Seeking Parsimony in Bankruptcy Studies Using Redundancy Analysis

Zahra Poorzamani[1], Azita Jahanshad[2]

Department of Accounting, Assistant Professor, Central Tehran Banch, Islamic Azad University, Tehran, Iran,
E- mail: zahra.poorzamani@yahoo.com
Department of Accounting, Assistant Professor, Central Tehran Banch, Islamic Azad University, Tehran, Iran
E- mail: az_jahanshad@yahoo.com

Abstract: Despite the long-standing criticism directed towards variable selection in financial distress studies, no research has so far examined the relationship between alternative variable sets used in such studies. Therefore, the main purpose of the present study is to employ canonical correlation analysis in order to examine the relationships that exist between variable sets employed in four bankruptcy studies and then to illustrate the value of applying the law of parsimony to canonical correlation analysis solutions. The primary purpose of parsimony is that the more parsimonious the solution is, the more replicable the model will be. In this study the goal was achieved by removing the three variables in variable sets employed in selected bankruptcy studies.

[Zahra Poorzamani, Azita Jahanshad**. Seeking Parsimony in Bankruptcy Studies Using Redundancy Analysis.** *Life Sci J* 2013;10(7s):70-78] (ISSN:1097-8135). http://www.lifesciencesite.com. 12

Keywords: financial distress, bankruptcy, canonical correlation analysis, law of parsimony, variable deletion strategy

1. Introduction

The development of the bankruptcy prediction model has long been regarded as an important and widely studied issue in the academic and business community. Numerous corporate failure prediction models have so far been developed, based on various modeling techniques and financial ratios. These methods all have their particular strengths and weaknesses in discrimination of failing and non-failing firms, but it is important that which ratios have a superior discriminating power (Leclere, 2006).

As suggested by Ball and Foster (1982), nearly 25 years ago, the selection of independent variables in financial distress studies provides convincing evidence that economic theory does not underlie the research. Variable selection in financial distress studies is commonly based upon suggestions in the literature, the success of variables in earlier studies, or the selection of a large set of variables with an accompanying data reduction procedure in order to maximize predictive ability or some other statistical criteria. Table 1 provides an overview of a large number of academically developed classic statistical business failure prediction models. The combination of the absence of economic theory and ad hoc variable selection together limit the ability of researchers to make generalizations as to the specific financial variables consistent with predictors of financial distress (Foster, 1986).

The purpose of the present study is to determine the relationships between variable sets used in bankruptcy studies.

2. An overview of the selected bankruptcy studies

The research selects four specific bankruptcy studies and examines the relationships between variable sets used. The studies chosen are Altman (1968), Zavgren (1983), Deakin (1972), and Ohlson (1980). These studies are representative of the genre, similar as to choice of statistical technique (discriminant analysis or logistic regression), and frequently cited by researchers. Variable selection has not improved in the years following the publication of these studies. Table 2 contains the independent variables employed in these studies.

Altman (1968) employed discriminant analysis to classify firms as failed or non-failed. The five ratios employed in the model were earnings before interest and taxes/total assets, market value of equity/book value of debt, retained earnings/total assets, sales/total assets, and working capital/total assets. In the year prior to the bankruptcy, the best model was 95 per cent effective in classifying the firms.

Deakin (1972) utilized the ratios of Beaver (1966; 1968) to build a discriminant model for predicting business failure. The ratios employed were cash/current debts, cash/sales, cash/total assets, cash flow/total debts, current assets/current debts, current assets/sales, current assets/total assets, net income/total assets, quick assets/current debts, quick assets/sales, quick assets/total assets, total debts/total assets, working capital/sales, and working capital/total assets. The model was 97 per cent effective in classifying the firms in the year prior to failure.

Ohlson (1980) attempted to provide some improvement in the area of financial distress research by employing some unique ratios in addition to traditional ratios. The ratios employed were change in income/sum of the absolute value of net income for two years (CHIN), current debts/current assets, size (measured as the log of total assets/GNP price-level index), net income/total assets, positive or negative owners' equity (OENEG), positive or negative earnings (INTWO), total debts/total assets, working capital/total assets, and working capital flow/total debts. The model was 96 per cent effective in classifying the firms in the year prior to failure.

Table 2 shows financial ratios and accounting variables employed in financial distress studies.

Table 1: Overview of classic statistical business failure prediction models

Method	Failure prediction model
Univariate analysis	Smith & Winakor (1935) Merwin (1942) Beaver (1967)
Risk index models	Tamari(1966) Moses & liao(1987)
MDA models	Altman (1968) Deakin (1972) Edmister (1972) Blum (1974) Altman et al. (1977) Deakin (1977) Taffler & tisshaw (1977) Springate (1978) Van frederikslust (1978) Bilderbek (1979) Dambolena & khoury (1980) Taffler (1983) Falmer (1984) Betts & Belhoul (1987) Declerc et al (1991) Laitinen (1992) Lussier & Corman (1994) Altman et al (1995) Ca – Score (1987) Shirata (1998)
Conditional probability models	Ohlson (1980) Swanson & Tybout (1988) Zavgren (1983) Zmijewski (1984) Gentry et al (1985) Zavgren (1985) Aziz et al(1988) Gloubos & Grammatikos (1988) Keasey & Mcguinness (1990) Platt & Plat (1990) Sheppard (1994) Lussier (1995) Mossman et al (1998) Grice (1998) Yang (2001) Becchetti & Sierra (2002) Charitou et al (2004))

Table 2: Financial ratios and accounting variables employed in financial distress studies

Ratio	Altman	Zavgren	Deakin	Ohlson
Cash/current debts			*	
Cash/sales			*	
Cash/total assets			*	
Cash flow/total debts			*	
Change in income/sum of the absolute value of net income for two years				*
Current assets/current debts			*	
Current assets/sales			*	
Current assets/total assets			*	
Current debts/current assets				*
Earnings before interest and taxes/total assets	*			
Log (total assets/GNP price level index)				*
Market value of equity/book value of debt	*			
Net income/total assets			*	*
Average inventories/net sales		*		
Average receivables/average inventories		*		
Owners' equity (positive or negative)				*
Positive or negative earnings				*
Long term debt/(total assets – short term debt)		*		
Net sales/(fixed assets+ working capital)		*		
Operational earnings/(total assets - Short term debt)		*		
Quick assets/current debts		*	*	
Quick assets/sales			*	
Quick assets/total assets		*	*	
Retained earnings/total assets	*			
Working capital/total assets	*		*	*
Working capital/ sales			*	
Sales/total assets	*			
Total debt / total assets			*	*

3. Research Hypothesis

Given the purposes in this research, the main hypothesis is as follows:

Much of the information contained in the Altman variable set is presented in the Deakin variable set

Much of the information contained in the Altman variable set is presented in the Zavgren variable set

Much of the information contained in the Altman variable set is presented in the Ohlson variable set

Much of the information contained in the Zavgren variable set is presented in the Deakin variable set

Much of the information contained in the Zavgren variable set is presented in the Ohlson variable set

Much of the information contained in the Ohlson variable set is presented in the Deakin variable set

4. Correlation test

Canonical correlation analysis is the most general case of the general linear model and a multivariate statistical technique employed to investigate the association between two sets of multiple variables (Thompson, 1984). Since Knapp (1978) demonstrated that canonical correlation analysis was the most general form of the general linear model, CCA has gained more in popularity. Thompson (1991) showed that CCA subsumes all other parametric methods including t-tests, point bisereal, analysis of variance (ANOVA), regression, discriminant analysis, and multivariate analysis of variance (MANOVA). Besides Knapp's demonstration, computer statistical packages have made its use more easily accessible to researchers. As Pedhazur (1997) has noted, canonical correlation matrix computation can become "prohibitive" and "complex". Modern statistical packages almost eliminate the need to create these matrixes.

Canonical correlation analysis examines the independent statistical relationships that exist between two variable sets by analyzing the sets simultaneously and identifying and quantifying the elements of one variable set most highly related to the elements of the other variable set (Kotz and Johnson, 1982; Thompson, 1984). This statistical technique can treat the two variable sets symmetrically or it can treat one variable set as the predictor set (independent or exploratory measures) and the other set as the criterion set (dependent measures). Furthermore multiple regression analysis could do the job if there were only one dependent variable; however, canonical analysis goes a step farther by allowing multiple dependent variables.

As mentioned in section 2, the four variable sets examined in this study are Altman, Zavgren, Deakin, and Ohlson. Four studies provide six pair wise comparisons of variables: Altman versus Zavgren, Altman versus Deakin, Altman versus Ohlson, Zavgren versus Deakin, Zavgren versus Ohlson, and Deakin versus Ohlson.

The research limited the computation of the ratios listed in Table 2 to manufacture firms listed in the Tehran Stock Exchange from 1997 to 2007. The research relies on a sample of 30 failed and 30 non-failed manufacturing firms. A sample of 30 manufacturing companies which had become bankrupt between 1997 and 2007 were identified from The Article No.141 of Commercial Law of Iran and matched to 30 non-failed companies on the basis simple Q- tubin.

The data for the canonical correlation analyses consisted of pooled observations for the six pair wise comparisons over the ten year period. The study performed the canonical correlation analysis for each of the six pair wise comparisons. It extracted the canonical structure and retained significant (0.05 level) canonical functions using Barlett's chi-square approximation to the distribution of Wilk's lambda. As different variable sets may appear to contain common information, the existence of similar variables in these variable sets most likely accounts for the appearance of common information. To determine whether this interpretation is correct, the study performs the analysis with the common variables in each pair wise comparison deleted (Leclere, 2006).

Canonical correlation analysis determines the extent of the relationship between two variable sets with redundancy coefficients. Redundancy coefficients indicate the degree of overlap between two sets of variables; more specifically, they are an index of the average proportion of variance in one variable set that is predictable from or shared with the canonical variates in the other set (Stewart and Love, 1968; Lambert et al.1988). Employing one set of variables to predict a second set of variables implies the second set is ''redundant'' upon knowing the first set. The examination of redundancy coefficients is either individually or pooled across canonical functions. Table 3 contains the redundancy coefficients and pooled redundancy coefficients for the significant canonical functions in each of the pair wise comparisons.

Table 3. Redundancy (R_d) and pooled redundancy ($\sum R_d$)

Function:	**1**	**2**	**3**
Panel A: Altman and Zavgren			
Altman			
R_d	0.19	0.16	0.13
$\sum R_d$	0.19	0.35	0.48
Zavgren			
R_d	0.16	0.12	0.07
$\sum R_d$	0.16	0.29	0.36
Panel B: Altman and Ohlson			
Altman			
R_d	0.20		
$\sum R_d$	0.20		
Ohlson			
R_d	0.19		
$\sum R_d$	0.19		
Panel C: Altman and Deakin			
Altman			
R_d	0.26	0.17	0.10
$\sum R_d$	0.26	0.43	0.52
Deakin			
R_d	0.16	0.05	0.07
$\sum R_d$	0.16	0.21	0.29
Panel D: Zavgren and Ohlson			
Zavgren			
R_d	0.11	0.10	
$\sum R_d$	0.11	0.20	
Ohlson			
R_d	0.15	0.07	
$\sum R_d$	0.15	0.22	
PanelE: Zavgren and Deakin			
Zavgren			
R_d	0.23	0.10	
$\sum R_d$	0.23	0.32	
Deakin			
R_d	0.13	0.06	
$\sum R_d$	0..13	0.19	
PanelF:Ohlson and Deakin			
Ohlson			
R_d	0.22		
$\sum R_d$	0.22		
Deakin			
R_d	0.31		
$\sum R_d$	0.31		

The only pair wise comparison that suggests substantial redundancy after the deletion of common variables is Altman vs Deakin and Altman vs Zavgren. Only one variable set, however, in each of these pairings suggests substantial redundancy.

The pooled redundancy coefficient of the Deakin set with respect to the Altman set is 0.29.The variable set employed by Deakin is not similar to the variable set employed by Altman. The Altman variable set is not a good predictor of the Deakin variable set. On the other hand, the pooled redundancy coefficient of the Altman variable set with respect to the Deakin variable set is still moderately high at 0.52. A large part of the Altman variable set is redundant to the Deakin variable set after dropping common variables. Much of the information contained in the Altman variable set is present in the Deakin variable set. Likewise, the pooled redundancy coefficient of the Altman variable set with respect to the Zavgren variable set is moderately high at 0.48 after deletion of common variables. The relative size of the Deakin and Zavgren variable sets allows them to capture the information contained in the smaller Altman variable set. With this one exception, the remaining variable sets do not contain much common information.

5. Variable deletion

The results of canonical analysis indicate that the pooled redundancy coefficient of the Altman set with respect to the Deakin variates is 0.52, the Deakin canonical variates account for 52 per cent of the variability among the Altman variables. One conclusion is the Deakin variable set would be a good predictor of the Altman variable set. To determine the common information between these two variable sets, one of the different deletion methods are delineated in the paper. To illustrate the deletion process, the results of full canonical analysis are compiled in Table 4.

Table 4: Initial Solution with Canonical Communality Coefficients Deletion Strategy

Variable statistic	FUNCTION1			FUNCTION2			FUNCTION1			h^2
	FUNC	r_s	r_s^2	FUNC	r_s	r_s^2	FUNC	r_s	r_s^2	
Sales / total assets	0.103	0,226	5.11%	-0.979	-0.963	92.74%	0.231	0.147	2.16%	100.0%
Market value of equity / book value of debt	0.032	0,095	0.90%	0.014	-0.042	0.18%	-0.052	-0.185	3.42%	4.50%
Earnings before interest and taxes / total assets	0.957	0,99	98.01%	0.27	0.119	1.42%	0.295	0.064	0.41%	99.84%
Retained earnings / total assets	0.076	0,334	11.16%	-0.141	-0.182	3.31%	-1.016	-0.923	85.19%	99.66%
Working capital / total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
adequacy			28.79%			24.41%			22.80%	
RD			26.20%			16.55%			9.53%	
$\sum R_d$			26.20%			42.75%			52.28%	
Rc2			91.00%			67.80%			41.80%	
RD			15.97%			5.29%			7.32%	
$\sum R_d$			15.97%			21.26%			28.57%	
adequacy			17.55%			7.80%			17.51%	
Quick assets/total assets	1.467	0.929	86.30%	0.783	0.166	2.76%	-0.316	0.212	4.49%	93.55%
Quick assets/current debts	-0.043	0.1	1.00%	0.083	-0.094	0.88%	-0.028	-0.546	29.81%	31.70%
Net income/total assets	0.066	0.538	28.94%	-0.076	-0.342	11.70%	-0.144	-0.574	32.95%	73.59%
Working capital/total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
Total debt / total assets	0.096	-0.196	3.84%	-0.251	-0.088	0.77%	0.275	0.564	31.81%	36.43%
Cash flow/total debts	0.282	0.539	29.05%	-0.065	-0.263	6.92%	-0.512	-0.59	34.81%	70.78%
Working capital/ sales	0.091	0.079	0.62%	0.026	-0.157	2.46%	-0.254	-0.547	29.92%	33.01%
Quick assets/sales	-0.633	0.792	62.73%	-0.546	0.332	11.02%	0.748	0.227	5.15%	78.90%
Current assets/total assets	-0.19	-0.03	0.09%	0.398	-0.237	5.62%	0.021	-0.346	11.97%	17.68%
Current assets/sales	0.034	-0.269	7.24%	-0.149	0.434	18.84%	0.027	-0.116	1.35%	27.42%
Current assets/current debts	0.168	-0.03	0.09%	-0.539	-0.072	0.52%	-0.357	-0.566	32.04%	32.64%
Cash/total assets	0.164	0216	4.67%	-2.721	-0.51	26.01%	-0.048	-0.168	2.82%	33.50%
Cash/sales	-0.027	0.066	0.44%	1.451	0.059	0.35%	-0.839	-0.259	6.71%	7.49%
Cash/current debts	-0.325	0.176	3.10%	0.969	-0.368	13.54%	1.105	-0.194	3.76%	20.40%

According to Humphries-Wadsworth (1998), canonical correlation analysis is a “rich tool for examining the multiple dimensions of the synthetic variable relationships” (p. 6). In addition to the standardized function coefficients and structure coefficients, three other coefficients are often examined and can facilitate interpretation: canonical communality coefficients, canonical adequacy coefficients, and canonical redundancy coefficients, however, see Roberts (1999) for discussion of the inadequacies of redundancy coefficients.

During the deletion process three coefficients will be consulted:

- r_s^2 - squared canonical structure coefficient - how much variance a variable linearly shares with a canonical variate (Thompson, 1984).
- h^2 – canonical communality coefficients - sum of all r_s^2; how much of the variance in a given observed variable is reproduced by the complete canonical solution (Thompson, 1991).
- Rc^2 - squared canonical coefficient– how much each function is contributing to the overall canonical solution (Thompson, 1991).

In this deletion strategy looked at the h^2s only. The process involved the following steps:

1. Look at all the h^2s
2. Find the lowest h^2 and delete the corresponding variable
3. Rerun the CCA and recalculate the h^2s
4. Check the change to the Rc^2 for each function
5. If there is little change to Rc^2 , find the next lowest h^2
6. Delete the variable with the corresponding lowest h^2 and repeat the process until the Rc^2 change is too great by researcher judgment.

Looking at Table 4, the predictor variables with the lowest h^2s was cash/sales (7.49%). This variable was quite a bit lower than the other twelve-predictor variables that ranged from 17.68% to 93.55%. Through this variable deletion strategy, the variable with the lowest h^2, cash/sales, was dropped first.

Table 5: Canonical Solution After Dropping Cash/Sales Based on Canonical Communality Coefficients Deletion Strategy

Variable statistic	FUNCTION1			FUNCTION2			FUNCTION1			h^2
	FUNC	r_s	r_s^2	FUNC	r_s	r_s^2	FUNC	r_s	r_s^2	
Sales / total assets	0.092	0.216	4.67%	-0.68	-0.714	50.98%	0.739	0.66	43.56%	99.21%
Market value of equity / book value of debt	0.032	0.095	0.90%	0.005	-0.118	1.39%	-0.142	-0.217	4.71%	7.00%
Earnings before interest and taxes / total assets	0.96	0.991	98.21%	0.381	0.127	1.61%	0.102	-0.004	0.00%	99.82%
Retained earnings / total assets	0.077	0.334	11.16%	-0.693	-0.673	45.29%	-0.74	-0.652	42.51%	98.96%
Working capital / total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
adequacy			28.73%			24.82%			22.70%	
RD			26.15%			12.96%			6.88%	
$\sum R_d$			26.15%			39.10%			45.98%	
Rc2			91.00%			52.20%			30.30%	
RD			15.91%			8.81%			3.58%	
$\sum R_d$			15.91%			24.72%			28.31%	
adequacy			17.49%			16.88%			11.82%	
Quick assets/total assets	1.486	0.93	86.49%	-0.244	0.254	6.45%	0.776	0.081	0.66%	93.60%
Quick assets/current debts	-0.043	0.1	1.00%	0.09	-0.359	12.89%	-0.19	-0.48	23.04%	36.93%
Net income/total assets	0.065	0.536	28.73%	-0.144	-0.616	37.95%	-0.07	-0.263	6.92%	73.59%
Working capital/total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
Total debt / total assets	0.094	-0.198	3.92%	-0.138	0.199	3.96%	0.561	0.639	40.83%	48.71%
Cash flow/total debts	0.283	0.537	28.84%	-0.345	-0.549	30.14%	-0.384	-0.35	12.25%	71.23%
Working capital/ sales	0.095	0.078	0.61%	-0.345	-0.425	18.06%	0.227	-0.393	15.44%	34.12%
Quick assets/sales	-0.652	0.794	63.04%	0.758	0.418	17.47%	-0.711	-0.041	0.17%	80.68%
Current assets/total assets	-0.186	-0.032	0.10%	0.347	-0.401	16.08%	-0.208	-0.117	1.37%	17.55%
Current assets/sales	0.029	-0.265	7.02%	0.097	0.348	12.11%	-0.31	-0.455	20.70%	39.84%
Current assets/current debts	0.162	-0.029	0.08%	-0.533	0.352	12.39%	-0.196	-0.488	23.81%	36.29%
Cash/total assets	0.128	0.212	4.49%	-1.753	-0.567	32.15%	0.508	0.268	7.18%	43.83%
Cash/sales	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
Cash/current debts	-0.319	0.173	2.99%	1.567	-0.445	19.80%	-0.016	0.116	1.35%	24.14%

Table 5 shows the canonical analysis after cash/sales was dropped. The Rc^2s were then examined for each function and there was only slight change. Function 1 did not change, Function 2 went from 67.8% to 52.2%, and Function 3 went from 41. 8% to 30.3%.

Table 6: Canonical Solution After Dropping Cash/Sales and Current Assets/Total Assets Based on Canonical Communality Coefficients Deletion Strategy

Variable statistic	FUNCTION1			FUNCTION2			FUNCTION1			h^2
	FUNC	r_s	r_s^2	FUNC	r_s	r_s^2	FUNC	r_s	r_s^2	
Sales / total assets	0.064	0.187	3.50%	-0.65	-0.69	47.61%	0.76	0.683	46.65%	97.75%
Market value of equity / book value of debt	0.021	0.081	0.66%	0.038	-0.091	0.83%	-0.189	-0.254	6.54%	7.93%
Earnings before interest and taxes / total assets	0.97	0.995	99.00%	0.355	0.095	0.90%	0.098	0.002	0.00%	99.90%
Retained earnings / total assets	0.065	0.32	10.24%	-0.732	-0.711	50.55%	-0.701	-0.618	38.19%	98.98%
Working capital / total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
adequacy			28.35%			24.97%			22.82%	
RD			25.57%			12.61%			6.82%	
$\sum R_d$			25.57%			38.18%			45.01%	
Rc2			90.20%			50.50%			29.90%	
RD			15.76%			8.37%			3.40%	
$\sum R_d$			15.76%			24.13%			27.53%	
adequacy			17.47%			16.37%			11.38%	
Quick assets/total assets	1.0351	0.939	88.17%	-0.022	0.222	4.93%	0.641	0.085	.72%	93.82%
Quick assets/current debts	-0.048	0.092	0.85%	0.115	-0.373	13.91%	-0.232	-0.478	22.85%	37.61%
Net income/total assets	0.087	0.525	27.56%	-0.202	-0.655	42.90%	-0.027	-0.225	5.06%	75.53%
Working capital/total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
Total debt / total assets	0.03	-0.196	3.84%	-0.011	0.223	4.97%	0.498	0.637	40.58%	49.39%
Cash flow/total debts	0.251	0.528	27.88%	-0.321	-0.587	34.46%	-0.393	-0.319	10.18%	72.51%
Working capital/ sales	0.028	0.071	0.50%	-0.225	-0.448	20.07%	0.16	-0.369	13.62%	34.19%
Quick assets/sales	-0.493	0.808	65.29%	0.491	0.389	15.13%	-0.571	-0.044	0.19%	80.61%
Current assets/total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
Current assets/sales	0.021	-0.253	6.40%	0.106	0.341	11.63%	-0.316	-0.464	21.53%	39.56%
Current assets/current debts	0.049	-0.035	0.12%	-0.363	-0.372	13.84%	-0.27	-0.472	22.28%	36.24%
Cash/total assets	-0.062	0.198	3.92%	-1.502	-0.576	33.18%	0.465	0.3	9.00%	46.10%
Cash/sales	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
Cash/current debts	-0.13	0.162	2.62%	1.323	-0.452	20.43%	0.044	0.138	1.90%	24.96%

The remaining canonical solution still contained current assets/total assets with an h^2 of 17.55%. That variable was considerably lower than the other variables in Table 5, therefore, current assets/total assets was dropped and little change was seen in the Rc^2s of each function as shown in Table 6. Function 1 changed from 91.0% to 90.2%, Function 2 changed from 52.2% to 50.5%, and Function 3 changed from 30.3%to 29.9%.

Table 7: Final Canonical Solution after Dropping Cash/Sales, Current Assets/Total Assets and Cash/Current Debts Based on Canonical Communality Coefficients Deletion Strategy

Variable statistic	FUNCTION1			FUNCTION2			FUNCTION1			h^2
	FUNC	r_s	r_s^2	FUNC	r_s	r_s^2	FUNC	r_s	r_s^2	
Sales / total assets	0.058	0.181	3.28%	0.663	0.705	49.70%	-0.757	-0.679	46.10%	99.08%
Market value of equity / book value of debt	0.026	0.084	0.71%	0.043	0.168	2.82%	0.191	0.259	6.71%	10.23%
Earnings before interest and taxes / total assets	0.974	0.996	99.20%	-0.346	-0.087	0.76%	-0.099	-0.002	0.00%	99.59%
Retained earnings / total assets	0.057	0.313	9.8%	0.706	0.701	49.14%	0.703	0.621	38.56%	97.50%
Working capital / total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
adequacy			28.25%			25.61%			22.84%	

RD			25.45%			12.14%			6.83%	
$\sum R_d$			25.45%			37.59%			44.42%	
Rc2			90.10%			47.40%			29.90%	
RD			15.54%			7.74%			3.37%	
$\sum R_d$			15.54%			23.28%			26.65%	
adequacy			17.25%			16.32%			11.29%	
Quick assets/total assets	1.365	0.941	88.55%	0.197	-0.217	4.71%	-0.634	-0.086	0.74%	94.00%
Quick assets/current debts	-0.039	0.09	.81%	-0.002	0.404	16.32%	0.237	0.481	23.14%	40.27%
Net income/total assets	0.098	0.521	27.14%	0.344	0.682	46.51%	0.033	0.229	5.24%	78.90%
Working capital/total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
Total debt / total assets	0.042	-0.195	3.80%	0.147	-0.241	5.81%	-0.493	-0.639	40.83%	50.44%
Cash flow/total debts	0.239	0.524	27.46%	0.225	0.615	37.82%	0.39	0.323	10.43%	75.71%
Working capital/ sales	0.051	0.068	0.46%	0.51	0.456	20.79%	-0.148	0.371	13.76%	35.02%
Quick assets/sales	-0.507	0.811	65.77%	-0.72	-0.395	15.60%	0.56	0.042	0.18%	81.55%
Current assets/total assets	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
Current assets/sales	0.019	-0.252	6.35%	-0.158	-0.379	14.36%	0.313	0.46	21.16%	41.87%
Current assets/current debts	0.006	-0.038	0.14%	-0.095	0.378	14.29%	0.254	.474	22.47%	36.90%
Cash/total assets	-0.185	0,194	3.76%	0.274	0.6	36.00%	-0.505	-0.296	8.76%	48.53%
Cash/sales	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%
Cash/current debts	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00	0.00	0.00%	0.00%

The remaining canonical solution still contained cash/current debts with an h^2 of 24.96%. That variable was lower than the other variables in Table 6, therefore, cash/current debts was dropped and little change was seen in the Rc^2s of each function in Table 7. Function 1 did not change, Function 2 went from 50.5% to 47.40 %, and Function 3 remained the same.

6. Conclusion

Variable selection in financial distress studies is commonly based upon suggestions in the literature, the success of variables in earlier studies, or the selection of a large set of variables with an accompanying data reduction procedure in order to maximize predictive ability or some other statistical criteria. The resulting consequence of ad hoc variable selection in financial distress studies is that consensus does not exist on a definitive set of variables that distinguish between distressed and non-distressed firms. If seemingly different variable sets exhibit a strong relationship and alternative variable sets can predict each other, then heterogeneous variable sets capture common information.

Despite the long-standing criticism directed towards variable selection in financial distress studies, no research has so far examined the relationship between alternative variable sets used in financial distress studies. Therefore, the main purpose of the present study was to employ canonical correlation analysis in order to examine the relationships that exist between variable sets employed in four bankruptcy studies and then to illustrate the value of applying the law of parsimony to canonical correlation analysis solutions.

Among four variable sets and six pair wise comparisons which are Altman versus Zavgren, Altman versus Deakin, Altman versus Ohlson, Zavgren versus Deakin, Zavgren versus Ohlson, and Deakin versus Ohlson, only two pooled redundancy coefficients indicated more similarity than dissimilarity. The relative size of both variable sets explains, in part, the fact that much of the Altman set is redundant to the Deakin set and Zavgren set .As the relative size of one variable set increases, spurious correlations serve to increase the pooled redundancy coefficient of the other variable set. The smaller size of the Altman variable set suggests cautiously interpreting the partial redundancy of both variable sets.

To determine the common information between these two variable sets, Altman and Deakin, one of the different deletion methods were delineated in the paper. The primary purpose of parsimony is that the more parsimonious the solution is, the more replicable the model will be. The goal of parsimony was achieved by removing the three variables and only a very small change was noted in either the communality coefficients or the squared canonical coefficients of each function.

Acknowledgement

This article is based on the research plan, "Applying Variable Deletion Strategies in Bankruptcy Studies to Capture Common Information and Increase Their Reality ", that profited from advice, encouragement and support of Central Tehran Branch, Islamic Azad University, Tehran, Iran.

References

1. Altman, E. 1968, ‘Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy’, *Journal of Finance*, vol. 23, no. 4, pp. 587-690.
2. Ball, R. & Foster, G. 1982, ‘Corporate financial reporting: a methodological review of empirical research’, *Journal of Accounting Research*, vol. 20, supplement, pp. 161-234.
3. Beaver, W. 1966, ‘Financial ratios as predictors of failure’, *Journal of Accounting Research*, vol. 4, supplement, pp. 71-111.
4. Beaver, W. 1968, ‘Alternative accounting measures as predictors of failure’, *The Accounting Review*, vol. 43, no. 1, pp. 113-122.
5. Deakin, E.B. 1972, ‘A discriminant analysis of predictors of business failure’, *Journal of Accounting Research*, vol. 10, no. 1, pp. 167-179.
6. Foster, G. 1986, *Financial Statement Analysis*, Prentice-Hall, New York.
7. Humphries-Wadsworth, T. 1998, *Features of published analyses of canonical results,* Paper presented at annual meeting of the American Educational Research Association, San Diego, CA, (ERIC Document Reproduction Service No. ED 418 125).
8. Knapp, T.R. 1978, ‘Canonical correlation analysis: A general parametric significance-testing system’, *Psychological Bulletin*, vol. 85, pp. 410-416.
9. Lambert, Z.V., Wildt, A.R. & Durand, R.M. 1988, ‘Redundancy analysis: an alternative to canonical correlation and multivariate multiple regression in exploring interest associations’, *Psychological Bulletin*, vol. 104, pp. 282-289.
10. Leclere, M. (2006), ‘‘Bankruptcy studies and ad hoc variable selection: a canonical correlation analysis’’, Review of Accounting and Finance, Vol. 5 No. 4, pp. 410-422.
11. Ohlson, J.A. 1980, ‘Financial ratios and the probabilistic prediction of bankruptcy’, *Journal of Accounting Research*, vol. 18, no. 1, pp. 109-131.
12. Pedhazur, E. J. 1997, ‘Canonical and discriminant analysis, and multivariate analysis of variance’, *Multiple regression in behavior research*, 3rd ed., pp. 924-979, Fort Worth, TX: Harcourt Brace.
13. Roberts, J.K. 1999, ‘Canonical redundancy (Rd) coefficients, they should (almost never) be computed and interpreted’, In B. Thompson (Ed.), *Advances in social science methodology*, vol. 5, pp. 331-333, Stamford, CT: J.A.I. Press.
14. Stewart, D. & Love, W. 1968, ‘A General Canonical Correlation Index’, *Psychological Bulletin*, vol. 70, no. 3, pp. 160-163.
15. Thompson, B. 1984, *Canonical Correlation Analysis*, Sage Publications, Thousand Oaks, CA.
16. Thompson, B. 1991, Methods, plainly speaking: A primer on the logic and use of canonical correlation analysis. *Measurement and Evaluation in Counseling and Development, 24* (2), 80-93.
17. Zavgren, C. 1983, ‘Assessing the Vulnerability to Failure of American Industrial Frms: a Logistic Analysis’, *Journal of Banking and Finance*, vol. 7, pp. 19-45.

1/8/2013