An Empirical Evaluation of Sampling Controversies in Ratio-Based Modelling of Corporate Collapse

Dr Ghassan Hossari
Professor Sheikh F. Rahman
and
Professor Janek Ratnatunga

Empirical investigations regarding ratio-based modelling of corporate collapse have been on going for decades. With any study of an empirical nature, a data sample is a necessary prerequisite. It allows testing the performance of the prediction model, thereby establishing its practical relevance. However, it is necessary to first ensure that the data sample used satisfies certain conditions, and these have lead to some choice controversies. This paper considers the controversial issues that arise in data sampling, provides a critical evaluation of these issues, and makes choice recommendations on the controversial aspects, by empirically examining the literature.

1. Introduction

Empirical approaches to signalling corporate collapse require the derivation of a statistical model that relies on financial ratios as predictors of collapse. The ratios are calculated from financial statements of a sample of companies. The data sample must satisfy certain conditions. However, a number of choice controversies have arisen in the literature in relation to these conditions. Data sampling controversies have arisen with regards to the selection of companies. Controversies have also arisen in relation to choice issues that are needed to satisfy certain statistical requirements of the methodological approach used, such as sample size and the inclusion (or not) of outliers (extreme values). Issues pertaining to the enhancement of empirical results have also had choice controversies arising in theory, especially pertaining to the reporting period, testing and validation procedures.

This paper provides resolutions to such controversies by examining the empirical evidence obtained from an extensive study of the published literature in the area. Explanations as to why a particular approach is being favoured in the modelling literature are also provided.

2. Data sampling issues REGARDING the selection of companies

The success of any model designed to signal corporate collapse is its ability to predict the event with a high degree of accuracy. The predictive ability of the model could be determined by testing it on a sample of companies. Ideally, the sample should be selected so that it contains both collapsed and non-collapsed companies, so that the model can separate (discriminate) between them. With regards to data sampling controversies in relation to the selection of such companies, of primary concern in the literature has been whether collapsed and non-collapsed companies should be paired, and if so what matching criteria should be used.

2.1 Paired sample design

Let us examine this particular choice controversy by first considering a sample where pairing does not take place. This could result in a sample of collapsed companies only. Such a sample is inadequate even if the model predicts the collapse of all the companies in it. This is because we do not know how the model performs when it comes to non-collapsed companies. Does it predict their survival with the same level of accuracy? The same argument applies to a sample of non-collapsed companies only, in which case the model predicts the survival of all the companies in it. Thus, a model that might perform well when it comes to signalling collapse, but may not necessarily be a good predictor of survival, and vice-versa.

Whilst a good model is one that signals both events with relatively high levels of accuracy, it could be argued that it is more costly to declare a company on the verge of collapse as healthy, than it is to declare a healthy company as being on the verge of collapse (Altman, 1968). Declaring a healthy company as being on the verge of collapse could in the worst case create uneasiness among the stakeholders of the company. However, they would incur no real financial losses. Such a prediction mistake is referred to as a Type II error (Altman, 1968). On the other hand, declaring a company on the verge of collapse as healthy would have wide-reaching repercussions. Stakeholders, who believe this, might end up paying dearly in terms of financial losses. Such a prediction mistake is referred to as a Type I error (Altman, 1968).

A good corporate collapse prediction model should ideally have a low occurrence of both Type I and Type II errors. However, if one has to choose between the two, it is preferable to maintain a lower occurrence of Type I relative to Type II errors (Altman, 1968).

Having established that a sample should contain both collapsed and non-collapsed companies, let us return to our initial choice controversy: i.e., should these companies be paired? A problem with pairing arises due to what is statistically known as *a-priori* probabilities. This refers to the relative occurrence of collapse in a population (Eisenbeis, 1977). In Australia, *a-priori* probabilities are around 0.33 percent per year among publicly listed companies. This is equivalent to 33 collapses per 10,000 companies.

Therefore, if one is to test the ability of a ratio-based model in signalling the collapse of a sample of 33 companies (all having collapsed), then 10,000 non-collapsed companies are required in order to maintain an *a-priori* probability of 0.33 percent collapses per year. This causes numerous sample size problems as discussed below.

First, there might not be 10,000 publicly listed companies to include in the sample. The number of publicly listed companies in Australia for example, is currently under 2,000. This is considerably below the 10,000 that would be required.

Second, even if such large numbers of publicly listed companies are available, it is not practical to include them all in the sample for two reasons: cost of obtaining data and statistical processing capability. A cost-benefit approach to this choice controversy is to weigh the costs of gathering and processing a large data set against the benefits gained in model performance[1]. For this we will need to examine the relative performance of models in studies that used larger vs. smaller samples. If a model based on a larger sample predicts collapse better, then there is a clear benefit that would outweigh the costs. This sample-size comparison is done in section 5 of this paper, where the empirical data is discussed.

As the number of companies that collapse may seem to be relatively small[2] when compared to the number of companies that continue to be financially viable, a question arises as to the economic usefulness of undertaking such modelling. However, the relatively small number of collapses per year does not necessarily underestimate the seriousness of the problem. After all, the fact that only a small percentage of babies suffer from cot death[3] has not precluded researchers from trying to determine its causes. Likewise, the seemingly small percentage of companies that collapse has not deterred researchers from developing models that signal corporate collapse before it happens. In fact, refereed empirical studies can be traced as far back as Winakor (1929) and extend to the immediate present with Hossari (2006a).

Returning to the question of what number of non-collapsed companies should there be for every collapsed company, if a-priori probabilities are disregarded then Meyer and Pifer (1970, p. 866) have argued that a ratio of 50:50 minimizes classification error. The reason is primarily statistical: whenever a screening procedure such as signalling collapse is used, the conditional probabilities generated by the prediction model may be inappropriate if the number of companies in the two groups are not equal. We will consider the evidence from the empirical analysis and make a choice recommendation in section 5 to determine if this pairing equality should be used or not.

2.2 Matching criteria in a paired sample design

Assuming that maintaining a pairing ratio is desirable when modelling corporate collapse, does this imply that companies in the sample should be paired based on some matching criteria, such as sales, assets, etc.? This question could be answered by considering whether the inclusion of matching criteria would add value or avoid potential problems in the data sample. If either is achieved, then collapsed companies should be paired with their non-collapsed counterparts in the ratio chosen, based on certain matching criteria.

Generally speaking, the matching principle surrounds us. For instance, in competitive sports, matching is used extensively. For example, a heavy weight boxer is matched with a heavy weight opponent (i.e. a 50:50 ratio), or a chess grandmaster is paired with (say) 100 amateurs in a demonstration competition (i.e. 1:100 ratio). Paired samples could be regarded in the same fashion.

When it comes to matching criteria in corporate collapse models, however, some have strongly argued *against* their necessity. For example, Ohlson (1980, p. 112) stated that there are certain problems associated with matching procedures that have typically been used in Multiple Discriminant Analysis (MDA). The criticism is that collapsed and non-collapsed companies are usually matched according to criteria such as size and industry, which due to reporting manipulations tend to be somehow arbitrary. Others appeared to be unsure whether or not matching was required. For instance, Taffler (1982) did not apply any matching criteria to a sample of collapsed and non-collapsed firms; whereas Taffler (1983) matched a paired sample based on industry classification and the size of assets.

Conversely, there is even stronger support *for* criteria-based matching in modelling theory, especially industry-based matching. Bird and McHugh (1977) argue that industry classification is a necessary matching criterion as financial ratios differ within industry sectors. A financial ratio such as 'Net Income/Total Assets' (NI/TA) for a healthy company in the telecommunication services industry differs significantly from that for a healthy company in the consumer staples industry. In this case, both companies are healthy (non-collapsed), however their NI/TA ratios differ considerably, and a statistical model for signalling collapse might interpret the difference as an indication that one of the companies is not healthy (i.e. collapsed). This leads to a wrong classification by the model. Thus, the supporters of industry-based matching argue that by adopting such criteria, any significant differences in the financial ratios of a pair of matched companies could be more likely attributed to variations in financial fragility, and that this in turn leads to more accurate predictions.

Assets can also be used as a matching criterion, as they provide a measure of possession.[4] No other items in the financial statements of a company describe possession (i.e. what we have) as 'Assets' do. Matching companies based on the size of their assets is like comparing wealthy individuals, e.g. Bill Gates to Warren Buffet. Liabilities can also be used as criteria, similar to how banks evaluate credit-worthiness based on one's liabilities. Also one can argue that the level of liabilities is an indicator of the riskiness of the company. Equity, the last major item on the balance sheet, can also be used, although here the book value of equity could differ *considerably* from its market value.

Items in the Income (Profit and Loss) Statement also could be considered for matching. For example, sales or net income could provide a measure of performance for matching. It could be argued that net income is a more accurate measure of performance than sales because not all companies that sell make a profit. Companies could be generating lots of sales, but at the same time making losses. However, sales have an advantage as a matching criterion in that it is as close to an unadjusted figure as we can get. There is a lot of accounting interpretation between sales (the first item on the Income Statement) and net income/loss (the last item on the Income Statement) that may distort the model's predictability.

In summarising the sample selection controversies with regards to pairing, it appears that matching pairs of companies (on some pairing-ratio) by industry classification is almost mandatory (Gentry et al., 1985; Platt et al., 1994) and further matching according to some balance sheet or profit and loss criteria, is desirable. Such matching provides measurable compatibility between each pair of companies. Whether to use assets, sales or any financial statement criteria is not as important as maintaining consistency (Lau, 1987; McGurr and Devaney, 1998). In Section 5 we will provide recommendations of which matching criteria are most used in the empirical analysis undertaken, and provide empirical justification for the choices made.

3. Data sampling issues necessary to satisfy certain statistical requirements

This section looks at issues necessary to satisfy certain statistical requirements of the methodological approach used. It considers how large the sample should be for proper model derivation and whether outliers among the financial ratios should be removed.

3.1 The optimal sample size

How large should the sample be for proper derivation of the corporate collapse prediction model? The answer depends on which methodological approach is used in deriving the model.

The year 1968 saw a major methodological shift in ratio-based modelling of corporate collapse mainly due to the pioneering work of (Altman, 1968). Up until 1968, signalling collapse was achieved by considering one ratio at a time (Beaver, 1966; Beaver, 1968a; Beaver, 1968b; Tamari, 1966). This gave rise to some serious problems, the most prevalent being the contradictory interpretation of results generated by the various *univariate* ratios.

The solution was to simultaneously consider a multitude of ratios when assessing the viability of a company (Altman, 1968), using a new statistical tool (at the time) called Multiple Discriminant Analysis (MDA) (Huber, 1964; Lachenbruch, 1967; Walter, 1959). Since the pioneering work of Altman (1968), the introduction of new multivariate statistical techniques have led to shifts (or some may argue advances) in the methodological approaches to modelling corporate collapse. These new approaches include Logit analysis (Ohlson, 1980), Neural Network analysis (Coats and Fant, 1993), Probit analysis (Gentry et al., 1985), ID3 (Kim and McLeod Jr., 1999), Recursive Partitioning Algorithm (Frydman et al., 1985), Rough Sets analysis (Dimitras et al., 1999), Decomposition analysis (Walker et al., 1979), Going Concern Advisor (Lenard et al., 1998), Koundinya and Puri judgmental approach (Clark et al., 1997), Tabu Search (Drezner et al., 2001), Mixed Logit analysis (Jones and Hensher, 2004) and Multi-Level Modelling (Hossari, 2006b).

Despite these methodological shifts, MDA still remains the dominant methodology. Its dominance is evident in two ways: first, being the primary approach in the early state of the literature (which has by and large stood the test of time); and second, being the dominant benchmark against which to compare new approaches. Therefore, this section of the paper discusses the optimal sample size in the light of MDA.

There is much controversy surrounding the question of sample size for the proper derivation of an MDA-based corporate collapse prediction models. One suggestion is that the sample size should be at least 3 times the number of discriminating variables (i.e. the financial ratios used in the model for signalling collapse) (Foley, 1972). However, there are counter arguments that such a sample size could be too small (Hecker and Wegener, 1978). One of the recommendations is to have 5 to 10 times as many samples (companies) as measurements (ratios) (Jain and Chandrasekaran, 1982, p. 852). On the conservative side, it is proposed that the smallest group should have a sample size that is at least 10 times the number of discriminating variables (Huberty, 1994). Another proposal is that the sample should exceed the number of discriminating variables by more than 2 (Klecka, 1982). In section 5 we will provide recommendations pertaining to sample size, and provide empirical justification for the choices made.

3.2 The proper treatment of outliers

From a purely technical statistical perspective it is advisable to eliminate outliers or extreme values because outliers may affect the performance of the statistical model (Aitkin and Wilson, 1980; Butler, 1986; Campbell, 1978; Collins, 1976; Collins, 1982; Collins and Portnoy, 1981; Collins and Wiens, 1985; Hampel et al., 1981; Huber, 1964; Johnson and Geisser, 1983; Maronna, 1976; McLachlan, 1992).

A counter argument is that it is not logical to remove outliers in the context of ratio-based modelling of corporate collapse, as any study in this area essentially focuses on outliers and abnormalities in the financial ratios of companies (Altman, 1973; Charitou et al., 2004; Laitinen, 1991; Zmijewski, 1984). Companies that are in financial trouble are expected to have ratios that are out-of-line from those of healthy firms. The elimination of outliers may lead to the removal of either the collapsed or the non-collapsed companies from the data set, which defies the objective of the whole exercise of signalling corporate collapse. Therefore, in the context of modelling corporate collapse, there is controversy as to if outliers should (or should not) be removed. In section 5 we will provide recommendations as to the treatment of outliers, and provide empirical justification for the choices made.

4. Data sampling issues for enhancing the empirical results

This section focuses on sampling issues that are desirable to enhance the empirical results, especially the appropriate length of the sample period; whether the results should be reported for single or multiple periods; and model validation issues.

4.1 The optimal sample period

How long should the sample period be? Is there a minimum requirement? If there is, what is it and what does it depend on? In answering these questions, the objective is to enhance the empirical results.

The answers to these choice controversies are contingent on the optimal sample size; and it was discussed earlier that the most stringent requirement found in modelling theory is that the number of companies in the data sample should be 10 times the number of financial ratios used in the corporate collapse prediction model (Huberty, 1994). This means that, if an MDA-based corporate collapse prediction model relies on 3 financial ratios as predictors of collapse - and if the above stringent requirements of sample size are to be adhered to - then the number of companies in the sample should be at least 30, divided equally (if a 50:50 pairing ratio is used) between collapsed (15) and non-collapsed (15) companies. If the number of collapsed companies in the sample should be at least 15, and given that the rate of collapse in Australia is around 5 companies per year; then a sample period of at least 3 years is necessary. This would be the minimum sample period for Australia. Other countries with larger numbers of listed companies could reduce the sample period. In section 5 we will provide recommendations as to the optimal sample period, and provide empirical justification for the choices made.

4.2 The optimal reporting period

The controversy here has to do with how far back a model should go in signalling corporate collapse. Statistically speaking, the reporting period is immaterial, because it has no impact on model derivation. Data sampling issues necessary to satisfy certain statistical requirements have already been discussed and they primarily relate to sample size and treatment of outliers. What influences an optimal reporting period though, is the *context* in which statistical tools – such as MDA - are applied (Karels and Prakash, 1987).

Although the incident of collapse is sudden, the process that leads to it is gradual and could span many years. Along the way signs and symptoms manifest themselves in the reported financial variables captured in the financial reports, from which financial ratios are calculated. Because the symptoms are time-variant, it is argued that the model should take the time factor into consideration, which necessitates testing the predictive ability of the model over some time period *prior* to the incident of collapse. If the reporting period were too short, then it would be too late to take corrective action and try to turn the company around, if at all possible. Likewise, if the reporting period were too long, then the prediction model might not detect any signs of impending collapse; this could be because the financials are still strong and no symptoms of financial fragility are apparent. Considering that there is no theoretical basis for answering this question of optimal period, it is necessary to reflect on the empirical results found in corporate collapse studies for an answer; something that is done in section 5 of this paper.

4.3 Validation procedures

The issue here deals with whether it is necessary to test the performance of the corporate collapse prediction model on a data sample other than the one used in its derivation. All corporate collapse models are *derived* based on a sample of collapsed and non-collapsed companies. Financial ratios are calculated from the financial statements of these companies over some sample period. These ratios are then analysed using some statistical procedure such as MDA. Ratios that are good predictors of collapse are included in the final model. Selection of these ratios is based on how well they signal the event of collapse *within* the sample of companies from which they are calculated. In other words, the predictive ability of the model is tested *within* the sample of companies from which it is derived.

Some studies have tested the predictive ability of their models on samples other than the ones used in deriving them. This process is referred to as *validation* (Huberty, 1994; Klecka, 1982). However, many studies have elected not to do such validation, and thus a controversy arises. Therefore, not only do we need to establish if undertaking validation provides a cost-benefit return, but also if it is not undertaken, the reasons why.

Technically, validation is not a statistical requirement, because it has no impact on model derivation as explained above. Instead, it is portrayed by those who use it, as a process that could enhance the analysis of their results (Altman, 1968; McGurr and Devaney, 1998; Neophytou and Molinero, 2004; Taffler, 1982). Therefore, in determining whether or not validation is necessary, it is best to establish whether or not it improves the performance (results) of the corporate collapse prediction model. If it does then it would be desirable, if not then why validate?

Studies that use validation, attempt to test the predictive ability of the *same* model on multiple samples. This always includes the original sample used in deriving the model. In addition, the ability of the model to predict collapse is also tested on some other sample or samples. A verdict is then announced as to whether the predictive accuracy of the model when tested on the original derivation sample, is similar to its predictive accuracy when tested on the validation sample(s).

What if the model does not perform as well when tested on the validation sample(s)? Does this mean it is useless? What is the alternative? The logical alternative would be to modify the model so that it performs equally on both the derivation as well as the validation sample(s). How could this be done? Statistically speaking, this could only be done if the validation sample becomes part of the original derivation sample and a new corporate collapse prediction model is derived based on this combined sample of companies (Huberty, 1994; Klecka, 1982).

However, this eradicates the validation sample. Therefore, the process of validation, though it might sound attractive, could become a circular exercise that goes nowhere especially if the objective is model enhancement. In section 5 we will provide recommendations as to if it is desirable (or not) to undertake validation, and provide empirical justification for the choice made.

5. Empirical evidence in relation to data sampling issues

This section considers empirical evidence in relation to the data sampling issues discussed in this paper, and provides recommendations on how to deal with the controversies that arise based on this evidence. A total of 7 issues were discussed. They dealt with paired sample design, matching criteria in a paired sample design, the optimal sample size, the proper treatment of outliers, the optimal sample period, the optimal reporting period and validation procedures. Appendix 1 contains the necessary data needed in considering the empirical evidence.

5.1 Empirical evidence in relation to paired sample design

The initial issue here is the maintenance of a-priori probabilities. An in-depth analysis of the empirical studies in the literature indicates that no study has set out to maintain a-priori probabilities and the reason appears to be one of cost-benefit, i.e. the empirical evidence suggests that the benefit (in terms of model performance) of corporate collapse prediction models in studies that used relatively large data samples was not any better than those that used smaller data samples (Boritz et al., 1995; Jones and Hensher, 2004; Micha, 1984; Ohlson, 1980; Richardson et al., 1998; Zmijewski, 1984). The recommendation is, therefore, to disregard the maintenance of a-priori probabilities.

Notwithstanding the disregarding of a-priori probabilities, the next issue is whether collapsed and non-collapsed companies should be paired, and if so, the ratio to be used. The empirical evidence indicates that in a majority of studies (61 percent), every collapsed company in the sample has been equally matched with one that is non-collapsed (i.e. a 50:50 ratio) (for example, Altman, 1968; Blum, 1974; Laitinen and Laitinen, 2000; Neophytou and Molinero, 2004; Sheppard and Fraser, 1994; Zmijewski, 1984). These results therefore, provide empirical justification regarding the need for an equally paired sample design in ratio-based modelling of corporate collapse, and this is a recommendation of our paper.

5.2 Empirical evidence in relation to matching criteria

A second controversy discussed in this paper was if a paired sample design is necessary in ratio-based modelling of corporate collapse (which we have now recommended), then what *matching criteria* should be used (e.g. industry sector, value of assets, liabilities, equity, or magnitude of sales)?

Notwithstanding the views against matching procedures, (such as in Ohlson (1980) and Taffler (1982)), the majority of the studies listed in Appendix 1 applied some matching criteria to their samples. Specifically, 29 studies, or 63 percent, did so. Among these, the matching criteria primarily included industry classification (90 percent), followed by size of assets (52 percent) and magnitude of sales (24 percent). Liabilities, Equity and Profits were not used at all.

Although matching is usually applied in the context of a paired sample design, there are 6 studies that did not adopt an equally paired sample design, but nevertheless attempted to match the two samples of collapsed and non-collapsed companies using some criteria (Altman et al., 1977; Bongini et al., 2000; Casey and Bartczak, 1985; Kim and McLeod Jr., 1999; Kyung et al., 1999; Lau, 1987). What these studies did was, in as much as possible; try to make the sample of collapsed companies comparable to that of the non-collapsed companies. In 4 out of the 6 studies, comparability was based on industry classification, whereby the two samples of collapsed and non-collapsed companies, though not paired, were chosen such that they belonged to the same industry classification. Table 1 provides a summary of the usage rates for each of the matching criteria across the 29 studies that adopted them.

Table 2 provides summary statistics similar to those presented in Table 1, but based on the entire pool of 46 studies. The number of studies corresponding to the various criteria remains the same, however the usage rates (expressed as percentages) differ.

Tables 1 and 2 indicate that there is consensus (defined as more than 50 percent) among the studies with respect to industry classification being the primary matching criterion. Although the second most popular matching criterion, i.e. size of assets, falls short of a consensus, its popularity is noticeable and therefore could not be ignored. The empirical analysis also indicates that matching by assets instead of

Table 1: Summary of the usage rates of the matching criteria across 29 studies (1968-2006)

Matching Criteria	Number of Studies Using a Particular Criterion	Usage Rate as a Percentage (n=29)
Industry Classification	26	90 percent
Size of Assets	15	52 percent
Magnitude of Sales	7	24 percent
Other*	5	17 percent

**The criterion 'other' includes number of employees, fiscal year, company age, as well as any unclear specification such as 'magnitude' where it was not clear what 'magnitude' referred to.*

Table 2: Summary of the usage rates of the matching criteria across 46 studies (1968-2006)

Matching Criteria	Number of Studies Using a Particular Criterion	Usage Rate as a Percentage (n=46)
Industry Classification	26	57 percent
Size of Assets	15	33 percent
Magnitude of Sales	7	15 percent
Other	5	11 percent

sales has its merits, probably because assets are more stable over time (Altman, 1968; Aly et al., 1992; Baldwin and Glezen, 1992; Charitou et al., 2004; Dambolena and Shulman, 1988; Darayseh et al., 2003; Fletcher and Goss, 1993; Ginoglou et al., 2002; Gombola et al., 1987; Hamer, 1983; Kim and McLeod Jr., 1999; Koh and Killough, 1990; Levitan and Knoblett, 1985; Neophytou and Molinero, 2004; Norton and Smith, 1979; Sharma and Mahajan, 1980; Zavgren, 1985). Therefore, based on the empirical evidence regarding matching criteria, the recommendation of this paper is to use industry classification as mandatory, and size of assets as a desirable secondary criterion.

5.3 Empirical evidence in relation to the optimal sample size

Earlier in this paper, the discussion considered two issues (and related controversies) necessary to satisfy certain statistical requirements of the methodological approaches used, i.e. sample size and treatment of outliers. Let us now make recommendations on these issues based on the empirical evidence.

With regards to sample size, we will focus on those studies that used MDA-based models; which require that the sample size (number of companies) be within a range. This could be anywhere from the number of predictor variables (financial ratios) plus 2, to 10 times the number of predictor variables. Table 3 lists the 29 studies that used MDA. The column with the heading 'n' indicates the number of financial ratios used in the final corporate collapse prediction model.[5] The columns with the heading 'SS' represent the sample size corresponding to each study listed.[6] The information is an extension of the results shown in Appendix 1. When a hash sign (#) appears it indicates that the corresponding study did not break down its sample size by period.

The average sample size is about 47 collapsed and 79 non-collapsed companies. The reason for this variation between the two groups of companies could be attributed to a number of factors. First, data on non-collapsed companies is more readily available. Second, not all studies used a paired sample design. Third, the number of collapsed companies is fairly constant

Table 3: Summary information regarding sample size and number of financial ratios used in signalling collapse across studies that used MDA (1968-2006)

Study	n	SS	Study	n	SS
Altman (1968)	5	33:33	Karels and Prakash (1987)	6	5:71
Deakin (1972)	14	32:32	Lau (1987)	2	#5:350
Edmister (1972)	11	24:24	Peel and Peel (1987)	3	#56:56
Altman (1973)	4	21:21	Coats and Fant (1993)	5	#94:188
Blum (1974)	2	#115:115	Poston and Harmon (1994)	7	#46:123
Elam (1975)	14	48:48	Clark et al. (1997)	9	#7:7
Altman et al. (1977)	5	#53:58	Lenard et al. (1998)	4	32:32
Ketz (1978)	9	#75:100	McGurr and Devaney (1998)	3	#66:66
Norton and Smith (1979)	5	30:30	Kim and McLeod Jr. (1999)	6	#58:57
Taffler (1982)	4	#23:45	Kyung et al. (1999)	4	#58:103
El-Hennawy and Morris (1983)	3	53:53	Gritta et al. (2000)	4	#16:16
Casey and Bartczak (1985)	6	#60:230	Zapranis and Ginoglou (2000)	5	20:20
Frydman et al. (1985)	16	#58:142	Drezner et al. (2001)	5	#185:185
Gentry et al. (1985)	1	#33:33	Ginoglou et al. (2002)	8	#20:20
Levitan and Knoblett (1985)	15	#35:35			

n = number of ratios
SS = sample size

from one year to another. Notwithstanding such variation, the overall average sample size is 63 companies.

The average number (n) of financial ratios based on the results in Table 3 is 6.4. Therefore, the average sample size of 64 is about 10 times the average number of financial ratios. This is equal to the conservative 10 times required by MDA as suggested by Huberty (1994). As such, there is empirical evidence to suggest that the optimal sample size should be 10 times the number of financial ratios used in the corporate collapse prediction model, and this is a recommendation of our paper.

5.4 Empirical evidence in relation to the proper treatment of outliers

Regarding the treatment of outliers, the discussion earlier in this paper mentioned that although the removal of outliers is desirable from a purely statistical perspective, a study of corporate collapse should focus on outliers as it is the abnormalities in the financial ratios that signal impending collapse. Let us, therefore, look at the empirical evidence in order to establish whether or not outliers should be removed, before making any recommendations.

Looking at Appendix 1, the results in the column heading 'EO' indicate that 39 out of the 46 studies listed (i.e. 85 percent), did not eliminate outliers from their data samples. Therefore, there is substantial empirical evidence to make the recommendation that outliers should not be eliminated when modelling corporate collapse.

5.5 Empirical evidence in relation to the optimal sample period

Table 4 facilitates answering the question regarding the optimal *sample* period. The discussion earlier in this paper stated that this period was variable depending on the size of the population and the rate of failure per year. In Australia, this worked out to 3 years. The results in Table 4 are based on the information presented in the column 'SP' (Sample Period) in Appendix 1. Only one study out of the 46 listed did not specify a sample period (Gritta et al., 2000). Therefore, it is excluded while compiling the results in Table 4. Moreover, all of the studies that specified a sample period were consistent in that they adhered to a single sample period throughout their analyses of their original derivation samples, with the exception of Taffler (1982) who used 2 sample periods: a 6-year period that extended from 1968 to 1973 for the collapsed companies and a 2-year period that extended from 1972 to 1973 for the non-collapsed companies. This was an unusual approach that was out of line with the norm and was adopted without justification. That being the case, Taffler (1982) is also excluded while compiling the results in Table 4. Therefore, the results in Table 4 are based on a total of 44 studies.

Table 4: Descriptive statistical measures for the sample period of the derivation sample in 44 studies (1968-2006)

	Sample Period of Derivation Sample (n=44)
Minimum Sample Period	1 year
Mean Sample Period	9 years
Median Sample Period	7 years
Modal Sample Period	7 years
Maximum Sample Period	32 years
Standard Deviation	6 years

Although the results in Table 4 indicate that the range between the minimum and maximum sample periods is seemingly large, this should not pose a concern, given the relatively small standard deviation and the fact that the mean, median and modal sample periods are very close to one another. Moreover, being so close to one another, the implication is that any one of these measures is a proper indicator of what is a generally acceptable sample period. The mode (7 years) is chosen here, because it is representative of the most recurring sample period. Therefore, at first glance, there is empirical evidence to support an optimal sample period of at least 7 years. However, as discussed before, the optimal period may vary depending on size of population and annual failure rate in the country/industry chosen.

5.6 Empirical evidence in relation to the optimal reporting period

Regarding the optimal *reporting* period, the question that was raised earlier in this paper asked how far back a model should go in predicting corporate collapse?

Upon an in-depth analysis of the 46 studies listed in Appendix 1 that adopted a ratio-based approach to signalling corporate collapse, a number of observations are noticeable. First, the predictive power of the models deteriorated as one moves further back from the year in which collapse occurred. Second, the models are incapable of signalling collapse when used beyond 5-years prior to the event. Therefore, it seems that the cut-off period beyond which it becomes ineffective to signal collapse should optimally be 5 years.

The empirical results in Table 5 confirm such a conclusion. These results are based on the information provided in the column 'RP' in Appendix 1. Studies that did not explicitly specify a reporting period, those where 'one period' represented something other than a year, and those that did not break down their results by period are excluded from the analysis. This is done in order to avoid contamination of the results, and therefore led to a reduction in the number of valid studies to 39.

Table 5: Descriptive statistical measures for the reporting periods used in 39 studies (1968-2006)

	Reporting Period of Derivation Sample (n=39)
Minimum Reporting Period	1.0 year
Mean Reporting Period	3.4 years
Median Reporting Period	3.0 years
Modal Reporting Period	5.0 years
Maximum Reporting Period	7.0 years
Standard Deviation	1.7 years

The results in Table 5 indicate that the modal reporting period tends to dominate both the mean and the median reporting periods. This makes the distribution skewed to the left. However, the standard deviation is, relatively speaking, not very large, which explains the tight distribution for the mean, median and mode. The mode is the most useful statistic here, because it reflects what the largest number of studies considered to be a suitable reporting period.

The results in Table 6 indicate the number of studies (out of the 39) that used a particular reporting period. The highest percentage (41 percent) corresponds to 5 years prior to collapse, which once again confirms an optimal reporting period of 5 years.

Table 6: Summary of choice of reporting period across 39 studies (1968-2006)

Reporting Period Prior to Collapse	Number of Studies Using Corresponding Reporting Period	Percentage of Studies Using Corresponding Reporting Period (n=39)
1 year	8	21 percent
2 years	5	13 percent
3 years	7	18 percent
4 years	2	5 percent
5 years	16	41 percent
6 years	0	0 percent
7 years	1	3 percent

5.7 Empirical evidence in relation to validation procedures

The discussion earlier in this paper argued that the process of validation could become a circular exercise that goes nowhere, and if so, validation should not be recommended. The discussion also highlighted several other problems with respect to validation procedures, all of which would make it less desirable. However, Appendix 1 demonstrates that a majority of studies (61 percent) used validation procedures. Let us therefore look at the empirical evidence in order to make recommendations with regards to the validation issue. We will specifically consider if the researchers in the studies that undertook validation, were consistent in the procedures chosen.

Although a majority result of 61 percent might indicate a bias in favour of carrying out validation, taking a closer look at the results of the 28 studies that undertook validation reveals a state of confusion and lack of consistency, as shown in Table 7. The empirical analysis demonstrates that 4 validation procedures were, more or less, equally used. These are *Lachenbruch jackknife, part of derivation sample, different sample* and *sub-period or different period*, as shown in Table 7.

The *Lachnbruch jackknife* procedure is a technique that tests a particular statistical model by removing items one at a time and classifying them into their respective groups using a model developed from the remaining sample items. The predictive accuracy of the model is then calculated by comparing the group classification determined by the jackknife technique with actual group classification. (Lachenbruch, 1967, pp. 264-265). In *part of derivation sample* validation, a subset of the original derivation sample is randomly selected for validation. In *different sample* validation, a sample different from the original derivation sample is used. In *different period* validation, the original derivation sample is used for validation, but from a different sample period or a sub-period of the original sample period.

It is clear from examining the literature that, even if validation is desirable, it is a process ridden with problems. First, it is sensitive to the size of the data sample; specifically, validation should not be used in small samples (Klecka, 1982). Besides, there is no agreement as to what constitutes a small sample or what portion of the data should be included in the validation sample (Chapman et al., 1977; Huberty, 1994; Lachenbruch and Mickey, 1968; LaRocco et al., 1977; Schaafsma and van Vark, 1979; Ware and Williams, 1977). Second, validation is usually applied to a sample (validation sample) other than the original sample (derivation sample). The predictive accuracy of the model in the validation sample may differ from that in the original derivation sample. Therefore, the classification rule that is usually used in carrying out validation is not the appropriate one, because it is based on a sub-set of the sample and not the entire sample (Lachenbruch and Mickey, 1968). Third, the size of the test sample has a bearing on the performance of the accuracy of the validation procedure. A large test sample provides poor but stable accuracy; whereas, a small test sample provides good but highly variable accuracy (Lachenbruch and Mickey, 1968). Given such problems, the recommendation of this paper, based on a detailed analysis of the empirical studies carried out, is that validation is not an exercise that provides value from a cost-benefit viewpoint.

Table 7 - Summary of the breakdown of validation procedures in 28 studies (1968-2006)

Validation Procedure	Number of Studies Using Corresponding Procedure	Percentage of Studies Using Corresponding Procedure (n=28)
Lachenbruch Jackknife	6	21 percent
Part of Derivation Sample	11	39 percent
Different Sample	5	18 percent
Sub-period or Different Period	5	18 percent
Other	2	7 percent

6. Conclusion and recommendations

Empirical approaches to signalling corporate collapse require the derivation of a statistical model. The data sample must satisfy certain conditions relating to the selection of companies, and the process of data analysis.

This paper provided a theoretical justification as to why collapsed and non-collapsed companies should be paired based on industry sector and size of assets. These recommendations were supported by empirical evidence presented from an in-depth analysis of 46 pertinent studies on corporate collapse.

Theoretical and empirical justification was also provided as to why the sample size should be at least 10 times the number of financial ratios used in the corporate collapse prediction model, and also why the removal of outliers is not desirable.

Empirical evidence was also presented in relation to data sampling issues needed to enhance the empirical results. Whilst the *sample* period depended on the size of the population and frequency of collapses, a period of 7 years was the mode. On the other hand, a *reporting* period of 5 years was the mode. Finally, due to the confusion regarding validation procedures when modelling corporate collapse, undertaking such procedures was not recommended from a cost-benefit viewpoint. **JARAF**

Dr Ghassan Hossari
Deakin University, Melbourne, Australia

Professor Sheikh F Rahman
Central Queensland University, Melbourne, Australia

Professor Janek Ratnatunga
Monash University, Melbourne, Australia

References

Aitkin, M. and Wilson, G. T. (1980), "Mixture Models, Outliers and the EM Algorithm", Technometrics, Vol. 22 (3), pp. 325-331.

Altman, E. I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", Journal of Finance, Vol. 23 (4), pp. 589-610.

Altman, E. I. (1973), "Predicting Railroad Bankruptcies in America", Bell Journal of Economics and Management Science, Vol. 4 (1), pp. 184-212.

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, Vol. 1 (1), pp. 29-54.

Aly, I. M., Barlow, H. A. and Jones, R. W. (1992), "The Usefulness of SFAS No. 82 (Current Cost) Information in Discriminating Business Failure: An Empirical Study", Journal of Accounting, Auditing and Finance, Vol. 7 (2), pp. 217-229.

Baldwin, J. and Glezen, G. W. (1992), "Bankruptcy Prediction Using Quarterly Financial Statement Data", Journal of Accounting, Auditing and Finance, Vol. 7 (3), pp. 269-285.

Beaver, W. H. (1966), "Financial Ratios as Predictors of Failure", Journal of Accounting Research, Vol. 4 (3), pp. 71-112.

Beaver, W. H. (1968a), "Alternative Accounting Measures as Predictors of Failure", Accounting Review, Vol. 43 (1), pp. 113-123.

Beaver, W. H. (1968b), "Market Prices, Financial Ratios, and the Prediction of Failure", Journal of Accounting Research, Vol. 6 (2), pp. 179-193.

Bird, R. G. and McHugh, A. J. (1977), "Financial Ratios - An Empirical Study", Journal of Business Finance and Accounting, Vol. 4 (1), pp. 29-46.

Blum, M. (1974), "Failing Company Discriminant Analysis", Journal of Accounting Research, Vol. 12 (1), pp. 1-26.

Bongini, P., Ferri, G. and Hahm, H. (2000), "Corporate Bankruptcy in Korea: Only the Strong Survive?", Financial Review, Vol. 35 (4), pp. 31-51.

Boritz, J. E., Kennedy, D. B. and Albuquerque, A. M. (1995), "Predicting Corporate Failure Using a Neural Network Approach", Intelligent Systems in Accounting, Finance and Management, Vol. 4 (2), pp. 95-111.

Butler, R. W. (1986), "Predictive Likelihood Inference with Applications", Journal of the Royal Statistical Society, Vol. 48 (1), pp. 1-38.

Campbell, N. A. (1978), "The Influence Function as an Aid in Outlier Detection in Discriminant Analysis", Applied Statistics, Vol. 27 (3), pp. 251-258.

Casey, C. and Bartczak, N. (1985), "Using Operating Cash Flow Data to Predict Financial Distress: Some Extensions", Journal of Accounting Research, Vol. 23, pp. 384-402.

Chapman, D. W., Holloway, R. and Kelly, E. (1977), "Using Ratings by Students to Predict High and Low Achievers in a PSI Course: A Discriminant Analysis", Educational Communication and Technology, Vol. 25 (4), pp. 411-422.

Charitou, A., Neophytou, E. and Charalambous, C. (2004), "Predicting Corporate Failure: Empirical Evidence for the UK", European Accounting Review, Vol. 13 (3), pp. 465-497.

Clark, C. E., Foster, P. L., Hogan, K. M. and Webster, G. H. (1997), "Judgemental Approach to Forecasting Bankruptcy", Journal of Business Forecasting Methods and Systems, Vol. 16 (2), pp. 14-19.

Coats, P. K. and Fant, L. F. (1993), "Recognizing Financial Distress Patterns Using a Neural Network Tool", Financial Management, Vol. 22 (3), pp. 142-156.

Collins, J. R. (1976), "Robust Estimation of a Location Parameter in the Presence of Asymmetry", Annals of Statistics, Vol. 4 (1), pp. 68-85.

Collins, J. R. (1982), "Robust M-Estimators of Location Vectors", Journal of Multivariate Analysis, Vol. 12, pp. 480-492.

Collins, J. R. and Portnoy, S. L. (1981), "Maximizing the Variance of M-Estimators Using the Generalized Method of Moment Spaces", Annals of Statistics, Vol. 9 (3), pp. 567-577.

Collins, J. R. and Wiens, D. P. (1985), "Minimax Variance M-Estimators in Epsilon-Contamination Models", Annals of Statistics, Vol. 13 (3), pp. 1078-1096.

Dambolena, I. G. and Shulman, J. M. (1988), "A Primary Rule for Detecting Bankruptcy: Watch the Cash", Financial Analyst Journal, Vol. 44 (5), pp. 74-79.

Darayseh, M., Waples, E. and Tsoukalas, D. (2003), "Corporate Failure for Manufacturing Industries using Firms Specific and Economic Environment with Logit Analysis", Managerial Finance, Vol. 29 (8), pp. 23-36.

Deakin, E. B. (1972), "A Discriminant Analysis of Predictors of Business Failure", Journal of Accounting Research, Vol. 10 (1), pp. 167-180.

Dimitras, A. I., Slowinski, R., Susmaga, R. and Zopounidis, C. (1999), "Business Failure Prediction Using Rough Sets", European Journal of Operational Research, Vol. 114 (2), pp. 263-280.

Drezner, Z., Marcoulides, G. A. and Stohs, M. H. (2001), "Financial Applications of a Tabu Search Variable Selection Model", Journal of Applied Mathematics and Decision Sciences, Vol. 5 (4), pp. 215-235.

Edmister, R. O. (1972), "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction", Journal of Financial and Quantitative Analysis, Vol. 7 (2), pp. 1477-1494.

Eisenbeis, R. A. (1977), "Pitfalls in the Application of Discriminant Analysis in Business, Finance and Economics", Journal of Finance, Vol. 32 (3), pp. 875-901.

El-Hennawy, R. H. A. and Morris, R. C. (1983), "The Significance of Base Year in Developing Failre Prediction Models", Journal of Business Finance and Accounting, Vol. 10 (2), pp. 209-224.

Elam, R. (1975), "The Effect of Lease Data on the Predictive Ability of Financial Ratios", Accounting Review, Vol. 50 (1), pp. 25-43.

Fletcher, D. and Goss, E. (1993), "Forecasting with Neural Networks: An Application Using Bankruptcy Data", Information and Management, Vol. 24, pp. 159-167.

Foley, D. H. (1972), "Consideration of Sample and Feature Size", IEEE Transactions on Information Theory, Vol. 18 (5), pp. 618-626.

Frydman, H., Altman, E. I. and Kao, D. (1985), "Introducing Recursive Partitioning for Financial Classification: the Case of Financial Distress", Journal of Finance, Vol. 40 (1), pp. 269-292.

Gentry, J. A., Newbold, P. and Whitford, D. T. (1985), "Classifying Bankrupt Firms with Funds Flow Components", Journal of Accounting Research, Vol. 23 (1), pp. 146-161.

Ginoglou, D., Agorastos, K. and Hatzigagios, T. (2002), "Predicting Corporate Failure of Problematic Firms in Greece with LPM, Logit, Probit and Discriminant Analysis Models", Journal of Financial Management and Analysis, Vol. 15 (1), pp. 1-16.

Gombola, M. J., Haskins, M. E., Ketz, J. E. and Williams, D. D. (1987), "Cash Flow in Bankruptcy Prediction", Financial Management, Vol. 16 (4), pp. 55-65.

Gritta, R. D., Wang, M., Davalos, S. and Chow, G. (2000), "Forecasting Small Air Carrier Bankruptcies Using a Neural Network Approach", Journal of Financial Management and Analysis, Vol. 13 (1), pp. 44-50.

Hamer, M. M. (1983), "Failure Prediction: Sensitivity of Classification Accuracy to Alternative Statistical Methods and Variable Sets", Journal of Accounting and Public Policy, Vol. 2 (4), pp. 289-307.

Hampel, F. R., Rousseeuw, P. J. and Ronchetti, E. (1981), "The Change-of-Variance Curve and Optimal Redescending M-Estimators", Journal of the American Statistical Association, Vol. 76 (375), pp. 643-648.

Hecker, R. and Wegener, H. (1978), "The Validation of Classification Rates in Stepwise Discriminant Analyses", Biometrical Journal, Vol. 20, pp. 713-727.

Hossari, G. (2006a), "A Dynamic Ratio-Based Model for Signalling Corporate Collapse", Journal of Applied Management Accounting Research, Vol. 4 (1), pp. 11-32.

Hossari, G. (2006b), A Ratio-Based Multi-Level Modelling Approach for Signalling Corporate Collapse: A Study of Australian Corporations, Swinburne University of Technology, Melbourne.

Huber, P. J. (1964), "Robust Estimation of a Location Parameter", Annals of Mathematical Statistics, Vol. 35 (1), pp. 73-101.

Huberty, C. J. (1994), Applied Discriminant Analysis, John Wiley, New York.

Jain, A. K. and Chandrasekaran, B. (1982), "Dimensionality and Sample Size Considerations in Pattern Recognition Practice", in P. R. Krishnaiah and L. N. Kanal (eds), Handbook of Statistics 2, North-Holland Publishing Company, New York, USA.

Johnson, W. and Geisser, S. (1983), "A Predictive View of the Detection and Characterization of Influential Observations in Regression Analysis", Journal of the American Statistical Association, Vol. 78 (381), pp. 137-144.

Jones, S. and Hensher, D. A. (2004), "Predicting Firm Financial Distress: A Mixed Logit Model", Accounting Review, Vol. 79 (4), pp. 1011-1038.

Karels, G. V. and Prakash, A. J. (1987), "Multivariate Normality and Forecasting of Business Bankruptcy", Journal of Business Finance and Accounting, Vol. 14 (4), pp. 573-594.

Ketz, J. E. (1978), "The Effect of General Price-Level Adjustments on the Predictive Ability of Financial Ratios", Journal of Accounting Research, Vol. 16 (3), pp. 273-285.

Kim, C. N. and McLeod Jr., R. (1999), "Expert, Linear Models, and Nonlinear Models of Expert Decision Making in Bankruptcy Prediction: A Lens Model Analysis", Journal of Management Information Systems, Vol. 16 (1), pp. 189-207.

Klecka, W. R. (1982), Discriminant Analysis, 5TH edn, Sage Publications, London.

Koh, H. C. and Killough, L. N. (1990), "The Use of Multiple Discriminant Analysis in the Assessment of the Going-Concern Status of an Audit Client", Journal of Business Finance and Accounting, Vol. 17 (2), pp. 179-192.

Kyung, S. T., Chang, N. and Lee, G. (1999), "Dynamics of Modeling in Data Mining: Interpretive Approach to Bankruptcy Prediction", Journal of Management Information Systems, Vol. 16 (1), pp. 63-85.

Lachenbruch, P. A. (1967), "An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant Analysis", Biometrics, Vol. 23 (4), pp. 639-645.

Lachenbruch, P. A. and Mickey, M. R. (1968), "Estimation of Error Rates in Discriminant Analysis", Technometrics, Vol. 10 (1), pp. 1-12.

Laitinen, E. K. (1991), "Financial Ratios and Different Failure Processes", Journal of Business Finance and Accounting, Vol. 18 (5), pp. 649-673.

Laitinen, E. K. and Laitinen, T. (2000), "Bankruptcy Prediction Application of the Taylor's Expansion in Logistic Regression", International Review of Financial Analysis, Vol. 9 (4), pp. 327-350.

LaRocco, J. M., Pugh, W. M. and Gunderson, E. K. (1977), "Identifying Determinants of Retention Decisions", Personnel Psychology, Vol. 30 (2), pp. 199-215.

Lau, A. H. (1987), "A Five-State Financial Distress Prediction Model", Journal of Accounting Research, Vol. 25 (1), pp. 127-139.

Lenard, M. J., Madey, G. R. and Alam, P. (1998), "The Design and Validation of a Hybrid Information System for the Auditor's Going Concern Decision", Journal of Management Information Systems, Vol. 14 (4), pp. 219-238.

Levitan, A. S. and Knoblett, J. A. (1985), "Indicators of Exceptions to the Going Concern Assumption", Auditing, Vol. 5 (1), pp. 26-40.

Maronna, R. A. (1976), "Robust M-Estimators of Multivariate Location and Scatter", Annals of Statistics, Vol. 4 (1), pp. 51-67.

McGurr, P. T. and Devaney, S. A. (1998), "A Retail Failure Prediction Model", International Review of Retail, Vol. 8 (3), pp. 259-277.

McLachlan, G. J. (1992), Discriminant Analysis and Statistical Pattern Recognition, John Wiley & Sons, Inc, New York, USA.

Meyer, P. A. and Pifer, H. W. (1970), "Prediction of Bank Failures", Journal of Finance, Vol. 25 (4), pp. 853-869.

Micha, B. (1984), "Analysis of Business Failures in France", Journal of Banking and Finance, Vol. 8 (2), pp. 281-291.

Neophytou, E. and Molinero, C. M. (2004), "Predicting Corporate Failure in the UK: A Multidimensional Scaling Approach", Journal of Business Finance and Accounting, Vol. 31 (5), pp. 677-710.

Norton, C. L. and Smith, R. E. (1979), "A Comparison of General Price-Level and Historical Cost Financial Statements in the Prediction of Bankruptcy", Accounting Review, Vol. 54 (1), pp. 78-94.

Ohlson, J. A. (1980), "Financial Ratios and the Probabilistic Prediction of Bankruptcy", Journal of Accounting Research, Vol. 18 (1), pp. 109-138.

Peel, M. J. and Peel, D. A. (1987), "Some Further Empirical Evidence on Predicting Private Company Failure", Accounting and Business Research, Vol. 18 (69), pp. 57-66.

Platt, H. D., Platt, M. B. and Pedersen, J. G. (1994), "Bankruptcy Discrimination with Real Variables", Journal of Business Finance and Accounting, Vol. 21 (4), pp. 491-515.

Poston, K. M. and Harmon, W. K. (1994), "A Test of Financial Ratios as Predictors of Turnaround versus Failure Among Financially Distressed Firms", Journal of Applied Business Research, Vol. 10 (1), pp. 41-57.

Richardson, F. M., Kane, G. D. and Lobinger, P. (1998), "The Impact of Recession on the Prediction of Corporate Failure", Journal of Business Finance and Accounting, Vol. 25 (1/2), pp. 167-186.

Schaafsma, W. and van Vark, G. N. (1979), "Classification and Discrimination Problems with Applications", Statistica Neerlandica, Vol. 33, pp. 91-126.

Sharma, S. and Mahajan, V. (1980), "Early Warning Indicators of Business Failure", Journal of Marketing, Vol. 44 (4), pp. 80-89.

Sheppard, J. P. and Fraser, S. (1994), "The Dilemma of Matched Pairs and Diversified Firms in Bankruptcy Prediction Models", Mid-Atlantic Journal of Business, Vol. 30 (1), pp. 9-26.

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

Taffler, R. (1983), "The Assessment of Company Solvency and Performance Using a Statistical Model", Accounting and Business Research, Vol. 13 (52), p. 295.

Tamari, M. (1966), "Financial Ratios as a Means of Forecasting Bankruptcy", Management International Review, Vol. 6 (4), pp. 15-21.

Walker, M. C., Stowe, J. D. and Moriarity, S. (1979), "Decomposition Analysis of Financial Statements", Journal of Business Finance and Accounting, Vol. 6 (2), pp. 173-187.

Walter, J. E. (1959), "A Discriminant Function for Earnings Price Ratios of Large Industrial Corporations", Review of Economics and Statistics, Vol. XLI, pp. 44-53.

Ware, J. E. and Williams, R. G. (1977), "Discriminant Analysis of Student Ratings as a Means for Identifying Lecturers who Differ in Enthusiasm or Information-Giving", Educational and Psychological Measurement, Vol. 37 (7), pp. 627-639.

Winakor, A. H. (1929), Standard Financial Ratios for the public Utility Industry, No. 26, Bureau of Business Research, The University of Illinois.

Zapranis, A. and Ginoglou, D. (2000), "Forecasting Corporate Failure with Neural Network Approach: The Greek Case", Journal of Financial Management and Analysis, Vol. 13 (2), pp. 11-21.

Zavgren, C. V. (1985), "Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis", Journal of Business Finance and Accounting, Vol. 12 (1), pp. 19-45.

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

APPENDIX 1: The data

Following are explanations of the headings in the table below, which include the data upon which the empirical results in this paper are based.

Study: This lists the 46 studies that were reviewed in gathering the relevant information. *The columns that follow deal with the original derivation samples, not the validation samples.*

P (Pairing): This indicates whether the sample of collapsed companies is paired with a sample of non-collapsed companies. The letters 'Y' and 'N' indicate if it was a paired sample or not.

MC (Matching Criteria): The following abbreviations are used: **I** (Industry); **S** (Sales); **O** (Other) indicates that some other matching criteria are applied; and a **dash** (-) indicates that no matching criteria are applied.

SS (Sample Size): The two sets of numbers represent, in order, the number of collapsed companies and the number of non-collapsed companies. Where the sample size differs from one reporting period to the other, the largest sample size is shown, in order to avoid counting a company more than once; particularly, that most of the companies in the sample appear in more than one reporting period. It is worth noting here that not all studies broke down their sample size by period. As a matter of fact, a substantial number reported an overall sample size, without breaking it down by period.

EO (Elimination of Outliers): Indicated by 'Y' (Yes) and 'N' (No).

SP (Sample Period): This is in years. Where a sample period is not specified in the corresponding study, a dash (-) is inserted.

RP (Reporting Period): In most cases one period represents one year; however, there are a few instances where one period represents something other than a year; in this case an asterisk (*) appears. Moreover, the acronym 'Ov' (Overall) might appear instead of a number, if the corresponding study reports overall results rather than by period. Where a reporting period is not specified in the corresponding study, a dash (-) is inserted.

VP (Validation Procedure): The following abbreviations are used: **LJ** (Lachenbruch Jackknife); **PS** (Part of Derivation Sample); **DS** (Different Sample); **DP** (Different Period); and **O** (Other). A dash (-) indicates that no validation is evident.

The table appears on the next page.

Appendix 1 Table 1: Summary of data sampling issues in 46 studies (1968-2006)

Study	P	MC			SS	EO	SP	RP	VP
Altman (1968)	Y	I	A		33:33	Y	20	5	PS
Deakin (1972)	Y	I	A		32:32	N	7	5	PS
Edmister (1972)	Y	-			24:24	N	16	3	PS
Altman (1973)	Y	I			21:21	N	32	2	PS
Blum (1974)	Y	I	S	O	115:115	Y	15	5	PS
Elam (1975)	Y	I	S		48:48	N	7	5	-
Altman et al. (1977)	N	I			53:58	N	7	5	LJ
Ketz (1978)	N	-			75:100	N	6	-	-
Norton and Smith (1979)	Y	I	A		30:30	N	5	4	-
Walker et al. (1979)	Y	I			8:8	N	10	3	-
Ohlson (1980)	N	-			105:2058	Y	7	2	-
Taffler (1982)	N	-			23:45	N	6&2	4	LJ
El-Hennawy and Morris (1983)	Y	-			53:53	N	12	5	PS
Mensah (1984)	Y	I	A		37:37	N	9	4*	PS
Zmijewski (1984)	N	-			40:800	N	7	7	DS
Casey and Bartczak (1985)	N	I			60:230	Y	12	5	-
Frydman et al. (1985)	N	-			58:142	N	11	1	PS
Gentry et al. (1985)	Y	I	A	S	33:33	N	12	3	-
Levitan and Knoblett (1985)	Y	I	A		35:35	N	3	3	-
Zavgren (1985)	Y	I	A		45:45	N	7	5	DP
Karels and Prakash (1987)	N	-			5:71	N	1	1	DP
Keasey and Watson (1987)	Y	I			73:73	N	14	1	DS
Lau (1987)	N	A			5:350	N	3	3	DP
Peel and Peel (1987)	Y	-			56:56	N	4	Ov	DS
Dambolena and Shulman (1988)	Y	I	A		50:50	N	4	2	PS
Flagg and Giroux (1991)	N	-			26:176	N	7	5	-
Coats and Fant (1993)	N	-			94:188	Y	20	3	PS
Platt et al. (1994)	N	-			35:89	N	7	1	-
Poston and Harmon (1994)	N	-			46:123	N	7	1	O
Sheppard and Fraser (1994)	Y	I	A	O	29:29	N	6	5	-
Wilson et al. (1995)	Y	-			40:40	N	8	1	-
Hill and Perry (1996)	N	-			75:182	N	11	1	-
Clark et al. (1997)	Y	I			7:7	N	6	5	DS
Lenard et al. (1998)	Y	-			32:32	N	2	2	-
McGurr and Devaney (1998)	Y	I	S		66:66	N	8	Ov	LJ, PS
Kim and McLeod Jr. (1999)	N	I	A		58:57	N	5	Ov	O
Kyung et al. (1999)	N	I	A	O	58:103	N	2	Ov	LJ
Bongini et al. (2000)	N	S			57:437	Y	1	1	-
Gritta et al. (2000)	Y	S			16:16	N	-	Ov	LJ
Laitinen and Laitinen (2000)	Y	I	S		200:200	Y	7	3	LJ
Zapranis and Ginoglou (2000)	Y	I	O		20:20	N	5	5	DP
Drezner et al. (2001)	Y	I	O		185:185	N	20	2	-
Ginoglou et al. (2002)	Y	I	A		20:20	N	5	5	-
Darayseh et al. (2003)	Y	I	A		100:100	N	8	5	DS
Jones and Hensher (2004)	N	-			110:4980	N	5	5	DP
Hossari (2006)	Y	I	A		37:37	N	14	5	-

Footnotes

1 It must be noted that even maintaining a-priori probabilities in a data sample of even a reasonable size can be a problem. If we consider a reasonable sample of 1500 companies (almost all Australian listed companies), then an a-priori probability of 0.33 percent collapsed companies is equivalent to just 5 collapsed companies per year.

2 In the way of comparison, 35,472 businesses filed for bankruptcy in the U.S in the year 2000 (Source: ABI World). This compares to a total of 5,652,544 firms that operated in the U.S. in the same year (Source: U.S. Census Bureau). Therefore, 0.62 percent of U.S. businesses have failed in the year 2000, which are nearly six firms in 1000.

3 The number of babies who died from cot death in the U.S. in the year 2000 is 0.62 per 1000, which is equivalent to 0.062 percent (Source: American SIDS Institute).

4 Note that many issues arise in asset valuations, such as the non-recognition of intangibles and fair-value accounting. Most corporate collapse models take asset values as given in the financial statements. This could have an impact on knowledge rich companies with most of their intangible assets off-balance sheet, which is why there is strong support to undertake an industry classification first.

5 Most studies start with a large number of ratios, but only a handful would make it to the final model.

6 These relate to the derivation samples, not the validation samples, as not all studies adopted validation samples.