_5)		(X_6)		(X_3)		
ca.sales	CURRENT ASSETS / SALES					(X_{11})				
ca.ta	CURRENT ASSETS / TOTAL ASSETS					(X_4)				
cash.cl	(CASH + MARK. SECURITIES) / CURRENT LIABILITIES					(X_{10})	(X_4)			(X_{10})
cash.sales	(CASH + MARK. SECURITIES) / SALES					(X_{14})				(X_9)
cash.ta	(CASH + MARK. SECURITIES) / TOTAL ASSETS					(X_7)	(X_3)			
cf.tl	CASH FLOW / TOTAL LIABILITIES (Cash Flow = NPAT + Depreciation)				(X_1)	(X_5)			(X_7)	(X_7)
cl.ca	CURRENT LIABILITIES / CURRENT ASSETS								(X_4)	
dumnpat	1 if NPAT is negative for the last two years, 0 otherwise								(X_8)	
dumtlta	1 if TL > TA, 0 otherwise								(X_5)	
ebit.ta	EBIT / TOTAL ASSETS	(X_3)	(X_3)	(X_1)						(X_3)
logebint	LOG (EBIT / INTEREST)			(X_3)						
logta	LOG (TOTAL ASSETS)			(X_7)						
logtagnp	LOG (TA / GNP PRICE LEVEL INDEX VALUE)								(X_1)	
ltl.tc	LTL / TOTAL CAPITAL						(X_6)			
mve.tl	MARKET VALUE OF EQUITY / TOTAL LIABILITIES	(X_4)								
npat.ta	NPAT / TOTAL ASSETS					(X_2)		(X_1)	(X_6)	(X_6)
npatch	(NPATt - NPATt-1) / (|NPATt| + |NPATt-2|)								(X_9)	
nwc.sales	NWC / SALES					(X_{13})				
nwc.ta	NWC / TOTAL ASSETS	(X_1)	(X_1)			(X_8)			(X_3)	(X_1)
oinc.tc	OPERATING INCOME / TOTAL CAPITAL (Total Capital = LTL + Equity)						(X_5)			
qa.cl	(CA - INV - OTHER CA) / CURRENT LIABILITIES					(X_9)				
qa.sales	(CA - INV. - OTHER CA) / SALES					(X_{12})				(X_8)
qa.ta	(CA - INV. - OTHER CA) / TOTAL ASSETS					(X_1)				
re.ta	RETAINED EARNINGS / TOTAL ASSETS	(X_2)	(X_2)	(X_4)						
sales.ta	SALES / TOTAL ASSETS	(X_5)	(X_5)							
salfanwc	SALES / (FIXED ASSETS + NWC)						(X_7)			
smve.tc	5-YEAR SMOOTHED (AVERAGE) MARKET VALUE EQUITY / TOTAL CAPITAL (Total Capital = Long-Term Liabilities + Equity)			(X_6)						
ste	STAND. ERROR OF X1 FOR 5-YEAR PERIOD			(X_2)						
tl.ta	TOTAL LIABILITIES / TOTAL ASSETS					(X_3)		(X_2)	(X_2)	(X_5)
f.nf	FAILURE STATUS (0 / 1)	0: Failed 1: Non-failed	0: Failed 1: Non-failed	0: Failed 1: Non-failed	0: Failed 1: Non-failed	0: Failed 1: Non-failed	1: Failed 0: Non-failed	1: Failed 0: Non-failed	1: Failed 0: Non-failed	0: Failed 1: Non-failed

NWC: Net Working Capital = Current Assets − Current Liabilities, EBIT: Earnings Before Interest and Tax, CA: Current Assets

Table II. Significance of Variables

| MODEL | OBSERVED SIGNIFICANCE LEVELS FOR THE RATIOS IN THE MODELS (%) UNIVARIATE ANALYSIS RESULTS | | | | | | | | | | | | | | VARIABLE(S) SIGNIFICANT AT <0,05 (UNIVARIATE) | VARIABLE(S) SIGNIFICANT AT <0,05 (STEPWISE) | Explanatory Power (R-Square Value-%) |
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}			
ALTMAN'S Z-SCORE MODEL	4,9	3,5	4,4	78,4	52,0	×	×	×	×	×	×	×	×	×	EBIT/TA RE/TA NWC/TA	NWC/TA	30,40
ALTMAN'S REVISED Z-SCORE MODEL	21,7	1,2	4,8	3,9	84,6	×	×	×	×	×	×	×	×	×	RE/TA EBIT/TA BVE/TL	NWC/TA, BVE/TL	36,10
ZETA MODEL	44,6	77,4	98,2	25,6	4,0	46,1	37,0	×	×	×	×	×	×	×	CA/CL	RE/TA	30,90
BEAVERS' UNIVARIATE ANALYSIS & DEAKIN'S DISCRIMINANT MODEL	0,8	1,0	1,1	1,3	3,8	87,9	4,2	86,1	58,2	45,2	3,3	3,2	58,0	19,3	CF/TL, NP/ATTA, TL/TA, CA/TA, QA/TA, CASH/TA CA/SALES QA/SALES	CF/TL, TL/TA	53,40
ZAVGREN'S BPR MODEL	18,8	18,2	64,0	35,3	62,9	49,7	38,8	×	×	×	×	×	×	×	None	CASH/TA	17,00
ZMIJEWSKI'S PROBIT MODEL	49,6	0,8	4,6	×	×	×	×	×	×	×	×	×	×	×	TL/TA, CA/CL	TL/TA, CA/CL	27,70
OHLSON' OSCORE MODEL	87,7	4,2	58,0	45,1	68,4	62,9	3,5	31,6	41,1	×	×	×	×	×	TL/TA, CF/CL	CF/TL, TL/TA	39,80
F-SCORE MODELS	21,2	65,0	59,2	45,1	4,7	73,0	13,1	79,1	87,6	46,7	×	×	×	×	TL/TA	CF/TL, TL/TA	42,20

The table values represent two-sided significance levels. A value larger than 0,05 suggests that the corresponding variable is not significant at the 5 % significance level. If below 0,05, it can be concluded that the variable can be considered statistically significant at 5 %.

Table III. Model Coefficients and Statistical Results

MODEL	SAMPLE SIZE						COEFFICIENTS															OVERALL SCORE	MODEL SCORE TO CLASSIFY FIRMS	CUT-OFF POINT FOR FAILURE
	MODELING			SECONDARY			β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}			
	Failed	Non-failed	TOTAL	Failed	Non-failed	TOTAL																		
ALTMAN'S Z-SCORE MODEL	29	30	59	13	33	46	0,339	2,298	1,963	-1,962	0,008	-0,182	.	.	.	.	.	.	.	.	.	Z-Score	Z-Score	< -0,01
ALTMAN'S REVISED Z-SCORE MODEL	35	35	70	20	35	55	-0,160	1,246	2,081	-1,725	0,361	-0,048	.	.	.	.	.	.	.	.	.	Z-Score	Z-Score	< 0,35
ZETA MODEL	23	27	50	9	27	36	-2,532	-2,190	-4,802	-0,016	2,415	0,855	0,489	0,481	.	.	.	.	.	.	.	Z-Score	Z-Score	< -0,05
BEAVER'S UNIVARIATE ANALYSIS & DEAKIN'S DISCRIMINANT MODEL	35	35	70	20	35	55	-1,457	-1,469	1,565	3,675	-4,059	3,302	0,129	-4,280	-0,179	0,736	0,545	1,282	1,456	0,034	Exc.	Z-Score	Z-Score	Beaver : < 0,13 Deakin : > 0
ZAVGREN'S BPR MODEL	35	35	69	19	34	53	-0,756	3,162	0,119	0,925	-1,593	-0,186	0,864	-0,151	.	.	.	.	.	.	.	Z-Score	I-Score I = 1 / (1+e²)	> 0,50
ZMIJEWSKI'S PROBIT MODEL	35	35	70	20	35	55	-2,065	-2,117	3,618	-0,337	.	.	.	.	.	.	.	.	.	.	.	Z-Score	P-Score P = Nd (Z-Score)	> 0,50
OHLSON' OSCORE MODEL	35	35	70	20	35	55	-4,582	-0,228	7,186	-0,073	0,613	-1,714	3,264	-4,187	0,438	-0,154	.	.	.	.	.	Z-Score	O-Score O = 1/(1+e²)	> 0,50
F-SCORE MODEL - DISCRIMINANT	35	35	70	20	34	54	0,780	2,095	1,146	-0,627	-0,382	-0,813	-0,967	1,126	-0,166	0,093	0,382	.	.	.	.	Z-Score	F-Score (Z-Score)	< 0
F-SCORE MODEL -MRA	35	35	70	20	34	54	0,695	0,525	0,287	-0,157	-0,096	-0,204	-0,242	0,282	-0,042	0,023	0,096	.	.	.	.	Z-Score	F-Score (Z-Score)	< 0,5
F-SCORE MODEL - PROBIT	35	35	70	20	34	54	2,408	3,436	1,389	-1,650	-0,920	-2,783	-1,776	2,542	0,151	-0,317	0,806	.	.	.	.	Z-Score	F-Score F = Nd (Z-Score)	< 0,5
F-SCORE MODEL - BINARY LOGISTIC REGRESSION	35	35	70	20	34	54	4,981	5,006	2,178	-4,481	-0,924	-7,076	-3,165	6,093	0,097	-0,804	2,492	.	.	.	.	Z-Score	F-Score F = 1 / (1+e²)	< 0,50

The numbers in the table represent the coefficients of the predictor variables. The final scores and cut-off points used for reclassification are summarized in the last three columns on the right side. The sample size data have been presented again.

Table IV. Overall Prediction Performances

MODEL NAME	ORIGINAL SAMPLE				SECONDARY SAMPLE				ENTIRE SAMPLE			
	Accuracy t-Value	OVERALL ACCURATE CLASSIFICATION (%)	Critical t-value	ACCURATE PREDICTION AT <0.05	Accuracy t-Value	OVERALL ACCURATE CLASSIFICATION (%)	Critical t-value	ACCURATE PREDICTION AT <0.05	Accuracy t-Value	OVERALL ACCURATE CLASSIFICATION (%)	Critical t-value	ACCURATE PREDICTION AT <0.05
ALTMAN'S Z-SCORE MODEL	4,01	76,30	2,01	Accurate	2,67	69,60	2,02	Accurate	4,76	73,30	1,99	Accurate
ALTMAN'S REVISED Z-SCORE MODEL	5,48	82,90	1,99	Accurate	2,58	67,30	2,01	Accurate	5,82	76,00	1,98	Accurate
ALTMAN'S ZETA MODEL	3,68	76,00	2,01	Accurate	2,01	66,70	2,05	Inaccurate	4,10	72,10	1,99	Accurate
BEAVER'S UNIVARIATE MODEL	5,00	80,00	1,98	Accurate	3,50	72,00	2,00	Accurate	6,00	76,80	1,96	Accurate
DEAKIN'S DISCRIMINANT MODEL	5,72	84,30	2,01	Accurate	2,82	69,10	2,02	Accurate	6,18	77,60	1,97	Accurate
ZAVGREN'S BPR MODEL	4,41	76,80	2,00	Accurate	0,95	56,60	2,01	Inaccurate	3,93	68,00	1,97	Accurate
ZMIJEWSKI'S PROBIT MODEL	5,00	80,00	1,99	Accurate	3,33	70,90	2,01	Accurate	5,82	76,00	1,96	Accurate
OHLSON'S O-SCORE MODEL	6,20	87,20	2,00	Accurate	3,98	75,00	2,00	Accurate	7,07	81,60	1,98	Accurate
F-SCORE MDA MODEL	5,72	84,30	2,01	Accurate	3,07	70,40	2,02	Accurate	6,31	78,20	1,99	Accurate
F-SCORE MRA MODEL	5,95	85,70	2,01	Accurate	3,07	70,40	2,02	Accurate	6,49	79,00	1,99	Accurate
F-SCORE PROBIT MODEL	6,20	87,20	2,01	Accurate	3,26	72,20	2,02	Accurate	6,82	80,60	1,99	Accurate
F-SCORE LOG-REG MODEL	6,42	88,50	2,01	Accurate	3,55	74,10	2,02	Accurate	7,27	82,30	1,99	Accurate

The significance of the models' prediction performances is tested by comparing the accuracy t-values for the original and secondary samples to the critical (table) t-values. A t-value larger than the critical t-value is a statistical proof to conclude that the performance of a model is statistically significant and accurate enough.

The Accuracy of Financial Distress Prediction Models in Turkey: A Comparative Investigation with Simple Model Proposals

45

Table VI. Original Performance Levels

MODEL	OVERALL PERFORMANCE (%)*
ALTMAN'S Z-SCORE MODEL	95
ALTMAN'S REVISED Z-SCORE MODEL	94
ALTMAN'S ZETA MODEL	93
BEAVER'S UNIVARIATE MODEL	87
DEAKIN'S DISCRIMINANT MODEL	97
OHLSON'S O-SCORE MODEL	96
ZAVGREN'S BPR MODEL MODEL	82
ZMIJEWSKI'S PROBIT	98
OHLSON'S O-SCORE MODEL	96

*All models are significant in predicting corporate failure at the significance level of 5 per cent.

Source: Derived from Zmijewski's Paper (1984), Table I, p. 61

Table V. Prediction Errors

MODEL	ORIGINAL SAMPLE		SECONDARY SAMPLE		ENTIRE SAMPLE	
	TYPE I ERROR	TYPE II ERROR	TYPE I ERROR	TYPE II ERROR	TYPE I ERROR	TYPE II ERROR
ALTMAN'S Z-SCORE MODEL	34,50	13,30	15,40	36,40	28,60	25,40
ALTMAN'S REVISED Z-SCORE MODEL	17,10	17,10	25,00	37,10	20,00	27,10
ALTMAN'S ZETA MODEL	30,40	18,50	11,10	44,80	25,00	29,60
BEAVER'S UNIVARIATE MODEL	20,00	20,00	20,00	32,00	20,00	25,70
DEAKIN'S DISCRIMINANT MODEL	17,10	14,30	35,00	28,57	23,60	21,40
ZAVGREN'S BPR MODEL	17,10	29,40	57,90	35,30	31,50	32,30
ZMIJEWSKI'S PROBIT MODEL	20,00	20,00	25,00	31,40	21,90	25,70
OHLSON'S O-SCORE MODEL	5,80	20,00	20,00	28,00	10,90	24,30
F-SCORE MDA MODEL	14,30	17,10	25,00	32,30	18,20	24,30
F-SCORE MRA MODEL	14,30	14,30	25,00	32,30	18,20	22,90
F-SCORE PROBIT MODEL	11,50	14,30	25,00	29,40	16,40	21,70
F-SCORE LOG-REG MODEL	11,50	11,50	25,00	26,50	16,40	18,80

Table VII. Prediction Performance for two to four years prior to failure

MODEL	YEAR 1			YEAR 2			YEAR 3			YEAR 4			STANDARD DEVIATION OF ACCURACY %	COEFF. OF VARIATION	AVERAGE OVERALL ACCURACY RATE %
	TYPE I	TYPE II	OVERALL	TYPE I	TYPE II	OVERALL	TYPE I	TYPE II	OVERALL	TYPE I	TYPE II	OVERALL			
F-SCORE LOG-REG MODEL	83,60	81,20	82,30	90,90	66,66	79,67	83,33	69,56	75,61	76,90	62,70	68,90	5,84	7,62	76,62
OHLSON'S O-SCORE MODEL	89,10	75,70	81,60	65,50	75,70	71,20	64,81	80,00	73,40	60,00	92,53	76,60	4,52	5,96	75,70
ALTMAN'S Z-SCORE MODEL	71,40	74,60	73,30	68,11	76,00	73,07	75,00	81,50	80,65	N/A	N/A	N/A	4,31	5,70	75,67
BEAVER'S UNIVARIATE MODEL	80,00	74,30	76,80	70,90	78,60	75,20	57,40	91,40	76,62	46,00	94,20	73,90	1,36	1,79	75,63
DEAKIN'S DISCRIMINANT MODEL	76,40	78,60	77,60	74,00	84,05	76,60	57,14	82,61	72,95	44,00	91,30	71,43	2,93	3,93	74,65
F-SCORE MRA MODEL	81,80	77,10	79,00	83,63	71,00	76,61	81,50	72,46	76,42	71,15	62,70	66,38	5,61	7,51	74,60
F-SCORE PROBIT MODEL	83,60	78,30	80,60	83,63	66,66	72,58	85,20	66,66	74,80	77,36	62,70	69,74	4,61	6,19	74,43
F-SCORE MDA MODEL	81,80	75,70	78,20	81,80	71,00	75,80	81,50	71,00	75,61	78,00	59,42	67,22	4,81	6,48	74,21
ZMIJEWSKI'S PROBIT MODEL	78,10	74,30	76,00	69,10	77,14	73,60	66,70	84,30	76,00	44,00	84,30	66,94	4,28	5,86	73,14
ALTMAN'S REVISED Z-SCORE MODEL	80,00	72,90	76,00	67,30	74,30	71,20	63,00	81,42	73,40	46,00	80,00	65,83	4,32	6,04	71,61
ALTMAN'S ZETA MODEL	75,00	70,40	72,10	64,50	65,31	65,00	25,00	78,60	71,88	N/A	N/A	N/A	4,04	5,80	69,66
ZAVGREN'S BPR MODEL	68,50	67,70	68,00	61,10	78,30	70,73	39,62	62,06	55,37	26,53	79,41	57,26	7,66	12,18	62,84

* The one-year model functions derived based on the first-year financial data were used in re-classification for two to four years prior to failure. The models are ordered according to their prediction performances in a hierarchical manner.

Table VIII. Comparative Performance Measures

MODEL	n (MODELING)	n1 (MODELING)	n2 (MODELING)	n (SECONDARY)	k	L_{UR}	L_R	Accuracy t-Value	RSS	SIMPLE R^2	EFFRON'S R^2	THEIL'S R^2	HOCKING'S S_P	AMEMIYA'S PC	AKAIKE'S AIC	LR R^2	THEIL'S U
Z-SCORE	59	29	30	46	5	0,99	0,77	4,76	9,50	0,369	0,36	0,18	0,18	11,26	11,64	0,01	1,04
REVISED Z-SCORE	70	35	35	55	5	1,00	0,66	5,82	10,40	0,454	0,41	0,16	0,16	11,00	11,34	0,01	0,95
ZETA	50	23	27	36	7	0,99	0,72	4,10	8,88	0,379	0,29	0,21	0,21	11,77	12,22	0,01	1,15
BEAVER & DEAKIN	70	35	35	55	13	1,00	0,70	6,18	7,85	0,519	0,55	0,14	0,14	11,44	11,72	0,01	0,92
ZAVGREN	69	35	34	53	7	1,00	0,85	3,93	9,53	0,354	0,45	0,15	0,16	11,69	12,02	0,00	1,47
ZMIJEWSKI	70	35	35	55	3	1,00	0,43	5,82	8,24	0,391	0,53	0,12	0,12	8,98	9,24	0,02	0,89
OHLSON	70	35	35	55	9	1,00	0,01	7,07&	6,18**	0,746*	0,65*	0,10**	0,10**	8,00**	8,22**	0,12*	0,82†
F-SCORE (MDA)	70	35	35	55	10	1,00	0,68	6,31	8,98	0,47	0,49	0,15	0,15	11,98	12,30	0,01	0,89
F-SCORE (MRA)	70	35	35	55	10	1,00	0,71	6,49	10,11	0,42	0,42	0,17	0,17	13,49	13,85	0,01	0,89
F-SCORE (PROBIT)	70	35	35	55	10	1,00	0,23	6,82	6,61	0,625	0,62	0,11	0,11	8,81	9,05	0,04	0,87
F-SCORE (LOG-REG)	70	35	35	55	10	1,00	0,03	7,27&	6,23**	0,745*	0,64*	0,10**	0,11**	8,31**	8,53**	0,10*	0,84†

*, & The highest two scores (Best), ** The lowest two scores (Best), n: Total sample size n1: Number of failed firms n2: Number of non-failed firms RSS:
Residual sum of squares k: Number of independent (predicting) variables, † A value closer to zero is better. The best two models according to all the criteria are
the O-Score and Log-Reg F-Score Models. It can be said that all the measures lead us to the same conclusions.

Table IX. Comparison between the F-Score Models with respect to the statistical techniques used

(Posterior Odd Ratios)

MODEL	SAMPLE SIZE	NUMBER OF PREDICTORS (k)	SSE (Sum of squared errors)	MODEL			
				F-SCORE (MDA)	F-SCORE (MRA)	F-SCORE (LOG-REG)	F-SCORE (PROBIT)
F-SCORE (MDA)	70	10	8,98	1,000000			
F-SCORE (MRA)	70	10	10,11	4433,194	1,000000		
F-SCORE (LOG-REG)	70	10	6,23	0,000194	0,000003	1,000000	
F-SCORE (PROBIT)	70	10	6,61	0,001540	0,000024	555,990651	1,000000

A posterior odds ratio larger than one suggests the superiority of a model to other model to which it is being compared. Larger values indicate greater superiority. The values presented in the above table show the extent to which each model in the columns can be considered to be superior to the models in the rows. For example, the value of 555,99 means that the F-Score Log-Reg Model is much superior to the F-Score Probit Model. On the other hand, F-Score MDA Model cannot be regarded as superior to Log-Reg and Probit Models because the values in the intersection cells are much less than one.

Table X. Composition of the Population and Samples

SIC CODE	INDUSTRY	POPULATION		ORIGINAL (MODELING) SAMPLE				SECONDARY (RE-TEST) SAMPLE			
		NUMBER OF FIRMS	%	Failed Firms	Non-Failed Firms	Total	%	Failed Firms	Non-Failed Firms	Total	%
1	Mining	1	0,35	0	0	0	0	0	0	0	0
2	Manufacturing Industry	173	60,25	25	25	50	71,43	15	32	47	83,93
3	Electricity, Gas, and Water	4	1,4	3	3$	6	8,57	1	1	2	3,57
4	Construction and Public Works	2	0,7	0	0	0	0	0	0	0	0
5	Wholesale and Retail Trade, Hotels and Restaurants	16	5,58	3	3	6	8,57	2	1	3	5,36
6	Transportation, Communication and Storage	5	1,75	0	0	0	0	0	0	0	0
7	Education, Health, Sports and Other Social Activities	3	1,05	0	0	0	0	0	0	0	0
8	Financial Institutions	76	26,48	3	3	6	8,57	1	1	2	3,57
9	Technology	7	2,44	1	1	2	2,86	1	1	2	3,57
	TOTAL	287	100	35	35	70	100	20	36	56	100

* Since the bankruptcy environment and financial reporting systems of the banking and financial services sectors are very different and this may bias the results, the firms operating in these sectors were excluded from the sample.

\$ Two of the failed firms in the relevant sector have been re-sampled as non-failed for the years when they were not actually in failure.