

財務危機預警模式之再探討

Reflections on Financial Distress Prediction Model

薛兆亨與李羿儒

國立高雄第一科技大學財務管理系助理教授(weissor@ccms.nkfust.edu.tw)

國立高雄第一科技大學財務管理系研究生(u9126805@ccms.nkfust.edu.tw)

摘要

財務危機預警模式之再探討

不同以往文獻著重於各方法之比較與探討，本文之研究重心關注在架構財務危機預警模式之三項議題，一為抽樣樣本誤差，二為財務變數之選擇，三為財務預警模式持續性；並使用日前新興起分類方法支撐向量機(SVM)及 Logit 迴歸為架構財務危機預警模式之方法，並加以比較之。變數選取方面以前向法略優於自選法，由 Logit 迴歸結果，以負債比率及營業利益率為最能偵測公司危機之預測變數，抽樣樣本誤差隨著隨機抽樣樣本數增加而增加，在樣本設計上，不宜選取過高正常公司樣本數，且不一定須進行樣本配對之設計，在驗證期間時，發現本文之模式持續性尚佳。在比較此兩種方式，以型一錯誤率低及正確率高二項原則為標準，支撐向量機(SVM)及前向法所建構之模式表現最佳。

關鍵詞：財務危機、支撐向量機、邏輯斯迴歸、抽樣樣本誤差

Abstract

Reflections on Financial Distress Prediction Model

This paper examines three issues-choice-based sample bias, predictor selecting, and consistency- in financial distress prediction. Support Vector Machine and Logit regression are proposed to predict financial distress. The selected techniques are Altman analysis and forward Wald selection method. In empirical results, the most important determinants of financial distress are leverage ratio and operating profit ratio. The choice-based sample bias is statistically significant. The predict results doesn't decreasing following time passing. The model with support vector machines and forward Wald selection method produces superior results for bankrupt companies.

Keywords: Financial distress; Choice-Based Sample Bias; Support Vector Machines; Logit regression

1 緒論

架構財務危機預警模式為 30 幾年來廣為學者所研究，自 Beaver(1966)以單變量之分類檢定法(dischotomous classification Test)求出最佳分類點事先預測危機企業及正常企業，著名 Z-score(Altman,1968)模式開始以多個財務比率進行預警模式之建構，Altman 建構預警模式所採用方法為區別分析 (Multiple Discriminant analysis)，然而區別分析包括嚴格資料分配之假設 (Ohlson,1980)，隨後，架構財務危機預警模式之方法包括 Logit 迴歸分析 (Ohlson, 1980)及 Probit analysis (Zmijewski, 1984)，且採用這兩種方法皆較採區別分析為優 (Collins and Green, 1982)，而 1990 年代，類神經網路亦應用於財務危機之預測(Altman et al., 1994)，類神經網路之優勢對於樣本資料之分配及統計性質並無嚴格要求，以往之研究，多於在預測方法之改善，以何種方式進行預測準確率最高。然而在財務危機模式建構之研究，除了使用何項方法，可以正確預測危機企業及正常企業外亦有數項研究主題，值得深入探討，分別說明如下：

主題一：抽樣樣本之偏差 (Choice-Based Sample Bias)

在財務危機預警模式之實證研究，正常公司之選取多採成對樣本(paired-sample)方式 (Beaver, 1966; Altman, 1968; Jo and Han, 1996)，採用成對樣本方式會導致過度估計危機公司被正確分類之比例 (Zmijewski, 1984；Platt and Platt, 2002)，因此在實證樣本之設計上，即有兩種方式，配對樣本方式可避免因危機公司樣本數過少而無法建構出區別率高的預警模式，但是亦有可能導致估計偏差，因此本文以成對樣本方式及隨機(Random)樣本方式進行樣本資料之選取，以檢驗樣本設計方式之優劣。

主題二：財務比率變數之選取方式

本文建構財務危機預警模式所採用之變數為財務比率變數 (Beaver, 1966; Altman, 1968; Ohlson, 1980; Zmijewski, 1984; Altman et al, 1994; Platt and Platt, 2002)，而在樣本選取方式上，回顧以往研究文獻，大致分為二項方式，一為以財務報表分析構面，再經由適當統計檢定，及考量財務比率變數與財務危機關係，選取特定變數 (Altman, 1968; Ohlson, 1980; Zmijewski, 1984; Altman et al., 1994; Lennox, 1999; Platt and Platt, 2002)，第二種方式為以

統計方法或人工智慧方法等進行財務比率變數的選取，統計方法採如多元逐步迴歸法、後向選取或是前向選取法，人工智慧方法有基因演算法(Genetic algorithm)等 (Back et al., 1996)。何種方法可以建構正確率較高的預警模式，卻是甚少在以往研究文獻中予以探討之，因此本文亦會深入研究。

主題三：預警模式持續性之研究

以危機前一年資料進行財務危機模式之架構是最具準確性地 (Altman et al., 1994)，然而目前所建構財務危機預警模式，是否可以提供危機下一年年繼續使用，其持續性如何，甚少研究進行說明，財務危機預警模式可否經得起時間的考驗？因此本文在以延長驗證年度方式，對財務預警模式持續性進行研究。

除上面三項研究主題外，類神經網路早在 1980 年代初期時就已經被應用在商業方面的問題上，例如：債券的評等、銀行或存款機構失敗預測、投資預測、企業失敗或財務危機預測，而且許多的研究也指出類神經網路在其他的問題上有優越表現，從過去的研究也顯示類神經網路較其他的歸納學習法佳(Jo and Han, 1996; Altman et al., 1994)。然而 Van Gestel et al. (2003) 發現支撐向量機 (support vector machine approach, 以下簡稱 SVM) 在預測企業之財務危機正確率比上述其他方法高，因此本文亦研究使用支撐向量機 (SVM) 建構財務危機預警模式，並與最為廣泛使用之 Logit 迴歸分析比較，並將考慮上述三項研究主題，進行實證過程之設計並加以檢定之。

2 文獻探討

Beaver (1966) 為研究財務危機之先趨，主要研究方法為採用 Dichotomous Prediction Test，為要進行預測，須先將預測變數之樣本資料進行升幕排序，進行檢視以取得最適臨界點，臨界點之決定在於最小化錯誤預測百分比，若財務比率低於此臨界點則該公司被區分為財務危機組，初始選取 30 項財務比率，最佳預測財務比率為現金流量對總負債比，這篇文章是率先以財務比率進行財務危機公司之判斷，而亦開啓日後財務危警研究之大門。

Altman (1978) 以多元區別分析(Multiple Discriminant analysis, 以下簡稱 MDA) 建構著名 Z-Score 模式，MDA 為依樣本資料之特性區別為數個群組之統計方法，其優點為考量相

近群組公司之特性及其交互影響，而 Beaver 所提出之單變量分析只考量單一變數；Z-score 模式公式如下：

$$Z=0.012X_1+0.014X_2+0.033X_3+0.006X_4+0.999X_5$$

X_1 = 營運資金對總資產比

X_2 = 保留盈餘對總資產比

X_3 = 稅前息前盈餘對總資產比

X_4 = 股東權益市值/負債帳面值

X_5 = 銷售對總資產

若 Zscores 大於 2.99 則為財務體質健全公司，若低於 1.81 則為財務危機公司，而在 1.81 及 2.99 之區間則為灰色領域，則進一步另行分析之。

Ohlson (1980) 則考量 MDA 之限制提出以 “condition logit model” 方法架構預警模型，由於 MDA 須確認各群組間預測變數之變異數—共變異數矩陣須相同，樣本資料須為常態分配，而採用 logit 分析法並不需要上述假設。Logit 分析法提供公司發生財務危機之機率，取一 “臨界機率” 並判斷是否為財務危機公司，係數之決定是採用最大化 log-likelihood 函數，Logit 分析法日後亦為被廣為使用，例如 Platt et al.(1990, 2002)。

Zmijewski (1984) 使用二元 probit 法 (bivariate probit model) 進行財務危機預警模式之建構，probit 法與 logit 法最大之差異在於轉換函數不同，Probit 為服從常態的累積分配函數，而 logit 是基於 logistic 分配之累積分配函數，其重要之財務比率預測變數，包括資產報酬率、負債比率 (槓桿作用) 及流動比率，並考慮抽樣樣本誤差，其模型之正確率隨著樣本數增加而隨之增加，由 92.6% 至 98.2%。

Odom and Sharda (1990) 以 Altman(1968) 提出之五個財務比率變數為變數，應用類神經網路 (neural networks) 於財務危機預警模式之研究，類神經網路為人工智慧方法，模擬人腦類神經元之運作所推演之數學運算方式，其特點為學習，而其權重係數 (Weighting coefficient) 須不斷進行調整，直到能適合每個樣本，本文之類神經網路架構一隱藏層包含五個節點，輸出層包括一節點，其研究期間 1975 年至 1982 年包括 65 個財務危機公司，64 家財務正常公司，74 家公司為訓練樣本，55 家公司是測試樣本，與區別分析相比較，測試

期間類神經之正確率（81.81%）是高於MDA(74.28%)。

Altman, Marco and Varetto (1994)比較區別分析及類神經網路 (neural networks, 以下簡稱 NN) 應用在財務危機之區別效果，在本文中，其不單只是考慮正確率，亦考量到類神經網路及區別分析之性質，認為須特別注意 NN 之黑箱 (black-box) 之問題，包括不具邏輯性 (illogical) 之權重及在訓練樣本過度正確 (overfitting) 之現象，所以 NN 並不全然是優於 MDA，而兩者之正確率皆超過 90%。

Back, Laitinen, and Sere (1996) 採用類神經網路應用於財務危機之研究，其在選擇預測變數則採用三種方式，分別為區別分析、logit 分析及基因演算法 (genetic algorithms)，基因演算法由達爾文適者生存之原理發展而來，藉由基因之基本運算子在每代間進行演化，以搜尋至適當問題之最佳解。選擇預測變數以基因演算法配合類神經網路之財務預警模式結果最佳。

Yang, Platt and Platt (1999) 應用機率型類神經網路 (Probabilistic Neural Network, 以下簡稱 PNN) 建構財務危機預警模式，將 PNN 與區別分析、倒傳遞網路 (back-propagation NN, 以下簡稱 BPN) 比較之，機率型類神經網路是由 D. E. Specht 提出的一種監督式 (supervised) 類神經網路架構，為三層 (three layered) 前饋式 (feedforward) 類神經網路，其主要特點在於它具有從一組訓練資料中快速學習的能力當有足夠的訓練資料時，PNN 已被證明將漸近地收斂至貝氏分類器 (Bayesian classifier)，而本文作者提出 PNN 相較於 BPN 而言，其較不呈現“黑箱作業”之現象，為另一項優點，以型一錯誤最低為判斷準則下，則以區別分析之結果最佳。

Lennox (1999) 比較三項統計方法，分別為區別分析、Probit 及 logit 模型，以 1987 年至 1994 年英國 949 家上市公司為樣本公司，其認為最重要預測變數為獲利能力、財務槓桿、現金流量、規模、產業及經濟景氣循環，而 logit 及 Probit 模型在辨認財務危機公司之正確性是優於區別分析法。

Fan and Palaniswami (2000) 應用 SVM 於財務危機之預測，支撐向量機為建構以結構風險最小化之最適區隔界面，使用線性規畫求得單一最佳解；並與三種方法比較，分別為區別分析、多層感知機 (Multi-layer perception) 及 Learning vector quantization (LVQ)，SVM

得到最佳預測結果 (70.97%)，次之則為多層感知機 (類神經網路) 之正確率為 66.11%，LVQ 之正確率為 62.71%，而之表現 MDA 最差，正確率為 61.39%。Van Gestel et al. (2003) 應用 SVM 於財務危機之預測，並與區別分析及 logit 分析比較，而其與 Fan et al. (2000) 不同之處在於其內部之 kernel 為 Fisher Discriminant Analysis，稱之為 LS-SVM，其實證結果發現 LS-SVM 優於區別分析及 logit 分析法。

Shin and Lee (2002) 運用基因演算法 (genetic algorithm) 萃取判斷規則及門檻值 (thresholds)，釋例如下：IF [Debt ratio > 1.5 and Quick ratio < 0.35] then Dangerous，使用者可輕易了解判斷理由類似於專家系統，研究期間為 1995 年 1997 年 528 家公司包括 264 家財務危機公司，90% 為訓練樣本，而 10% 為驗證樣本，其選取變數的方式採取二項步驟，一為因素分析 (factor analysis) 選取九項變數後，第二項步驟為採逐步分析法及單變量分析決定其臨界值，並萃取 5 項判斷規則，判斷正確率在訓練及測試樣本皆為 80% 左右。Kim and Han (2003) 亦採用基因演算法進行財務危機預測，其與 Shin and Lee (2002) 不同在於考量專家在風險衡量過程所使用問題解決知識 (problem-solving knowledge)，利用基因演算法萃取專家之知識，並與歸納學習法 (inductive learning methods) 及類神經網路相比較，基因演算法在正確率及涵蓋率上是優於歸納學習法及類神經網路。

上述文獻在於預測方法之改變，期以更適當方法建構財務危機預警模式，除以之外亦有其他學者探討其他相關議題，例如：抽樣樣本誤差 (choice-based sample biases) (Zmijewski, 1984)，考量產業因素之財務預警模式 (Platt and Platt, 1990, 1991) 等。

Zmijewski (1984) 討論抽樣樣本誤差這項議題時，認為以往研究以配對樣本方式為樣本設計之主要方法，即分為財務危機及正常公司兩組之後，再隨即以財務危機公司之樣本性進行正常公司之選取，此方式並不為隨機抽取樣本法，因此其使用 Probit 方法進行實證以釐清抽樣樣本誤差對樣本模型之影響，然而其實證結果認為未調整 probit 模型下確實發生誤差，但是在使用調整後 Probit (Weighted exogenous sample maximum likelihood probit) 可避免偏差之產生，且其認為不會影響統計之推論結果與整體模型判斷正確率，在使用調整後 Probit 預測模式中財務危機公司之預測準確性不高僅為 45% 至 55% 左右。Zmijewski 認為除非考量所有分配，否則估計係數將產生偏差，即造成財務危機公司正確估計率過度高

估的現象。Platt and Platt (2002)亦對此項議題進行探討，採用 logit 法進行實證，在驗證抽樣誤差時，其模擬 50 條 logit 迴歸式，以所估計之平均係數 $\hat{\beta}$ 與母體係數 (population parameter) β 比較之，即以 $(\hat{\beta} - \beta) / \sigma_{\hat{\beta}}$ 檢驗是否產生偏差。其取樣樣本不若 Zmijewski (1984) 採用全部產業，其僅採用單一產業為汽車上下游產業，其實證結果，整體模型正確率高達 98% 以上，所有樣本數須全部放入，否則將產生樣本偏差之現象，因此當危機對正常公司比率增加時，則樣本偏差亦會增加。

Platt and Platt (1990, 1991) 以依產業平均水準調整後之財務比率為預測變數，其認為產業調整過後之預測變數優於未調整過之產業變數，主要原因之一為其排除產業產業之影響，二為減低時間之影響，使得預測變數較為穩定，其預測正確率亦優於未經產業調整之模式。

3 研究設計及方法

3.1 研究流程

本文主要之研究研究架構分為五項程序說明如下。

1. 樣本資料進行產業水準調整

依產業標準化公司財務比率是較佳之預測指標 (Platt et al., 1990)，而公司個別之財務比率在不同產業間存在著相當差異，不利於跨產業財務危機預警模式之建構，因此本文參酌 Platt et al. (1990) 提出之方式進行財務比率之調整，公式為公司財務比率除以產業平均水準。

2. 進行研究變數選擇程序

Altman (1968) 選擇變數的過程，首先就依據過去文獻最常被列入的及與此研究有相當關連財務比率選出 22 個財務比率，然而 Altman 並非使用全部 22 個比率來做為預測財務危機之變數，其採用四項準則選出 5 項財務比率，建立 Z-score 模式。判斷變數之四項準則為一為觀察並檢定 (F 檢定及 t 檢定) 各個變數分配及統計性質，二為評估在各變數之間的相關性，三為觀察變數趨勢圖以了解其預測正確程度，四為由分析者

自行主觀判斷；由上面四項準則，決定 5 項財務比率，以建構財務危機預警模式。

本文對於變數之選擇採用二種方式，一為以 Altman 選取方式為主，稱為自選法，而近年來部分文章以多元逐步回歸或是人工智慧方法例如基因演算法(Back et al., 1996)等方式進行變數選取，本文之第二種方式為採用逐步選取法中之前向選取法進行變數選取。

3. 訓練及驗證資料整理

研究期間為 87 年至 92 年，以財務危機前一年為樣本資料，訓練期間為 3 年為 86 至 88 年，驗證期間為 3 年為 89 年至 91 年，此方式進行實證模式持續性之探討。

4. 抽樣樣本資料處理

訓練期間及驗證期間之抽樣方式，分為二項方式，一為成對配對方式，二為隨機抽樣方式，進行抽樣樣本誤差實證研究。

5. 架構財務危機預警模式及檢定

利用 SVM 及 Logit 迴歸架構財務危機預警模式，並進行驗證期間正確率之判斷及檢定。

3.2 樣本資料

本文之研究樣本以台灣證券交易所上市公司為主，主要原因其財報表資料之正確性及完整性較高，危機公司之定義依據台灣證券交易所營業細則第 49、50、50-1 條規定變更原有交易方法者，為了使財務危機發生時點能予以客觀認定，研究期間為 87 年至 92 年，危機公司共 72 家，扣除缺漏資料及其他產業內危機公司，則危機公司為 67 家，以往研究發現以財務危機發生前一年之資料建構財務危機預警模式正確率最高 (Beaver, 1966; Altman, 1968; Altman et al., 1994; Back et al., 1996)，因此本文亦以財務危機發生前一年之財務資料進行模式之建構，樣本公司之財務比率資料來源為台灣經濟新報之財務資料庫，財務危機公司係依據台灣證券交易所每月出版「證交資料」。

為了探討抽樣樣本誤差研究主題，本文將研究樣本設計二種方式，一為配對樣本，二為隨機抽樣，配對樣本方式為採相同產業及總資產規模相似二項原則，依 1:1 之比例予以

選取之；隨機抽樣方式，以危機公司之倍數進行抽樣，抽取 1 倍以跟配對進行比較之，另抽取 2 倍、4 倍、6 倍、8 倍及 10 倍進行抽樣樣本誤差之研究。

3.3 研究變數之定義

本文以財務比率為預測變數，過去文獻所選用 (Dimitras et al., 1996) 及與研究有相當關連之財務比率，優先選取 18 個財務比率，分為五個財務構面，分別為短期償債能力、長期償債能力、獲利能力、經營績效及現金流量方面。Platt and Platt (2002)認為財務危機與財務槓桿 (financial leverage) 成正相關，但是與獲利能力、短期償債能力、現金部份、成長力與經營績效成反比。在本文則認為財務危機與短期償債能力、獲利能力、經營績效與現金流量成反比，而與長期償債能力成正比，各比率公式列示於表 1。

表 1 研究變數操作定義

類別	代碼	變數名稱	公式	單位
短期償債能力	S1	1. 流動比率	流動資產/流動負債	%
	S2	2. 速動比率	(現金+短期投資+應收款項)/流動負債	%
	L1	1. 負債比率	負債總額/總資產	%
長期償債能力	L2	2. 權益比率	股東權益/總資產	%
	L3	3. 負債對股東權益比	總負債/股東權益	%
	L4	4. 長期資金適合率	(股東權益+長期負債)/固定資產+長期投資	%
	L5	5. 利息保障倍數	(稅前淨利+利息費用)/利息費用	%
	P1	1. 總資產報酬率(稅後息前)	營業利益/銷貨淨額	%
獲利能力	P2	2. 股東權益報酬率(稅後)	稅後純益/銷貨淨額	%
	P3	3. 營業利益率	(稅後純益+利息費用*(1-稅率))/平均總資產	%
	P4	4. 稅後純益率	稅後純益/平均股東權益	%
	T1	1. 總資產週轉率	銷貨淨額/平均總資產	次
經營績效	T2	2. 應收帳款週轉率	銷貨淨額/平均應收帳款	次
	T3	3. 存貨週轉率	銷貨淨額/平均存貨	次
	T4	4. 固定資產週轉率	銷貨淨額/平均固定資產	次
	T5	5. 股東權益週轉率	銷貨淨額/平均股東權益	次
	現金流量	C1	1. 現金流量比率	營業活動淨現金流量/流動負債
C2		2. 現金再投資比率	(營業活動淨現金流量-現金股利)/固定資產毛額+長期投資+其他資產+營運資金)	%

3.4 研究方法

3.4.1 資料特性檢定

首先針對依產業調整過之財務比率，進行常態檢定，有助於接續實證之進行，爲了瞭解財務危機公司與正常公司在財務比率之差異性，以利於財務變數之選擇，將進行 t 檢定或是 Mann-Whitney U 檢定，在進行 logit 迴歸模式時，亦需考慮各變數間共線性問題，因此在預測變數之選擇上，亦利用 Pearson 檢定或 Spearman 檢定進行相關性分析。

3.2 支撐向量機 (support vector machine)

SVM 爲解決分類問題而發展出來，是以統計學習運算理論爲基礎所發展之機器運算系統，由 Coter and Vapnik (1995) 率先提出，其爲減少錯誤風險之邊界 (Bound of the misclassification risk)。

在訓練樣本 $\{(x_i, d_i)\}_{i=1}^N$ 之下，爲 x_i 輸入樣本型態，爲 d_i 所對應目標輸出， d_i 爲二分類輸出，一組爲表示 $d_i = +1$ ，一組爲表示 $d_i = -1$ ，表示分隔超平面之方程式如式(1)：

$$w^T x + b = 0 \quad (1)$$

x 表示輸入向量， w 爲調整權重向量， b 則爲偏差項。

式(2)可清楚了解在樣本區分界面之關係。

$$\begin{aligned} w^T x + b &\geq 0 && \text{for } d_i = +1 \\ w^T x + b &< 0 && \text{for } d_i = -1 \end{aligned} \quad (2)$$

區別函數(discriminant function)如式(3)：

$$g(x) = w_0^T x + b_0 \quad (3)$$

從 x 至最適超平面之距離代數爲式(4)：

$$x = x_p + r \frac{w_0}{\|w_0\|} \quad (4)$$

爲了求出最適邊界，其爲最佳化之問題，在訓練樣本爲 $\{(x_i, d_i)\}_{i=1}^N$ ，找出一權重向量 w 及偏差 b 之最佳值，須滿足限制式即權重 w 最小化成本函數，以式(5)說明之：

$$d_i(w^T x_i + b) \geq 1 \quad \text{for } i=1,2,\dots,N$$

$$\Phi(w) = \frac{1}{2} w^T w \quad (5)$$

在同一樣本資料之平面上，可能發生資料點位於分隔界面內，或是資料點位於不同邊，因此需要針對式(3)進行調整，以式(6)表示之。

$$d_i(w^T x_i + b) \geq 1 - \xi \quad \text{for } i=1,2,\dots,N \quad (6)$$

ξ_i 稱為沈滯變數(slack variables)，式(6)同時定義最適直線邊界及非區別樣本，找到一最適分隔超平面及最小化其錯誤分類錯誤率是 SVM 之主要目標，而最小化權重 w 及錯誤分類率 ξ_i ，可被重新定義式(7)。

$$\begin{aligned} \Phi(\xi) &= \sum_{i=1}^N \xi_i \\ \Phi(w, \xi) &= \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i \end{aligned} \quad (7)$$

係數 C 為控制在支向機之複雜性與非區別樣本數之間的抵換 (trade-off) 關係，類似於“調節”係數 (regularization parameter)，是一項 user-specified 正向係數。

為了解決此最佳化之問題使用 Lagrange 乘數法(Lagrange multipliers)，為非區別型態之雙元問題 (dual problem)。

在訓練樣本 $\{(x_i, d_i)\}_{i=1}^N$ ，以 Lagrange 乘數 $\{\alpha_i\}_{i=1}^N$ 最大化目標函數。

$$\begin{aligned} Q(\alpha) &= \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j x_i^T x_j \\ \text{subject to (a)} & \sum_{i=1}^N \alpha_i d_i = 0 \quad \text{(b)} \quad 0 \leq \alpha_i \leq C \quad \text{for } i=1,2,\dots,N \end{aligned} \quad (8)$$

由 $K(x_i, x_j)$ (inner-product kernel) 代替 $x_i^T x_j$ 。

$$K = \{K(x_i, x_j)\}_{(i,j)=1}^N$$

式(8)可以被重寫為式(9)：

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j K(x_i, x_j) \quad (9)$$

核心部份 $K(x, x_i)$ 需要滿足 Mercer's 理論 (mercer's theorem)，一般常被使用三項核心部份，即選擇了何種非線性轉換方式，以映照至高維度之平面，三種核心方法介紹如下：

1. Polynomial learning machine

$$(x^T x_i + 1)^p$$

2. Radial-basis function

$$\exp\left(-\frac{1}{2\sigma^2} \|x - x_i\|^2\right)$$

3. Two-layer perceptron

$$\tanh(\beta_0 x^T x_i + \beta_1)$$

3.3 Logit 迴歸分析

Logit 迴歸之理論依據、模式參數估計方法及本文所使用之檢定進行說明，logit 迴歸從多元迴歸發展而來，主要為因應應變數為 2 分類變數 (dichotomous) 之情形，如式(10)。

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon_i = \alpha + \sum_{j=1}^p \beta_j X_j + \varepsilon_i \quad (10)$$

當 $Y=1$ 時當事件發生，當 $Y=0$ 表示事件未發生，則 ε_i 在不同情況下可由式(11)示之。

$$\begin{aligned} \varepsilon_i &= 1 - \alpha - \beta_1 X_1 - \beta_2 X_2 - \dots - \beta_j X_j & Y=1 \\ \varepsilon_i &= -\alpha - \beta_1 X_1 - \beta_2 X_2 - \dots - \beta_j X_j & Y=0 \end{aligned} \quad (11)$$

ε_i 不會趨近常態分配，式(10)並不適宜於兩元迴歸模式，故令 $E(Y_i | X_i) = \pi_i$ ，此時 $P(Y_i=1)$ ，表示 π_i 在給定 X_i 下之機率， X_i 極大值時，則 π_i 幾乎為 1，若 X_i 極小時，則 π_i 幾乎為 0，則 π_i 與 X_i 之間的關係曲線則為一 S 曲線，以式(12)表示之。

$$\pi(x) = \frac{\rho^{(\alpha + \beta_1 X_1 + \dots + \beta_j X_j)}}{1 + \rho^{(\alpha + \beta_1 X_1 + \dots + \beta_j X_j)}} \quad 0 \leq \pi_i \leq 1$$

$$Y = \pi(x) + \varepsilon \quad (12)$$

式(13)是由邏輯轉換(logistic transform)而來, 稱為 logit 迴歸模式, $\log(\frac{\pi}{1-\pi})$ 為 log odds ratio, $\pi=P(Y=1)$, $(1-\pi)=P(Y=0)$, 因此 $\frac{\pi}{1-\pi}$ 為此 2 項機率之比率, 稱為 odds ratio。在 logit 迴歸模式內之係數估計計算, 則利用最大概似法(maximum likelihood estimation, MLE), 以疊代方式得到最大概似估計值 α 及 β_i 。

$$\begin{aligned} \pi/(1-\pi) &= \exp(\alpha + \sum_{j=1}^p \beta_j X_j) \\ \log(\frac{\pi}{1-\pi}) &= \alpha + \sum_{j=1}^p \beta_j X_j \end{aligned} \quad (13)$$

在 logit 迴歸模式中利用 Wald 檢定去檢驗係數之顯著性, 以式(14)表示之。

$$W = (\hat{\beta}_k / SE_{\hat{\beta}_k})^2 \quad SE_{\hat{\beta}_k} \text{ 爲 } \hat{\beta}_k \text{ 之標準差} \quad (14)$$

虛無假設 $H_0 : \beta_k = 0$

3.4 前向選取法

選擇變數之第二個方法為前面選取法, 該方法源自多元迴歸, 在所有可能影響到應變數的自變數中, 根據所設之顯著標準將變數加入模式中或是模式中剔除, 而本文所選取前向選取法, 即以常數項模式為基礎, 將符合所定顯著水準之自變數一次一個加入模式內, 而顯著水準則 Wald 統計量為篩選標準。

4 實證結果

4.1 敘述統計

表 2 為樣本公司所分佈的產業狀況, 發生財務危機最多之產業為建設業, 次之為食品業, 再次之為電子業, 而食品業及鋼鐵業發生財務危機之家數亦超過五家, 財務危機發生年度則以 88 年最多, 而多集中分布在 89 年、90 年及 91 年。觀察每年財務危機公司佔上市公司總家數之比率, 多為 1.2%至 1.8%間, 所佔比率相當低, 以 87 年年發生比率最高, 92 年最低, 而扣除其他產業及缺漏資料後樣本資料則危機公司為 67 家。

表 2 樣本公司產業分佈表

財務危機		產業別													總計	上市公司	比率	
年	前	汽	觀	建	玻	食	紡	造	貿	塑	電	電	電	鋼	其	(a)	總家	(a/b)
度	一	車	光	設	璃	品	織	紙	易	膠	子	器	機	鐵	他	數	(b)	
	年				陶				百		電							
					瓷				貨		纜							
87	86			2		2		1			1			1	1	8	437	1.83%
88	87	1		2	1	2	2			2	2	1		3	1	17	462	1.73%
89	88		1	1		5			1		1		1	1		11	531	1.51%
90	89			1		2	5				3		1	1	2	15	584	1.37%
91	90			7	1		2					1		2		13	638	1.25%
92	91			2	1	1					3	1				8	669	1.20%
	總計	1	1	15	3	12	9	1	1	2	10	3	2	8	4	72	3321	0.24%

在表 3 為樣本敘述統計結果，流動比率在財務危機群低於產業標準，而財務正常群則高於產業標準，速動比率與流動比率呈現相同結果，負債比率在財務危機群則高於產業標準之 1.4 倍左右，而財務正常群低於產業標準，但是在股東權益比率及總資產週轉率方面財務危機群擁有較高股東權益比率，負債對股東權益比方面財務危機群高於產業標準甚多，長期資金適合率在財務危機群仍低於產業標準及財務正常群，表示其財務結構較不妥適，在利息保障倍數方面，不論是財務危機群或是財務正常群皆高於產業標準，惟財務危機群仍低於財務正常群甚多，總資產報酬率卻在財務危機群表現優於正常公司群，值得進一步探討，股東權益報酬率、營業利益率、應收帳款週轉率及存貨週轉率四項比率則在財務危機群呈現負值，表現不佳遠低於產業標準及財務正常群甚多，而在稅後淨利率方面，財務危機群與財務正常群較太大之差異，固定資產週轉率方面財務正常群表現優於財務危機群，在股東權益週轉率、現金流量比率及現金再投資比率方面，財務危機群表現皆劣於財務正常群，並低於產業標準甚多。

變數名	財務危機 (N=67)		財務正常 (N=67)		財務正常 (N=3158)	
	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
S1	0.634164	0.871973	1.318549	0.966875	1.3601159	1.1501107
S2	0.507666	1.226918	1.421739	1.53699	1.5547765	2.4766053
L1	1.449342	0.587074	0.989649	0.363056	0.975181	0.3871781
L2	2.06097	6.788322	1.009557	0.274709	1.0291548	0.2742057
L3	3.214008	4.434253	1.094868	0.63593	1.1784966	1.1600884
L4	0.894419	0.914002	4.390666	23.598901	1.9848258	6.3779446
L5	1.994493	19.94397	48.799737	510.716231	130.111144	1740.953287
P1	23.70026	93.357354	2.890468	13.435132	1.4565902	8.869226
P2	-6.035109	51.745639	1.310476	3.913846	1.1749421	6.6974429
P3	-60.312427	366.301777	6.024078	102.062969	0.2840203	15.6641903
P4	5.295084	78.427306	1.200261	5.008172	3.4352256	228.6809184
T1	3.007578	9.238123	1.279996	0.87639	1.3954422	0.9011264
T2	-1.566672	30.82155	2.7122	7.991949	1.6139076	5.7345161
T3	-597.912552	5657.297192	1.379561	1.418722	4.1158475	92.834866
T4	1.260298	1.683938	5.835641	32.378567	3.4735321	13.0487216
T5	-9.34746	124.04101	1.2927	0.893261	1.5364297	1.4419082
C1	-1.333648	16.370782	4.710544	20.707168	2.7486253	28.027306
C2	-21.059371	437.561882	-0.318701	16.744205	0.8512164	25.1404051

4.2 財務比率選擇

本節財務比率選擇之實證結果，分為二部份，一為自選法以 Altman (1968) 之選取方式進行預測變數之選取，二為前向選取法。

4.2.1 自選法

本文先針對產業化財務比率進行常態檢定，在表 4 發現各財務比率皆為非常態，僅負債比率為常態，因此本文在進行區別檢定時，則採 Mann-Wanney U 檢定，由表 5 發現，除總資產報酬率、總資產週轉率、應收帳款週轉率、存貨週轉率、股東權益週轉率及現金再投資比率外，皆呈現顯著，顯示大多數的比率在危機群及正常群具相當區別能力。採用 Spearman 等級和檢定以檢定各預測變數間之相關性，表 6 列舉各變數下相關

變數，發現這 18 個預測變數互有關係之程度很高，流動比率與其他比率之間相關性程度最高，除總資產週轉率、應收帳款週轉率、存貨週轉率及現金再投資比率外，因此在選擇預測變數時，各變數仍具有一定之相關程度。

本文以各變數之區別能力作為選擇變數主要考量，同時參酌前人之研究成果與財務報表分析構見之完整性，分別選擇五項比率，一為流動比率(S1)，二為負債比率(L1)，三為營業利益率(P3)，四為固定資產週轉率(T4)，五為現金流量比率(C1)。

4.2.2 前向選取法

在使用前面選取法時，各個不同樣本群，造成不同的選取變數 (Back et al.,1996)，因此本文之實證研究將產生 7 組預測變數，詳細列表結果將於 4.3Logit 迴歸模式進行詳細說明。

表 4 One-Sample Kolmogorov-Smirnov Test			表 5 Mann-Wanney U Test		
變數名	Kolmogorov-Smirnov Z	Sig.	變數名	Z	Sig.
S1	2.679	0 **	S1	-7.429	0 **
S2	2.429	0 **	S2	-6.588	0 **
L1	0.938	0.343	L1	-5.493	0 **
L2	5.166	0 **	L2	-5.596	0 **
L3	3.075	0 **	L3	-5.569	0 **
L4	4.976	0 **	L4	-4.457	0 **
L5	4.961	0 **	L5	-2.89	0.004 **
P1	4.073	0 **	P1	-1.346	0.178
P2	3.612	0 **	P2	-2.076	0.038 *
P3	5.172	0 **	P3	-2.565	0.01 **
P4	3.699	0 **	P4	-2.232	0.026 *
T1	4.352	0 **	T1	-1.629	0.103
T2	5.609	0 **	T2	-0.107	0.915
T3	6.046	0 **	T3	-0.167	0.867
T4	5.075	0 **	T4	-2.997	0.003 **
T5	5.915	0 **	T5	-1.324	0.186
C1	4.791	0 **	C1	-4.884	0 **
C2	5.401	0 **	C2	-1.106	0.269

表 6 變數相關表

變數名稱	S1	S2	S3	L1	L2	L3	L4	P1	P2	P3	P4	T1	T2	T3	T4	T5	C1	C2	
顯著相關之變數	S2	S1	S1	S1	S1	S1	S1	S1	S1	S1	S1	S2	P3	S2	S1	S1	S1	P3	
	L1	L1	S2	S2	S2	S2	S2	L5	L5	S2	L2	L1	T1	L1	S2	T1	S2	T1	
	L2	L2	L2	L1	L1	L1	L1	P2	P1	L3	L5	L2	T3	L2	L4	T2	L1	T2	
	L3	L3	L3	L3	L2	L2	L2	P3	P3	L5	P1	L3	T5	L3	T1	T3	L2	T5	
	L4	L4	L4	L4	L4	L3	L3	P4	P4	P1	P2	T2	C1	P4	T3	T4	L3	C1	
	L5	L5	L5	L5	L5	T4	P1	C1	C1	P2	P3	T3	C2	T1	T5	C2	L5		
	P1	P3	T1	P4	P3		P2			P4	T3	T4		T2			P1		
	P2	T1	T3	T1	T1		P3			T2	C1	T5		T4			P2		
	P3	T3	C1	T3	T3		P4			C1		C1		T5			P3		
	P4	T4		C1	C1		C1			C2		C2					P4		
	T4	C1																T1	
	T5																		T2
	C1																		C2

4.3 Logit 迴歸模式實證結果

以 Logit 迴歸 (以下簡稱 LR) 建立之預警模式，應變數之屬性設定上，本文將發生財務危機之公司設為 0，若係數為正表示其與財務危機為負相關，若係數為負則與財務危機為正相關，於表 7 列示配對及隨機樣本下，採用自選法及前向選取法，所有 LR 模式中所建構之 β 係數。

自選法下所選擇之五項變數中，以負債比率係數在所有樣本群中皆呈現顯著，係數為負值顯示若負債比率愈高，愈容易發生財務危機，營業利益率則為次重要變數，除在隨機 2 倍及隨機 8 倍樣本群中未呈現顯著外，其他樣本群之係數值為正且呈現統計顯著性，顯示營業利益率愈高愈不易發生財務危機。

在前向選取法所選出之變數中，最為重要之變數亦為負債比率，不論是配對或是隨機樣本群皆選入，係數值為負值且呈現統計上顯著，另一項重要變數則為營業利益率，除隨機 8 倍未選入外，其他皆樣本群皆選入，係數值為正且呈現統計顯著現象。另一個有趣現象為總資產週轉率，總資產週轉率除在隨機 1 倍及隨機 8 倍未選入外，其他樣本群皆選入，但是係數為負值且呈現統計顯著，若總資產週轉率愈高愈容易發生財務危機，一般而言，總資產週轉率愈高，表示運用資產效率愈佳，愈不易發生財務危機，但是獲利性

不佳之危機公司則會因為總資產週轉率高而導致其加快倒閉之速度。

判斷財務危機預警模式建構是否良窳，模式預測辨別之正確率為一重要指標，表顯示 LR 模式正確率、型一及型二錯誤率之一覽表，以 0.5 機率為區別機率 (Zmijewski,1984)，型一錯誤為當公司發生財務危機但被預測為財務體質健全公司，型二錯誤則為公司為財務健全公司卻被預測為發生財務危機之公司，對債權人及投資人而言，型一錯誤所產生之隱含損失是高於型二錯誤，將 LR 模式之型一錯誤、型二錯誤及正確率之一覽表列示於表 8。由表 9 及圖 1 發現，前向選取法訓練期間組之正確率略高於自選法訓練期間，這符合預期，因為其選入之變數為最適合此一模式，有助於提高正確率，在驗證期間前向選取法亦略優於自選法，兩者差距不大。前向選取法 LR 之整體正確率平均值為 81.11%，而自選法 LR 模式正確率平均值為 80.06%，正確率皆達到 80% 以上，整體而言，以 Logit 迴歸法所架構危機預警模式之預測準確性尚稱優良。

表 7 LR 係數表

	變數名	自選法				前向選取法				
		B	Wald	Sig.	變數名	B	Wald	Sig.		
配對	S1	0.33	1.203	0.273	L1	-5.363	10.278	0.001	**	
	L1	-3.308	9.889	0.002	**	P3	0.542	8.568	0.003	**
	P3	0.134	5.666	0.017	**	T1	-0.485	7.89	0.005	**
	T4	0.302	1.036	0.309	Constant	6.953	11.87	0.001	**	
	C1	0.268	4.32	0.038	*				**	
	Constant	2.571	4.41	0.036	*				**	
隨機 1倍	S1	0.332	1.235	0.266	L1	-2.106	6.609	0.01	**	
	L1	-1.553	4.349	0.037	*	P1	0.071	4.517	0.034	*
	P3	0.094	4.939	0.026	*	P3	0.064	10.252	0.001	**
	T4	0.164	0.468	0.494	P4	0.114	7.717	0.005	**	
	C1	0.182	2.648	0.104	Constant	2.687	7.677	0.006	**	
	Constant	1.171	1.176	0.278					**	
隨機 2倍	S1	0.201	0.408	0.523	L1	-2.471	12.092	0.001	**	
	L1	-1.593	6.499	0.011	**	P3	0.172	8.496	0.004	**
	P3	0.08	3.681	0.055	T1	-0.266	2.74	0.098		
	T4	0.278	2.383	0.123	T5	-0.022	2.794	0.095		
	C1	0.16	2.375	0.123	Constant	4.386	18.263	0	**	
	Constant	1.877	3.763	0.052					**	

隨機 4倍	S1	0.204	0.328	0.567		L1	-2.075	12.937	0	**
	L1	-1.839	10.404	0.001	**	P3	0.069	12.336	0	**
	P3	0.051	4.775	0.029	*	P4	0.068	6.48	0.011	**
	T4	0.358	5.075	0.024	**	T1	-0.371	7.035	0.008	**
	C1	0.036	0.498	0.481		Constant	4.757	29.499	0	**
	Constant	2.81	9.919	0.002	**					**
隨機 6倍	S1	0.403	0.743	0.389		L1	-2.023	16.088	0	**
	L1	-1.948	10.625	0.001	**	P3	0.155	14.794	0	**
	P3	0.079	6.996	0.008	**	P4	0.023	5.153	0.023	**
	T4	0.297	3.513	0.061		T1	-0.326	7.998	0.005	**
	C1	0.046	0.825	0.364		Constant	4.959	42.583	0	**
	Constant	3.213	10.202	0.001	**					**
隨機 8倍	S1	0.743	2.102	0.147		L1	-1.633	11.187	0.001	**
	L1	-2.001	10.071	0.002	**	L3	-0.23	2.339	0.126	
	P3	0.043	5.024	0.025		Constant	4.3	58.801	0	**
	T4	0.314	4.293	0.038	*					**
	C1	0.038	0.386	0.535						**
	Constant	3.167	8.844	0.003	**					**
隨機 10倍	S1	0.469	0.824	0.364		L1	-1.972	17.777	0	**
	L1	-2.099	12.155	0	**	P1	0.073	5.037	0.025	
	P3	0.051	4.989	0.026	*	P3	0.084	11.339	0.001	**
	T4	0.283	3.791	0.052		T1	-0.354	9.532	0.002	**
	C1	0.042	0.449	0.503		Constant	5.351	59.894	0	**
	Constant	3.846	13.414	0	**					**

表 8 LR 實證結果表

變數選 取法	時間	自選法					前向選取法				
		86年-88 年	89年	90年	91年	平均值	86年-88 年	89年	90年	91年	平均值
		Train	+1	+2	+3		Train	+1	+2	+3	
配對	Type I	17.65%	30.77%	46.15%	42.86%	34.36%	17.65%	84.62%	61.54%	71.43%	58.81%
	Type II	11.76%	38.46%	38.46%	42.86%	32.89%	11.76%	0.00%	23.08%	0.00%	8.71%
	正確率	85.29%	65.38%	57.69%	57.14%	66.38%	85.29%	57.69%	57.69%	64.29%	66.24%
隨機 1倍	Type I	17.65%	46.15%	46.15%	42.86%	38.20%	23.53%	53.85%	53.85%	42.86%	43.52%
	Type II	20.59%	30.77%	23.08%	42.86%	29.32%	17.65%	30.77%	30.77%	42.86%	30.51%
	正確率	80.88%	61.54%	65.38%	57.14%	66.24%	79.41%	57.69%	57.69%	57.14%	62.98%
隨機 2倍	Type I	55.88%	69.23%	61.54%	42.86%	57.38%	44.12%	46.15%	61.54%	42.86%	48.67%
	Type II	13.24%	3.85%	19.23%	0.00%	9.08%	8.82%	7.69%	7.69%	0.00%	6.05%

	正確率	72.55%	74.36%	66.67%	85.71%	74.82%	79.41%	79.49%	74.36%	85.71%	79.74%
隨機4倍	Type I	70.59%	84.62%	69.23%	42.86%	66.82%	58.82%	69.23%	53.85%	42.86%	56.19%
	Type II	1.47%	0.00%	3.85%	10.71%	4.01%	2.94%	1.92%	3.85%	3.57%	3.07%
	正確率	84.71%	83.08%	82.42%	82.86%	83.26%	85.88%	84.62%	89.01%	88.57%	87.02%
隨機6倍	Type I	70.59%	92.31%	69.23%	42.86%	68.75%	64.71%	84.62%	76.92%	42.86%	67.28%
	Type II	0.98%	1.28%	0.00%	7.14%	2.35%	1.97%	0.00%	3.85%	4.76%	2.64%
	正確率	89.08%	85.71%	85.71%	87.76%	87.06%	89.03%	87.91%	85.71%	89.80%	88.11%
隨機8倍	Type I	73.53%	92.31%	76.92%	57.14%	74.98%	85.29%	84.62%	84.62%	57.14%	77.92%
	Type II	0.74%	0.00%	0.00%	1.79%	0.63%	0.74%	0.00%	0.00%	0.00%	0.18%
	正確率	91.18%	89.74%	88.03%	93.65%	90.65%	89.84%	90.60%	90.60%	93.65%	91.17%
隨機10倍	Type I	76.47%	100.00%	100.00%	57.14%	83.40%	67.65%	84.62%	76.92%	57.14%	71.58%
	Type II	0.29%	0.00%	0.00%	1.43%	0.43%	0.59%	1.54%	0.77%	1.43%	1.08%
	正確率	92.78%	90.91%	90.91%	93.51%	92.03%	93.30%	90.91%	92.31%	93.51%	92.51%

表9 LR 正確率一覽表

變數選取法	時間	配對	隨機1倍	隨機2倍	隨機4倍	隨機6倍	隨機8倍	隨機10倍	平均值
自選法	86年-88年 Train	85.29%	80.88%	72.55%	84.71%	89.08%	91.18%	92.78%	85.21%
	89年 +1	65.38%	61.54%	74.36%	83.08%	85.71%	89.74%	90.91%	78.68%
	90年 +2	57.69%	65.38%	66.67%	82.42%	85.71%	88.03%	90.91%	76.69%
	91年 +3	57.14%	57.14%	85.71%	82.86%	87.76%	93.65%	93.51%	79.68%
	平均值	66.38%	66.24%	74.82%	83.26%	87.06%	90.65%	92.03%	80.06%
前向選取法	86年-88年 Train	85.29%	79.41%	79.41%	85.88%	89.03%	89.84%	93.30%	86.02%
	89年 +1	57.69%	57.69%	79.49%	84.62%	87.91%	90.60%	90.91%	78.42%
	90年 +2	57.69%	57.69%	74.36%	89.01%	85.71%	90.60%	92.31%	78.20%
	91年 +3	64.29%	57.14%	85.71%	88.57%	89.80%	93.65%	93.51%	81.81%
	平均值	66.24%	62.98%	79.74%	87.02%	88.11%	91.17%	92.51%	81.11%
LR	平均值	66.31%	64.61%	77.28%	85.14%	87.59%	90.91%	92.27%	80.59%

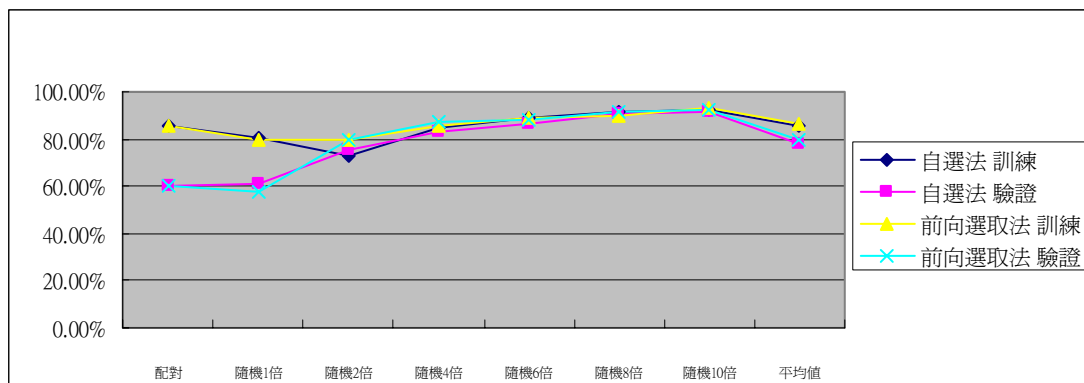


圖1 LR 正確率平均值圖

4.4 SVM 實證結果

衡量分類方法之優劣，最重要的關注重點在於分類之正確率（Van Gestel et al., 2002; Fan et al., 2000; Altman et al., 1994），使用 RBF 函數，在使用 RBF 函數時，須測試參數 γ 及成本參數 C ($C=2^{-5}, 2^{-3}, \dots, 2^{15}$; $\gamma=2^{-15}, 2^{-13}, \dots, 2^3$) 以促使分類正確率能提高，決定最佳分類模式，本文採用 LIB SVM¹實驗室所發展之工具進行參數選取及 SVM 之實證，參數表如表 10。

對於使用者而言，正確率、型一錯誤及型二錯誤皆十分重要，SVM 模式正確率、型一錯誤及型二錯誤列示於表 11，由表 12 及圖 2 發現，自選法訓練期間組之正確率略高於前向選取法訓練期間，在驗證期間逐步法略優於自選法，前向選取法 SVM 之整體正確率平均值為 83.01%，而自選法 SVM 之正確率平均值為 82.01%，正確率達到 80% 以上，整體而言，以 SVM 所架構危機預警模式之預測準確性尚稱優良。

表 10 參數表

實證樣本組	參數 C	參數 γ	最佳正確率(10%)
配對	2^{13}	2^{-13}	85.29
隨機 1 倍	2^{11}	2^{-13}	83.82
隨機 2 倍	2^5	2^{-7}	76.47
隨機 4 倍	2^5	2^{-7}	81.77
隨機 6 倍	2^{11}	2^{-7}	85.71
隨機 8 倍	2^5	2^{-7}	87.26
隨機 10 倍	2^3	2^{-9}	90.91

表 11 SVM 實證結果表

變數選取法		自選法					逐步法				
實證樣本	時間	86 年-88 年 Train	89 年 +1	90 年 +2	91 年 +3	平均值	86 年-88 年 Train	89 年 +1	90 年 +2	91 年 +3	平均值
配對	Type I	5.88%	0.00%	7.69%	0.00%	3.39%	17.65%	23.08%	15.38%	14.29%	17.60%
	Type II	8.82%	61.54%	69.23%	57.14%	49.18%	14.71%	30.77%	69.23%	57.14%	42.96%
	正確預測率	92.65%	69.23%	61.54%	71.43%	73.71%	83.82%	73.08%	57.69%	64.29%	69.72%
隨機 1 倍	Type I	11.76%	15.38%	15.38%	42.86%	21.35%	14.71%	0.00%	0.00%	0.00%	3.68%
	Type II	14.71%	46.15%	46.15%	14.29%	30.32%	2.94%	23.08%	69.23%	14.29%	27.38%
	正確預測率	86.76%	69.23%	69.23%	71.43%	74.16%	91.18%	88.46%	65.38%	85.71%	82.68%

¹ <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

隨機 2 倍	Type I	29.41%	61.54%	61.54%	71.43%	55.98%	35.29%	30.77%	23.08%	42.86%	33.00%
	Type II	1.47%	19.23%	19.23%	14.29%	13.55%	2.94%	3.85%	26.92%	0.00%	8.43%
	正確預測率	89.22%	66.67%	66.67%	66.67%	72.30%	86.27%	74.36%	74.36%	85.71%	80.18%
隨機 4 倍	Type I	35.29%	92.31%	76.92%	100.00%	76.13%	26.47%	84.62%	92.31%	100.00%	75.85%
	Type II	0.74%	1.92%	2.56%	3.57%	2.20%	0.00%	1.92%	2.56%	3.57%	2.01%
	正確預測率	92.35%	80.00%	86.81%	77.14%	84.08%	94.71%	81.54%	84.62%	77.14%	84.50%
隨機 6 倍	Type I	47.06%	100.00%	100.00%	100.00%	86.76%	41.18%	100.00%	92.31%	100.00%	83.37%
	Type II	0.49%	0.00%	0.00%	0.00%	0.12%	0.00%	0.00%	1.28%	0.00%	0.32%
	正確預測率	92.86%	85.71%	85.71%	85.71%	87.50%	94.12%	85.71%	85.71%	85.71%	87.82%
隨機 8 倍	Type I	41.18%	100.00%	92.31%	100.00%	83.37%	79.41%	84.62%	84.62%	57.14%	76.45%
	Type II	0.00%	0.00%	0.00%	0.00%	0.00%	0.74%	0.00%	0.00%	0.00%	0.18%
	正確預測率	95.42%	88.89%	89.74%	88.89%	90.74%	90.52%	90.60%	90.60%	93.65%	91.34%
隨機 10 倍	Type I	61.76%	100.00%	100.00%	100.00%	90.44%	50.00%	100.00%	100.00%	100.00%	87.50%
	Type II	0.00%	0.00%	0.77%	0.00%	0.19%	0.00%	0.00%	0.77%	0.00%	0.19%
	正確預測率	94.39%	90.91%	90.21%	90.91%	91.60%	95.45%	90.91%	90.21%	90.91%	91.87%

表 12 SVM 正確率一覽表

變數選取法	時間	配對	隨機 1 倍	隨機 2 倍	隨機 4 倍	隨機 6 倍	隨機 8 倍	隨機 10 倍	平均值	
自選法	86 年-88 年	Train	92.65%	86.76%	89.22%	92.35%	92.86%	95.42%	94.39%	91.95%
	89 年	+1	69.23%	69.23%	66.67%	80.00%	85.71%	88.89%	90.91%	78.66%
	90 年	+2	61.54%	69.23%	66.67%	86.81%	85.71%	89.74%	90.21%	78.56%
	91 年	+3	71.43%	71.43%	66.67%	77.14%	85.71%	88.89%	90.91%	78.88%
		平均值		73.71%	74.16%	72.30%	84.08%	87.50%	90.74%	91.60%
前向選取法	86 年-88 年	Train	83.82%	91.18%	86.27%	94.71%	94.12%	90.52%	95.45%	90.87%
	89 年	+1	73.08%	88.46%	74.36%	81.54%	85.71%	90.60%	90.91%	83.52%
	90 年	+2	57.69%	65.38%	74.36%	84.62%	85.71%	90.60%	90.21%	78.37%
	91 年	+3	64.29%	85.71%	85.71%	77.14%	85.71%	93.65%	90.91%	83.30%
		平均值		69.72%	82.68%	80.18%	84.50%	87.82%	91.34%	91.87%
SVM	平均值		71.72%	78.42%	76.24%	84.29%	87.66%	91.04%	91.74%	83.01%

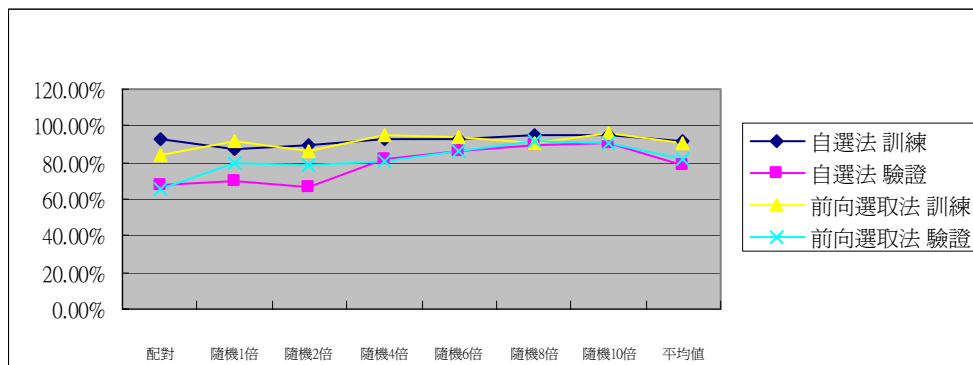


圖 2 SVM 正確率平均值圖

4.5 變數選擇、抽樣誤差及持續性之探討

針對三項議題進行探討，分別為變數選擇方法，二為抽樣誤差，三為模式持續性，分就如下說明之：

4.5.1 變數選擇

本文以 Altman 提出之架構決定預測變數，採用此方式，考量財務比率之區別性及財務報表分析結構，而另一個方式則使用迴歸模式選取變數之前向選取法，完全取決於統計方法上之結果，如何決定那種方式較佳，本文除正確率外，另需考量型一及型二錯誤率，由表 13 發現，配對以自選法結果最佳，在隨機樣本中則以前向選取法為佳，整體而言，前向選取法仍優於自選法其型一及型二錯誤率低於自選法，若以 SVM 之正確率為選取變數之依據，配對樣本下，以自選法之結果優於前向選取法，隨機樣本下則前向選取法為優，自選法之型一錯誤略高於前向選取法，本文認為前向選取法優於自選法。

表 13 變數選取正確率簡表

樣本群	自選法			前向選取法		
	Type I	Type II	正確率	Type I	Type II	正確率
配對	18.88%	41.04%	70.04%	38.20%	25.84%	67.98%
隨機 1 倍	29.78%	29.82%	70.20%	23.60%	28.95%	72.83%
隨機 2 倍	56.68%	11.32%	73.56%	40.83%	7.24%	79.96%
隨機 4 倍	71.48%	3.10%	83.67%	66.02%	2.54%	85.76%
隨機 6 倍	77.76%	1.24%	87.28%	75.32%	1.48%	87.96%
隨機 8 倍	79.17%	0.32%	90.69%	77.18%	0.18%	91.26%
隨機 10 倍	86.92%	0.31%	91.81%	79.54%	0.64%	92.19%
平均值	60.09%	12.45%	81.04%	57.24%	9.55%	82.56%

4.5.2 抽樣誤差

抽樣誤差以隨機抽樣樣本群進行探討，由表 14 列示各組隨機樣本型一錯誤、型二錯誤及正確率之平均值，依表 14 繪出圖 3，由表 14 及圖 3 可清楚發現隨著抽樣樣本之增加造成型一錯誤率及正確率隨之遞增，而型二錯誤率隨之遞減，在隨機一倍之型一錯誤率及型二錯誤率之較易令人接受，但是在隨機 2 倍型一錯誤率大幅增加至 48.76%，

型一錯誤率太高，表示易於將財務危機公司判別錯誤，本文認為型一錯誤率較型二錯誤率重要，若一模式型一錯誤率過高，即正確率表現也不錯，仍不認為這模式預測能力佳，在隨機 4 倍、6 倍、8 倍及 10 倍皆發生此種現象。

為了確認樣本量大小與型一錯誤率、型二錯誤率及正確率關係，將各實證模式所得結果進行相關性檢定 (Zmijewski, 1984)，採用 Spearman test 列示於表 15，B 為抽樣樣本量，1 倍為 1，2 倍為 2，4 倍為 3，6 倍為 4，8 倍為 5，10 倍為 6，T1 為型一錯誤率，T2 為型二錯誤率，C 為正確率，T1 與 B 呈正相關且呈現顯著，顯示抽樣樣本愈大型一錯誤率亦隨之增加，C 與 B 之相關性與 T1 相同，T2 與 B 呈負相關且呈現顯著，顯示抽樣樣本愈大型二錯誤率亦隨之降低，因此在進行財務危機模式之建構時，須考慮危機樣本群與正常樣本群之數量比例，以獲致較佳預測結果。

表 14 平均值一覽表

樣本群	Type I	Type II	正確率
隨機 1 倍	26.69%	29.39%	71.52%
隨機 2 倍	48.76%	9.28%	76.76%
隨機 4 倍	68.75%	2.82%	84.72%
隨機 6 倍	76.54%	1.36%	87.62%
隨機 8 倍	78.18%	0.25%	90.98%
隨機 10 倍	83.23%	0.47%	92.00%

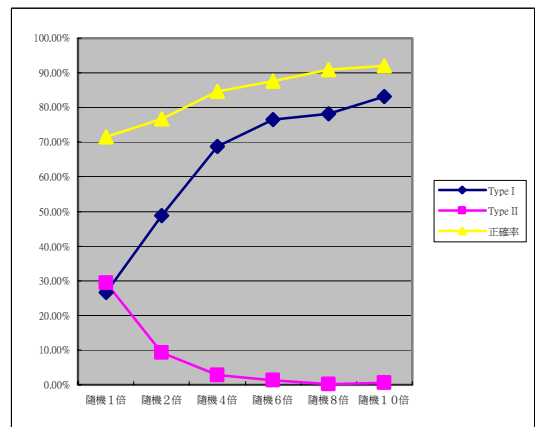


圖 3 型一錯誤率、型二錯誤率及正確率圖

表 15 樣本量大小與型一、型二錯誤率及正確率之間相關性

Correlations			B	T1	T2	C
Spearman's rho	B	Correlation Coefficient	1.000	.654**	-.749**	.786*
		Sig. (2-tailed)	.	.000	.000	.000
		N	96	96	96	96
	T1	Correlation Coefficient	.654**	1.000	-.560**	.186
		Sig. (2-tailed)	.000	.	.000	.070
		N	96	96	96	96
	T2	Correlation Coefficient	-.749**	-.560**	1.000	-.697*
		Sig. (2-tailed)	.000	.000	.	.000
		N	96	96	96	96
	C	Correlation Coefficient	.786**	.186	-.697**	1.000
		Sig. (2-tailed)	.000	.070	.000	.
		N	96	96	96	96

** . Correlation is significant at the .01 level (2-tailed).

4.5.3 模式持續性

預警模式除了正確預測下一年度外，是否能持續長時間加以應用，並維持其預測之正確性，亦為使用者十分關心之議題，預期 91 年即第三年正確率即為最低，從圖 4 可清楚觀察 91 年之正確率未低於 89 年及 90 年，利用 Spearman test 進行時間 (Time) 與正確率 (C)、型一錯誤率 (T1) 及型二錯誤率 (T2) 之相關性檢定，由表 16 發現時間與這三項變數相關性檢定皆不顯著，顯示時間愈長並不能影響預警模式驗證結果。

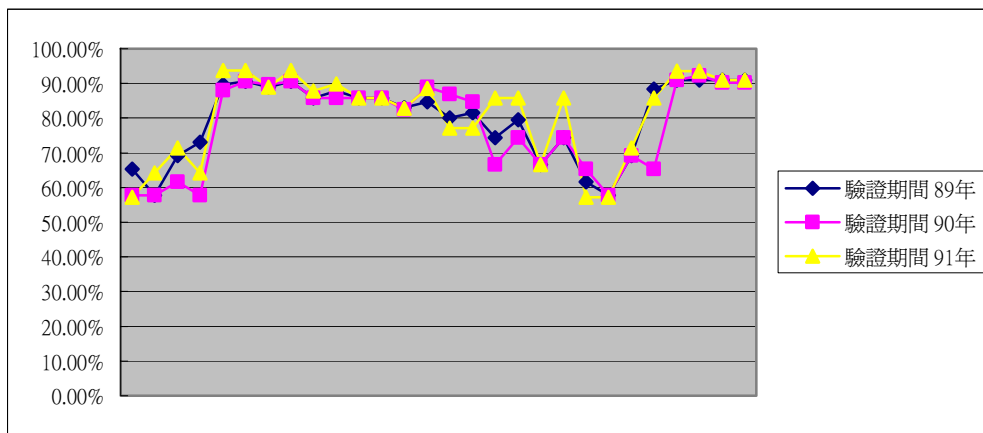


圖 4 驗證期間正確率圖

表 16 樣本量大小與型一、型二錯誤率及正確率之間相關性

Correlations			TIME	C	T1	T2
Spearman's rho	TIME	Correlation Coefficient	1.000	.059	-.165	.047
		Sig. (2-tailed)	.	.592	.133	.673
		N	84	84	84	84
C		Correlation Coefficient	.059	1.000	.322*	-.380*
		Sig. (2-tailed)	.592	.	.003	.000
		N	84	84	84	84
T1		Correlation Coefficient	-.165	.322*	1.000	-.694*
		Sig. (2-tailed)	.133	.003	.	.000
		N	84	84	84	84
T2		Correlation Coefficient	.047	-.380*	-.694*	1.000
		Sig. (2-tailed)	.673	.000	.000	.
		N	84	84	84	84

** . Correlation is significant at the .01 level (2-tailed).

4.6 LR 及 SVM 之比較

比較 LR 及 SVM 兩種方式所建構預警模式之良窳，主要以正確率高低及型一錯誤率為主 (Back et al., 1996)，在表 17 中，不論訓練期間或是驗證期間，SVM 正確率平均值高於 LR，但在配對樣本中 LR 之表現優於 SVM，SVM 隨機樣本群之訓練期間型一錯誤率是低於 LR 甚多，但在隨機樣本 4 倍以上之樣本群型一錯誤率則高於 LR，再依變數選取方法分為自選法及前向選取法進行分析，列示於表 18 及表 19，由正確率及型一錯誤率之平均值來看，SVM 仍略優於 LR，在訓練期間 SVM 之表現皆優於 LR，除配對樣本群外，但在驗證期間時，SVM 之預測則呈現較為分歧狀況，若考慮抽樣誤差情形，排除隨機 4 倍以上之樣本群，則 SVM 之表現是優於 LR。而在選取樣本方式之搭配上，以前向選取法略優於自選法，但是相差幅度不大。

表 17 LR 及 SVM 實證結果表

		配對	1 倍	2 倍	4 倍	6 倍	8 倍	10 倍	平均值
正確率	訓練期間 LR	85.29%	79.41%	79.41%	85.88%	89.03%	89.84%	93.30%	86.02%
	訓練期間 SVM	83.82%	91.18%	86.27%	94.71%	94.12%	90.52%	95.45%	90.87%
	驗證期間 LR	67.40%	69.96%	66.67%	81.32%	85.71%	89.17%	90.68%	78.70%
	驗證期間 SVM	65.02%	79.85%	78.14%	81.10%	85.71%	91.62%	90.68%	81.73%
型一錯誤率	訓練期間 LR	17.65%	23.53%	44.12%	58.82%	64.71%	85.29%	67.65%	51.68%
	訓練期間 SVM	17.65%	14.71%	35.29%	26.47%	41.18%	79.41%	50.00%	37.82%
	驗證期間 LR	2.56%	24.54%	64.84%	89.74%	100.00%	97.44%	100.00%	68.45%
	驗證期間 SVM	17.58%	0.00%	32.23%	92.31%	97.44%	75.46%	100.00%	59.29%

表 18 自選法 LR 及 SVM 實證結果表

		配對	1 倍	2 倍	4 倍	6 倍	8 倍	10 倍	平均值
正確率	訓練期間 LR	85.29%	80.88%	72.55%	84.71%	89.08%	91.18%	92.78%	85.21%
	訓練期間 SVM	92.65%	86.76%	89.22%	92.35%	92.86%	95.42%	94.39%	91.95%
	驗證期間 LR	60.07%	61.36%	75.58%	82.78%	86.39%	90.48%	91.77%	78.35%
	驗證期間 SVM	66.48%	70.33%	66.67%	81.98%	85.71%	89.32%	90.56%	78.72%
型一錯誤率	訓練期間 LR	17.65%	17.65%	55.88%	70.59%	70.59%	73.53%	76.47%	54.62%
	訓練期間 SVM	5.88%	11.76%	29.41%	35.29%	47.06%	41.18%	61.76%	33.19%
	驗證期間 LR	39.93%	45.05%	57.88%	65.57%	68.13%	75.46%	85.71%	62.53%
	驗證期間 SVM	3.85%	29.12%	66.48%	88.46%	100.00%	96.15%	100.00%	69.15%

表 19 前向選取法 LR 及 SVM 實證結果表

		配對	1 倍	2 倍	4 倍	6 倍	8 倍	10 倍	平均值
正 確 率	訓練期間 LR	85.29%	79.41%	79.41%	85.88%	89.03%	89.84%	93.30%	86.02%
	訓練期間 SVM	83.82%	91.18%	86.27%	94.71%	94.12%	90.52%	95.45%	90.87%
	驗證期間 LR	67.40%	69.96%	66.67%	81.32%	85.71%	89.17%	90.68%	78.70%
	驗證期間 SVM	65.02%	79.85%	78.14%	81.10%	85.71%	91.62%	90.68%	81.73%
型 一 錯 誤 率	訓練期間 LR	17.65%	23.53%	44.12%	58.82%	64.71%	85.29%	67.65%	51.68%
	訓練期間 SVM	17.65%	14.71%	35.29%	26.47%	41.18%	79.41%	50.00%	37.82%
	驗證期間 LR	2.56%	24.54%	64.84%	89.74%	100.00%	97.44%	100.00%	68.45%
	驗證期間 SVM	17.58%	0.00%	32.23%	92.31%	97.44%	75.46%	100.00%	59.29%

5 結論

不同以往文獻著重於各方法之比較與探討，本文之研究重心關注在架構財務危機預警模式之三項議題，一為抽樣樣本誤差，二為財務變數之選擇，三為財務預警模式持續性；並使用日前新興起分類方法支撐向量機(SVM)及以為廣為使用 Logit 迴歸為架構財務危機預警模式之方法，並加以比較之。

自選法及前向選取法之變數選取法，以前向選取法略優於自選法，由 Logit 迴歸結果，在台灣上市公司以負債比率及營業利益率為最能偵測公司危機之預測變數，隨著隨機抽樣樣本數增加，型一錯誤率及正確率亦隨之增加，這是因為型二錯誤率隨之減少，這樣趨勢並不樂見，因為型一錯誤率過高，損失將大於型二錯誤，而隨機樣本 1 倍之實證結果亦優於配對樣本，因此在樣本設計上，不宜選取過高正常公司樣本數，並不一定須進行樣本配對，在驗證期間時，發現本文之模式持續性尚佳，第三年之預測結果並未較第一年大幅減少。

運用 Logit 迴歸法建構預警模式，可了解各變數與財務危機之關係，惟須考量統計上性質，而支撐向量機(SVM)之分類正確率高，惟為黑箱作業，並無法了解其中變數之間關係，在比較此兩種方式，以型一錯誤率低及正確率高二項原則為標準，支撐向量機(SVM)及前向選取法所建構之模式表現最佳。

參考文獻

- Altman, E (1968), Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance*, 23, 589-609.
- Altman, E., G. Macro and F. Varetto (1994), Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience), *Journal of Finance*, 18, 505-529.
- Back, B., T. Laitinen and K. Sere (1996), Neural networks and genetic algorithms for bankruptcy predictions, *Expert Systems with Applications*, 11, 407-413.
- Beaver, W. (1966), Financial ratios as predictors of failure, *Empirical Research in Accounting: Selected Studies, Supplement to Journal of Accounting Research*, 4, 71~111.
- Collins, R.A. and R.D. Green (1982), Statistical methods for bankruptcy forecasting, *Journal of Economics and Business*, 34, 349-354.
- Cristianini, N. and J. Shawe-Taylor (2000), *An Introduction to support Vector Machines*, Cambridge University Press.
- Dimitras, A.I., S.H. Zanakis, and C. Zopounidis (1996), A survey of business failures with an emphasis on prediction methods and industrial applications, *European Journal of Operational Research*, 90, 487-513.
- Fan, A. and M. Palaniswami (2000), Selecting bankruptcy predictors using a support vector machine approach, *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, 6, 354-359.
- Jo, H. and I. Han (1996), Integration of case-based forecasting, neural network, and discriminant analysis for bankruptcy prediction, *Expert Systems with Applications*, 11, 415-422.
- Kim, M.J. and I. Han (2003), The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms, *Expert Systems with Applications*, 25, 637-646.
- Lennox, C. (1999), identifying failing companies: a reevaluation of the logit, probit and DA approaches, *Journal of Economics and Business*, 51, 347-364.
- Ohlson, J.A. (1980), Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research*, 28, 109-131.
- Odom, M., and R. Sharda (1990), A neural networks model for bankruptcy prediction, *Proceedings of the IEEE International Conference on Neural Network*, 2, 163-168.
- Platt, H.D. and M.B. Platt (1990), Development of a class of stable predictive variables: the case of bankruptcy prediction, *Journal of Business Finance and Accounting*, 17, 31-51.
- Platt, H.D. and M.B. Platt (1991), A note on the use of industry-relative ratios in bankruptcy prediction, *Journal of Banking and Finance*, 15, 1183-1194.
- Platt, H.D. and M.B. Platt (2002), Predicting corporate financial distress: reflections on choice-based sample bias, *Journal of Economics and Finance*, 26, 184-199.
- Shin K.S. and Y.J. Lee (2002), A genetic algorithm application in bankruptcy prediction modeling, *Expert Systems with Applications*, 23, 321-328.

- Van Gestel, T., B. Baesens, J. Suykens, M. Espinoza, D. E. Baestaens, J. Vanthienen, and B. De Moor (2003), Bankruptcy prediction with least squares support vector machine classifiers, *IEEE International Conference on Computational Intelligence for Financial Engineering*, 1-8.
- Vapnik, V. (1998), *Statistical Learning Theory*, New York, Wiley.
- Vapnik, V. (1995), *The Nature of Statistical Learning Theory*, New York, Springer Verlag.
- Yang, Z.R., M.B. Platt and H.D. Platt (1999), Probabilistic neural networks in bankruptcy prediction, *Journal of Business Research*, 44, 67-74.
- Zmijewski, M. E. (1984), Methodological issues related to the estimation of financial distress prediction models, *Journal of Account Research*, 22, 59-82.