

房屋抵押貸款客戶違約預測模式之比較研究

黃嘉興* 謝永明** 劉宗哲***

(收稿日期：91年11月27日；第一次修正：92年4月2日；第二次修正：92年8月14日；
第三次修正：92年10月17日；接受刊登日期：93年11月4日)

摘要

本研究以國內某銀行一南部分行的房屋貸款客戶為研究對象，運用 logistic 迴歸與區別分析兩種統計方法，尋求最佳預測變數組合以建立房屋貸款客戶是否違約之預測模式，並與五種預測模式（四種模式來自過去研究及一種來自樣本銀行的個人信用評估表）進行比較。此外，亦特別蒐集已經認列損失或拍賣抵押房屋之違約房貸戶資料，同時檢驗模式的預測準確度與可降低錯誤分類損失之能力。

實證結果發現，本研究所建構模式在第一階段分析中，其準確率與模式配適度均較其他預測模式為佳。但由於諸多預測變數未符合常態分配，區別分析的結果將受到違反基本假設的限制，故而在後續降低錯誤分類損失能力之檢驗上，僅採用 logistic 迴歸進行分析。在第二階段的分析中，本文所建構模式對違約客戶的分類準確率為 60.7%。進一步分析後發現，模式違約預測準確率較高者，其所能降低錯誤分類成本（或損失）通常也較大。整體而言，本研究所建立的預測模式未來可作為銀行房屋抵押放款之參考，以協助銀行有效降低其潛在違約損失。

關鍵詞：房屋抵押貸款、預測模式、信用評等、貸款違約、logistic 迴歸

* 國立雲林科技大學財務金融系副教授。

** 東吳大學會計學系講師，台灣大學會計學系博士候選人。

*** 台灣中小企業銀行領組。

壹、前言

民國八十一年後金融業因應新銀行成立之競爭壓力，以及為滿足購屋者之需求，大量承做房屋抵押貸款，並採取高額度、低資格授信策略以求市場之高占有率及巨額獲利。事實上高占有率並不一定代表高獲利，必須同時配合較低的逾放比例。長久以來，銀行業以擔保品為決定融資額度的主要考量因素，對於授信的個人因素往往因業績壓力而較少考量，這與國外不但強調擔保品，更重視個人信用評價有相當大的差距。根據近幾年銀行放款貸後追蹤發現，許多借款戶之擔保品雖大幅貶落，卻依然正常履約，顯示擔保品並非影響貸款違約的唯一因素。因此銀行有必要改變過去以擔保品評價為主軸之徵信方式，在審核貸款之初，即按借款人、保證人之個人資料，評估借款人之信用風險，積極預測借款人之償還能力及償還意願，以求能事先發現其違約跡象，確實將逾放比例控制在較低的水準。

本研究以國內某銀行一南部分行的房屋貸款客戶為研究對象，運用 logistic 迴歸與區別分析兩種統計方法，尋求最佳預測變數組合以建立房屋貸款客戶是否違約之預測模式，並與過去研究之預測變數組合所建立的預測模式進行比較，期能建立預測銀行房貸客戶違約風險之最佳模式，俾有利於銀行未來在房屋抵押貸款之徵、授信工作。除此之外，為瞭解本研究所建立之預測模式是否優於利用樣本銀行之「個人信用評等表」所形成之預測模式，本研究亦對兩者之預測準確度進行比較。最後本研究特別蒐集已認列損失或拍賣抵押房屋之違約房貸客戶資料，除做為預測準確度的檢驗之外，也同時檢驗預測模式可降低錯誤分類損失之能力。

實證結果發現，本研究所建構模式在第一階段分析中，其準確率與模式配適度均較其他預測模式為佳。但由於諸多預測變數未符合常態分配，區別分析的結果將受到違反基本假設的限制，故而在後續降低錯誤分類損失能力之檢驗上，僅採用 logistic 迴歸進行分析。在第二階段的分析中，本文所建構模式對違約客戶的分類準確率為 60.7%，若銀行徵、授信人員採用此模式事先判斷這些客戶會違約，而做成拒絕貸、放款的決策或提早要求還款，即可避免或減少大部份呆帳損失。進一步分析後發現，模式違約預測準確率較高者，其所能降低錯誤分類成本（或損失）通常也較大。整體

而言，本研究所建立的預測模式未來可作為銀行房屋抵押放款之參考，以協助銀行有效降低其潛在違約損失。

貳、過去相關研究之探討

Steenackers 及 Goovaerts(1989)利用逐步 logistic 迴歸方法建立個人貸款的信用評分模式(credit scoring model)，其研究發現申貸人的年齡、是否有電話、居住現址的時間長度、地區別、職業、是否在公家機關上班、月薪、住宅所有權、先前貸款次數、貸款期間等個人資料，與將來是否正常履約有顯著關係。

江百信及張金鶚（民 84）於購屋貸款放款條件之研究中，以民國 78 年至民國 82 年間台灣土地銀行 69 家分行共 2445 筆放款案例，選用借款人之年齡、職業、所得、不動產狀況（擔保品之座落、價格、面積等）與銀行往來實績、放款金額、放款利率及放款期限等變數，以最小平方法(ordinary least squares method, OLS)進行迴歸分析。結果發現擔保品（物的因素）是銀行決定融資額度的最主要因素，其他如借款人所得、學歷、借款、往來實績等因素（人的因素）亦為銀行放款所重視，但影響程度有限。

杜慶麟及張瑞芬（民 87）在銀行抵押貸款授信決策之研究中，利用類神經網路的訓練學習，嘗試將資深授信人員或信用分析主管的經驗及判斷以模糊集合(fuzzy sets)量化，進而建立一具有授信決策經驗之類神經網路專家系統。他們選擇 19 個授信決策考量變數，其中含 16 個評估變數及 3 個貸放變數（利率減碼、貸款成數及貸款年限），評估變數再依據類神經網路(artificial neural network, ANN)選用變數原則—獨立性、代表性、完整性、可用性等綜合成 8 個變數值做為 ANN 的輸入向量。這 8 個變數包括：(1)抵押品現值：買賣價或淨現值；(2)抵押品狀況：屋齡、位置、屋況；(3)個人資料：年齡、教育或專業程度、居住情況；(4)家庭背景：家中依賴人口數、有無配偶；(5)財務狀況：年收入（含配偶）、其它動產及不動產；(6)職業狀況：服務年資、職業類型；(7)保證人資料：保證人綜合評價；(8)其他：與本行往來情形、經辦面談之評價。以某銀行台北某分行民國 83 年間的 584 個房屋抵押貸款案件為研究樣本，其中 484 個案件作為訓練實例，100 個案

件作為測試範例。其結論指出，經過訓練後的ANN確能將授信專家經驗轉化至ANN系統內，且此一ANN決策系統對測試範例中的申貸案件，除準確度較佳外，亦能以較客觀、公平及一致的基礎進行審核。

李桐豪及呂美慧（民89）以某金融機構民國85及86年362個核貸案件（其中正常貸款案件258件、列入催收案件104件）作為研究對象，運用logistic迴歸分析探討金融機構房貸客戶授信評量模式。其結果發現，在不考慮借款人通信區域時，房貸客戶之婚姻狀況、學歷、金融往來關係、貸款期間、借款人與擔保人關係以及借款人通信地址與擔保品位置的相對關係等，是影響房屋抵押貸款品質好壞的主要因素。

王泰裕（民90）應用五層式模糊神經網路(fuzzy artificial neural network)，建構消費貸款評估模式，期能協助銀行評估授信風險，並改善傳統信用評等方法所產生的不精確結果，使授信決策者在面對不確定的狀況下，能做出正確的評判。他將消費性貸款信用評等之主要因素，歸類為年齡、家族狀況、住宅狀況、居住年數、職業、銀行往來關係、服務年數、年收入等八項，使用中部某銀行30個實例案件，另外加入模擬案例30筆，隨機抽取其中50個案例作為訓練範例，其餘則做為測試案例。研究中發現，在案例數足夠的情況下，模糊類神經網路所架構出的個人信用評等模式具有相當的正確性、可更新性及解釋能力。

參、研究方法

一、預測模式統計分析方法的選擇

過去研究「企業財務危機」預測模式或銀行授信客戶違約風險之預測模式，最常用的統計方法，首推區別分析(discriminant analysis)與logistic迴歸。近年來，則有學者開始嘗試使用無特定結構訊息的類神經網路，建立類似的預測模式（陳錦村、許通安及林蔓蓁，民85；杜慶麟及張瑞芬，民87）。雖然過去陳錦村（民86）和李桐豪及呂美慧（民89）的研究結果顯示，logistic分析法優於區別分析法，但在陳錦村、許通安及林蔓蓁（民85）和Tam及Kiang (1992)的研究結果中，logistic法並非全然優於區別分析法，有時區別分析法的預測準確率也會高於logistic迴歸法。故運用區別分析與

logistic 分析兩種統計方法所得預測結果，彼此之間並無確切證據顯示孰優孰劣，端視資料屬性與分析方法基本假設之吻合情況而定。

由於本研究之主要目的，在比較不同的房屋貸款違約預測模式，重點在找尋較佳的預測變數組合，而非比較何種分析方法較佳，故本研究僅使用傳統的區別分析與二元 logistic 迴歸分析(binary logistic regression)兩種統計方法。

(一)區別分析法

區別分析類似於多元迴歸，是一種最常用的分類方法，然兩者主要差異在於自變數的屬性不同，多元迴歸的自變數多為連續變數，而區別分析的自變數則為次序或類別變數。區別分析根據資料特性，將資料分類至不同群組中。通常，須先建立一個線性區別（或分類）函數，用以計算個別資料的分數。其區別函數形式如下（假設每個資料有 p 種特性）：

$$Z = w_0 + w_1 X_1 + w_2 X_2 + \dots + w_p X_p \quad (1)$$

其中， Z 即為區別函數，且對每一類資料而言， w_i 為 X_i 的係數， w_0 則為常數項， $i=1, \dots, p$ 。區別分析的目的即是要辨識此一區別函數的各個權數(w_i)，以求能極大化組間差平方和(between-group sum of squares, SS_b)對組內差平方和(within-group sum of squares, SS_t)的比值，亦即極大化下式之 Λ 值：

$$\Lambda = \frac{SS_b}{SS_t} \quad (2)$$

線性區別函數為資料觀察值在不同特性上的加權線性組合，加權的權數(w_i)則代表各種特性(X_i)的相對重要程度或影響力。基於此一區別函數，資料便能被適當地分類。

區別分析的功用如同一般的迴歸分析，主要用於解釋與預測兩種途徑。在解釋方面，區別分析可以分析並瞭解那些變數可以區分資料為不同的組別，以及這些不同組別間又有何差異。在預測用途方面，則可根據訓練資料所得的區別函數，將觀察資料的各個變數值代入以求得一個預測值，再對其進行分類，事後則可據以驗證其預測的準確度。不過，運用區別分析時仍須注意下列限制（施能仁、方南芳，民 86）：(1)預測用變數（即自變

數) 必須符合多元常態分配; (2) 樣本中不同類別之資料均具有相同的共變異數矩陣; (3) 當預測變數不為常態分配時, 區別分析估計式將不具統計一致性, 亦即估計係數可能因喪失統計一致性而會產生偏誤; (4) 未考慮變數自我相關的問題; (5) 預測變數的選用必須取決於研究者主觀的抉擇, 可能會忽略重要變數。

(二) Logistic 迴歸法

logistic 迴歸模式的基本形式與一般線性迴歸 (如: 複迴歸) 大致相似, 但是 logistic 模式之因變數 (被解釋變數) 的結果必須轉換成介於 0 到 1 之間的機率值 ($0 \leq P \leq 1$), 因此其因變數未必符合常態分配假設, 此即與一般線性迴歸法最主要差異之處。若模式中因變數的值為 0 或 1 兩種結果之一, 則應採用二元 logistic 迴歸法進行分析。而若因變數為多元順序尺度 (multiple ordinal scale) 或類別尺度的變數, 例如: 可將客戶信用或違約程度分為 1, 2, 3, 4 及 5 等五個等級, 等級愈高者表示信用愈不好 (或愈好), 或違約程度愈嚴重 (或愈不嚴重), 則可將此五個等級以 0 到 1 之間的值表之, 1 表示 $0 \leq P < 0.2$, 2 表示 $0.2 \leq P < 0.4$, 3 表示 $0.4 \leq P < 0.6$, 4 表示 $0.6 \leq P < 0.8$, 5 表示 $0.8 \leq P < 1$ 。此種多元尺度之因變數的 logistic 迴歸法, 在 SPSS 統計軟體中即稱為多元 logistic 迴歸 (multinomial logistic regression)。Logistic 迴歸的基本模式如下:

$$P_i = F(Z_i) = \frac{1}{1 + e^{-Z_i}} \quad (3)$$

$$Z_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n \quad (4)$$

其中, P_i 為事件 i (違約或未違約) 發生的機率, $F(\cdot)$ 為 Logistic 的累積機率分配函數 (cumulative density function), X_n 為迴歸式中第 n 個預測變數, 預測變數個數的多寡則視預測模式而定, 而 β_n 則為迴歸式中第 n 個預測變數的係數。

陳錦村 (民 86) 比較層級分析法 (analytic hierarchy process, AHP), 區別分析法與 Logistic 模式對銀行授信客戶信用的區別能力, 結果顯示 Logistic 模式的整體區別能力優於其他兩種方法。Espahibodi (1991) 的研究也發現, 在解釋變數不符合常態分配假設時, 採用 logistic 迴歸的預測準確率較佳。過去研

究的結論，大多認為由於 logistic 迴歸對解釋（或預測）變數的分配並無特定限制，當解釋變數中同時有離散(discrete)及連續(continuous)變數時，logistic 迴歸法所建立之預測模式的預測準確率通常較高。此外，logistic 迴歸模式除可預測分類組別之外，尚可計算事件發生的機率（李桐豪及呂美慧，民 89）。實際運用此類預測模式時，決策者可視當時經濟景氣、違約頻率、或銀行資金狀況與承受違約損失能力等，自行調整分類的臨界機率值高低（一般是以 0.5 為臨界值），以達到更佳的預測效果。

二、研究變數的選定

研究變數的選擇方式上，本研究主要是參酌過去文獻，並經與多位專家進行訪談，以及考量取得資料的可行性等。首先，從研究主題與本研究較為相近之相關文獻中，選擇其中四篇文獻所採用之預測變數（彙整於表 1），並按變數屬性將其分為客戶特性變數與房屋貸款契約特性變數兩類。其次，研究期間作者曾訪問於金融機構服務期間超過二十年，或徵授信經驗超過十年之資深經理及徵授信人員。以銀行評估信用五項原則：借款戶 (people)、資金用途 (purpose)、還款來源 (payment)、債權保障 (protection)、授信展望 (perspective) 進行實務上之討論。最後，再根據訪談結果及資料取得之可行性，對表 1 中的變數略作增刪，本研究最終選用表 2 中的預測變數。表 2 除了按客戶特性與房貸契約特性分別列示各項變數名稱外，並包括本研究在實證分析時使用各變數的操作性定義，以及預期各變數對客戶發生違約可能性的影響方向。

表 1 各相關研究所採用之預測變數彙整

預測變數	Steenackers 及 Goovaerts (1989)	江百信及 張金鶚 (民 84)	杜慶麟及 張瑞芬 (民 87)	李桐豪及 呂美慧 (民 89)
客戶特性變數：				
抵押品現值：買賣價			X	
抵押品現值：淨現值			X	
抵押品狀況：屋齡			X	
抵押品狀況：位置	X		X	
抵押品狀況：屋況			X	
通信處與抵押品同區域				X
住宅所有權	X			
年齡	X	X	X	X
教育或專業程度／學歷			X	X
居住情況／地區別			X	X
是否有電話	X			
居住現址與在現職工作的時間長度	X			
家庭背景：家中依賴人口數			X	
家庭背景：婚姻狀況（有無配偶）			X	X
財務狀況：收入（含配偶）或所得	X	X	X	X
財務狀況：不動產		X	X	
財務狀況：其它動產			X	
年利息支出佔年收入比例				X
職業狀況：服務年資			X	X
職業狀況：職業類型	X	X	X	X
職業狀況：是否在公家機關工作	X			
保證人=借款人自己				X
保證人=借款人配偶				X
對保證人之綜合評價			X	
與本行往來情形			X	
經辦面談之評價			X	
與銀行往來實績		X		X
先前貸款次數	X			
房貸特性變數：				
利率減碼			X	
貸款成數			X	
貸款年限／期間／期數	X	X	X	X
貸款金額		X		
是否先繳息一段時間再攤還本息				X

資料來源：本研究整理

表 2 本研究所採用的預測變數

客戶特性變數：	操作性定義	對違約發生的影響方向
1. 抵押品價值：買賣價	銀行鑑價與買賣價孰低	Uncertain
2. 抵押品現值：淨現值	銀行現在重估可能之市價	Uncertain
3. 抵押品狀況：屋齡	幾年	正向
4. 抵押品狀況：位置	1. 都會區 2. 都市邊緣區 3. 鄉鎮中心 4. 鄉鎮邊緣區	正向
5. 抵押品狀況：屋況	1. 都會區透天厝 2. 鄉鎮透天厝 3. 都會區公寓 4. 鄉鎮區公寓	正向
6. 抵押品狀況：面臨道路	1. 10 米以下巷弄 2. 10-20 米 3. 20 米以上道路	負向
7. 抵押品與所有權人居住關係	1. 住宅現為所有權人居住地 2. 住宅與所有權人現在居住地同一鄉鎮（區）不同村（里） 3. 住宅與所有權人現在居住地同一縣市不同鄉鎮（區） 4. 住宅與所有權人居住地不同縣市	正向
8. 抵押品所有權人	1. 借款人本人 2. 借款人之配偶 3. 借款人之父母 4. 借款人之子女 5. 借款人之其他親戚 6. 借款人之同事 7. 借款人之朋友	正向
9. 抵押品提供者	1. 借款人 2. 保證人	正向
10. 年齡	貸放時之歲數	正向
11. 性別	1. 男性 2. 女性	正向
12. 教育或專業程度／學歷	1. 博碩士 2. 大學 3. 專科 4. 高中（職） 5. 國中以下。	正向
13. 通訊處與抵押品位置是否同一區域碼	1. 是 0. 否	負向
14. 家庭背景：家中依賴人口數	子女人數	正向
15. 家庭背景：婚姻狀況（有無配偶）	1. 有婚姻關係 2. 無婚姻關係	正向
16. 財務狀況：收入（含配偶）或所得	年收入	負向
17. 財務狀況：不動產	1. 尚有其他不動產 2. 已無其他不動產	正向
18. 年利息支出佔年收入比例	以銀行借款利息支出佔收入之百分比	正向
19. 職業狀況：服務年資	年數	負向
20. 職業狀況：職業類型	1. 政府機關 2. 醫療、金融保險業 3. 買賣、服務業 4. 製造業 5. 交通運輸業 6. 農林漁牧業 7. 建築、裝潢、礦業土石業 8. 餐飲業 9. 自由業（如家庭代工）	正向
21. 職業狀況：職稱	1. 軍公教金融機關主管級以上或相當職位人員或專院校講師級以上及具專門技術自由業者或意代表及上市、上櫃公司經理級以上人員。 2. 軍公教金融機關編制內正式職員或中小型民營企業事業主或大型民營企業課長級以上人員。 3. 大型民營企業之正式職員及中小型民營企業之基層主管以上 4. 中小型民營企業之正式人員或軍公教機關臨時雇員、其他自由業（農、林、魚、牧）、一般技術人員（美髮美容、油漆水泥工、水電工） 5. 無固定職業者（如家管）。	正向
22. 職業狀況：是否在公家機關工作	1. 是 0. 否	負向

表 2 本研究所採用的預測變數 (續)

客戶特性變數：	操作性定義	對違約發生的影響方向
23. 保證人與借款人關係	1. 無保證人 2. 配偶 3. 父母 4. 子女 5. 其他親戚 6. 同事 7. 朋友	正向
24. 保證人是否=本人	1. 是 0. 否	負向
25. 保證人是否=配偶	1. 是 0. 否	負向
26. 保證人評價	1. 差 2. 稍差 3. 普通 4. 佳 5. 頗佳	負向
27. 與本行往來實績	1. 本行授信戶滿三年往來正常或活期性存款最近 6 個月平均餘額 20 萬元以上或定期性存款 100 萬元以上之一。 2. 本行授信戶滿二年往來正常或活期性存款最近 6 個月平均餘額 10 萬元以上或定期性存款 50 萬以上。 3. 本行授信戶滿一年往來正常或活期性存款最近 6 個月平均餘額 5 萬元以上。 4. 往來無不良紀錄之本行客戶。 5. 其他尚未與本行往來者。	正向
28. 經辦面談之評價	1. 差 2. 稍差 3. 普通 4. 佳 5. 頗佳	負向
29. 與銀行往來情形	0. 在其它銀行無借款 1. 在其它銀行有借款	正向
30. 先前貸款次數	曾於本行貸款次數	負向
房貸特性變數：		
31. 貸款成數	貸放金額／銀行鑑價	正向
32. 貸款年限／期間／期數	月數	負向
33. 貸款金額	實際撥貸金額	正向
34. 拍賣後收回金額	法院拍定分配款	uncertain
35. 拍賣後認列損失金額	本行認列損失 (包括未收回之本金、利息、違約金)	uncertain

正向：表示數值越大越容易違約

負向：表示數值越小越容易違約

uncertain：表示尚難確定其影響方向

肆、實證結果與分析

一、選樣與敘述統計

本文之研究對象為國內某銀行一南部分行的房屋貸款客戶，從該分行民國 80 年 9 月至民國 90 年 6 月間所有申請獲准之房貸客戶資料中，以隨機方式選出 300 個客戶資料做為建立預測模式的樣本，其中在核貸後曾發生違約者計有 100 戶 (本文稱之為違約戶)，未曾發生違約者則有 200 戶 (正常戶)。樣本中所有客戶資料，除「抵押品現值」為事後 (蒐集研究資料當時已知) 的資料外，其餘均為申貸當時的資料。由於本研究是以實質上放款能否收回為判定依據，故違約者係指發生下列情況之一者：(1) 已屆清償期而未受清償「或未辦理展期」之各項放款及其他授信款項；(2) 對於各

項放款及其他授信款雖未屆清償期，但銀行已依契約通知喪失期限利益，請求提前清償，視同到期者；(3)各項放款及其他授信款項除另有規定外利息未繳超過三個月者。

另外，本研究為比較不同預測模式的準確度，特從已認列損失或拍賣抵押房屋之違約房貸客戶資料中，隨機篩選出28個客戶資料做為測試樣本(testing or hold-out sample)進行分析，這些資料的核准貸放日期介於民國80年8月至民國86年10月間，經拍賣後認列損失時間則介於民國85年1月至民國90年12月間。本研究選用此一測試樣本資料，除做為預測準確度的檢驗之外，也特別考慮同時檢驗預測模式之準確度與可降低錯誤分類損失之程度。根據Nanda及Pendharkar(2001)的研究指出，好的企業破產預測模式不僅要有較高的預測準確率，更重要的是，要能極小化錯誤分類的成本(minimize misclassification costs)。而銀行對客戶房屋貸款之徵、授信決策，也有相似的情形，若要建立良好的房貸客戶違約預測模式，除了要求較高的預測準確率，同時也應考慮能否使徵、授信決策錯誤的損失降至最低。

表3列示上述300戶預測模式樣本與28戶(純為違約客戶)測試樣本的敘述統計資料，且300戶資料又再區分為正常戶與違約戶，分別做敘述統計分析。而300戶資料中，並無「拍賣後收回金額」及「拍賣後認列損失金額」的資料，因為只有經過法定程序處分後的28戶違約客戶資料，才有這兩項資料。值得注意的是，28戶測試樣本的資料屬性似乎與預測樣本中100違約戶有頗大差異，例如在「抵押品買賣價」、「抵押品現值」、「抵押品屋齡」、「年收入或所得」、「服務年資」、「職業類型」、「貸款成數」及「貸款金額」等項目上。這些差異隱含某種意義，表示28戶測試樣本可能有代表性不足的問題。

表3 預測變數之敘述統計分析

預測變數	300 戶樣本						28 戶樣本	
	全體		200 戶正常戶		100 戶違約戶		平均數	標準差
	平均數	標準差	平均數	標準差	平均數	標準差		
1. 抵押品價值：買賣價	499.09	269.33	518.00	289.93	461.29	218.98	529.50	281.24
2. 抵押品現值：淨現值	336.64	163.11	359.13	172.68	291.65	131.7	4234.82	117.60
3. 抵押品狀況：屋齡	8.70	7.94	7.95	8.02	10.20	7.59	3.00	3.96
4. 抵押品狀況：位置	3.28	0.68	3.25	0.69	3.34	0.65	3.18	0.76
5. 抵押品狀況：屋況	2.16	0.64	2.10	0.57	2.28	0.75	2.57	0.90
6. 抵押品狀況：面臨道路	1.38	0.66	1.34	0.67	1.47	0.63	1.36	0.67
7. 抵押品與所有權人居住關係	1.80	1.06	1.78	1.05	1.86	1.09	2.46	1.21
8. 抵押品所有權人	1.16	0.69	1.09	0.33	1.30	1.09	1.21	0.77
9. 抵押品提供者	1.08	0.27	1.07	0.26	1.10	0.30	1.11	0.31
10. 年齡	38.77	9.63	37.82	8.96	40.68	10.64	37.36	8.73
11. 性別	1.32	0.47	1.31	0.46	1.34	0.48	1.43	0.49
12. 教育或專業程度／學歷	4.07	0.91	3.93	0.94	4.36	0.76	4.50	0.78
13. 是否同一區域碼	0.76	0.43	0.77	0.42	0.73	0.45	0.50	0.50
14. 家庭背景：家中依賴人口數	1.79	1.25	1.74	1.23	1.91	1.29	1.86	1.12
15. 婚姻狀況（有無配偶）	1.17	0.38	1.17	0.37	1.20	0.40	1.21	0.41
16. 收入（含配偶）或所得	74.38	47.40	79.78	55.04	63.59	22.77	72.57	34.81
17. 財務狀況：不動產	1.82	0.38	1.85	0.36	1.77	0.42	1.71	0.45
18. 年利息支出佔年收入比例	38.96	18.46	34.74	16.69	47.40	19.00	51.39	19.69
19. 職業狀況：服務年資	9.17	6.28	8.78	6.13	9.94	6.53	4.29	2.70
20. 職業狀況：職業類型	4.86	2.29	4.28	1.94	6.04	2.47	7.07	2.30
21. 職業狀況：職稱	3.44	1.00	3.21	1.02	3.89	0.80	4.21	0.72
22. 是否在公家機關工作	0.07	0.26	0.10	0.29	0.03	0.17	0.00	0.00
23. 保證人與借款人關係	2.83	1.48	2.55	1.17	3.39	1.85	3.18	1.89
24. 保證人是否=本人	0.02	0.14	0.02	0.12	0.03	0.17	0.04	0.19
25. 保證人是否=配偶	0.66	0.48	0.74	0.44	0.50	0.50	0.61	0.49
26. 保證人評價	2.82	0.85	2.98	0.84	2.50	0.79	2.46	0.68
27. 與本行往來實績	3.65	1.58	3.49	1.62	3.98	1.46	4.32	1.20
28. 經辦面談之評價	3.15	0.68	3.30	0.69	2.85	0.58	2.79	0.49
29. 與銀行往來情形	0.16	0.40	0.14	0.40	0.20	0.40	0.29	0.45
30. 先前貸款次數	0.62	0.99	0.66	1.01	0.54	0.96	0.50	1.24
31. 貸款成數	4.95	1.52	4.88	1.59	5.09	1.39	6.37	0.86
32. 貸款年限／期間／期數	18.02	3.86	17.90	4.04	18.26	3.49	19.64	1.29
33. 貸款金額	243.35	139.35	245.76	139.81	238.52	139.02	330.82	161.94
34. 拍賣後收回金額	-	-	-	-	-	-	206.64	100.39
35. 拍賣後認列損失金額	-	-	-	-	-	-	168.75	98.57

二、本研究預測模式之建構與分析

本研究先以 SPSS 軟體中之二元(binary) logistic 迴歸分析，運用 300 個客戶資料之樣本進行逐步迴歸，來建構房屋貸款客戶違約之預測模式。SPSS 總共提供六種逐步迴歸的方法，包括：forward：likelihood ratio、forward：conditional、forward：wald、backward：likelihood ratio、backward：conditional 以及 backward：wald，故本研究即先以這六種逐步迴歸的結果做為基礎，尋找最佳預測變數組合，以建立本研究初步的預測模式。在逐步迴歸過程中，本研究設定選定進入變數的顯著機率門檻為 0.05，移除門檻為 0.10。逐步迴歸分析的結果顯示，backward：wald 方法所得的預測模式，其平均預測準確率為 85.0%，是六種逐步迴歸方法中最高的。其餘方法之預測準確率介於 83.3% 至 84.7% 之間，預測效果的差異並不大。針對實際違約客戶的預測能力，六種方法的結果均在 70% 或 71% 間，而 backward：wald 方法針對實際違約客戶的預測能力，也達到較高的 71%。表 4 即列示 backward：wald 法所得預測模式的各項預測變數、其迴歸係數、係數之標準差、wald 統計量與顯著性 (P 值) 之數據等，此模式即為本研究所構建的房屋貸款客戶違約預測模式之首選。

在此預測模式之預測變數組成上，表 4 顯示「抵押品買賣價」、「抵押品跌價」（為抵押品買賣價與抵押品現值的差額）、「年利息支出占年收入比率」、「保證人與借款人關係」、「保證人評價」及「貸款成數」等預測變數均極其顯著 (Wald 值大於 10 且 P 值均小於 0.002)，「抵押品屋齡」、「年收入」、「職業類型」、「經辦面談之評價」及「貸款金額」等亦有很顯著的影響 (P 值小於 0.05)。而其中大多數預測變數的係數符號方向均與預期相同，並與前述相關文獻的研究結果相吻合，故不再贅述。「抵押品買賣價」對違約發生有負向的影響，亦即抵押品買賣價愈低愈可能發生違約。這或許是因為違約客戶之經濟能力較差，無力購買高價房屋有關。「抵押品跌價」則對違約發生有正向影響，亦即抵押品跌價愈多愈可能發生違約，這合乎一般徵、授信人員的預期。至於「年收入」與「貸款成數」的係數符號方向與預期不同，這是較值得深入討論的問題。

表 4 以二元 Logistic 迴歸所建立之預測模式

預測變數	β係數	β之標準差	Wald 值	P 值
抵押品買賣價	-0.019	0.0042	6.897	0.000
抵押品跌價*	0.012	0.004	10.016	0.002
抵押品屋齡	0.099	0.034	8.589	0.003
年收入或所得	0.016	0.006	7.239	0.007
尚有不動產	-0.944	0.520	3.297	0.069
年利息支出佔年收入比率	0.107	0.020	29.526	0.000
職業類型	0.198	0.096	4.193	0.041
職稱	0.464	0.277	2.799	0.094
保證人與借款人關係	0.525	0.138	14.393	0.000
保證人評價	-0.817	0.258	10.017	0.002
經辦面談之評價	-0.880	0.385	5.228	0.022
與銀行往來情形	0.818	0.477	2.947	0.086
貸款成數	-0.967	0.268	13.024	0.000
貸款金額	0.016	0.005	8.801	0.003
常數項	3.584	2.464	2.116	0.146
備註：	-2 Log likelihood = 203.802 Cox 及 Snell R Square = 0.448 Nagelkerke R Square = 0.622			

*抵押品跌價 = 抵押品買賣價 - 抵押品現值

樣本資料中所有 300 個房貸客戶年收入的平均值為 74.4 萬元，200 戶正常戶年收入的平均值為 79.8 萬元，100 戶違約戶年收入的平均值則為 63.6 萬元。由此可知樣本資料內的平均值與預期結果相同，這表示年收入愈高的房貸客戶發生違約的可能性應該較低，亦即按常理推斷年收入之係數應為負數（表示負向影響）。但是，〈表 4〉的結果卻顯示年收入的係數為 0.016 的正值，本研究認為這可能是以連續性變數資料執行二元 logistic 迴歸的結果。因為正常戶中高收入者較多（大於 80 萬元者占 33.5% (=67 戶 / 200 戶)），收入金額較大被賦予的係數權重也較大；反之，違約戶中高收入者較少（大於 80 萬元者僅佔 13% (=13 戶 / 100 戶)）。因此，本研究認為此項結果，應是樣本資料分佈偏頗所造成的影響。就「貸款成數」的係數符號方向而言，樣本中 300 個房貸客戶貸款成數的平均值為 4.95（即貸放金額除以銀行對抵押房屋之鑑定價格的 49.5%），200 正常戶貸款成數的平均值為 4.88，100 戶違約戶貸款成數的平均值則為 5.09。由此亦可知樣本內的平均值與預期相

同，表示貸款成數愈高的房貸客戶發生違約的可能性較高，亦即貸款成數之係數按理應為正數（表示正向影響）。但是，表 4 的結果卻顯示貸款成數的係數為-0.967 的負值，因此本研究認為此項結果，亦是資料分佈偏頗所造成的影響。

若採用區別分析法來構建房屋貸款客戶違約之預測模式，SPSS 軟體中之區別分析功能（在分類(classify)功能之下），提供五種逐步迴歸的方法，包括：wilks' lambda、未解釋變異(unexplained variance)、M 氏距離(mahalanobis distance)、最小 F 比率法(smallest F ratio)、以及 Rao 氏 V 值法(Rao's V)等，故本研究即以這五種逐步迴歸的結果，來尋找區別能力最佳之預測變數組合，期能改善本研究的預測模式。在區別分析之逐步迴歸過程中，本研究亦設定選擇進入變數的顯著機率門檻為 0.05，移除門檻為 0.10。區別分析之逐步迴歸的結果顯示，以五種逐步迴歸方法所得的區別分析結果完全相同，其係數未標準化之區別函數式如下：

$$\begin{aligned} Z = & -0.521 - 0.004*「抵押品現值」 + 0.048*「抵押品屋齡」 - 0.551* \\ & 「尚有不動產」 + 0.041*「年利息支出佔年收入比率」 + 0.22*「職 \\ & 業類型」 + 0.264*「保證人與借款人關係」 - 0.379「保證人評價」 \end{aligned} \quad (5)$$

區別分析的結果對整體樣本的平均預測準確率為 83%（請參閱表 5），略低於 logistic 分析所得的預測模式(85%)，主要是因對正常戶的預測準確率只有 85%（低於 logistic 預測模式的 92%），但對違約戶的預測準確率則為 79%，高於 logistic 預測模式的 71%。區別分析所得預測模式中的預測變數僅有「抵押品現值」、「抵押品屋齡」、「尚有不動產」、「年利息支出佔年收入比率」、「職業類型」、「保證人與借款人關係」以及「保證人評價」等七項，這些變數除「抵押品現值」外，均出現在 logistic 分析所得的預測模式中。此一模式的 Wilks' Lambda 值為 0.606，卡方值為 147.392，表示預測變數的綜合區別能力非常顯著（P 值 < 0.001）。

由(5)式可知，除了「尚有不動產」一項變數外，其餘之預測變數的係數符號方向均與預期相同，並且符合過去相關文獻的預測。「尚有不動產」的係數符號原預期應為正向，但在 logistic 分析（參見表 4）與此處之區別分析的結果均為負向，顯示由本研究資料的實證分析結果傾向支持，尚有其

他不動產之客戶發生違約的可能性較高。

然而，經本研究進一步以單樣本 kolmogorov-smirnov(K-S)檢定該區別函數式中之預測變數是否均符合常態分配時，結果卻發現「抵押品現值」、「抵押品屋齡」、「尚有不動產」、「年利息支出占年收入比率」、「職業類型」、「保證人與借款人關係」以及「保證人評價」等七項預測變數的K-S檢定Z值分別為：2.69 (p值< 0.001)、2.89 (p值< 0.001)、8.66 (p值< 0.001)、1.44 (p值=0.032)、5.61 (p值< 0.001)、6.72 (p值< 0.001)以及4.19 (p值< 0.001)，均顯著拒絕這些變數符合常態分配。由於這些預測變數均不合乎常態分配，區別分析的結果將受到違反基本假設的限制。

三、與其他預測模式預測準確度之比較

為評估本研究前一節所建立的預測模式，是否為預測效果良好的預測模式，必須找尋合理且公平的比較基準，而與過去研究之預測模式相互比較，應為最佳的選擇。但為求比較基礎一致，首先，必須以本研究的訓練樣本做為過去研究所建立預測模式的測試樣本，檢驗過去研究預測模式對本研究對象之樣本資料的預測效果。本研究根據Steenackers及Goovaerts(1989)、江百信及張金鶚(民84)、杜慶麟及張瑞芬(民87)、李桐豪及呂美慧(民89)等四篇研究所選用之預測變數組合，形成四種不同的預測模式，而後利用本研究之樣本資料進行二種統計方法的分析，其結果列示於表5。其中表內李桐豪及呂美慧(民89)的預測模式，係採用他們文章中的模式V(李桐豪及呂美慧，民89，第14頁)，根據他們所報導的分析結果，該模式對正常戶(正常貸款)的預測準確率為100%，而在違約戶(非正常貸款)的預測準確率也高達93.27%。但是，以本研究的樣本資料分析後發現，李桐豪及呂美慧(民89)以logistic迴歸分析所建立之預測模式，在正常戶的預測準確率雖高達92.5%，但在違約戶的預測準確率卻僅有31%。由此一結果可知，李桐豪及呂美慧(民89)之預測模式對資料的敏感性較大，意即其模式的相對穩定性較為不足，當測試的樣本資料不同時，其預測準確度的差異較大。

表 5 本研究與過去研究預測模式之預測準確率比較

預測模式	二元 Logistic 迴歸分析				區別分析		
		預測		預測 準確率	預測		預測 準確率
		正常戶	違約戶		正常戶	違約戶	
本研究	正常戶	184	16	92.0%	170	30	85.0%
	違約戶	29	71	71.0%	21	79	79.0%
	全體			85.0%			83.0%
Steenackers 及 Goovaerts (1989)	正常戶	175	25	87.5%	159	41	79.5%
	違約戶	53	47	47.0%	37	63	63.0%
	全體			74.0%			74.0%
江百信及張金鵬 (民 84)	正常戶	182	18	91.0%	153	47	76.5%
	違約戶	52	48	48.0%	39	61	61.0%
	全體			76.7%			71.3%
杜慶麟及張瑞芬 (民 87)	正常戶	181	19	90.5%	157	43	78.5%
	違約戶	44	56	56.0%	28	72	72.0%
	全體			79.0%			76.3%
李桐豪及呂美慧 (民 89)	正常戶	185	15	92.5%	131	69	65.5%
	違約戶	69	31	31.0%	38	62	62.0%
	全體			72.0%			64.3%
以樣本銀行「個人 信用評等表」建立 之預測模式	正常戶	180	20	90.0%	147	53	73.5%
	違約戶	47	53	53.0%	31	69	69.0%
	全體			77.7%			72.0%

過去四篇研究之預測模式中，預測效果最佳者應為杜慶麟及張瑞芬（民 87）的預測模式，在 logistic 迴歸法下，其全體樣本的平均預測準確率為 79%，在正常戶的預測準確率為 90.5%，而在違約戶的預測準確率則有 56%。在區別分析法下，其全體樣本的平均預測準確率為 76.3%，在正常戶的預測準確率為 78.5%，而在違約戶的預測準確率則有 72%。不過，杜慶麟及張瑞芬（民 87）的原始預測模式是經由類神經網路的訓練學習，將資深授信人員或信用分析主管的經驗及判斷以模糊集合 (fuzzy sets) 量化的加權值，而建立成為類神經網路專家系統。由於本研究僅是選用杜慶麟及張瑞芬（民 87）模式之預測變數組合來建立其預測模式，基於分析方法與其原始模式並不相同，故本研究所得預測準確率結果，並不宜與其原始結果做比較。然而，

值得一提的是，本研究所採用的杜慶麟及張瑞芬（民 87）模式的預測變數組合，共包含十八個預測變數，或許是因所選用預測變數較多，而使其預測準確率較高。

欲比較不同預測模式的優劣，除了比較其預測的準確率之外，也應檢視不同預測模式本身的配適度(goodness of fit)，方能確保預測模式的有效性。例如，以二元 logistic 迴歸分析的結果來比較，本研究所建構之預測模式與過去四篇研究的各項配適度檢驗結果顯示，各預測模式的-2 Log Likelihood 值，本研究預測模式的值為 203.802（如表 4 備註中所示），而過去四篇研究之預測模式的值則介於 275.026 至 339.829 之間。由於該值愈小表示模式的配適度愈好，因此本研究的預測模式顯然優於其他模式。

若以另兩項指標來比較，本研究的 Cox 及 Snell R 平方值為 0.448，且 Nagelkerke R 平方值為 0.622，均顯著高於其他四個模式的 Cox 及 Snell R 平方值（介於 0.131 和 0.3 之間）與 Nagelkerke R 平方值（介於 0.182 和 0.416 之間）。由於 R 平方值愈接近 1 表示模式的配適度愈好，本研究之預測模式的兩項 R 平方值，亦優於其他模式。

本研究資料來源銀行（以下稱為樣本銀行）目前對客戶的徵、授信決策，通常會參考該銀行「個人信用評等表」評量的結果。為瞭解本研究所建立之預測模式是否優於該「個人信用評等表」模式，表 5 中亦比較利用該銀行「個人信用評等表」之預測變數組合，所建立預測模式之預測準確度，是否與本研究之預測模式有所差異。結果顯示，雖然「個人信用評等表」之預測變數組合所建立的預測模式，其整體與個別的預測準確率仍低於本研究所建立的預測模式，但在平均上仍優於某些過去研究所建立的模式。

四、各預測模式降低錯誤分類成本能力之分析與比較

由於 Nanda 及 Pendharkar (2001) 認為，好的預測模式不僅要有較高的預測準確率，也要能極小化錯誤分類的成本。然而，對銀行而言，放棄一位正常客戶只是少賺，並無虧損的問題。若有因拒絕正常客戶而損失應賺而未賺的貸放款利息收入，也將因缺乏實際資料而無法具體衡量，故事實上亦無從以實證方法檢驗之。因此，本研究將比較的重點放在檢驗預測模式對違約客戶的區別能力上，僅以事後實際認列損失或拍賣抵押房屋之違約房

貸客戶資料進行分析。為此，本研究特別蒐集該銀行分行因每一客戶違約而實際發生的損失金額資料，28 個客戶的違約損失（含本金與利息損失）總計為 NT\$4,725 萬元。

本研究係以 logistic 迴歸的預測模式進行分析，比較各預測模式可降低錯誤分類損失之程度。由於 logistic 迴歸法在計算與實務運用上較為簡便，故本研究僅選用 logistic 迴歸法而不用區別分析法。實際應用 logistic 迴歸之預測模式時，只要先將每一個觀察對象（房屋貸款客戶）的每一筆記錄資料，根據對映於預測模式之預測變數組合，分別代入(4)式中，因為各預測模式的 β 係數已知，故將各預測變數(X_n)的實際值代入(4)式後，即可求出 Z_i 的值。之後，再將 Z_i 值代入(3)式中，便可求得 P_i ，再根據 P_i 是否大於 0.5 來分類，預測其是否較可能發生違約。最後再按分類結果，計算此一預測模式可降低錯誤分類之個別損失與總計。

表 6 各種預測模式之違約預測準確率與可降低錯誤分類損失程度之比較

預測模式	分類結果		違約預測 準確率	原違約預測 準確率	可降低錯誤 分類損失金額
	正常戶*	違約戶			
本研究(Logistic)	11	17	60.7%	71.0%	\$2,426 萬元
樣本銀行「個人信用評等表」	12	16	57.1%	53.0%	\$2,720 萬元
Steenackers 及 Goovaerts(1989)	13	15	53.6%	47.0%	\$2,399 萬元
江百信及張金鵬（民 84）	10	18	64.3%	48.0%	\$2,899 萬元
杜慶麟及張瑞芬（民 87）	2	26	92.9%	56.0%	\$4,402 萬元
李桐豪及呂美慧（民 89）	19	9	32.1%	31.0%	\$1,231 萬元

*分類至正常戶表示此一分類結果為錯誤分類。

表 6 彙總比較各預測模式之違約預測準確率與可降低錯誤分類損失總金額，即為本研究分別按上述步驟計算所得的結果。在表 6 中，可發現違約預測準確率最高(92.9%)，且可降低錯誤分類成本最多者（高達 \$4,402 萬元）為杜慶麟及張瑞芬（民 87）之模式。實際預測效果最差者為李桐豪及呂美慧（民 89）之模式，其違約預測準確率僅有 32.1%，可降低錯誤分類損失金額亦只達 \$1,231 萬元。而樣本銀行「個人信用評等表」之模式雖僅有 57.1%的違約預測準確率，但卻可降低錯誤分類損失金額達 \$2,720 萬元，其

預測效果比本研究原採二元 logistic 逐步迴歸 (backward : wald 方法) 所建立之模式稍佳，本研究所建立模式之違約預測準確率為 60.7%，可降低錯誤分類損失金額則為 2,426 萬元。由此可知，模式違約預測準確率較高者，其所能降低錯誤分類成本 (或損失) 通常也較大。但或許由於測試樣本較少，先前由訓練樣本所推演出原違約預測準確率最高 (71%) 之預測模式，對測試樣本的違約預測準確率，卻未必亦獲致最佳的結果。未來或後續研究，可設法增加測試樣本來改善此一問題。

在檢視資料的過程中亦發現，「貸款金額」與「拍賣後認列損失金額」兩者間的相關係數高達 0.8189。由此可知，貸款金額愈高，其日後發生的實際損失確實也愈大。基於此，當金融機構之徵、授信人員面對貸款金額較大之案例時，其徵、授信決策制定過程應更加謹慎才是。

實務應用上，銀行可利用本研究之最佳預測模式配合信用評等制度建構一完善之房屋抵押貸款違約預警系統，使每個房屋抵押貸款案件得到較客觀、公平且一致性的審核，避免審核中由於人為的錯誤判斷、人情的關說壓力或不良操守等因素造成授信品質的低落，本研究所建立的預測模式可以有效的找出客戶違約的機率，使銀行對違約風險偏高之客戶於申貸之初即可主動婉拒，對於違約風險較低之客戶可以積極招攬，主管人員或有經驗的徵信人員只須針對邊際客戶詳加查核，減少審核的時間，大幅提升徵、授信人員的工作效率，並且因個人信用資料量化的結果可以隨時經由電腦及統計方法計算出風險變化情形，即使銀行誤將不良違約戶誤判為正常戶，經由本研究之違約預測模式過濾也能使損失真正極小化。銀行若能運用前述所建立之房屋抵押貸款違約預警系統，即可在申貸之時發現偏高的違約風險，並因制度的建立得以阻止管理階層的人為操作。

伍、結論與建議

本研究有別於過去文獻之主要貢獻在於：(1)重新建構一個適合南部地域性的房屋貸款客戶違約之預測模式，(2)同時檢驗此類預測模式的預測準確率與可降低錯誤分類損失能力，這是過去研究未曾涉獵的範圍。由實證結果顯示，本研究所建構模式在第一階段分析中，其準確率與模式配適度

均較其他預測模式為佳。但由於諸多預測變數未符合常態分配，區別分析的結果將受到違反基本假設的限制，故而在後續降低錯誤分類損失能力之檢驗上，僅採用 logistic 迴歸進行分析。

此外，在第二階段的分析中，以 28 個已認列損失或拍賣抵押房屋之違約房貸客戶資料，同時檢驗預測模式之準確度與可降低錯誤分類損失之程度。經分析後發現，模式違約預測準確率較高者，其所能降低錯誤分類成本（或損失）通常也較大。但或許由於測試樣本較少（只有 28 筆觀察資料），可能存在代表性不足問題，原先由訓練樣本所推演出準確率最高之預測模式，對測試樣本的預測準確率，卻未必亦獲致最佳的結果。

在運用本研究所建立之預測模式前，亦須注意些許限制。首先，由於本研究樣本的代表性可能不足，本文的研究結果未必能一般化至適用所有銀行的房屋抵押貸款決策。其次，本研究所採用的統計分析方法均可歸類為線性迴歸模式，若資料的實際特性偏離線性迴歸模式的假設很多，則本研究結果或許會受到影響。未來研究或可採用非線性分析模式，或是類神經網路模式，其結果可能更臻理想。再者，在可降低錯誤分類損失金額之分析中，本研究的測試樣本只有 28 個觀察值，或許會因樣本規模較小，而使研究結果有所偏誤。未來研究除可增加測試樣本（應該拒絕卻被錯誤接受之客戶資料）規模之外，亦可考量將正常戶（應該接受卻被錯誤拒絕之客戶資料）錯誤分類所造成的損失納入分析模式之中。最後，過去研究所建立之預測模式對資料的敏感性或許較高，例如：李桐豪及呂美慧（民 89）之預測模式，意謂其模式的相對穩定性較為不足，當測試樣本資料不同時，其預測準確度的差異可能較大。故銀行界在實務上應用此類模式時，應注意其模式適用的範圍或對象，甚至要考慮其在地域上的適用性。

參考文獻

1. 王泰裕 (民 90), 模糊類神經網路在消費性貸款之應用, 成功大學工業管理學系碩士論文。
2. 江百信及張金鵬 (民 84), 「我國購屋貸款放款條件之研究」, *住宅學報*, 第三卷第一期, 頁 1-20。
3. 李桐豪及呂美慧 (民 89), 「在金融機構房貸客戶授信評量模式—以 Logistic 迴歸分析」, 第一卷第一期, *台灣金融財務季刊*, 頁 1-20。
4. 杜慶麟及張瑞芬 (民 87), 「銀行授信決策應用神經網路之研究—抵押貸款之實證研究」, *模糊系統學刊*, 第四卷第一期, 頁 31-44。
5. 施能仁及方南芳 (民 86), 「以類神經網路建立台灣儲蓄戶助社財務危機預警模式」, *台灣經濟*, 第 247 期, 頁 34-80。
6. 陳錦村 (民 86), 「銀行授信客戶之信用評等與模式比較」, *輔仁管理評論*, 第四卷第一期, 頁 145-72。
7. 陳錦村、許通安及林蔓蓁 (民 85), 「銀行授信客戶違約風險之預測」, *管理科學學報*, 第十三卷第二期, 頁 173-95。
8. Espahibodi, P. (1991), "Identification of Problem Banks and Binary Choice Models." *Journal of Banking & Finance*, 15, No. 1, pp. 53-71.
9. Nanda, S. and P. Pendharkar (2001), "Linear Models for Minimizing Misclassification Costs in Bankruptcy Prediction." *International Journal of Intelligent Systems in Accounting, Finance & Management*, 10, No. 3, pp. 155-68.
10. Steenackers, A. and M. J. Goovaerts (1989), "A Credit Scoring Model for Personal Loans." *Insurance: Mathematics and Economics*, 8, No. 1, pp. 31-4.
11. Tam, K. Y. and Y. Kiang (1992), "Managerial Applications of Neural Networks: The Case of Bank Failure Predictions." *Management Science*, 38, No. 7, pp. 926-47.

Soochow Journal of Economics and Business

No.48 (March 2005) : 103-126.

A Comparative Study of Predictive Models for Mortgage Loan Default

Chia-Hsing Huang^{*} Yung-Ming Hsieh^{} Tsung-Je Liou^{***}**

(Received: November 27, 2002; First Revised: April 2, 2003; Second Revised: August 14, 2003;
Third Revised: October 17, 2003; Accepted: November 4, 2004)

Abstract

The debtors' default on mortgage loan is an important issue for Taiwan local banks. It does not only deteriorate the capital structure of banks but also damage their profit margin especially during financial crisis or economic depression. Therefore, a default prediction model for banks to improve the mortgage loan decisions and/or to reduce their potential losses is highly needed. This study tries to establish a better prediction model of the debtors' default for bank's mortgage loan decision by searching a better combination of predicting variables.

Using two statistical techniques (binary logistic regression and discriminant analysis), this study analyzes the accuracy rate of prediction and the goodness of fit for two models built in this study and five other models (four from past research and one from the "individual credit evaluation form" of sample bank) and then compares their performance with these indicators. In addition, the abilities for all models to minimize misclassification costs are also compared. The result shows that the combination of predicting variables formulated by this study is a great model either in the performance indicators or in the ability to minimize the misclassification costs.

Keywords: Mortgage Loan, Prediction Model, Credit Rating, Debtors' Loan Default, Logistic Regression

* Associate Professor, Department of Finance, National Yunlin University of Science and Technology.

** Instructor, Department of Accounting, Soochow University; Ph.D. Candidate, Department of Accounting, National Taiwan University.

*** Assistant Banking Officer, Taiwan Business Bank.
