

學術論著

金融機構住宅房屋貸款信用評分系統之建構研究

The Building of Credit Scoring System on the Residential Mortgage Finance

莊瑞珠* 陳穆貞**

Rwei-Ju Chuang*, Mu-jean Chen**

摘要

本研究以金融機構分析住宅房屋貸款授信風險的方向，評估最能衡量借款戶信用、償債能力的預測變數，由發掘房貸授信風險評估因素，並依照各個因素對於信用狀況的影響程度給予不同的權重，利用勝算比的估計，決定迴歸估計係數的權重，建立較完整評估系統與事前的風險量化研究模型，藉以評定授信條件策略，以期提高信用良好比例，減少銀行呆帳的發生。本研究顯示分別以月付比例、過去信用狀況、貸款成數、借保關係為住宅房屋貸款信用好壞的顯著相關因素。在線性條件預測模型下，建立邏輯斯迴歸分析模型，並依此結果製成業界所能直接應用的評分卡，揭露依評分模型所計算出評分分數等級與所屬百分位點，協助資訊使用者瞭解評分所代表的實際風險意義，並評估風險模型適用性。因每一借款者違約所造成金融機構的呆帳損失並不相同，應用決策樹方法偵測不同風險程度的信用組合，作為建構以住宅房屋貸款信用評分分級為基礎的授信審核政策之參考。

關鍵詞：住宅房屋貸款，授信風險，評分系統，決策樹，邏輯斯迴歸模型

ABSTRACT

The volume of credit business in the residential mortgage finance has greatly expanded and the use of credit scoring through the evaluation of large credit portfolio becomes crucial to guard against any management risk. The objective of this study is to devise a credit scoring system for finance granting decisions. We describe statistical method to create scorecards and show how the result of the model is applied to calculate score point weights. Scorecards are built using the logistic regression method which estimates the relationship between the individual characteristics and the log of the odds (risk) so that the score point weights can be calculated directly from the regression coefficients. The model performance is usually monitored by the model validation and classification error. We propose an alternative measure for power of model discriminations and credit-granting decisions.

Key words: residential mortgage, credit risk, credit scoring system, decision tree, logistic regression model

(本文於2006年11月17日收稿，2007年4月20日審查通過，實際出版日期2007年5月)

* 輔仁大學管理學院統計資訊學系副教授。

Associate Professor, Department of Statistics and Information Science, College of Management, Fu Jen Catholic University, Taiwan, R.O.C.

** 成功大學公共衛生學系碩士。

Master, Department of Public Health, College of Medicine, National Cheng Kung University, Taiwan, R.O.C.

一、緒論

金融產業主要經營業務為傳統之存款收受及放貸兩業務，兩者間利差(spread)為其重要利潤來源。而台灣於西元90年代開放新商業銀行加入經營，新競爭者不斷加入，並允許各銀行增加分支機構據點與逐漸引導外國銀行來台設立分行，促使了銀行業自由競爭日益激烈；而金融經營環境的不確定性增加，造成銀行業經營風險大幅提高。然而截至今國內過去五年內平均存放款利率利差下降比率異常大(2.54%下滑至2.07%)在資金市場相對充裕情況下，造成利基損失。其中有關房地產相關放款包括房屋貸款與建築融資兩主要業務，出現產業過熱訊：本國銀行總放款約為17兆元，其中住宅貸款和房屋修繕貸款與營建業放款所佔比例26% (為其它一般產業四倍)，顯示銀行對房屋建築融資單一產業的放款比重，已存在過高風險。而且現行銀行對相關貸款削價競爭，房貸平均利率普遍較低，利潤已近微薄，銀行其實承擔不起任何潛在逾放損失。同時銀行亦為競爭業務爭取客戶，對房屋融資貸款成數曾高達百分之百甚至更高，導致客戶購屋貸款超過實際房價，引發另一種信用過度擴張的危機。根據西元2007年台灣金融統計顯示，此類貸款成數過高房貸的寬限期(只繳利息不用返還本金的期限)多為一至三年，且相當集中今上半年到期須開始還本金，房貸支出立即增加三倍以上；一旦中央銀行利率持續上揚，加重繳息負擔，對於這些過度擴張信用的房貸消費群，存在延滯繳款可能性。而對經營房貸的銀行界，已預期將承受相當程度的逾放損失可能風險。如何執行有效風險管理，將成為金融機構核心能力與策略優勢的競爭關鍵。整體而言，金融產業當然為影響經濟發展之重要產業。早期基於金融穩定之考量，各政府對金融業一向採取嚴格的管制措施，然而，自西元1970年起，由於經濟環境的變遷，加上世界潮流趨勢的影響，各國的銀行逐漸走向自由化與國際化的發展方向。同時各國政府也為因應金融業高度競爭，於是若干國際上共同資本管制的協定的需求，其中以「巴塞爾資本協定」(The Basel Capital Accord)最具有代表性。主要是規範金融機構的信用風險，為全球銀行監理所需遵循的最低共同標準，其中要求銀行必須針對其授信資產依規定權數計算『信用風險性資產總額』，並據以計提一定比率的最低適足資本以吸納其風險。資本適足率為決定銀行機構是否倒閉的重要指標。最新協定Basel II於西元2006年起實施，其中二項主要要求內容為1.信用風險資本的計提，改用外部信用評等結果，以決定適用風險權數大小。2.增加監理機關覆核程序及市場制約機能，以避免銀行濫用內部模型。其中規定各金融機構準備至少三年以上的有效風險評估資料，並展示有效風險評估模型，這意味著銀行至少必須在2007年底前完成風險管理資料蒐集相關系統的建置，同時確保相關資料的完整性與數字的即時性。而風險管理與評估原就是銀行營運管理的必要環節，但隨著Basel II施行期限的迫近，更讓風險管理躍升為各銀行的首要重點工作。房地產業及金融產業息息相關，而且台灣銀行業經過現金卡信用卡雙卡風暴後，個人消費金融貸款業務全面萎縮轉向衝刺房地產相關放款，讓在雙卡風暴中利潤受傷銀行得以喘息機會，雖難預測房地產未來後期市場趨勢，但當台灣金融政策，開始執行全面監控銀行Basel II對風險控管能力，銀行業界亦開始緊縮放款，並針對該高風險房貸，增提擔保債務呆帳，對國內的房地產業及本身金融業影響不容忽視。

對經營房貸業務的銀行界而言，購屋貸款是一種消費者金融產品，非常仰賴經營績效行銷、行銷手段與服務技巧。「風險」原本就是預測的概念，要做到預測和管理風險，銀行

必須參考過去的歷史資料，正確分析並提早規劃。然而，銀行的資料原本就散佈在不同部門，這不僅提高了資料蒐集和分析的困難度，且當資料逐日劇增時，銀行有更多的決策需要擬定，但回應市場與客戶的時間卻相對變少，也讓銀行愈難精確評估風險，因而暴露在高風險。台灣市場購屋貸款同質競爭高，經營利基已限於極有微薄的空間。所以，不論是回應Basel II的要求、滿足業務競爭需求，或提升資產品質，現行較可能成功的經營策略，必需將購屋貸款產品區隔多樣化與分眾化，產生差異化策略，且各個產品需維持相當程度的經營規模與良好的經營效率。另一方面，過去標準管理制度，涵蓋應用於大部分客戶的機制，不再適用市場發展；必需針對產品區隔及信用等級不同，發展獨立的風險管理模型。

雖然台灣主要房貸銀行已開始建立信用評分系統，但因為需要投入相當統計技術研發與成本，加以新消費金融產品不斷推出，能擁有完善的評分系統的發卡銀行仍屬不多。在Basel II要求下，需升級資料攫取與模型建立的準確度。即使已有某些內部評等系統，仍缺乏新協定規定蒐集客戶風險資訊來形成風險組合的系統資本適足要求。金融機構用以評斷信用風險的準則概分為三種：經驗法則(judgmental approach)、信用評等準則(credit ranking)為與專家系統準(expert system)準則。其中經驗法則係傳統上指依據授信及回收經驗在試驗過程後，列出拒絕授信對象；或經由徵信人員個人之經驗，以主觀判斷作成授信決定。此法簡單易行缺乏客觀標準制度不健全，易生弊端；僅適用於小額信用或在未建立客觀制度前，小量的申請作業程序。信用評等將客戶信用品質細分為若干項目，然後有系統的對每個個別項目分別考評等級，再綜合等級用以代表該顧客的信用評價。本研究所探討的信用評分制度(credit scoring)，不同於信用評等，在於需對評等項目系統化區隔並予以權重標準評分，亦可同時選定決定信用風險的重要解釋變異組合，並加總各評等項目下的分數，而且製作信用評分表格。而專家系統由於為一種「自動」評核系統，用以輔助或取代授信專業人員的審核工作，其中包括傳統知識庫系統、範例學習法、類神經網路等，亦可設定總評分的審視標準。唯對各評等項目的權重(係數)，無法個別給予統計顯著性的檢定。但對於處理大量申請者資料庫，有其專業上一定的準確性與即時性，尤其對於應用於解釋變數與反應變數之間的非線性關係更加適用。值得注意的是，在大部分操作風險測量中，任何量化方法必須克服用極少觀察線索資料(例：違約事件發生在不同時間點的重複案例不夠多)估計風險機率分配的統計問題。在操作風險評估方法仍不一致，無法準確反映風險。

故在本研究中，我們綜合信用評等與評分制度的方法，發展能顯著衡量購屋貸款債信能力的統計建構模式，以作為事前風險量化研究的基礎。本文主要內容分第貳部分探討有關常見信用評分方法優缺點比較之文獻探討，第參節中本研究所選擇使用的類別相關分析方法以探求與購屋貸款顧客信用風險(credit risk)的相關解釋因素，並依此建立業界能直接使用之信用評分線性模型。在第肆節中對所建立模式討論適合度診斷方式，尤其針對不同授信標準下評估計分卡的區分能力。最後第伍節綜合以非線性假設下討論個案發卡銀行如何偵測顧購屋貸款客在評分下的風險組合，以提供銀行業者在發展購屋貸款市場上有更精確的策略。

二、文獻探討

當消費者向銀行申辦房貸時，銀行為建購房貸信用評分的房貸客戶資料庫，會對申辦內容有主要兩方面考量，：主要是標的物的擔保品價值，銀行會透過個別不動產鑑價系統

(collateral appraisal system)，考量擔保品環境、交通、價位、格局及實際使用情況等因素，作為評估貸款額的依據。國內研究(江百信、張金鵬，1995)對個案銀行購屋貸款放款條件研究中亦指出，不論建築融資或購屋貸款，標的物的價值皆為銀行決定融資額度的主要原因。其二就是貸款人的信用狀況，係指申請人在金融機構之支票存款、借款與信用卡等往來不良紀錄(例如延滯、逾期、催收與呆帳等)且民間負債等條件，通常輔以貸款人基本社會經濟變數(social demographics variables)評估借款人償債能力與意願。若已有某些信用不良紀錄但經加徵核貸條件(如加徵具有正當職業或擁有不動產保證人)考慮降低貸款額度或以縮短放款期間、加速還款等條件後得以承做者。這部分因為牽涉到個人償款能力，必須收集包括負債、金融往來狀況、年齡、收入、職業…等個人財務變數與基本資料，評定消費者除具有穩定收入來源，以期在貸放後繳款能正常化，不需經過處分擔保品確保銀行債權。就違約之影響因素探討，此類貸款人基本社會經濟變數雖不是直接顯著因素(Herzon & Earley, 1970)，但亦造成抵押遲付的重要解釋變數，造成銀行業交易成本之增加(Canner et al., 1991)。因此本研究基於以上文獻探討與考慮業界實用價值，除以上兩大類解釋違約風險因素，加入第三類經核准房屋貸款種類的特徵，其中包含借款人權益與償還能力因素，影響違約的重要因素(Lawrence et al., 1992) (例如貸款金額、貸款成數、每月償還本利和佔家庭所得之比例等貸款條件)，共18項解釋違約風險變數列入為驗證評等模型範疇。

理論上成熟的信用授信市場，常基於假設消費者在某一特定期間內的行為會遵循過去的某些規則性的變動，可利用過去決策的經驗，來合理評估分析現行的申請行為。個別信用放款機構在消費性金融市場所面臨的主要問題，是在於如何對申請人進行信用風險及償債能力的評估。因每一借款者違約所造成個別金融機構的呆帳損失並不相同，信用組織系統中的信用部門，必須蒐集到足夠個人徵信資訊，並對資料完整循序挖掘潛在客戶的整體信用品質的變數，以為信用風險評估的基礎且依影響信用重要性的大小，給予權重的決定，可分別量化顧客風險等級及估計相對風險(機率)，並就決策的自動化可提供直接應用的模型，則稱統計模型下的信用評等(credit scoring)制度。信用評分基本原理是藉由分析技術，由過去資料研發能預測未來授信客戶表現的分數，假設在授信審核時已知的客戶特性，將與授信客戶未來是否準時還款有關，一旦找出關連的規律性，即可套用現有資料做為未來預測。此種假設和早期授信人員基於主觀審核經驗授信客戶行為的情形基本上類似，所不同者是藉由軟體與演算法的輔助，可使預測更精確。但通常經過愈久時間，信用評分模型預測力會愈差，所以信用評分模型建立完成後需持續監控以確保有效解釋信用的風險。

Mays(1998, 2001)簡敘建立申請信用評分制度(application scoring)的過程，係由為數眾多(兩年間狀況下)對單一同性質授信產品申請族群中，獲取其相關的資料(如：申請書上所記錄的資料、申請人的信用記錄)。通常申請人的信用記錄期間可設定為1年至2年間。由於消費者的違約率乃是消費者為該機構客戶的時間函數，且為該機構客戶的時間超過12個月之後，違約率才會開始平穩。因此，任何太短的時間間隔都會低估違約率，而且無法反應出用以預測違約的完整特徵性質。但是，如果時間間隔若大於24個月，信用評分系統會由於母體特徵性值的分配隨著時間變動而改變。一般而言，通常為連續3個月未繳款者為違約顧客，另可視個別信用卡發行機構需求設定不同條件。因此此類評估制度的操作期間(operational life)通常為5年：即以前2年為收集過去顧客資料，其餘3年間發展及測試顧客風險等級之統計模型。信用

評估資料分析，通常使用在一固定階段時間內全部申請者的資料，因此有關利率變動、金融環境、提前清償、二胎抵押貸款或不同地區間的房貸研究情因素之考量，不納入本研究考範疇。因此資料來源極少來自定義完整母體下隨機抽樣機制，屬於個案產品資料預測問題而非一般性推論研究。

一般而言，驗證評等模型涵蓋範圍相當廣泛，可依據不同的對象與應用方向，將信用評等概分為四大類：

- (一) 申請信用評等：針對新客戶，依據其基本資料、工作情況、職位高低、居住區域與其他可資證實之財務資料，來預測未來是否會發生信用問題。
- (二) 逾期收款信用評等：針對已發生信用延遲付款之客戶群，進行可能還款之預測，藉以集中心力處理較高可能性之可回收個案。
- (三) 行為評等：針對客戶之消費行為(如：刷卡消費行為)進行瞭解，並建立未來行為預測模型。預測項目包括找出對銀行高度偏好的高貢獻客戶群、對該行其他金融商品有興趣，以及對某特定商品具有高度消費可能性的族群，有助於擬定跨售與行銷策略。
- (四) 行銷反應評等：針對過去對行銷活動有反饋的客戶群進行瞭解，以便掌握並預測未來行銷時客戶可能的反應。可作為行銷計畫執行前之效果評估參考，讓銀行得以用最精省的花費進行最有效的行銷活動。

信用評等應用在歐美已盛行多年，大部份歐美銀行都擁有內部的信用評等機制，自行開發以客戶行為偏好度預測為主的信用評分模型，應用在包括車貸、現金卡、企業破產信用評分和企業期中審查等方面。國內信用評等的應用則還在萌芽階段，大部份銀行仍零星購買國外的已建購信用評等模型，應用於消費金融信用為主，尚未自行研發建置。因此基於以上的限制，若需全面性完整的信用決策分析，需先利用此研究模型所篩選誤判群體，另結合徵信中心所提供必需付費的外部負債資(通常區分為五大類資料：繳款類、負債類、其他信用申請類、信用長度類與信用型態類)，對更完整與精緻的資料庫作交叉分析，進行該機構獨立全面性之風險權重評斷。根據歐美銀行的經驗，即時且精準的信用風險評估能有效降低5%到15%的壞帳率，同時降低5%到10%的重覆核貸作業成本。以積極面來看，銀行可進一步將未提撥的資本準備金做最佳運用以創造更高獲利。

由於一般評分卡的主要作用探討影響繳交購屋貸款項發生逾期之因素，並將探討的因素作為分析模式的預測變數，藉以建構一個統計上的迴歸評分模式。分析主要目的，是找出隱含於資料的相關(correlations)、通則(patterns)、集群(clusters)與趨勢(trends)，亦常需要經過資料轉換中較複雜的數學運算，結合數種統計方法。國外發展的重要預測技術，包括統計模型中常見的線性判別函數、羅吉斯迴歸、類神經網路及決策樹，僅提出較具代表成功適用業界的內容加以討論。其中Srinivisan & Kim(1987)最先使用羅吉斯迴歸，依據各個變數之權重給定信用分數，並可用來篩選重要變數，尤其對解釋變數之資料尺度不受限制，可考慮解釋變數之間共線性並進行預測的資料。Epley et al.(1996)研究抵押貸款違約分析時，也同樣利用邏輯斯迴歸分析、probit分析及區別分析三種統計方法，個案分析結果顯示，此三種統計方法不但在分類正確性上有顯著的差異，並且影響抵押貸款違約的顯著因素及係數正負符號，也會有不同的結果；但模型區分能力而言，以邏輯斯迴歸分析之群組分類正確率為最高(95%)，其他兩種方法相較較低(例如該區別分析之群組分類正確率為80%)。Henley(1996)利用線性判別

函數，可依據各個變數權重給定信用分數可用來篩選出重要變數，無法處理解釋變數之間共線性問題，應用層面有極大限制。Boyle(1992)使用決策樹方法，找出目標變數與解釋變數存在互動效果，並區分兩個以上的目標群體。雖然資料尺度不受限制，亦可找對目標群體解釋區分的重要因素組合，終究仍無法給定權重，以進行評Yobas(2000)以較少量的資料應用類神經網路方法，應用於解釋變數與目標變數為非線性關係資料分析，但對業界實用目的應須考量學習模擬訓練次數，以避免過度配適。

由於評分卡是決定信用申請人特徵的相對加權總分之制度，並以相對風險的觀念，給較高風險者較低分數；反之給較高分數。授信市場中適用的評分制度，必需對評等項目系統化區隔並予以權重標準，亦需同時選定決定信用風險的重要解釋變數組合，並加總各評等項目下的分數而且製作信用評分表格(score cards)。然而，線性判別函數為早期用於區分顧客信用風險群的方法，但受限制解釋變數之資料尺度(包含類別型)與共變數矩陣相等，而且無法顯示信用卡基本變數間強烈的共線性問題。決策樹在其自動選擇變數和表現變數間交互關係有強大的功能；但無法給定權重，以進行評分。因此，單一決策樹方法的應用無法產生實務上所需的評分卡，應佐以複合統計方法(Hybrid model)應用。而羅吉斯迴歸方法為其中一種不受限制解釋變數之資料尺度，找出顯著變數線性組合與相對權重來預測風險，鑑別區分信用風險群體的重要有效方法，並可同時進行必要的統計顯著性檢定，但必須以周詳的類別相關分析方法以事先選出與顧客信用風險顯著解釋因素，才能顯示出模型之鑑別度。(Capon, 1982; Radding, 1992; Royston, 1994; Grablowsky & Kalley, 1981; Srinivisan & Kim, 1987; Steenackers & Govaerts, 1989; Gilbert et al., 1990; Carter & Catlett, 1987; Davis et al., 1992; Eisenbeis, 1978; Makowski, 1985)相對於購屋貸款市場的成長快速，國內專對於房貸的信用風險管理的研究多屬探討影響房屋抵押貸款風險主要因素在單一統計方法的應用(如logit或probit迴歸評分模式，江百信、張金鶚，1995；李桐豪、呂美惠，2000；周建新、于鴻福等，2004)或著重於方法論等學術性討論議題(丁正中，2004)且少見專對於信用評等建構之完整定價區隔評估系統的研究分析。

以本文以預測信用風險的方向，利用羅吉斯迴歸模式進行分析，以預測機率值(乘以10)以作為評估顧客信用分數形成評分基礎，並以決策樹分析中CHAID (Chi Square automatic interaction detector) Bonferroni同時修正卡方檢定法(Miller, 1991)將預測變數的反應水準進行分割與合併，以求得以最顯著的預測變數的最小分群數。其目的為偵測個案銀行顧客，經逐次搜索樹狀分析下最能解釋信用風險變化因素研究，結合羅吉斯迴歸方法及決策樹，有效應用兩種方法的優點，希望建立能兼具實證研究與實際應用目的之信用評分卡制度。

三、研究方法

(一) 研究變數與操作性定義

建購房貸信用評分的房貸客戶資料庫，首先依顧客消費/償還信用行為分三群：好(有授信用價值)顧客、壞(無授信用價值)顧客，以及無法判定價值的顧客(通常指尚未產生足夠繳款記錄)。信用評估卡制度的建立主要是為了區分好壞顧客的特性，故第三類屬於無法判定價值顧客群的資料，通常無法直接使用。因為此類資訊可能同時包含部份好壞顧客，無法形成有助

於判讀顧客行為的資料庫，但可用為後續對模型預測能力的參考驗證資料來源。

當借款者未正常償還其貸款時，金融機構根本無法得知此未償還是屬於暫時性的或永久性的，所以當借款者第一次未照規定時間來償還其貸款時，金融機構皆先將此遲付視為暫時性。金融機構通常在借款者遲付10天後展開催繳行動，催收階段可分為超過正常償還日30天、60天及90天三種。如果借款者在90天內將逾期未還之部分加以償還，則該遲付行為稱之為「暫時性的遲付」，可能原因在於借款者有健忘的習慣或一時資金週轉不靈導致無法按期償還貸款本息，但情況不致於嚴重到違約(Rakes, 1973)。此時，金融機構雖不需立即執行抵押品贖償權利；但是遲付行為已導致金融機構及借款者兩方皆必須付出處理遲付問題的交易成本。實務操作上亦發現有遲付記錄借款者的違約機率是正常借款者高。因此，借款者在償還本息過程中，如果曾有遲付情形發生，金融機構應盡快的決定及評估借款者遲付的原因及其目前的財務狀況，以避免借款者發生真實違約的行為而造成金融機構之呆帳損失。當遲付行為已超過90天時，這種遲付已不再是暫時性的遲付行為，此種問題放款案件會進入違約處理程序，此時金融機構本身會考量可能產生的違約成本(default costs)與目前借款者抵押貸款的負債餘額兩者，來決定是否以延長貸款期限等各種方法來協助借款者渡過難關。任何經測試後正式使用的評估卡制度，亦須隨時因經濟變動(例：全球性不景氣、美國911攻擊事件等)、市場規模(例：同業間激烈的競爭策略等)，甚至社會價值觀改變，對預測變數的反應能力再測試與評估，才能即時減少企業損失。

因為以過去顧客行為為樣本的評估準則，將無可避免產生判斷現有顧客風險的誤差，因此真正適用的評估卡制度，必須掌握研究目標群體的特性，同時將樣本分為訓練樣本與測試樣本，模型研發完成後需驗證，以避免研究人員抽樣時產生偏誤或忽略部分重要因素，造成模型實際運用時產生落差。故此資料分析的範圍包括研究銀行自西元1995年1月起至1996年12月底之2年內所有顧客族群其中資料筆數共508筆，依變數為是否有違約，解釋變數共18項，包含非間斷型尺度變數7項及非間斷型尺度變數11項。在此研究中操作性定義如下詳列：

反應變數(Y = 1/0)：正常/違約顧客，其中違約定義：貸款者在貸款期間未繳本息超過180天或本金超逾約定清償期90天以上未辦理轉期或清償者。

解釋變數(X)：為整體呈現目標房屋貸款顧客群的特性，將影響是否正常償還其貸款的風險因素，共有18變數分為三大方面：第一類為貸款者個人特徵變數，包含性別、婚姻狀況、教育程度、申貸年齡、職業、過去信用狀況、服務年資、月收入。第二類為貸放款類型相關變數，包含有無政府優貸、月付比例、貸款成數、貸款金額、貸款產品類型、保證人狀況、借保關係。第三類為房屋擔保品價值相關變數，包含房屋類型、房屋使用用途、屋齡。

(二) 統計分析方法

1. 類別關聯性分析

此關聯性分析是以削減錯誤比例(proportional-reduction-in-error, PRE)的概念，來降低預測誤差，彌補卡方檢定無法說明行列變數之間的關聯程度與方向的缺點。由於本研究的資料特性因業界實務需求設計為間斷變數(類別變數)，其中PRE越大，代表變數之間關聯強度越大，並且可以檢定預測變數之間的關聯程度方向。採用由PRE概念中較穩定關聯性測量統計量為Goodman & Kruskal's λ (1954)做分析(註1)。

2. 邏輯斯迴歸模型(Logistic Regression Model)

在實務運用上，選用邏輯斯迴歸模型來建構評分模型，結果可直接產生評分卡，在模型實行上成本較少也較快速。邏輯斯迴歸模式是用來處理依變數屬於類別變數的一種統計分析方法，因依變數可能包含多種可能狀態，常被用來分析一個二元的反應變數。其特性在於利用logistic變數轉換，使反應變數轉換為介於0到1之間的機率值，其中定義反應變數Y為1(代表事件發生)和0(代表事件不發生)。

若假設有p-1獨立的解釋變數，令其向量定義為 $x' = (x_1, x_2, \dots, x_{p-1})$ ，則反應變數的條件機率定義為 $P(Y = 1|x) = \pi(x)$ 為發生事件之機率，

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}}, \text{ 其中 } 0 \leq \pi \leq 1$$

邏輯斯迴歸模式通常表示為：

$$g(x) = \ln \left[\frac{\pi(Y = 1|x)}{1 - \pi(Y = 1|x)} \right] = \beta_0 + \beta_1 x_1 + \dots + \beta_{p-1} x_{p-1}$$

經過logit的轉換後， $g(x)$ 是參數的線性組合，範圍是從 $(-\infty, \infty)$ ，與變數X呈線性關係且為單調遞增/遞減特性，更能處理 $P(Y = 1|x) = \pi(x)$ 發生事件之機率範圍限制的問題。迴歸係數最大似估計式，具有統計一致性與有效性的優點。

3. 評估模型配適程度方法

在線性迴歸中，通常都是利用反應變數之配適值與實際觀察值之差異(即殘差)，來建構檢定適合度的統計量。在此研究中，利用Hosmer-Lemeshow統計量(1980, 1982) (註2)討論整體適合度及Deviance統計量(註2)亦可由個別殘差值 $(Y - \hat{Y})$ 差異作檢定計算，可獲得極端值或影響值(influence value)的部分訊息。

4. 評估配適羅吉斯模式區別能力之方式

(1) K-S兩個獨立樣本統計檢定

信用評等模型應能區分違約授信與正常授信之間的差異，驗證此模型是否具有足夠能力區分信用狀況的好壞，並評估模型的誤差率是否在合理可接受範圍內。建議的方法包括K-S兩個獨立樣本統計檢定(Kolmogorov-Smirnov two-sample test) (註3)。

當信用評分模型建置完成後，利用兩樣本累積相對次數分配之K-S檢定，驗證違約授信戶的評等分數分配是否與正常授信戶的評等分數分配是否一致，以瞭解所建置的模型是否能區分兩種不同的授信客戶。當K-S值愈大時，愈能證明兩種不同的授信客戶的評等分數分配並不一致，因此愈能利用評分去區分違約授信與正常授信之間的差異。

(2) 判別表(Classification Table)

判別表是利用配適邏輯斯迴歸模式所得到的預測機率值，與事先設定切斷點作比較，形成原始觀察機率值與預測機率值交叉二維表，透過分類表中的正確判斷率，可判斷模式各個情況誤差。切斷點的不同設定與兩群體樣本數相對個數，會影響正確判斷率。一般最常使用的切斷點為0.5，以表示事件發生的機率是均等的，即假設發生違約的機率為均等，但實務上發生機率顯見並不相等。同時邏輯斯迴歸模式所配適的預測機率值，屬於連續數值，此分類表僅利用切斷點二分法做區分，已失去部分預測能力訊息。若僅以單一判別表評估模型全面

區分效力，會隱藏預測能力不佳情形，因此需同時檢視各個情況下的誤差。故基以以上兩個原因，此研究中我們考慮多種不同的切斷點，並依實務對偏差加以調整，可使模型的預測更趨精確。

四、實證結果與分析

(一) 敘述統計與二元關聯性分析

本研究中影響是否正常償還其貸款的風險因素，共有18變數，其中，申貸年齡、服務年資、月收入、月付比例、貸款成數、貸款金額、屋齡皆為非間斷型尺度變數，除了保留原始尺度下訊息完整分析外，為便於觀察與測量此兩變數與信用好壞是否有差異與關聯程度大小，也將此類變數分組，使組內變異小而組間變異大(DH分層原則，註4)，可利於觀察違約行為並直接應用於評分制度基準。綜合分析結果於下表一，並且解釋初步敘述統計與條列可能關連性分析。

表一 影響是否正常償還其貸款的風險因素敘述統計

		樣本數	分佈比例	違約數	違約比例
性別	男	251	49.5%	19	7.6%
	女	257	50.5%	32	12.5%
狀況 婚姻	已婚	385	75.8%	33	8.6%
	未婚	123	24.2%	18	14.6%
教育程度	國中小	120	23.6%	18	15.0%
	高中	185	36.4%	19	10.3%
	大專	189	37.2%	14	7.4%
	碩士以上	14	2.8%	0	0%
職業	公務員	24	4.7%	3	12.5%
	五師	14	2.8%	0	0%
	上市櫃職員	4	0.8%	0	0%
	私人機構負責人	139	27.4%	15	10.8%
	私人機構職員	226	44.5%	19	8.4%
	自由業	60	11.8%	9	15.0%
	家管與其他	41	8.1%	5	12.2%
保證人數	無	21	4.1%	6	28.6%
	1人	442	87.0%	36	8.1%
	2人	43	8.5%	7	16.3%
	3人	2	0.4%	2	100%
借保關係	直系親屬	380	74.8%	29	7.6%
	非直系親屬	66	13%	8	12.1%
	其他關係	44	8.7%	9	20.5%
	無關係	18	3.5%	5	27.8%
優貸 政府	有	102	20.1%	5	4.9%
	無	406	79.9%	46	11.3%
狀況 房屋	自住屋	455	89.6%	45	9.9%
	非自住屋	53	10.4%	6	11.3%

表一 影響是否正常償還其貸款的風險因素敘述統計(續)

		樣本數	分佈比例	違約數	違約比例
貸款類型	新購屋	182	38.5%	32	17.6%
	他行平轉	49	9.6%	5	10.2%
	舊客戶增貸	79	15.6%	3	3.8%
	他行增貸	106	20.9%	10	9.4%
	舊客戶平轉	32	6.3%	0	0%
	原有房屋貸款	60	11.8%	1	1.7%
狀況信用	信用正常	460	90.6%	41	8.9%
	曾經違約	48	9.4%	10	20.8%
房屋類型	非小坪數	427	84.1%	46	10.8%
	小於20坪	56	11.0%	4	7.1%
	國宅	9	1.8%	1	11.1%
	工業區	26	5.1%	0	0%
申貸年齡	21-35歲	211	41.5%	23	10.9%
	36-40歲	121	23.8%	7	5.8%
	41-45歲	76	15.0%	8	10.5%
	46-55歲	80	15.7%	8	10.0%
	56以上	20	3.9%	5	25.0%
服務年資	0-3年	194	38.2%	29	14.9%
	4-7年	133	26.2%	9	6.8%
	8-11年	90	17.7%	2	2.2%
	12-19年	39	7.7%	3	7.7%
	20-30年	52	10.2%	8	15.4%
月收入	2萬-5.4萬	131	25.8%	23	17.6%
	5.5萬-9.4萬	213	41.9%	22	10.3%
	9.4萬-40萬	164	31.3%	6	3.7%
貸款金額	0-200萬	138	27.2%	7	5.1%
	201-300萬	138	27.2%	15	10.9%
	301-450萬	139	27.4%	14	10.1%
	451-600萬	58	11.4%	11	19.0%
	601萬以上	35	6.9%	4	11.4%
成貸數	0-60%	123	24.2%	4	3.3%
	61-100%	385	75.8%	47	12.2%
屋齡	0-3年	164	32.3%	22	43.1%
	4-11年	114	22.4%	14	27.5%
	12-17年	111	21.9%	7	13.7%
	18-23年	86	16.9%	5	9.8%
	24年以上	33	6.5%	3	5.9%
比例月付	0-29.5%	153	30.1%	1	2%
	29.6%-44.5%	190	37.4%	13	25.5%
	44.6%以上	165	32.5%	37	72.5%

1. 性別特性：

女性貸款不論在貸款金額、貸款成數與表現違約行為比例均高於男性，係可能基於個案。實證結果中顯示，女性表現為平均較低教程度(高中以下佔70%，而男性則少50%)，且符合較低服務年資與收入的趨勢。

2. 婚姻狀況中：

已婚貸款人多居於有一人以上的得證人(九成以上)與借保關係多為直系血親(二者比例各九成)是假設為正常繳款得可能因素，但相較於未婚者，明顯在收入與服務年資較高條件下，有較低但非顯著差異(檢定 $P > 0.05$)的違約比例。交叉分析顯示是否已婚對於購屋行為/用途、貸款來源與過去信用記錄並無顯著差異(檢定 $P > 0.05$)。

3. 在教育程度上：

呈現教育水準愈低，違約比例愈高趨勢，而多為申貸年齡較高的族群，合理解釋台灣教育日益普及的現象。

4. 職業方面的分佈

證券金融上市櫃職員申貸比例最低，而其它私人機構人員比例最高(各為0.8%與45%)。除了常見之「五師」之稱職業種類與金融從業人員較低違約比例外，其他職業別間無顯著差異性。

5. 貸款購屋用途與類型

購屋用途為自住型多為較低年齡層，平均收入歸於中低，但多為過去無不良信用狀況之貸款人。其購屋類型坪數多大於20坪以上、屋齡較低，而飛自住型則相當程度呈現明顯相反趨勢；顯示非自住型可能存在投資行為，但自住型多傾向於貸款成數較高，亦需監控其付款行為。

6. 貸款金額、成數與月付比例上

貸款金額、成數與月付比例上，皆呈現和違約比例正相關性，且違約比例明顯在貸款金額大於450萬、成數大於六成且月付比例高於60%較高，是另一值得注意授信風險的區隔族群。

整體有關關聯性分析，則利用卡方檢定，Goodman, Kruskal λ 統計量的不對稱關聯性，以檢定各個解釋變數之水準變動與申請人違約關聯性並由兩種檢定中得到一致的關聯性分析：教育程度、過去信用狀況、月收入、月付比例、貸款成數、貸款金額及借保關係解釋變數(初步表示 $p\text{-value} < 0.05$ 結果)可能與有關，但因考慮上述關聯性測量統計量數值皆偏低(介於0.1與0.4)雖亦為類別型關聯性分析常見觀察現象，但僅將此部份關聯性分析視為篩選風險變數的初步過程，並檢定變數間交互作用的共線性(註5)且依統計線性模型原則加以查證以確立邏輯斯模型(莊瑞珠，2007)。

(二) 邏輯斯迴歸模型分析

本研究雖非大型資料，但若以全部有效樣本數來建立模式，則難以取得其他類似機密性資料集來驗證本研究建構模型是否能有效預測違約行為。因此依比例配置原則，將全部有效樣本數隨機分為訓練組、受測樣本兩群。分群原則依兩群樣本之群內差異大與群間差異小之

特性(得到結果依訓練樣本：測試樣本=70%：30%原則)其中以訓練樣本來建立模式，以測試樣本集以確立邏輯斯迴歸模型預測能力。

同時為避免隱藏模型預測能力不佳情形，需同時檢視各個情況下的誤差，我們考慮多種不同的切斷點下模型預測率(表二)，在實務成本考量下，依3種預測正確比率(依次為94.9%，63.3%與90.8%)比較基準，選擇以切點為0.6的邏輯斯迴歸預測模型。

因此，在此個案實證分析之最適多元邏輯斯模式(表三)可得到與違約行為最有關的4個變數，分別為月付比例、貸款成數、信用狀況及借保關係(模式依逐步進入法達顯著差異($\alpha = 0.05$))，其中估計指數勝算比(odds)用來解釋相對違約機率的關係：例如，過去信用狀況存在不良紀錄客戶之違約/正常比為信用狀況存無不良紀錄1.098倍；借保關係為非直系親屬關係之違約/正常比為關係為直系親屬關係

五、建構住宅房屋貸款信用評分系統

(一) 評分卡模型結果分析

由邏輯斯迴歸模式，其在各變數水準組合下所得的預測機率值 \hat{p} (即預測該顧客違約機

表二 受測樣本在各機率切點下邏輯斯迴歸模型預測能力正確率

切點機率	預測正常客戶群 正確率	預測違約客戶群 正確率	整體預測正確率
0.2	99.3%	6.7%	90.1%
0.3	99.3%	6.7%	90.1%
0.4	98.5%	33.3%	92.1%
0.5	96.4%	40.0%	90.8%
0.6	94.9%	63.3%	90.8%
0.7	86.9%	66.7%	84.9%
0.8	80.3%	73.3%	79.6%
0.9	63.5%	93.3%	68.4%

表三 最適多元邏輯斯模式估計係數與指數勝算比率

邏輯斯模式違約行為之顯著變數	參數估計值	估計指數勝算比
月付比例(連續尺度)	0.052	1.053
貸款成數(連續尺度)	0.066	1.068
過去信用狀況_不良紀錄	0.094	1.098
借保關係_非直系親屬關係	0.498	1.645
借保關係_非親屬關係	1.138	3.121
借保關係_其他關係	1.541	4.669

註：參考組設定分別為過去信用狀況無不良紀錄與借保關係為直系親屬關係的1.645倍，若借保關係為非親屬關係比率更高達3.12倍。

率)，利用 $1-p$ 乘上10以轉換成申請人之信用分數。此分數之意義為在該顧客之申請基本條件下，所可獲得的分數。再以10分數為間隔分類，藉以觀察顧客群在不同的評分分數下，違約比例之分佈狀況。

如表四可得知，正常/違約累積比例差異最大在51分至未滿60分，從總人數統計方面，可知大部分顧客的信用分數落在60分至75分與75分以上兩分數區間。

(二) 計分卡預測模型配適度評估

計分卡區分能力評估則以Hosmer-Lemeshow統計量及Deviance統計量(註2)顯示邏輯斯迴歸模式(以預設切斷點0.6)所建構之計分卡，依照所建構出正常/違約顧客之信用分數分佈，有明顯的差異(2種統計量檢定 $p\text{-value} < 0.01$)顯示預測模型配適良好。

(三) 計分卡模型區分效力分析

由以上模式所建構之評分卡，考量正常/違約比例與業界誤判風險單位成本，配合KS統計量(註3)中正常/違約累積百分比差異最大(40.27%)在61分至75分之結果，以信用分數為60分作為此案例中正常/違約的得分切割點。此切割點可當作潛在客戶群風險試算判別界線標準值，事先預防並採取措施，以期降低授信逾放比率。

由表五可得知整體正確預測百分比分別為90.8%；正常繳款的正確預測比例為94.9%；違約行為的正確預測比例為63.3%。違約客戶被誤判為正常之比率偏高的原因，為資料基本結構的不完備與設定研究目的應用。由資料的取得，為係基於授信銀行在最邊際風險管控時間

表四 正常/違約比例與評分分數統計表

分數區間	繳息正常 樣本數	違約 樣本數	繳息正常 比例	違約 比例	Fa(X)	Fb(X)	差異
11~20	1	2	0.22	3.92	0.22	3.92	3.70
21~30	0	0	0.00	0.00	0.22	3.92	3.70
31~40	2	7	0.44	13.73	0.66	17.65	16.99
41~50	4	0	0.88	0.00	1.53	17.65	16.12
51~60	10	5	2.19	9.80	3.72	27.45	23.73
61~75	23	11	5.03	21.57	8.75	49.02	40.27
75以上	417	26	91.25	50.98	100.00	100.00	0.00

表五 預測模型判別表：以0.6為預測機率切斷點

實際繳款狀態	預測行為狀態		正確預測百分比
	正常	違約	
正常	130	7	94.9%
違約	7	8	63.3%
總計	137	15	整體正確率=90.8%

成本下，研究目的建構的個人信用評分結果應用為金融機構進行評估的初步標準；只包含個人與金融機構往來時申請信用時所提供之填寫基本正面資料，如收入、職稱、服務年資、動產狀況等，為產生本信用評分的唯一基礎，而且整體的模式預測力在預定達成目標內(90%以上)。業界可參考發卡目的與成本需求做調整界定信用好壞準則(表二)，亦可選擇其他分數作為評估的準則，得到不同的判別率。

對模型中觀察已違約被誤判為可正常授信客戶，需加以進一步偵測，就其還款行為採取即刻性的信用觀察回饋評分，藉以掌握信用預測模式之效度，並有效進行風險控制。綜合誤判分析顯示，違約被誤判為可正常授信客戶7人中，大部分(85.7%)違約狀況為暫時性遲繳行為但在事後不同時期多已償還並繳清欠款，非屬風險操作時強烈預警的客戶群，其他嚴重違約需列入壞帳準備人數極稀少。另一方面，此違約族群自本研究所建構評分卡中得到權重分數，相對高於其他本國債務人履行財務承諾之能力的評級，應仍屬合理預測方向。但仍需分析顯示上述基本條件組合，若僅以制式觀察標準，不易定義是否為正常繳款顧客群，在引用此計分卡時，須特別注意這些組合。

(四) 信用結構組合與評級定義說明

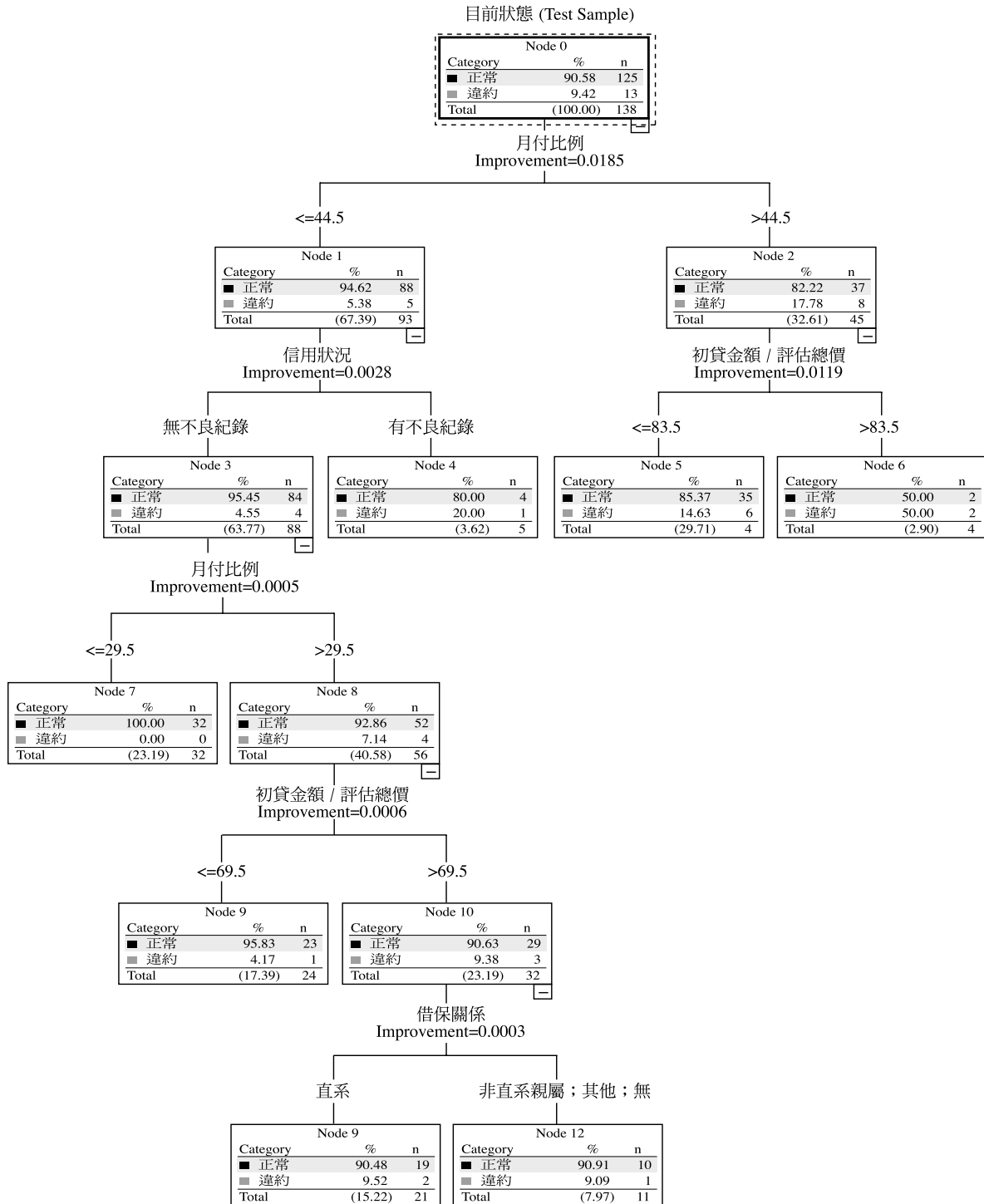
評分模式固然可以預測未來發生客戶違約的機率，但由於信用卡金融商品的特色，預測能力並未必一定代表獲利基礎，必需在風險與獲利的取得風險控制管理的平衡點。此處解釋變數為分類變數，因此由卡方自動互動檢視法(CHAIID)(註6)加以檢視分類結構。下圖一結果顯示：月付比例、貸款成數、信用狀況及借保關係與正常/違約繳款行為有互動結構存在並達成與上述邏輯斯迴歸預測模式一制的結論。

由各變數不同水準組成種組合與其相應評分等級(表六)。並另表(表七)說明該房貸債務人，相較於其他同性質房貸債務人履行財務承諾之能力的評級及建議所實施方案中相對客戶比例。在此個案研究中，信用等級的最佳組合為月付比例低於30%且過去無不良信用紀錄者，等級最差指為月付比例大於45%且貸款成數高於85%，為最該列為風險操作時強烈預警的客戶群。

六、結論與建議

(一) 結論

台灣經歷雙卡冰風暴重創後，金融市場進入所謂的「後卡債時代」，各家銀行紛紛緊縮部分消費性貸款業務，包括信用卡、現金卡以及小額信貸都陷入前所未有的低潮銀行獲利大幅縮水，但是同是屬於金融費性貸款業務的購屋貸款以及和房貸有關的貸款卻因最近幾年資金市場充裕，房貸利率惡性殺價嚴重，在低利刺激下，購屋住宅貸款迅速攀升。此措施固然將使銀行經營短期內獲利帳面趨於穩健，但也將很迅速衝擊購屋貸款市場。尤其不少銀行業界前些年在房貸等授信業務上競爭不惜成本，積極推出百分之百購屋貸款。此種比原本標準風險操作下房貸額度約八成相比，多出的這兩成額度，幾乎由更高風險的個人信用貸款取代，加上房貸利息寬限期加長至七年(通常為兩年的房貸寬限期—債務人在寬限期內，只須還利息，寬限期屆滿才還本金)屆時還款壓力湧現，對銀行營運而言，放款風險確實也跟著提高。反觀美國相對操作成熟第二順位房貸也提前發生違約狀況嚴重相關業務出現緊縮，而台灣部分銀行甚至將第二順位房貸歸類為房貸或列為個人信用貸款項下，被視為是下一波整



圖一 正常/違約繳款行為樹狀分類結構組合

表六 信用等級結構組合表

組合	變數特徵	月付比例	不良信用記錄	貸款成數	借保關係	信用等級	百分比
1		小於30%	無			A-1	23.19
2		大於30%	無	小於7成		A-2	17.39
3		大於30%	無	7至8成5	直系親屬	A-3	15.22
4		大於30%	無	7至8成5	非直系親屬	A-4	7.07
5		大於45%		大於8成5		B	29.71
6		大於45%	有			C	3.62
7		大於45%		大於8成5		D	2.90

表七 信用評級使用表

信用等級	評分	信用評級定義說明	百分比
(A-1)	75以上	有較高評級之履行財務承諾之能力。	67.5%
(A-2)	60-75	履行財務承諾之能力滿意，相對於高等級可能受到環境與經濟條件變動之影響。	15.2%
(A-3)	50-60	履行財務承諾之能力具保障性，相對於高等級容易受到環境與經濟條件變動之影響。	7.5%
(B)	40-30	該債務人相較於其他本國短期債務人，如期履行財務承諾之能力，將顯著地受到不利的企業、金融或經濟條件所影響。	3.7%
(C)	30-20	償債能力令人質疑	3.9%
(D)	20以下	無法履行債務	2.2%

註：參考中華信用評等公司對短期債務所評等之等級定義說明，表示該房貸債務人相較於其他房貸債務人履行財務承諾之能力的評級。

體購屋貸款金融市場的未爆彈。因此值得進一步觀察影響台灣房地產市場交易。根據金管會規定，金融機構應參酌可貸資金成本、營運成本、預期損失、金融市場資金現況及其他經營策略考量等因素，訂定合理的放款定價政策，董事會並應定期審視、檢討。金融機構內部稽核單位也應將執行放款定價政策情形，納入查核項目。一旦發現金融機構未依規定辦理，如放款利率的訂定未合理評估成本，發現預期損失的成本有低估時，金管會還可能要求金融機構增提損失準備。即時與精準的信用風險評估可協助銀行業者深入了解客戶，同時有效控制風險以增加資金應用的靈活度。要達到上述目的，銀行業者必須擁有自行規劃與建置信用評等機制的能力，並確認模式的預測力仍在合理範圍內。若模式預測力與發展模式有顯著差異時，必須調整評分模式的結果或重新建構新的評分模式。要有效管理所有可能的信用風險組合，並就目前可能面臨的風險提出完整的報告，選用內部評等法來評估銀行交易對象的違約機率，需建構多種模型，才足以區別不同顧客組合結構，在不同資產產品的輪廓。

因此本研究之結果，可提供金融業者可依上述信用組合結構(表六)探討，設計定價不同及服務內容不同的商品，既可辨識顧客風險來源，訂定購屋貸款風險指標與作業程序，亦應將購屋貸款各項收費應事先揭露力求透明化，以利消費者參考與多樣性選擇。更應制定購屋貸款專用界訂定義，就不同的遲繳天數，鑑別不同程度違約行為。

由於研究主要方向，為預測未來授信顧客兩年發生違約的機率，但違約事件有許多不同定義，模型應用目的不同，資料來源也有差異；同時愈複雜的定義，資料取得愈困難。而本研究資料來源是基於最邊際風險管控時間成本下，取得只包含個人與金融機構往來時申請信用時所提供之填寫基本正面資料，如收入、職稱、服務年資、動產狀況等，為產生本信用評分的唯一基礎。因此，性質上屬於上述第一類的申請信用評等，也是較易執行但信用評分模型預測力會愈差，因為僅就申請者基本條件變數，但不包括個人與其他金融機構的整體貸款情形、其他票信與信用查詢資料等可能造成負面信用的訊息，其重要性可能更勝於已收集的基本資訊。近期台灣銀行公會理監事會為管理分歧的鑑價方式，將在聯徵中心(JCIC)建置「不動產成交行情及鑑價資訊平台」。目前香港也有這套系統，未來若順利建置，將對台灣房貸市場的過度競爭與放款級距可望大幅改善。此套系統中不動產成交行情將由仲介業提供，鑑價資訊平台由銀行端提供，此系統正式運作後可大幅提高銀行的風險管理。因為目前國內各銀行的房屋核貸成數，部分採買賣價，部分採鑑價方式承做，兩者核貸的金額落差大，若經由此系統，各銀行可獲得透明化資訊，產生制衡效果，使房貸核貸標準趨於平均。

(二) 建議

1. 對金融界的建議

以台灣的金融環境而言，銀行在審核放貸時主要是依賴財金資訊股份有限公司、金融聯合徵信中心所調查之資料，但是在審核系統方面金融環境卻相較國外較弱。在系統方面，近年來許多機構以專家系統迅速審核放貸，而本研究結果發現以邏輯斯迴歸統計方法去建立評核系統會有更準確的預測結果，尤其是對違約上的預測能力會較高。建議銀行可將應用本研究所研發的模式發展個別辨識分析的顧客分佈，持續紀錄並追蹤，機動調整信用利率與額度，取得風險與獲利的平衡點。

一般而言，規劃信用評等機制時必須注意以下事項：

- (1) 建立一個容易存取與整合不同系統資源的智慧架構，讓內部使用者能夠輕鬆存取與謀和必要的協力廠商、申請表單、帳款償還紀錄等資訊，以做出更精確的市場區隔和行銷活動組合。
- (2) 發展更快速、精省且彈性的計分卡，讓銀行能立即發揮資訊投資效益，藉由更精準的市場區隔快速擬定風險容忍範圍內的價格策略，以有效控管壞帳、提升客戶服務和申請信用評分流程的效率。
- (3) 提供完整報告，讓使用者能立即察覺客戶信用評分狀況與信用評分機制的穩定性，以及早將成果報告提供給管理階層參考。

對於同樣的客戶，風險管理人員與市場營業人員可能會有不一致評斷好壞顧客的標準。審核應該要以重要影響因子優先，且對於貸款戶的基本資料應該確實掌握，尤其是還款資金來源，來源的多寡、是否穩定、借貸者之背景等，都需要一制標準評估，共同決定是否放

款。配合徵信人員的工作即是針對授信人員提供的客戶資料進行信用調查，評估額度可行與否。然而信用管理的首要工作是了解客戶特徵，依此建立內部健全客戶信用資料庫，包括付款紀錄、財務狀況、銀行往來情形，不但有助於信用額度的日後訂定，亦可作進一步風險分析，將購屋住宅貸款者依照風險與交易金額綜合比對，定義風險權重的高低，使業界能及早找出風險高組合。若僅以制式觀察標準(一般購屋住宅貸款標準)，不易定義是否為正常繳款顧客群，不論是信用或操作風險，風險管理是需要持續監控的。因此，真正成功的風險管理，必須以如何創造更好的銀行營運為出發點來思考。即使在時間壓力下，必須將風險管理視為日常營運管理和業務流程的一部分，讓組織內的每個人都了解風險管理的重要性。完善的資料蒐集、分析和報告工具，將能幫助銀行有效集中管理資料並呈現風險評估結果，讓組織內的每個人，以共同語言來為控制風險和締造收益。

2. 對購屋住宅貸款消費者建議

個別購屋住宅貸款消費者應多了解自己所屬的風險狀況與信用評分。良好的信用評分，可以幫助該個人申請貸款、信用卡、保險甚至應徵工作等；相對的，違約的信用紀錄，也可能造成某些限制或不便。認知『信用』是個人最有用的資源，不慎使自己信用貶落，可能成個人最沉重的負擔，因此，謹慎維護個人信用紀錄，是任何消費者必須重視的課題。尤其，目前國內各金融機構，依規定都必須定期將個人貸款、信用卡、現金卡等各項紀錄(包括還款紀錄)，送到「財團法人金融聯合徵信中心」建檔，切勿輕忽任何應付的最低繳款額及繳款期限，一旦延遲繳款或付款不足，即會列入個人信用不良紀錄，信用額度被緊縮或取消，而成為金融機構拒絕往來的對象。故消費者須建立「謹慎理財、信用無價」觀念，學習量入為出、一切消費以還款能力為依歸，並重視保護個人的信用紀錄。應認知自我權利與義務，選擇公平合理的定型化契約，注意逾期繳款費用與個人資料之保護。目前政府機關規定：信用往來資料以公開5年內交易紀錄為限，若其中未繳消費超過1,000元以上，則該紀錄公開長達7年；若一旦持卡人清償消費款，該紀錄從清償日後開始算起，揭露年限便減3年。只要信用貸款不論是房貸、信用卡、小額信貸等繳款時間在1個星期內繳交，在聯合徵信中心的記錄即可保持良好，銀行會積極爭取這類客戶業務。Basel II實施前，銀行已著手差別購屋住宅貸款利率定價，推出房貸利率優惠，卻僅限於軍公教、前1,000大企業員工，負債比不能太高；至於一般民眾房貸利率差別利率高約1.5個百分點。Basel II實施後，收入狀況不佳或信用不好的客戶，房貸利率的差距倍數調整差距。受限於中低收入民眾，通常資金運用彈性度較低，可優先考慮政府政策性辦理的優惠房貸、青年貸款、勞宅貸款或公教貸款等；由於利率加碼方式固定，在還款的財務規劃上不會產生暴增的突發個人財務危機。而礙於資格不符無法申辦或其他不足的部分，再佐以銀行自辦的房貸方案申請辦理。以本研究結果建議，把握自備款為房屋總價的3成，每月房貸不超過家庭收入的1/3的穩健購屋計畫亦可提高信用評分。以本研究結果亦顯示房貸成數貸款總額以七成以內為有利，申辦超過8成5房貸容易陷入以債養債的循環。消費者必須針對自己的資金狀況做好還款及理財規劃。除了利率的選擇外，調整房貸利率、還款方式更是另一個資金運用的重要技術。

對購屋住宅貸款消費者市面上常見的房貸類型包括指數型房貸、固定型房貸及理財型房貸等，可供選擇。指數型房貸利率前半年可提供較民營銀行優惠的利率，可貸額度較低。固定型房貸，適合擔心利率上升者，或希望每月固定繳費，希望長期慢慢還款的客戶，這類

房貸存在壽險公司居多，主要是壽險公司擁有長期投資運用的資金。理財型房貸則以消費者可以預先設定好一個額度，然後將以償還的房貸本金轉換為消費者的貸款循環額，一旦消費者需要資金轉做投資用途或其他需求，就可以借出這些金額，以日計息，如果不動用則不計息，利率比小額信貸或現金卡低很多。如果已償還的房貸本金越多，能夠動用的彈性資金當然也越多，但要考量一旦動用的這筆資金，投資報酬率是否能高於循環利息的利率。以往低利率時代設計的指數型房貸半年後的利率將呈現階梯式調升，每三個月調整一次情況下，低利率時代的指數型房貸，利率平均多期數不一定如其他房貸商品較低。不要被表面的低利率蒙蔽，可比較各房貸商品後，算出平均利率，再選擇對自己最有利的商品申貸。如果暫時繳不出房貸，可向銀行提出展延只還利息的貸款寬限期的要求，並與銀行討論，讓每月需償還的本金下降，把原本房貸延長可暫緩難關。若已收到銀行依法律程序催收者，屋主應由自己賣屋來清償貸款，因為若房地產價格居於高點時，至少還可發揮房屋的剩餘價值；如果等到銀行向法院提出拍賣，房屋價格更不划算。如果有保證人或是第三人願意替逾期戶代償，這段逾期所造成的違約金、逾期費用等等，可以透過與銀行協商，給予減免或直接不必付的優惠。因為銀行在意的是金額龐大的房貸本金，在實務上，手續費是談判籌碼，銀行不會強迫繳納，亦可降低個別貸款者支出成本。若已負債應積極付清債務而不是四處搬家躲避，因為付清已送至討債公司的債務無法清除您信用報告上的討債記錄，將該記錄已轉註為已付清，仍可獲得正向的積分。

3. 對後續研究建議

後續研究應可就不同的遲繳天數或其他定義去判定為行為，對模式判別能力與風險經營管理成本影響。就研究目的之不同，將購屋住宅貸款客戶信用狀況再加以詳細區分成多類(例如：信用良好、普通、不佳…類推)，利用多元邏輯斯迴歸(multinomial logistic regression)，亦可做多族群的區別性判斷分析。目前對於購屋住宅貸款資料集甚少，我們只藉由初步審核資料來分析，完成初步申請信用評等與部分的逾期收款信用評等。進一步可延伸探討行為評分表(behavioral scoring)消費能力與消費習慣，針對客戶之消費行為進行瞭解，並建立未來行為預測模型。預測項目包括找出對銀行高度偏好的高貢獻客戶群、對該行其他金融商品有興趣，以及對某特定商品具有高度消費可能性的族群，有助於擬定跨售與行銷策略。而對於過去對行銷活動有反饋但無法判定價值(通常指無足夠刷卡消費記錄)的客戶群進行瞭解，以便掌握並預測未來行銷時客戶可能的反應，可作為行銷計畫執行前之效果評估參考，銀行得以用最精省的花費進行最有效的行銷活動，更完整的行銷資訊。

註 釋

註1：當我們預測行變數(C)或列變數(R)若假設C、R之間關係為以下其一；

(1) C, R獨立(C不為R之函數)

(2) C, R不獨立(C為R之函數)

PRE的定義為：
$$\text{PRE} = \frac{\text{不正確猜測(1)機率} - \text{不正確猜測(2)機率}}{\text{不正確猜測(1)機率}}$$

在此分析中以解釋風險的基本變數為列變數(R)，反應變數為行變數(C)，則已知R來預測C相對不用R來預測C相對改進比率即為PRE的數值。

Goodman & Kruskal's λ 相關強度可分兩預測方向；

$$\hat{\lambda}_{(C|R)} = \frac{\sum_i n_{im} - \max(C_i)}{N - \max(C_j)}$$
 當列已知，預測行減少錯誤百分比，

其中 $\sum n_{im}$ 為在*i*列中最大 n_{ij} 總和， $\max(C_i)$ 為最大列邊際值。

$$\hat{\lambda}_{(R|C)} = \frac{\sum_j n_{mj} - \max(R_j)}{N - \max(R_j)}$$
 當行已知，預測列減少錯誤百分比，

其中 $\sum n_{mj}$ 為在*j*中最大 n_{ij} 總和， $\max(R_j)$ 為最大行邊際值。

Goodman & Kruskal's λ 檢定方法；

$H_0 : \lambda_{(C|R)} = 0$

$H_1 : \lambda_{(C|R)} \neq 0$

檢定統計量抽樣分配型式；
$$\tilde{Z} \approx \frac{(\hat{\lambda}_{C|R})}{\sqrt{\text{Var}(\hat{\lambda}_{C|R})}}$$
 當 $N \rightarrow \infty$ ，

其中
$$\text{Var}(\hat{\lambda}_{C|R}) = \frac{(N - \sum n_{im})(\sum n_{im} + \max(C_i) - 2\sum n'_{im})}{[N - \max(C_j)]^2}$$

且 $(\sum n'_{im})$ 在 $\max(C_i)$ 行中最大的 n_{ij} 。

註2：首先定義

p：解釋變數數目。

X：代表客戶解釋變數所形成的向量 (x_1, x_2, \dots, x_p) 。

J：代為具有相同x^ˆ向量的組合數。J < n。

m_j ：代表擁有相同x值的數目 $\sum_{j=1}^J m_j = n$ ，且n為所有客戶數目。

y_j ：代表在 m_j 個中，其Y值呈現正反應的數目。 $\sum_{j=1}^J y_j = n_1$ ，且 n_1 代表在所有目標個體中，反應變數Y = 1的個數。

在線性迴歸中，通常都是利用準則變數之配適值與實際觀察值之差異(即殘差)，來建構檢定適合度的統計量。而在羅吉斯迴歸中，有數種不同的方式用來測量配適值(由模式所產生的預測值)與實際觀察值之間的差異。介紹如下：

(1) 偏離值統計量(Deviance)

定義：配適值 $\hat{y}_j = m_j \hat{\pi}_j = m_j \frac{e^{\hat{g}(x_j)}}{1 + e^{\hat{g}(x_j)}}$ ，

其中 $\hat{\pi}_j$ 為在第j種解釋變數組合型態下的Y = 1之預測機率值， $\hat{g}(x_j)$ 為在第j種解釋變數組合型態下之估計的logit。

在邏輯斯迴歸模式下Deviance residual定義為

$$d [y_j, \pi_j] = \pm \left\{ 2 \left[y_j \ln \left(\frac{y_j}{m_j \hat{\pi}_j} \right) + (m_j - y_j) \ln \left(\frac{m_j - y_j}{m_j (1 - \hat{\pi}_j)} \right) \right] \right\}^{1/2}$$

則檢定統計量為 $D^2 = \sum_{j=1}^J d^2(y_j, \hat{\pi}_j) \sim \chi^2(J-p-1)$ 。

D² 統計量測量觀察值的概似函數與配似模型後概似函數差距。一般而言，單獨觀察值下的d_j小，則表示配適較佳。

(2) Hosmer-Lemeshow統計量

Hosmer-Lemeshow統計檢定量為一個在統計軟體(例如：SPSS, SAS)中較常使用的適合度檢定量。分別在1980、1982提出利用羅吉斯迴歸模式所得到估計機率來進行分組(grouping)。此處為了討論的緣故，假定J = n，且認定有n行相對於n個估計機率值，且第1行估計機率值為最小，第n行估計機率值位最大。分組的方式乃是根據估計機率值之百分位數(quantile)來進行分組。即令g = 10，即根據估計機率值來分成10組，第1組包含n'₁ = (n/10)個目標個體數為擁有最小估計機率值，而第10組包含n'₁₀ = (n/10)個目標個體數為擁有最大估計機率值。則可以得到一Hosmer-Lemeshow 適合度統計量 \hat{C} ，如下：

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - n_k \bar{\pi}_k)^2}{n_k \bar{\pi}_k (1 - \bar{\pi}_k)}$$

，則 $\hat{C} \sim \chi^2(g-2)$ ，若J ≈ n。

其中n'_k：代表在第k組內總計客戶人數，

c_k：代表在第k百分位數時，解釋變數所形成的型態組合數(covariate patterns)，

o_k = $\sum_{j=1}^{c_k} y_j$ ：代表在所有解釋變數所形成的型態組合(j = 1...c_k)下所有總計的反應變數個數，

$\bar{\pi}_k = \sum_{j=1}^{c_k} \frac{m_j \hat{\pi}_j}{n_k}$ ：m_j $\hat{\pi}_j$ 表示第j種之解釋變數所組合的型態下，顧客個數乘以在第j種解釋變數所組合的型態下事件發生的機率，即為第j種解釋變數型態下發生事件的預期次數，

$\sum_{j=1}^{c_k} m_j \hat{\pi}_j$: 所有解釋變數所組合型態下發生事件的預期次數，

$\sum_{j=1}^{c_k} \frac{m_j \hat{\pi}_j}{n_k}$: 表示發生事件的加權平均機率。

此方法的優點為可提供單一的、簡單的解釋但缺點容易受到估計機率之分組多寡的影響，故在使用時要小心解釋。

註3：KS兩個獨立樣本統計檢定(Kolmogorov-Smirnov two-sample test)

以好壞顧客為兩組獨立樣本； (x_1, x_2, \dots, x_n) 代表正常客群分數， (y_1, y_2, \dots, y_m) 代表違約客群分數，其累積分配各為 $F_x(x)$ 與 $F_y(y)$ 其檢定方法為：

$$H_0 : F_x(x) = F_y(y)$$

$$H_1 : F_x(x) \neq F_y(y)$$

若定義 x 代表一特定信用分數，則 $F_x(x)$ 代表分組正常客群少於此特定信用分數的比例。

同理， $F_y(y)$ 代表分組違約客群少於此特定信用分數的比例。計算 $D = \max_k |F_x(x) - F_y(y)|$ 為 k 組正常客群與違約客群累積比例最大差異的統計量。

決策準則：當兩群體樣本屬於小樣本時($n, m < 20$)，根據所需顯著水準值，可由KS表得知臨界值。其他則使用大樣本常態趨近(large-sample approximation) (Smirnov, 1948)。在本研究中，相對較高的D統計量表示該信用評分分配有較好的區分能力。

$$w_h = \int_{y_{h-1}}^{y_h} f(t) dt$$

$$\mu_h = \frac{1}{w_h} \cdot \int_{y_{h-1}}^{y_h} t \cdot f(t) dt$$

$$S_h^2 = \int_{y_{h-1}}^{y_h} (t - \mu_h)^2 \cdot \frac{1}{w_h} \cdot f(t) dt$$

註4：DH最小變異數分層原則(Dalenius and Hodges)

令 y 之分配為連續，其pdf為 $f(y)$ ， $a \leq y \leq b$ 欲將其分配分成L層，取層內變異數極小，層間變異數極大下，若任一層 h 的相對次數為 w_h ，平均數 μ_h 、變異數 S_h^2 ，則：

欲找出 y_h ， $h = 1, 2, 3, \dots$ 使得 $(\sum_{i=1}^L W_h S_h^2)^2$ 極小的原則：證明在 $y \in U(y_{h-1}, y_h)$ ，利用等分

$\sqrt{f(y)}$ 累積次數 $F(y)$ 求得L層的最適邊界點(optimum stratification)。

註5：所謂交互作用(interaction or confounding)存在於兩變數間，表示當其一變數水準變動時，對於另一變數的影響，屬非固定常數的效應。首先比較單變量與多變量分析下，解釋變數X別的係數估計值。計算只含主要效果之初步多元邏輯斯迴歸模型，並比較個別解釋變數X，是否有明顯差異，以檢查可能存在的共線性。比較單變量與多變量分析下各解釋變數的X係數估計值，並無在正負號關係明顯差異，且分類變數下係數大小的關係仍趨一致，顯著性亦沒有改變($p\text{-value} < 0.05$)。以下步驟模型將不考慮此類交互項目。

註6：自動互動檢視法(Automatic Interaction Detector，簡稱AID)是一種逐步搜尋的過程，由Sonquist和Morgan在1960s初期所發展出來的一種逐次分析工具。在分析過程中各變數出現的順序並不代表變數的相對重要性，而未出現的變數也不一定不重要。CHAID是一卡方檢定將同值的樣本歸為一類，利用伴層次逐步搜尋而完成的分割過程，只適用在類別資料。

列聯表分析雖可了解變數與變數間的相關與否，但無法將變數間的交互關係作表示，所以當自變數間有交互關係或有一個以上的自變數存在時，單純的自變數交叉乘積是不易處理這種複雜的交互關係，為了處理這種自變數間的交互關係，1980年Perreault及Barksdale提出卡方自動互動檢視法，以Bonferroni方式調整個別的卡方 α ，作為分割樣本集群的依據。以卡方檢定使同性質樣本歸唯一類，逐次搜尋完成合併，分割的程序。主要是利用關聯性高的變數將樣本劃分為若干個具有同性質的組，以同樣程序進行分割，直到分群結果不顯著，以致無法計算有意義的機率估計為止。但被選入結構中的變數並不代表其重要性，而未被選入的變數也並不代表不重要。CHAID是屬於資料分析的程序，其目的在觀察變數間相互關聯的結構，並可檢測預測值的類別集合，卡方交互作用自動檢視法有三個運算程序，包括合併、分割、停止，在應用上主要著重於幫助了解資料結構的組合及互動特性。

參考文獻

丁正中

2004 〈消費新榮信用風險研究—信用評分概述〉《信用資訊》十一月號：1-32。

李桐豪、呂美惠

2000 〈金融機構房屋客戶授信評量模式分析—Logistic 回歸之應用〉《台灣金融財務季刊》1(1)：1-20。

江百信、張金鶚

1995 〈我國購屋貸款放款條件之研究〉《住宅學報》3：1-20。

莊瑞珠

2007 〈邏輯斯迴歸模型運用於女性信用卡評分制度之研究〉《輔仁管理評論》14(1)：127-154。

周建新、于鴻福、陳進財

2004 〈銀行業房貸授信風險評估因素之選擇〉《中華管理評論》7(2)：77-103。

Boyle, M., J.N. Crook, R. Hamilton & L.C. Thomas

1992 “Methods for Credit Scoring Applied to Slow Payers,” in *Proceedings of the IMA Conference on Credit Scoring and Credit Control*. 75-90. ed. Thomas, Crook & Edelman, Oxford: Clarendon Press.

Canner, G.B., S.A. Gabriel & J.M. Wooley

1991 “Race, Default Risk and Mortgage Lending: A Study of the FHA and Conventional Loan Markets,” *Southern Economic Journal*. 249-262.

Capon, N.

1982 “Credit Scoring Systems: A Critical Analysis,” *Journal of Marketing*. 46: 82-91.

Carter, C. & J. Catlett

1987 “Assessing Credit Card Applications Using Machine Learning,” *EEE Expert*. 2:71-79.

Davis, R.H., D.B. Edelman & A.J. Gammerman

1992 “Machine-learning Algorithms for Credit-card Applications,” *IMA Journal of Mathematics Applied in Business and Industry*. 1(4): 43-52.

Eisenbeis, R.A.

1978 “Problems in Applying Discriminant Analysis in Credit Scoring Models,” *Journal of Banking and Finance*. 2: 875-900.

Epley, D.R., K. Liano & R. Haney

1996 “Borrower Risk Signaling Using Loan-to-Value Ratios,” *The Journal of Real Estate Research*. 11(1): 71-86.

Gilbert, L.R., K. Menon & K.B. Schwartz

1990 “Predicting Bankruptcy for Firms in Financial Distress,” *Journal of Business Finance Accountings*. 17(1): 161-171.

Goodman, L.A. & W.H. Kruskal

1954 “Measures of Association for Cross-Classifications,” *Journal of the American Statistical*

- Association*. 49: 732-764.
- Grablowsky, B.J. & W.K. Kalley
 1981 "Probit and Discriminant Functions for Classifying Credit Applicants: A Comparison," *Journal of Economics and Business*. 33: 254-261.
- Henley, W.E. & D.J. Hand
 1996 "A k-NN Classifier for Assessing Consumer Credit Risk," *The Statistician*. 65: 77-95.
- Herzon, J.P. & J.S. Earley
 1970 *Home Mortgage Delinquency and Foreclosure*. New York: National Bureau of Economic Research.
- Hosmer, D.W. & S. Lemeshow
 1980 "A Goodness-of-Fit Test for the Multiple Logistic Regression Model," *Communication in Statistics*. 10: 1043-1069.
- Hosmer, D.W. & S. Lemeshow
 1982 "The Use of Goodness-of-Fits Statistics in the Development of Logistic Regression Models," *American Journal of Epidemiology*. 115: 92-106.
- John, A.S. & N.M. James
 1964 *The Detection Of Interaction Effects*. University of Michigan. Monograph.
- Mays, E.
 1998 *Credit Risk Modeling: Design and Application*. New York: AMACOM.
- Mays, E.
 2001 *Handbook of Credit Scoring*. New York: AMACOM.
- Makowski, P.
 1985 "Credit Scoring Branches out," *The Credit World*. 75: 30-37.
- Miller, D.C.
 1991 *Handbook of Research Design and Social Measurement*. 5th ed. Newbury Park, CA.: Sage Publications.
- Radding, A.
 1992 "Credit Scorings: New Frontier," *Bank Management*. 68: 57-62.
- Rakes, G.K.
 1973 "A Numerical Credit Evaluation Model for Residential Mortgages," *Quarterly Review of Economics and Business*. 73-84.
- Royston, P. & D.G. Altman
 1994 "Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modeling (with Discussion)," *Applied Statistics*. 43: 429-467.
- Srinivasn, V. & Y.H. Kim
 1987 "Credit Granting: a Comparative Analysis of Classification Procedures," *Journal of Finance*. 42: 665-683.
- Steenackers, A. & M.J. Goovaerts
 1989 "A Credit Scoring Model for Personal Loans," *Insurance Mathematics Economics*. 8:

301-325.

William, D.P., J.C. Hiram & J. Barksdale

1980 "A Model-Free Approach for Analysis of Complex Contingency Data in Survey Research," *Journal of Marketing Research*. 17(4): 503-515.

Yobas, M.B., J.N. Crook & P. Ross

2000 "Credit Scoring Using Neural and Evolutionary Techniques," *IMA Journal of Management Mathematics*. 11: 111-125.