Grammatical Evolution and Corporate Failure Prediction

Anthony Brabazon
Dept. Of Accountancy,
University College Dublin, Ireland.

Michael O'Neill
Dept. Of Computer Science And Information Systems,
University of Limerick, Ireland.

Robin Matthews
Center for International Business Policy
Kingston Business School, London

Conor Ryan
Dept. Of Computer Science And Information Systems,
University of Limerick, Ireland.

Abstract

This study examines the potential of Grammatical Evolution to uncover a series of useful rules which can assist in predicting corporate failure using information drawn from financial statements. A sample of 178 publically quoted, failed and non-failed US firms, drawn from the period 1991 to 2000 are used to train and test the model. The preliminary findings indicate that the methodology has much potential.

1 Introduction

The objective of this study is to determine whether an evolutionary automatic programming methodology, Grammatical Evolution, is capable of uncovering useful structure in financial ratio information which can be used to predict corporate failure.

Corporate failure is an essential component of an efficient market economy, allowing the recycling of financial, human and physical resources into more productive organisations [10] [32]. However, many parties including shareholders, providers of debt finance, employees, suppliers, customers, managers and auditors have an interest in the financial health of organisations as corporate failure can impose significant private costs on all these groups. Even where total failure can be averted by firm reorganization, the costs of major restructuring can be as high as 12% to 19% of firm value [41]. If a trajectory leading to corporate failure can be identified sufficiently early to allow successful intervention, these costs can be reduced. [14] suggest that indicators of corporate failure can be present up to ten years prior to final failure, providing an opportunity for construction of models which predict corporate failure.

Corporate failure can arise for many reasons. It may occur due to a single catastrophic event or it may be the end result of a lengthy process of decline. Under the second perspective, corporate failure is a process which starts with management defects, leading to poor decisions, leading to financial deterioration and finally resulting in corporate collapse [3] [20]. Most attempts to predict corporate failure implicitly assume that management decisions critically impact on firm performance [5] [20]. The premise of this paper is that a series of poor decisions lead to a deterioration in the financial health of the firm and finally to its demise. Although the decisions are not directly observable, their consequent affect on the financial health of the firm can be observed.

Previous studies have utilised a wide variety of explanatory variables in the construction of corporate distress models[1], including variables drawn from the financial statements of firms, from financial markets, general macro-economic variables [29], and non-financial, firm-specific information, including director turnover [27]. In this study, we limit our attention to information drawn from the financial statements of firms.

1.1 Potential for application of evolutionary automatic programming

There are a number of reasons to suppose that the use of an evolutionary automatic programming (EAP) approach such as Genetic Programming (GP) or GE can prove fruitful in the prediction of corporate failure. The problem domain is characterised by a lack of a strong theoretical framework, with many plausible, competing explanatory variables. The selection of quality explanatory variables and model form represents a high-dimensional combinatorial problem, giving rise to potential for an EAP methodology. Use of EAP also facilitates the development of complex fitness functions including discontinuous, non-differentiable functions. This is of particular importance in a prediction domain as fitness criteria may be complex. Generally, the cost of misclassifications of failing / non-failing firms will be asymmetric. Another useful feature of an EAP approach is that it can produce human-readable rules that have the potential to enhance understanding of the problem domain.

[1] [3] and [22] provide good reviews of the development of empirical research in bankruptcy prediction.

1.2 Motivation for study

This study was motivated by a number of factors. Although a substantial volume of research utilising traditional statistical modelling techniques has been undertaken in the corporate failure domain, to date only a limited number of studies have applied GA / GP methodologies [39] [19] [6]. This study builds on these initial studies and adopts a novel evolutionary automatic programming approach.

1.3 Structure of paper

This contribution is organised as follows. Section 2 provides a short discussion of prior literature in the corporate failure domain and outlines the definition of corporate failure employed in this study. Section 3 provides an introduction to Grammatical Evolution. Section 4 describes both the data utilised, and the model development process adopted in this paper. Section 5 provides the results of the constructed model. Finally, conclusions and a discussion of the limitations of the contribution are provided in Section 6.

2 Background

Formal research into the prediction of corporate failure has a long history [12] [35] [16]. Early statistical studies such as [7], adopted a univariate methodology, identifying which accounting ratios had greatest classification accuracy when identifying failing and non-failing firms. Although this approach did demonstrate classification power, it suffers from the shortcoming that a single weak financial ratio may be offset (or exacerbated) by the strength (or weakness) of other financial ratios. [1] addressed this issue by employing a linear discriminant analysis (LDA) model, which utilised both financial and market data concerning a firm, and this was found to improve the classification accuracy of the developed models. The discriminant function which produced the best classification performance in Altman's 1968 study was:

$$Z = .012X_1 + .014X_2 + .033X_3 + .006X_4 + .999X_5$$

where:
X_1 = working capital to total assets
X_2 = retained earnings to total assets
X_3 = earnings before interest and taxes to total assets
X_4 = market value of equity to book value of total debt
X_5 = sales to total assets

LDA assumes both multi-variate normality and the equality of the covariance matrices of each classification group. Generally, these assumptions do not hold for financial ratio data. Other statistical methodologies which have been applied include logit and probit regression models [13] [42] [24]. In recent times, methodologies applied to this problem domain have included neural networks [34] [33] [40], genetic algorithms [39] [19] and hybrid neural network / genetic algorithm models [6].

2.1 Definition of Corporate Failure

No unique definition of corporate failure exists [3]. Possible definitions range from failure to earn an economic rate of return on invested capital given the risk of the business, to legal bankruptcy followed by liquidation of the firm's assets. Any attempt to uniquely define corporate failure is likely to prove problematic. While few publicly quoted companies fail in any given year[2], poorer performers are liable to acquisition by more successful firms. Thus, two firms may show a similar financial trajectory towards failure, but one firm may be acquired and 'turned-around' whilst the other may fail.

The definition of corporate failure adopted in this study is the entry of a firm into Chapter 7 or Chapter 11 of the US Bankruptcy code. The selection of this definition provides an objective benchmark as the occurrence, and date of occurrence, of either of these events can be determined through examination of regulatory filings. Chapter 7 covers corporate liquidations and Chapter 11 covers corporate reorganizations, which usually follow a period of financial distress. Under Chapter 11, management is required to file a reorganisation plan in bankruptcy court and seek approval for this plan. When the court grants approval for the plan the firm is released from Chapter 11 bankruptcy and continues to trade. In most cases, Chapter 11 reorganisations involve significant financial losses for both shareholders [30] and creditors [11] of the distressed firm. [23], in a study of the outcomes of Chapter 11 filings, found that 'there were few sucessful reorganisations' (p. 125), despite a perception that some management teams were using Chapter 11 filings as a deliberate strategy for dealing with certain firm specific events such as onerous labor contracts or produce liability claims[3].

2.2 Explanatory variables utilised in prior literature

Five groupings of explanatory variables, drawn from financial statements, are given prominence in prior literature [4]:

i. Liquidity

ii. Debt

iii. Profitability

iv. Activity / Efficiency

v. Size

[2][42] reports that this rate is less than 0.75% in the US and [22] suggests that the rate is below 2% in the UK.

[3][23] report that out of a sample of 73 firms entering Chapter 11 between 1980 and 1986, only 44 were sucessfully reorganized with only 15 of these firms emerging from Chapter 11 with more than 50% of their prebankruptcy assets.

Liquidity refers to the availability of cash resources to meet short-term cash requirements. Debt measures focus on the relative mix of funding provided by shareholders and lenders. Profitability considers the rate of return generated by a firm, in relation to its size, as measured by sales revenue and/or asset base. Activity measures consider the operational efficiency of the firm in collecting cash, managing stocks and controlling its production or service process. Firm size provides information on both the sales revenue and asset scale of the firm and also provides a proxy metric on firm history. The groupings of potential explanatory variables can be represented by a wide range of individual financial ratios, each with slightly differing information content. The groupings themselves are interconnected, as weak (or strong) financial performance in one area will impact on another. For example, a firm with a high level of debt, may have lower profitability due to high interest costs. Whatever modelling methodology is applied, the initial problem is to select a quality set of model inputs from a wide array of possible financial ratios, and then to combine these ratios using suitable weightings in order to construct a high quality classifier. Given the large search space, an evolutionary automatic programming methodology has promise.

2.3 Results of Prior Literature

Earlier studies [1] [2] [3] [4] [9] [6] have suggested that the classification accuracy of failure models increases rapidly as the date of final failure approaches. Generally, results indicate that the most significant deterioration in financial ratios occurs in the third year prior to eventual failure. Although sample sizes, dates and methodologies differ between studies, these findings have been replicated in a broad series of studies. In Altman's 1968 study [1], the developed LDA model correctly identified (in-sample) 95% of failing firms one year prior to failure. The classification accuracy fell to 72% and 48% in the second and third year prior to failure. [2] demonstrated a classification accuracy (in-sample) of approximately 93% in the year prior to failure declining to 68% four years prior to failure. [40] report in-sample classifications of 98.7% for a neural network model, one-year prior to failure, and 95% for a logit model on the same data. [39] reports in-sample classification accuracy for a GA based model of approximately 97%, one year prior to failure. In-sample classification accuracies provide a limited assessment of model generalisability. Enhanced in-sample classification accuracies could result from data-mining. Hence, in this study, developed models are solely assessed based on classification performance on out-of-sample data.

A description of the evolutionary automatic programming system used to evolve rules for prediction of corporate failure is provided in the next section.

3 Grammatical Evolution

Grammatical Evolution (GE) is an evolutionary algorithm that can evolve computer programs in any language. Rather than representing the programs as syntax trees, as in traditional GP [18], a linear genome representation is adopted. A genotype-phenotype mapping process is used to generate the output program for each individual in the population. Each individual, a variable length binary string, contains in its codons (groups of n-bits, where n equals 8 here) the information to select production rules from a Backus Naur Form (BNF) grammar. The BNF is a plug-in component to the mapping process, that represents the output language in the form of production rules. It is comprised of a set of non-terminals that can be mapped to elements of the set of terminals, according to the production rules.

An example excerpt from a BNF grammar is given below. These productions state that S can be replaced with either one of the non-terminals `expr`, `if-stmt`, or `loop`.

```
S ::= expr      (0)
    | if-stmt   (1)
    | loop      (2)
```

The grammar is used in a generative process to construct a program by applying production rules, selected by the genome, beginning from the start symbol of the grammar.

In order to select a rule in GE, the next codon value on the genome is read and placed in the following formula:

$$Rule = Codon\ Value\ MOD\ \#Rules$$

If the next codon integer value was 4, given that we have 3 rules to select from as in the above example, we get $4\ MOD\ 3 = 1$. S will therefore be replaced with the non-terminal `if-stmt`.

Beginning from the left hand side of the genome codon integer values are generated and used to select rules from the BNF grammar, until one of the following situations arise:

i. A complete program is generated. This occurs when all the non-terminals in the expression being mapped, are transformed into elements from the terminal set of the BNF grammar.

ii. The end of the genome is reached, in which case the *wrapping* operator is invoked. This results in the return of the genome reading frame to the left hand side of the genome once again. The reading of codons will then continue unless an upper threshold representing the maximum number of wrapping events has occurred during this individual's mapping process. This threshold is currently set to ten events.

iii. In the event that a threshold on the number of wrapping events is exceeded and the individual is still incompletely mapped, the mapping process is halted, and the individual assigned the lowest possible fitness value.

GE uses a steady state replacement mechanism, such that, two parents produce two children the best child replacing the worst individual in the current population if the child has a greater fitness. In the case where both children have the same fitness and are better than the current population worst, a child is chosen at random. The standard genetic operators of point mutation, and crossover (one point) are adopted. It also employs a duplication operator that duplicates a random number of codons and inserts these into the penultimate codon position on the genome. A full description of GE can be found in [26] [25] [31].

4 Problem Domain & Experimental Approach

This section describes both the data utilised by, and the model development process adopted in, this study.

4.1 Sample Definition and Model Data

A total of 178 firms were selected judgementally (89 failed, 89 non-failed), from the Compustat Database [8][4]. The criteria for selection of the failed firms were:

i. Inclusion in the Compustat database in the period 1991-2000

ii. Existence of required data for a period of three years prior to entry into Chapter 7 or Chapter 11

iii. Sales revenues must exceed $1M

The first criterion limits the study to publicly quoted, US corporations. The second criterion injects an element of bias into the sample in that companies without a three year financial history prior to entering Chapter 7 or Chapter 11 are omitted. Twenty-two potential explanatory variables, are collected for each firm for the three years prior to entry into Chapter 7 or Chapter 11[5]. For every failing firm, a matched non-failing firm is selected. They are matched both by industry sector and size (sales revenue three years prior to failure)[6]. The set of 178 matched firms are randomly divided into model building (128 firms) and out-of-sample (50 firms) datasets. The dependant variable is binary (0,1), representing either a non-failed or a failed firm.

[4]Firms from the financial sector were excluded on grounds of lack of comparability of their financial ratios with other firms in the sample.

[5]The date of entry into Chapter 7 or Chapter 11 was determined by examining regulatory filings for each firm.

[6]It is recognised that the use of an equalised, matched sample entails sampling bias and eliminates firm size and industry nature as potential explanatory variables (see [22] for a detailed discussion of these points). It is noted that utilising an unmatched sample imposes its own bias.

The choice of explanatory variables is hindered by the lack of a clear theoretical framework which explains corporate failure [5] [38] [40]. Most empirical work on corporate failure adopts an ad-hoc approach to variable selection. Prior to the selection of the potential explanatory variables for inclusion in this study, a total of ten previous studies were examined [7] [1] [2] [9] [24] [33] [17] [6] [37] [21]. These studies employed a total of 58 distinct ratios divided amongst the five classifications noted by [4]. A subset of 22 of the most commonly used financial ratios was selected for this study. The selected ratios were:

i. EBIT / Sales

ii. EBITDA / Sales

iii. EBIT / Total Assets

iv. Gross Profit / Sales

v. Net Income / Total Assets

vi. Net Income / Sales

vii. Return on Assets

viii. Return on Equity

ix. Return on Investment

x. Cash / Sales

xi. Sales / Total Assets

xii. Inventory / Cost of Goods Sold

xiii. Inventory / Working Capital

xiv. Fixed Assets / Total Assets

xv. Retained Earnings / Total Assets

xvi. Cash from Operators / Sales

xvii. Cash from Operations / Total Liabilities

xviii. Working Capital / Total Assets

xix. Quick Assets / Total Assets

xx. Total Liabilities / Total Assets

xxi. Leverage

xxii. EBIT / Interest

5 Results

Accuracy of the developed models is assessed based on the overall classification accuracy arising in both the model-building and out-of-sample datasets. For simplicity, the cost of each type of classification error is assumed to be symmetric in this study. The fitness function could be easily altered to bias the model development process to minimise a specific type of classification error if required, and later studies will address this issue.

The classification problem which plays an important role in decision-making, consists of assigning observations to disjoint groups [28]. The decision scenario faced in this study comprises a binary classification. In general, the construction of classifier systems such as linear discriminant analysis, logit or ANN models consists of two components, the determination of a valuation rule which is applied to each observation, and the determination of a 'cut-off' value. The grammar adopted in this study is given below and its output is interpreted using a fixed 0.5 cut-off value to produce a classification.

```
lc : output = expr ;
expr : ( expr ) + ( expr )
     | coeff * var
var : var1[index]   | var2[index]   | var3[index]
    | var4[index]   | var5[index]   | var6[index]
    | var7[index]   | var8[index]   | var9[index]
    | var10[index]  | var11[index]  | var12[index]
    | var13[index]  | var14[index]  | var15[index]
    | var16[index]  | var17[index]  | var18[index]
    | var19[index]  | var20[index]  | var21[index]
    | var22[index]
coeff :  ( coeff ) op ( coeff )
       | float
op : + | - | *
float : 20 | -20 | 10 | -10 | 5 | -5 | 4 | -4
       | 3 | -3 | 2 | -2 | 1 | -1 | .1 | -.1
```

The above grammar generates classifiers of the form:

$$output = (< some\ expression > *varX)\ +\ ..$$
$$+\ (< some\ expression > *varY)$$

Any combination of the twenty two explanatory variables can be exploited by an evolved classifier, including zero or more occurrences of any one variable. This is in contrast to an LDA approach where classifiers would generally utilise all the explanatory variables within the expression. In the LDA case, of course some of those variables could be *switched off* by multiplying their value by zero.

Three series of models were constructed using explanatory variables drawn from one, two and three years (T1, T2 and T3) prior to failure. For each set of models, 30 runs were conducted using population sizes of 500, running for 100 generations, adopting one-point crossover at a probability of 0.9, and bit mutation at 0.01, along with a steady state replacement strategy.

A plot of the mean average and mean best fitness values over the 30 runs for each time period can be seen in Figure 1.

Years Prior to Failure	In Sample	Out Of Sample
1	85.9%	80%
2	82.8%	80%
3	75.8%	70%

Table 1: The best classification accuracies reported for each of the three years prior to failure.

The best individuals evolved for each period are reported in Table[7] 1. In-sample it can be seen that the performance of the models generated falls off gracefully as we move out each year. It is interesting to note that out-of-sample there is no performance difference between the evolved models in periods T1 and T2, both giving 80% correct classifications.

The best classifiers evolved for each period are given in Table 2.

5.1 Discussion

The classification results of the evolved models show promise. Despite drawing a sample from a wide variety of industrial sectors, the models demonstrate a high classification accuracy in and out-of-sample, which degrades gracefully rather than suddenly in the third year prior to failure. Although the evolved models were free to select from twenty-two potential explanatory variables, it is notable that each model only employed a small subset of these. This lends support to the proposition that many financial ratios have similar information content and that classification accuracy is not enhanced through the construction of models with a large number of these ratios. It is also notable that each model has (approximately) included one variable drawn from the four main categories of explanatory variables suggested in the corporate failure literature (Liquidity, Debt, Profitability, and Activity/Efficiency), lending empirical support to earlier work[8].

The risk factors suggested by each model differ somewhat and contain some less-intuitive but nonetheless plausible findings.

Examining the best classifer evolved for T1 suggests that risk factors include low return on assets, low retained earnings and a high ratio of total liabilities to total assets, which concords with financial intuition. Less obviously, a high ratio of inventory to net liquid assets (inventory+receivables+cash-payables) is also a risk factor, possibly resulting from depletion of cash or build-up of inventories as failure approaches.

Risk factors for firms at T2 include low return on assets and a low ratio of earnings to interest costs. Less intuitive risk factors indicated are a low ratio of fixed assets to total assets and a high ratio of sales to total assets. The former could indicate firms with a lower safety cushion of saleable resources which could be sold to stave-off collapse, the latter could be serving as a proxy variable for firms with rapid sales growth. Over-rapid sales growth can be a danger signal, indicating that management resources are being spread too-thinly.

Finally, risk factors indicated for firms at T3 include low return on assets, a low ratio of profit to interest charge, a low level of cash generated from operations and as for T2, a high ratio of sales to total assets.

Although each model is evolved seperately, the general form of each model appears consistent with the hypothesis that

[7] Calculation of Press's Q statistic [15] for each of these models rejects a null hypothesis, at the 5% level, that the out-of-sample classification accuracies are not significantly better than chance.

[8] Size, the fifth category, is not considered in this study due to the matching process utilized.

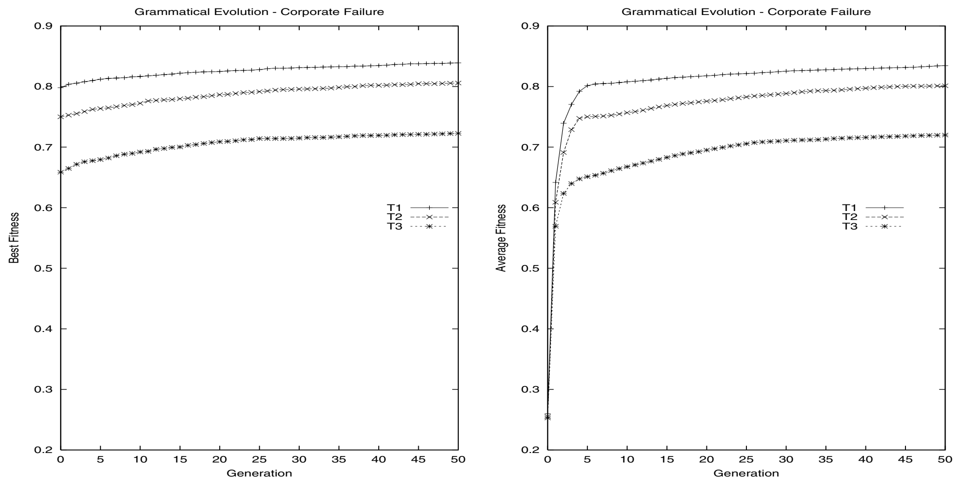

Figure 1: A comparison of the mean best fitness between T1, T2, and T3 (left), and of the mean average fitness values (right) for the same time periods.

Years Prior to Failure	Best Classifier
1	output = -3*var7-5*var8+3*var17-20*var19+4*var24
2	output = -2*var8+10*var15-10*var18-2*var25
3	output = -4*var8+20*var15-72.9*var20-10*var25

Table 2: The best classifiers evolved for each of the years analysed.

there is a financial trajectory towards failure. Low profits and high interest payments as a percentage of profits in periods T3 and T2 indicate a firm in financial difficulties, with an erosion of the safety cushion provided by high levels of (saleable) fixed assets indicated in the risk factors at T2. The final year prior to failure sees additional risk factors indicated by high levels of debt and reducing cash balances / inventory build-up.

6 Conclusions & Future Work

GE was shown to successfully evolve useful rules for prediction of corporate failure with a performance equivalent to that reported in prior studies. In assessing the performance of the developed models, a number of caveats must be borne in mind. The premise underlying this paper (and all empirical work on corporate failure prediction) is that corporate failure is a process, commencing with poor management decisions, and that the trajectory of this process can be tracked using accounting ratios. This approach does have inherent limitations. It will not forecast corporate failure which results from a sudden environmental event. It is also likely that the explanatory variables utilised contain noise. Commentators [5] [36] have noted that managers may attempt to utilise creative accounting practices to manage earnings and / or disguise signs of distress. Additionally, financial data is produced on a time-lagged basis. Although not undertaken in this preliminary study, the incorporation of non-financial qualitative explanatory variables or variables related to the firm's share price performance could further improve classification accuracy. Another limitation of all models of corporate distress is that the underlying relationships may not be stationary [4] [17]. Accounting standards and economic environment faced by firms will vary over time. Finally, the firms sampled in this study are relatively large and are publically quoted. Thus, the findings of this study may not extend to small businesses.

Despite these limitations, the high economic and social costs of corporate failure imply that models which can indicate declining financial health will have utility. Given the lack of a clear theory underlying corporate failure, empirical modelling usually adopts a combinatorial approach, a task for which GE is well suited. The results of this preliminary study indicate that GE has useful potential for the construction of corporate failure models.

References

[1] Altman, E. (1968). 'Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy', *Journal of Finance*, 23:589-609.

[2] Altman, E., Haldeman, R. and Narayanan, P. (1977). 'ZETA Analysis: A new model to identify bankruptcy risk of corporations', *Journal of Banking and Finance*, 29-54.

[3] Altman, E. (1993). *Corporate Financial Distress and Bankruptcy*, New York: John Wiley and Sons Inc.

[4] Altman, E. (2000). 'Predicting Financial Distress of Companies: Revisiting the Z-score and Zeta models', *http://www.stern.nyu.edu/ ealtman/Zscores.pdf, October 2001*.

[5] Argenti, J. (1976). *Corporate Collapse: The Causes and Symptoms*, London: McGraw-Hill.

[6] Back, B., Laitinen, T., Sere, K. and van Wezel, M. (1996). 'Chosing Bankruptcy Predictors using Discriminant Analysis, Logit Analysis and Genetic Algorithms', *Technical Report no. 40, Turku Centre for Computer Science, Turku School of Economics and Business Administration.*.

[7] Beaver, W. (1968). 'Financial Ratios as Predictors of Failure', *Journal of Accounting Research - Supplement: Empirical Research in Accounting*, 71-102.

[8] Compustat Database. http://www.compustat.com

[9] Dambolena, I. and Khoury, S. (1980). 'Ratio Stability and Corporate Failure', *Journal of Finance*, 35(4):1017-1026.

[10] Easterbrook, F. (1990). 'Is Corporate Bankruptcy Efficient?', *Journal of Financial Economics*, 27:411-417.

[11] Ferris, S., Jayaraman, N. and Makhija, A. (1996). 'The Impact of Chapter 11 filings on the Risk and Return of Security Holders, 1979-1989', *Advances in Financial Economics*, 2:93-118.

[12] Fitzpatrick, P. (1932). *A Comparison of the Ratios of Successful Industrial Enterprises with Those of Failed Companies*, Washington: The Accountants' Publishing Company.

[13] Gentry, J., Newbold, P. and Whitford, D. (1985). 'Classifying Bankrupt Firms with Funds Flow Components', *Journal of Accounting Research*, 23(1):146-160.

[14] Hambrick, D. and D'Aveni, R. (1988). 'Large Corporate Failures As Downward Spirals', *Administrative Science Quarterly*, 33:1-23.

[15] Hair, J., Anderson, R., Tatham, R. and Black, W. (1998). *Multivariate Data Analysis*, Upper Saddle River, New Jersey: Prentice Hall.

[16] Horrigan, J. (1965). 'Some empirical bases of financial ratio analysis', *The Accounting Review*, July 1965, 558-568.

[17] Kahya, E. and Theodossiou, P. (1996). 'Predicting Corporate Financial Distress: A Time-Series CUSUM Methodology', *Review of Quantitative Finance and Accounting*, 13:71-93.

[18] Koza, J. (1992). *Genetic Programming*. MIT Press.

[19] Kumar, N., Krovi, R. and Rajagopalan, B. (1997). 'Financial decision support with hybrid genetic and neural based modelling tools', *European Journal of Operational Research*, 103:339-349.

[20] McRobert, A. and Hoffman, R. (1997). *Corporate Collapse: An early warning system for lenders, investors and suppliers*, Roseville:NSW, McGraw-Hill (Australia).

[21] Moody's (2000). 'RiskCalc For Private Companies: Moody's Default Model', *http://riskcalc.moodysrms.com/us/research/crm/56402.pdf*, July 2001.

[22] Morris, R. (1997). *Early Warning Indicators of Corporate Failure: A critical review of previous research and further empirical evidence*, London: Ashgate Publishing Limited.

[23] Moulton, W. and Thomas, H. (1993). 'Bankruptcy As a Deliberate Strategy: Theoretical Considerations and Empirical Evidence', *Strategic Management Journal*, 14:125-135.

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

[25] O'Neill, M. (2001). Automatic Programming in an Arbitrary Language: Evolving Programs in Grammatical Evolution. PhD thesis, University of Limerick, 2001.

[26] O'Neill M., Ryan C. (2001) Grammatical Evolution, *IEEE Trans. Evolutionary Computation*. 2001.

[27] Peel, M., Peel, D. and Pope, P. (1986). 'Predicting Corporate Failure: Some Results for the UK Corporate Sector', *Omega International Journal of Management Science*, 14:5-12.

[28] Pendharkar, P. (2001). An empirical study of design and testing of hybrid evolutionary-neural approach for classification, *Omega*, 29:361-374.

[29] Rose, P., Andrews, W.and Giroux, G. (1982). 'Predicting Business Failure: A Macroeconomic Perspective', *Journal of Accounting, Auditing and Finance*, Fall:20-31.

[30] Russel, P., Branch, B. and Torbey, V. (1999). 'Market Valuation of Bankrupt Firms: is there an anomaly?', *Quarterly Journal of Business and Economics*, 38:55-76.

[31] Ryan C., Collins J.J., O'Neill M. (1998). Grammatical Evolution: Evolving Programs for an Arbitrary Language. *Lecture Notes in Computer Science 1391, Proceedings of the First European Workshop on Genetic Programming*, 83-95, Springer-Verlag.

[32] Schumpeter, J. (1934). *The Theory of Economic Development*, Cambridge, MA: Harvard Business Press.

[33] Serrano-Cina, C. (1996). 'Self organizing neural networks for financial diagnosis', *Decision Support Systems*, 17:227-238.

[34] Shah, J. and Murtaza, M. (2000). 'A Neural Network Based Clustering Procedure for Bankruptcy Prediction', *American Business Review*, 18(2):80-86.

[35] Smith, R. and Winakor, A. (1935). 'Changes in the Financial Structure of Unsuccessful Corporations', *University of Illinois, Bureau of Business Research, Bulletin No. 51*.

[36] Smith, T. (1992). *Accounting for Growth*. London: Century Business.

[37] Sung, T., Chang, N. and Lee, G. (1999). 'Dynamics of Modelling in Data Mining: Interpretative Approach to Bankruptcy Prediction', *Journal of Management Information Systems*, 16(1):63-85.

[38] Trigueiros, D. and Taffler R. J. (1996). 'Neural Networks and Empirical Research in Accounting', *Accounting and Business Research*, 26(4):347 - 355.

[39] Varetto, F. (1998). 'Genetic algorithms in the analysis of insolvency risk', *Journal of Banking and Finance*, 22(10):1421 - 1439.

[40] Wilson, N., Chong, K. and Peel, M. (1995). 'Neural Network Simulation and the Prediction of Corporate Outcomes: Some Empirical Findings', *International Journal of the Economics of Business*, 2(1):31-50.

[41] Wruck, K. (1990). 'Financial Distress, Reorganization and Organizational Efficiency', *Journal of Financial Economics*, 27(2):419-444.

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