Forecasting Financial Distress: A Comparative Study of Statistical and Artificially Intelligent Models with and without Financial Theory

Şaban Çelik[1]   Bora Aktan[2]   Berna Tektaş Sivrikaya[3]

Short Abstract

The framework of predicting financial distress and/or bankruptcy can be categorized into three layers. The first type of models develops and/or proposes a conceptual framework in predicting financial distress and/or bankruptcy[4]. In this case, the conceptual framework determines which constructs (factors) and/or variables are appropriate in predicting financial distress and/or bankruptcy. The second type of models uses statistical methods for differentiating distressed firm from others without relying on a theoretical justification[5]. The third category involves artificially intelligent models[6]. These models resemble to statistical-based models in the sense that they do not rely on a theoretical foundations. In a dissimilar way, these types of models apply different sets of algorithms (neural networks, decision tress, etc.) to classify or differentiate the bankrupt and non-bankrupt firms. Although there are many studies comparing statistical and artificially intelligent models for forecasting financial distress and bankruptcy, there is a need to combine these three types of forecasting methodology within one research setting in order to compare their efficiency and effectiveness with and without financial theory. In the current literature there is no a single study to cover all in one. Therefore, the main distinctiveness of the present study is to forecast financial distress by statistical and artificially intelligent models with and without a financial theory.

Keywords: financial distress, bankruptcy, statistical models, artificially intelligent models, financial theory

---

[1] Assist.Prof.Dr. Izmir Katip Çelebi University, saban.celik@ikc.edu.tr
[2] Assoc.Prof.Dr. University of Bahrain, gbora@uob.edu.bh
[3] Assist.Prof. Dr. Izmir Katip Çelebi University, berna.tektas.sivrikaya@ikc.edu.tr
[4] Gambler's Ruin Theory (Wilcox, 1971), Failing Company Model (Blum, 1974), Balanced Sheet Decomposition Model (Theil, 1969), Cash Management Theory (Laitinen and Laitinen, 1998) and Contingent Claim Model (Merton, 1974) are example of this kind.
[5] Multiple Discriminant Analysis (Altman, 1968; Altman, et.al., 1977; Taffler, 1982; Appetiti, 1984, among others), Logit (Ohlson, 1980; Zavgren, 1985; Keasey and McGuinness, 1990, among others) and Probit Models (Zmijewski, 1984; Gentry, et al., 1985; Skogsvik, 1990, among others) are example of this kind.
[6] Recursive Partitioning Decision Trees (Frydman et al., 1985; Messier and Hansen, 1988; Tam and Kiang, 1992; Pompe and Feelders, 1997; Sung et al., 1999; Beynon and Peel, 2001; Hu and Ansell, 2007, among others), Case Based Reasoning (Jo et al., 1997; Park and Han, 2002; Li and Sun, 2008, among others), Neural Networks (Coats and Fant, 1993; Altman et al., 1994; Zhang, et al., 1999; Lin, 2009, among others), Genetic Algorithms (Varetto, 1998; McKee and Lensberg, 2002; Shin and Lee, 2002; Min and Jeong, 2009, among others) are example of this kind.

EXTENDED ABSTRACT

The evaluation of any model can be judged by several ways. First of all, time dimension is the primary step to judge a model. That is, a model should be effective in the long run. Most of the statistical based models fail in this step. Altman (1968; 1977), the well-know contributor of this field, proposed two models for predicting bankruptcy. These are called Z-Score Model and ZETA Model. Both models contain different variables whereas both models are used for the same purpose. The main reason is that such way of constructing models (not relying on a theoretical framework) is subject to time effect in which the data are collected. The second step is about sample characteristics. When we construct a model depending mainly upon sample characteristics, then it is logical to expect that the model will be needed to modify. This is the case for almost all Statistical based Models and Artificially Intelligent Systems (AIES) based Models. The third step is about the structure of the model. If another construct (factor) or variable is added to the model, then the marginal contribution of the mentioned variable should be negligible. However, it is the case for almost all models in which different variables were used. The fourth step is about how the models reflect financial health of the firms. This requires a deep understanding of financial theory of the firm. Statistical based Models and AIES based Models are all failed in this step. The fifth step is about sector or country specification. Most of the models do contain different set of variables depending upon sector or country. The last but not least, the models should be flexible to reflect life cycle of the firms. This means that all firms are not at the same level of their life. Some may be at growing stage or some may be at mature stage. As a result of all these reasons, there is no superior performance of any of these models as depicted in Table 1-4 (see appendices).

The main purpose of the present study is to use a theoretical model that incorporates the dynamics of the firms with bankruptcy process and estimate it via statistical and AIES based model. In this stage, we follow a recent conceptualization of Celik (2013a) as theoretical framework for predicting financial distress by conducting statistical and AIES based models. At the second stage, we will apply statistical and AIES based models on the same dataset of distressed and non-distressed firms without any theoretical framework. As a result, we will compare these three types of models in terms of efficiency and effectiveness. This research setting is unique in the related literature (Çelik, 2013b).

Researchers have used different sets of variables in predicting bankruptcy. Financial ratios are the oldest and most applied variables in this manner. The early studies used financial ratios extensively. In addition, trend variables, statistical variables and dummy variables are employed to increase efficiency of predictions. We will compare the models with respect to type 1 errors classifying failed firms as non-failed, and type 2 errors classifying non-failed firms as failed in addition to the overall performances of the model.

In this context, the main testable proposition is that how to differentiate distress and non-distress firms with and without the scope of the model. In order to perform the required tests, univariate and

multivariate statistical analyses are conducted. In the context of univariate analysis, parametric and non-parametric independent sample tests are applied depending upon the normality test of the variables. In the context of statistical models, multivariate logistic and probit regressions will be conducted for the purpose of determining the variable that affect the probability of belonging the specified sample. In the context of AIES based models, Neural networks will be executed for the same purpose.

The analysis will be conducted on manufacturing firms listed in Istanbul Stock Exchange (ISE) for the period from 2007 to 2016. The analyses are carried out within the structure of cross-sectional framework due to the nature of prediction. The data including financial statements and their footnotes, stock prices, special reports, annual reports, etc. are derived mainly from the websites of ISE, Public Disclosure Platform (PDP), Capital Markets Board of Turkey and the sample firms.

Evaluation processes of estimated models are carried out at four stages. The first stage gives an examination of overall accuracy of classification, Type I and Type II Error rates. The second stage examines significance of coefficients of the estimated models. The third stage evaluates signs of coefficients of the estimated model with respect to the proposed model. Finally at the last stage, the overall model fit is analyzed.

The main contribution of the present study is to forecast financial distress with and without the framework of a conceptual model that incorporates the firm dynamics with value addition and dilution process of the firms. Therefore, we want to show the efficiency and effectiveness of prediction.

The main research implication of the present study is that predicting financial distress and/or bankruptcy may be biased due mainly to research setting. Therefore, all market players including executives, investors, creditors, auditors and all others may benefit from the findings.

Appendices:

Table 1: Findings based Review

		Number of treatment	OPA (mean)	Number of treatment	Type I Error (%) (mean)	Number of treatment	Type II Error (%) (mean)
Models	Theory Based	13	83.96	11	28.22	11	17.17
	Statistic Based	118	85.73	90	15.66	89	11.99
	AIES Based	50	85.82	23	13.05	13	12.37

Source: Celik (2013a,b)

Note: NA treatments are excluded; OPA: Type I Error (%): classifying failed firms as non-failed; Type II Error (%): classifying non-failed firms as failed; Overall Performance Accuracy; AIES: Artificially Intelligent Expert Systems.

Table 2: Findings based Review on Theory based Models

		Number of treatment	OPA (mean)	Number of treatment	Type I Error (%) (mean)	Number of treatment	Type II Error (%) (mean)
Theory Based Models	Gambler's Ruin	3	95.06	NA	NA	NA	NA
	FCM	1	88.6	1	15	1	10
	BSDM	2	85.1	2	10.5	2	8.5
	CASH	3	68.18	3	22.82	3	33.82
	Contingent Claim	4	85.75	5	41.13	5	12.09

Source: Celik (2013a,b)

Note: NA (not available) treatments are excluded; Type I Error (%): classifying failed firms as non-failed; Type II Error (%): classifying non-failed firms as failed; OPA: Overall Performance Accuracy; AIES: Artificially Intelligent Expert Systems; FCM: Failing Company Model; BSDM: Balanced Sheet Decomposition Model; CMT: Cash Management Theory.

Table 3: Findings based Review on Statistical based Models

		Number of treatment	OPA (mean)	Number of treatment	Type I Error (%) (mean)	Number of treatment	Type II Error (%) (mean)
Statistical Based Models	Univariate	4	79.26	4	25.55	3	22.31
	MDA	50	87.56	38	12.55	38	10.11
	LPM	2	86.33	2	3.28	2	11.11
	LOGIT	47	86.38	35	16.87	35	12.79
	PROBIT	8	87.33	7	15.45	7	10.38
	Cluster	5	65.26	2	36.45	2	24.55
	CUSUM	1	82.5	1	18	1	17
	HAZARD	NA	NA	NA	NA	NA	NA
	ZPP	NA	NA	1	34.76	1	18.84
	QRA	1	88	NA	NA	NA	NA

Source: Celik (2013a,b)

Note: NA (not available) treatments are excluded; Type I Error (%): classifying failed firms as non-failed; Type II Error (%): classifying non-failed firms as failed; OPA: Overall Performance Accuracy; MDA: Multivariate Discriminant Analysis; LPM: Linear Probabilistic Model; CUSUM Par.Adj.: Cumulative Sum Partial Adjustment; ZPP: Zero-Price Probability Model; QRA: (binary) Quantile Regression Approach.

Table 4: Findings based Review on AIES based Models

		Number of treatment	OPA (mean)	Number of treatment	Type I Error (%) (mean)	Number of treatment	Type II Error (%) (mean)
AIES Based Models	RPDT	7	84.6	4	21.3	4	12.4
	NN	20	85.36	8	8.82	8	13.21
	GA	5	85.42	2	15.13	2	7
	CBR	5	87.34	NA	NA	NA	NA
	RS	3	85.4	3	14.17	3	14.8
	PDA	2	81.58	2	13.16	2	23.69
	MCDA	1	99.5	1	0	1	1
	DT	2	85.75	1	4	1	6
	SMO	1	90.24	NA	NA	NA	NA
	DEA	3	88.17	2	20.71	2	8.2
	SOM	1	82.73	NA	NA	NA	NA

Source: Celik (2013a,b)

Note: NA (not available) treatments are excluded; Type I Error (%): classifying failed firms as non-failed; Type II Error (%): classifying non-failed firms as failed; OPA: Overall Performance Accuracy; AIES: Artificially Intelligent Expert Systems; RPDT: Recursive Partitioning Decision Trees; MCDA: Multi-Criteria Decisions Aid; CBR: Case Based Reasoning; NN: Neural Networks; GA: Generic Algorithm; RS: Rough Set; PDA: Preference Disaggregation Analysis; DA: Data Mining; SMO: Sequential Minimal Optimization; DEA: Data Envelop Analysis; SOM: Self-Organizing Map.

REFERENCES

Altman, E.I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*. 23(4): 589-609.

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

Altman, E.I., Marco, G. and Varetto, F. (1994). Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (The Italian Experience). *Journal of Banking and Finance*. 18(3): 505-529.

Appetiti, S. (1984). Identifying Unsound Firms in Italy: An Attempt to Use Trend Variables. *Journal of Banking and Finance*. 8(2): 269-279.

Beynon, M.J. and Peel, M.J. (2001). Variable Precision Rough Set Theory and Data Discretisation: An Application to Corporate Failure Prediction. *OMEGA International Journal of Management Science*. 29(6): 561-576.

Blum, M.P. (1974). The Failing Company Doctrine. *Boston College Law Review*. 16(1-1): 75-113.

Celik, S. (2013a), "*Micro Financial Credit Risk Metrics: A Proposed Model For Bankruptcy And Its Estimation*", Unpublished PhD Thesis, Dokuz Eylul University, Institute of Social Sciences.

Celik, S. (2013b) Micro Credit Risk Metrics: A Comprehensive Review. *Intelligent Systems in Accounting Finance and Management,* 20(4): 233–272.

Coats, P.K. and Fant, L.F. (1993). Recognizing Financial Distress Patterns Using a Neural Network Tool. *Financial Management*. 22(3): 142-155.

Frydman, H., Altman, E.I. and Kao, D.L. (1985). Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress. *The Journal of Finance*, 40(1): 269-291.

Gentry, J.A., Newbold, P.and Whitford, D.T. (1985). Predicting Bankruptcy: If Cash Flow's Not the Bottom Line, What Is? *Financial Analysts Journal*, 41(5): 47-56.

Hu, Y.C. and Ansell, J. (2007). Measuring Retail Company Performance Using Credit Scoring Techniques. *European Journal of Operational Research*. 183(3): 1595–1606.

Jo, H., Han, I. and Lee, H. (1997). Bankruptcy Prediction Using Case-Based Reasoning, Neural Networks, and Discriminant Analysis. *Expert Systems With Applications*, 13(2): 97-108.

Keasey, K. McGuinness, P. and Short, H. (1990). Multilogit Approach to Predicting Corporate Failure Further Analysis and the Issue of Signal Consistency. *OMEGA International Journal of Management Science*. 18(1): 85-94.

Laitinen, E.K. and Laitinen, T. (1998). Cash Management Behavior and Failure Prediction. *Journal of Business Finance & Accounting,* 25(7&8): 893-919.

Li, H. and Sun, J. (2008). Ranking-Order Case-Based Reasoning for Financial Distress Prediction. *Knowledge-Based Systems.* 21(8): 868–878.

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.* 72(16-18): 3507–3516.

McKee, T.E. and Lensberg, T. (2002). Genetic Programming and Rough Sets: A Hybrid Approach to Bankruptcy Classification. *European Journal of Operational Research.* 138(2): 436-451.

Merton, R.C. (1974). Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science.* 4(1): 141-183.

Messier, Jr. W.F. and Hansen, J.V. (1988). Inducing Rules for Expert System Development: An Example Using Default and Bankruptcy Data. *Management Science*, 34(12): 1403-1415.

Min, J.H. and Jeong, C. (2009). A Binary Classification Method for Bankruptcy Prediction. *Expert Systems with Applications.* 36 (3-1): 5256–5263.

Ohlson, J.A. (1980).Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research.* 18(1):109-131.

Park, C.S. and Han, I. (2002). A Case-Based Reasoning with the Feature Weights Derived by Analytic Hierarchy Process for Bankruptcy Prediction. *Expert Systems with Applications.*22(3): 255-264.

Pompe, P.P.M. and Felders, A.J. (1997). Using Machine Learning, Neural Networks, and Statistics to Predict Corporate Bankruptcy. *Microcomputers in Civil Engineering.* 12 (4): 267–276.

Shin, K.S. and Lee, Y.J. (2002). A Genetic Algorithm Application in Bankruptcy Prediction Modeling. *Expert Systems with Applications.* 23(3): 321-328.

Skogsvik, K. (1990). Current Cost Accounting Ratios as Predictors of Business Failure: The Swedish Case. *Journal of Business Finance & Accounting*, 17 (1): 137-160.

Sung, T.K., Chang, N. and Lee, G. (1999). Dynamics of Modeling in Data Mining: Interpretive Approach to Bankruptcy Prediction. *Journal of Management Information Systems.* 16(l): 63-85.

Taffler, R.J. (1982). Forecasting Company Failure in the UK Using Discriminant Analysis and Financial Ratio Data. *Journal of the Royal Statistical Society*. 145(3): 342-358.

Tam, K.Y. and Kiang, M.Y. (1992). Managerial Applications of Neural Networks: The Case of Bank Failure Predictions. *Management Science*, 38(7): 926-947.

Theil, H. (1969). On the Use of Information Theory in Concepts in the Analysis of Financial Statements. *Management Science*. 15(9): 459-480.

Varetto, F. (1998). Genetic Algorithms Applications in the Analysis of Insolvency Risk. *Journal of Banking and Finance*. 22(10-11): 1421-1439.

Wilcox, J.W. (1971). A Simple Theory of Financial Ratios as Predictors of Failure. *Journal of Accounting Research.* 9(2): 389-395.

Zavgren, C.V. (1985). Assessing the Vulnerability to Failure of American Industrial Firms: a Logistic Analysis. *Journal of Business Finance & Accounting*, 12(1): 19-45.

Zhang, G., Hu, M.Y., Patuwo, B.E. and Indro, D.C. (1999). Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-Validation Analysis. *European Journal of Operational Research.* 116(1): 16-32.

Zmijewski, M.E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*. 22(Studies on Current Econometric Issues in Accounting Research): 59-82.