

# 樣本選擇偏誤於企業財務危機預警模型之研究：以台灣上市公司為例

林郁翎、張大成、黃士賓\*

## 摘要

傳統財務危機預警模型建立，皆是以已通過審核申請者樣本建立模型，忽略未通過審核申請者樣本，然而以這些樣本所建構出來的模型，因為不能反映母體的變動程度與變數間的相互影響效果，亦未考慮樣本選擇偏誤問題，故會影響模型的配適度與預測能力。本文則是加入拒絕推論技術建立修正後 Heckman 兩階段樣本選擇模型，以台灣上市公司為例，與傳統財務危機預警模型進行比較；研究結果發現，財務危機模型建構中的審核模型與違約模型兩階段間存在顯著之相關，若不採取樣本選擇模型，將對模型預測結果產生很大的偏誤；而觀察模型的配適度與預測能力後亦可發現，修正後 Heckman 兩階段樣本選擇模型的配適度與預測能力確實優於傳統的財務危機預警模型。

關鍵詞：樣本選擇、拒絕推論、財務危機預警模型、Probit 模型  
JEL 分類代號：C10, G20, G32

---

\* 三位作者分別為聯絡作者：林郁翎，德明財經科技大學財務金融學系助理教授，台北市內湖區環山路一段 56 號，電話：02-26585801 轉 5182，E-mail: [yulinglin@mail.takming.edu.tw](mailto:yulinglin@mail.takming.edu.tw)。張大成，東吳大學國際經營與貿易學系教授，台北市中正區貴陽街一段 56 號，電話：02-23111531，E-mail: [dachen@scu.edu.tw](mailto:dachen@scu.edu.tw)。黃士賓，陽明海運專員，基隆市七堵區明德一路 271 號，電話：02-24559988，E-mail: [benz12354@hotmail.com](mailto:benz12354@hotmail.com)，作者由衷感謝編輯委員與兩位匿名評審之協助與寶貴建議，然而文章中若仍有疏誤，均屬作者之責任。  
投稿日期：民國 98 年 3 月 24 日；修訂日期：民國 98 年 5 月 13 日；  
接受日期：民國 98 年 10 月 12 日。

## 1. 緒論

企業財務危機預警模型的建立一直是學術界與實務界相當關心的議題之一。在實務上，為因應新巴塞爾資本協定之實施，銀行業者無不積極引進或開發風險管理系統，由於信用風險是銀行風險的主要來源，因此建立相關模型以準確地衡量貸款申請者的信用品質或是違約機率是不可或缺的，尤其是在新巴塞爾資本協定下，信用風險內部評等法之建立，不僅可節省法定資本額計提，亦可降低金融業客戶違約的風險，達到其追求最大利潤之目標；除此之外，評等公司對於受評企業的信用等級歸類，亦須對企業危機預測有相當程度之涉入，故而建立一套合適的財務危機預警模型，實為刻不容緩之議題。在學術界裡，自從 Altman (1968) 採用多變量區別分析法 (multivariate discriminant analysis, MDA)，Ohlson (1980) 採用 Logit 迴歸分析法 (logistic regression analysis)，Zmijewski (1984) 採用 Probit 迴歸分析法 (probit regression analysis) 進行企業危機預測後，相關研究一直持續進行中，這些研究主要精神在於利用企業之歷史財務變數為解釋變數，並以隔年 (或未來幾年) 企業是否發生財務危機為被解釋變數，建構多變量之區別函數，或以 Logit、Probit 迴歸模型，進行企業危機預測，這樣的分析架構文獻上稱之為會計基礎 (accounting based) 的企業危機預測模型，後續的研究包括：Belkai (1980)、Ederington (1985)、Back et al. (1996) 與張大成 (2003) 等文。另外，晚近幾年學術界的研究則是逐漸朝向利用 Black and Scholes (1973) 與 Merton (1974) 所提出利用選擇權評價模式，考慮股東之有限責任 (limited liability) 類似買權 (call option) 之下檔風險有限的觀念，並引用股價資料進行企業危機預警模式之建立，由於分析過程中採納了股價資訊，因此相對於會計基礎分析法，此一分析法文獻上稱之為市場基礎 (market based) 的企業危機預警模型，相關文獻包括：Charitou and Trigeorgis

(2000)、Vassalou and Xing (2004)、Farmen et al. (2004) 與 Hillegeist et al. (2004) 等人皆曾以市場基礎法進行企業危機預警模型建立，也都獲得不錯的實證結果。

然而，無論是學術界或是實務界，對於企業財務危機預警模型之建立，多僅使用部分樣本訊息進行模型建構，換言之，即是以已通過銀行審核之申貸企業樣本進行模式建構，而將被拒絕申請之企業樣本排除在模型建構外。但財務危機預警模型乃是要應用於所有申請者之評估上的，包括信用良好，以及違約風險較高的申請者；因而，傳統的財務危機預警模型，由於忽略未通過審核申請者資訊，僅使用核貸者資訊作為預警模型樣本並據以進行模型建構，如此所建構出來的模型，無論是在變數的選取、權重的設計、對未來用於審核所有貸款申請人是否給予貸款的決策上，均容易產生偏誤，此乃因申請者母體分配與通過審核之申請者樣本分配差異下的樣本選擇偏誤 (sample selection bias) 所導致。相關的研究亦證實了此項論點，如：Hand (1998) 指出，若是信用評分模型無法隨著母體的變動與變數效果的改變而更新，則原始的評分模型將會喪失其預測能力；而若僅使用被接受的申請者資料來更新模型，則樣本選擇問題會對新模型的有效性產生影響。Greene (1998) 亦指出，根據被接受的申請者樣本所建構之預警模型，除會降低模型的預測能力外，也會導致決策者做出錯誤的決定。Verstraeten and van den Poel (2004) 則就特定的樣本規模進行觀察，結果發現樣本選擇偏誤對於消費者信用評分模型之績效與金融機構之獲利性有顯著不利影響。因此，若能利用一些計量方法推論出那些被拒絕的貸款申請者信用狀況，則模型的建構將會有一個較完整的樣本訊息，除可避免前述的樣本選擇偏誤情形外，並可提升預警模型的預測能力。毫無疑問地，既往被廣泛應用於勞動經濟學研究的樣本選擇模型是解決此類問題的方法；而過去於信用評分領域中，針對樣本選擇偏誤問題所發展出來的解決方法即稱為拒絕推論技術，透過此類技術，在建構財務危機預警模型時，可推估被拒絕申請者訊息並納入樣本

訊息集中，利用被接受與被拒絕之申請者訊息建構財務危機預警模型，如此不僅可提升模型的預測正確性，亦可解決樣本選擇偏誤問題。

Chen and Astebro (2001) 曾將應用於信用評分模型之拒絕推論技術區分為三種類型。第一種為理想的技術，也就是樣本可以完全代表整個母體之情況；換言之，銀行在一開始須接受所有申請者的貸款申請，然後隨時間經過，觀察這些申請者的還款行為，以瞭解他們違約情形，由於此類技術之抽樣資料即為真實母體，因此就理論而言是完美的，但在執行上卻須花費鉅額成本。第二種類型為假設被接受申請者的樣本分配經過某些觀察與假設後可被擴展至被拒絕申請者分配；此類方法假設被拒絕申請者中信用良好的比例與被接受申請者中信用良好的比例相同；然而 Hsia (1978) 則提出原始的信用評分模型未必有足夠的能力以接受/拒絕決策來區分申請者是否有違約的可能性；若是原始的評分模型無法由壞的貸款申請人中區分出較好的貸款申請者時，則原始評分模型便無法用來當作衡量之標準；為了避免此不合實際的假設，Hand (1998) 提出當原始評分模型使用的特性集合若應用於新的信用評分模型時，透過拒絕推論技術之外推法 (extrapolation)，可將該假設合理的應用於新的評分模型上，但使用這類技術仍然無法得知可降低多少的偏誤及能增加多少的模型預測能力；Crook and Banasik (2004) 則利用外推法與擴增法 (augmentation) 兩種拒絕推論技術，觀察是否拒絕推論技術可增加信用評分模型的預測績效，實證結果顯示，使用這兩種拒絕推論技術對於申請者日後的違約預測績效相較於僅使用被接受申請者樣本以 Logit 迴歸所建立之模型還要來的差。第三種類型則假設被接受申請者之母體分配不同於被拒絕申請者之母體分配；而最為著名的模型即為 Heckman (1976, 1979) 所提出的兩階段樣本選擇模型，在信用評分模型的應用上，也就是將貸款審核決策與違約模型描述為樣本部分觀察下的兩階段模型；此類方法的假設相較前兩種類型較廣為學者們所接受，如 Poirier (1980)、Van de Ven

and van Praag (1981) 等人均曾對此方法進行相關討論；Meng and Schmidt (1985) 則是認為僅使用部份觀察到之樣本建立模型會導致模型估計不精確結果，因此在模型建立上應盡可能的蒐集額外訊息；Boyes et al. (1989) 設計兩個二元 Probit 模型：貸款同意決策與違約模型，利用兩模型間的關係修正因樣本選擇所導致的違約方程式估計偏誤；Copas and Li (1997) 則是擴展此類拒絕推論技術於非隨機樣本上；後續的實證研究如 Greene (1998)、Chen and Astebro (2001)、Banasik et al. (2003) 與 Jacobson and Roszbach (2003) 與 Kim and Sohn (2007) 等人亦延伸 Heckman (1976, 1979) 的兩階段樣本選擇模型於信用評分或財務危機預警之領域上。

其中的 Jacobson and Roszbach (2003) 則是利用二元 Probit 模型建立信用評分模型，以 1994 年 9 月至 1995 年 8 月間的瑞典某貸款機構的 13,338 個消費金融貸款申請者資料進行實證分析；實證結果發現，被貸款機構拒絕的申請者樣本分配不同於被接受申請者樣本分配，且貸款審核階段與貸款違約階段間有高度的關連性，因此，會影響審核階段申請人之變數也會對於申請者未來是否違約造成影響。Kim and Sohn (2007) 則是使用 1997 年至 2002 年間向韓國技術發展基金申請資金的企業資料進行實證分析，實證進行觀察後，得出考慮樣本偏誤修正的 Probit 模型相對於僅使用通過申請者樣本所建立之模型預測財務危機績效有提升之效果。

然而，觀之過去相關的財務危機預警模型建立之研究，不是未考慮申請者母體分配與通過審核申請者樣本分配差異下之樣本選擇偏誤問題，如：Altman (1968)、Ohlson (1980) 與 Zmijewski (1984) 等；便是聚焦於對消費金融領域之探討，如：Greene (1998)、Chen and Astebro (2001)、Banasik et al. (2003)、Jacobson and Roszbach (2003) 與 Crook and Banasik (2004) 等；雖然 Kim and Sohn (2007) 曾利用向韓國技術發展基金申請資金的企業進行拒絕推論技術之應用，然而此類樣本性質，在融資需求母體，及財務體質上與至金融機構借款之企業客戶母體並不完全相同。職是之故，本文認為在我

國已於 2006 年底開始與國際同步實施新巴塞爾資本協定後，有關金融業的信用風險控管與分析已是一項刻不容緩之議題，且當前國內各企業為提升其本身的競爭能力，無不積極的取得資金以進行相關投資，故對於本國金融機構之營運而言，企業金融業務實為相當重要的一環，若能建立預測準確率高的企業違約預警模型，將可適度的降低企業貸款違約的風險，提高銀行的獲利能力。

雖然，就金融機構而言，以個人或家庭為授信對象之消費金融放款業務，其授信對象之財務狀況極易受到失業、經濟景氣等因素影響，所面對之風險亦不可謂不大。然而，在當前相關文獻的研究中，已有部分國外學者針對此議題進行相關研究探討，故本文僅以法人機構為授信對象之企業金融放款業務為探討對象，觀察傳統應用於勞動經濟學的樣本選擇分析法，應用至財務領域之信用風險模式建構上，是否能夠提供金融業在新巴塞爾資本協定實施後，建立一套完善的信用風險衡量制度，以作為計算風險性資產與資本計題之依據。因此，本文將由企業金融角度，以台灣上市公司為例，考量過去財務危機預警模型未考慮通過審核申請者分配與母體分配差異下之樣本選擇偏誤缺失，以 Heckman (1976, 1979) 所提出之兩階段方法為本文研究方法基礎，對其進行修正以建立本文修正後 Heckman 財務危機預警模式，並進行相關的實證分析。故本文不僅延伸國內外對企業財務危機預警模型之既有研究 (Altman, 1968; Ohlson, 1980; Zmijewski, 1984 等)，並藉由計量方法的應用，修正過去文獻中所忽略之申請者母體分配與通過審核申請者樣本分配差異下的樣本選擇問題，以建立更完善的企業財務危機預警模型。本文首先根據過去多數財務危機預警模型非以全樣本資料建立之方式 (如：Altman, 1968; Ohlson, 1980; Zmijewski, 1984 等)，應用 Heckman 樣本選擇模型於非全樣本資料建立本文修正後 Heckman 財務危機預警模型；其次則採重複抽樣方法，探討本文模型在不同樣本組合下的穩定性；最後亦採全樣本資料，進行 Heckman 樣本選擇模型應用，並探討本文模型在不同樣本規模與樣本組合下的穩定性。本

文實證結果顯示本研究之修正後 Heckman 兩階段樣本選擇模型，其配適度與預測能力皆較傳統僅考慮通過審核申請者樣本訊息所建立之傳統財務危機預警模型為佳。故透過本文立基於嚴謹的理論架構與計量模型之上，可精確地對企業財務危機進行預警，並完備當前台灣財務危機預警之研究，俾達成強化金融業對台灣企業放款之有效信用風險管理目標。

最後，則就本文章節安排說明如下，除本節說明研究背景、動機、目的與相關文獻之探討外；第二節研究方法部份，主要介紹本文研究模型建構與傳統效力驗證方法；第三節研究設計部份，說明本文模擬樣本與研究變數之選取；第四節為實證結果分析；最後則為總結本文研究結果之結論。

## 2. 研究方法

由於銀行在貸款審核結束後，無法觀察被拒絕申請人之樣本，因此我們無法判斷這些樣本最後的違約情形。除此之外，在銀行審核階段中，通過貸款核准的申請人，其後違約與否往往會受到貸款審核階段中核貸與否之因素所影響，兩個階段間可能存在高度關聯性，因此我們必須針對傳統僅以核貸通過之申請者樣本所建立之財務危機預警模型進行修正，考慮樣本選擇偏誤對財務危機預警模型之影響。在一般的情況下，被接受之申請者母體與被拒絕申請者母體是不同的，因此，本文將使用考慮樣本選擇偏誤的二元 Probit 模型進行拒絕推論，利用修正後的 Heckman 兩階段樣本選擇模型來進行分析，<sup>1</sup> 故以下先介紹傳統 Probit 財務危機預警模型，接著介紹本文修正後的 Heckman 兩階段樣本選擇模型，最後則介紹模型

---

<sup>1</sup> Heckman (1976, 1979) 的兩階段樣本選擇模型，其第二階段模型的被解釋變數係探討女性勞工工作時數的連續變數，而本文第二階段的違約與否模型，其被解釋變數則係探討違約或非違約的間斷變數，故本文須針對此項差異進行相關修正。

效力驗證之方法，詳如下述。

## 2.1 傳統 Probit 模型

Probit 模型假設違約事件發生機率服從標準常態分配函數，每一樣本皆存在一組變數  $X$ ，且這些變數之線性組合可使每個樣本都得到一個分數  $Y_i^*$ ，即：

$$Y_i^* = \beta' X_i + \varepsilon_i, \quad (1)$$

其中， $\beta$  為待估參數； $X_i$  為解釋變數； $\varepsilon_i \sim N(0, \sigma^2)$ ，為隨機誤差項；而  $Y_i^*$  代表通過貸款審核之申請者信用品質，為無法觀察到的變數，一般稱為潛伏變數 (latent variable)，但可利用觀察得到的虛擬變數  $Y_i$  作為  $Y_i^*$  的替代變數，當  $Y_i^*$  大於 0，則表示企業未來可能會發生違約情形，如下所示：

$$Y_i = \begin{cases} 1, & \text{if } Y_i^* > 0 \\ 0, & \text{if } Y_i^* \leq 0 \end{cases}, \quad (2)$$

根據上式，可以定義當  $Y_i = 1$  時的機率如下：

$$P_i = P(Y_i = 1 | X_i) = P(Y_i^* > 0 | X_i) = P(\varepsilon_i > -\beta' X_i | X_i) = \Phi(\beta' X_i), \quad (3)$$

其中： $\Phi$  為標準常態累積機率密度函數。利用最大概似估計法即可估計參數  $\beta$ ，概似函數 (likelihood function) 則如下所示：

$$L = \prod_{Y_i=1} P_i \cdot \prod_{Y_i=0} (1 - P_i) = \prod_{Y_i=1} \Phi(\beta' X_i) \times \prod_{Y_i=0} \Phi(-\beta' X_i). \quad (4)$$

## 2.2 修正後的 Heckman 兩階段樣本選擇模型

假設被接受申請者母體分配不同於被拒絕申請者母體之分配，且  $Y_{1i}^*$ 、 $Y_{2i}^*$  為無法觀察到的隨機變數：



$$Y_{1i}^* = \beta_1' X_{1i} + \varepsilon_{1i} , \quad (5)$$

$$Y_{2i}^* = \beta_2' X_{2i} + \varepsilon_{2i} , \quad (6)$$

其中， $\varepsilon_{1i} \sim N(0, \sigma_1^2)$ 、 $\varepsilon_{2i} \sim N(0, \sigma_2^2)$ 、 $\text{corr}(\varepsilon_{1i}, \varepsilon_{2i}) = \rho$ ， $X_{1i}$  與  $X_{2i}$  分別為兩階段模式之解釋變數，(5) 式為貸款審核模型，(6) 式為貸款違約模型。而  $Y_{1i}^*$  為貸款申請者的信用品質， $Y_{2i}^*$  為通過貸款審核之申請者信用品質，兩者皆為無法觀察到的變數，一般稱之為潛伏變數，但可利用觀察得到的虛擬變數  $Y_{1i}$  與  $Y_{2i}$  作為  $Y_{1i}^*$  與  $Y_{2i}^*$  的替代變數，例如當  $Y_{1i}^*$  大於某一門檻值  $k$  時，表示企業申請之貸款會通過銀行審核；而當  $Y_{2i}^*$  大於 0，則表示通過貸款審核之企業未來將會有違約的可能，如下所示：

$$Y_{1i} = \begin{cases} 1, & \text{if } Y_{1i}^* > k \\ 0, & \text{if } Y_{1i}^* \leq k \end{cases} , \quad (7)$$

$$Y_{2i} = \begin{cases} 1, & \text{if } Y_{2i}^* > 0 \\ 0, & \text{if } Y_{2i}^* \leq 0 \end{cases} . \quad (8)$$

然而，實際上只有當  $Y_{1i} = 1$  時，才可觀察到  $Y_{2i}$  之情形，換言之，只有當企業通過貸款申請後，方能知道他們日後的違約情形。因此，可將這種情形表示如下：

$$Y_{1i} \text{ 在所有情況下，皆可被觀察到；} \quad (9)$$

$$Y_{2i} \text{ 只有在 } Y_{1i} = 1 \text{ 的情況下，可被觀察到。} \quad (10)$$

根據 (9)、(10) 兩式，可將所有申請貸款的企業母體以及可得之訊息以表 1 表示。

表 1 可觀察到之訊息

	非違約	違約	總計
被接受	觀察到	觀察到	觀察到
被拒絕	無法觀察到	無法觀察到	觀察到

本文目的係期望能觀察  $P(Y_{2i} = 1 | X_{2i})$  之情形，但由於傳統財務危機模型只能觀察到在  $Y_{1i} = 1$  之下的申請者特性，也就是，傳統模型只能在  $Y_{1i} = 1$ （貸款申請者核貸通過）之下，觀察是否  $Y_{2i} = 1$  或 0（違約或未違約）。在假設  $\varepsilon_{2i}$  為標準常態分配下，傳統財務危機預警模型之 Probit 方法，因僅使用被接受申請者資訊估算 (6) 式之參數，故其模式將如下所示：

$$P(Y_{2i} = 1 | X_{2i}) = P(Y_{2i}^* > 0 | X_{2i}) = P(\varepsilon_{2i} > -\beta_2' X_{2i} | X_{2i}) = \Phi(\beta_2' X_{2i}) \quad (11)$$

然而，若可將未通過貸款審核之申請者訊息納入模式估計中，此時模型將為：

$$\begin{aligned} P(Y_{2i} = 1 | X_{2i}, Y_{1i} = 1) &= P(Y_{2i} > 0 | X_{2i}, Y_{1i} = 1) \\ &= P(\varepsilon_{2i} > -\beta_2' X_{2i} | X_{2i}, Y_{1i} = 1) \end{aligned} \quad (12)$$

在  $\varepsilon_{1i}$  與  $\varepsilon_{2i}$  為二元標準常態分配之假設下，可能的聯合機率結果將為：

$$Y_{1i} = 1, Y_{2i} = 1: P(Y_{1i} = 1, Y_{2i} = 1) = \Phi_2(\beta_1' X_{1i}, \beta_2' X_{2i}, \rho) \quad (13)$$

$$Y_{1i} = 1, Y_{2i} = 0: P(Y_{1i} = 1, Y_{2i} = 0) = \Phi_2(\beta_1' X_{1i}, -\beta_2' X_{2i}, -\rho) \quad (14)$$

$$Y_{1i} = 0: P(Y_{1i} = 0) = \Phi(-\beta_1' X_{1i}) \quad (15)$$

其中， $\Phi_2$  為二元常態累積機率分配函數， $\Phi$  則為單變量常態累積機率分配函數， $\rho$  為貸款審核模式與違約模式間的相關係數。因

此，概似函數為：

$$\begin{aligned}
 L &= \prod_{\text{good loans}} P(\text{good loan}) \times \prod_{\text{bad loans}} P(\text{bad loan}) \times \prod_{\text{no loans}} P(\text{no loan}) \\
 &= \prod_{i=1}^{n_1} \Phi_2(\beta'_1 X_{1i}, -\beta'_2 X_{2i}; -\rho) \times \prod_{i=n_1+1}^n \Phi_2(\beta'_1 X_{1i}, \beta'_2 X_{2i}; \rho) \\
 &\quad \times \prod_{i=n+1}^m \Phi(-\beta'_1 X_{1i}), \tag{16}
 \end{aligned}$$

其中， $1 \sim n_1$  為通過審核之申請者未違約數目， $n_1 \sim n$  為通過審核之申請者違約數目， $n \sim m$  則為一開始未通過審核之申請者數目；故在此修正模式下的違約機率為：

$$P(Y_{2i} = 1 | X_{2i}, Y_{1i} = 1) = \frac{\Phi_2(\beta'_1 X_{1i}, \beta'_2 X_{2i}; \rho)}{\Phi(\beta'_1 X_{1i})} \tag{17}$$

在既有的傳統模型中，係直接估計 (3) 式或 (11) 式，然而，經由 (17) 式可知，惟有當  $\rho = 0$  時，在樣本選擇之過程中才不會導致  $\beta_2$  的偏誤估計，而 (17) 式將會等於 (3) 式或 (11) 式之結果。但若  $\rho \neq 0$  時，則不宜直接以 (3) 式或 (11) 式進行違約機率之估算，因為此時貸款審核模式與違約模式誤差項間具有相關性，僅以 (3) 式或 (11) 式之傳統財務危機預警模式進行違約機率估算，將會造成很大的參數估計偏誤情形。

參酌 Greene (2003) 所說明之 Heckman (1976, 1979) 兩階段估計過程，本文的兩步驟估計過程為：第一階段，先利用最大概似法估計二元 Probit 模型，以獲得  $\beta_1$  之估計值，並且對選擇的樣本中每一個估計值，計算其對應之選擇性矯正因子 (inverse Mills ratio 或稱 Mills lambda)，以  $\lambda_i$  表示之，計算方式為  $\lambda_i = \phi(\hat{\beta}'_1 X_{1i}) / \Phi(\hat{\beta}'_1 X_{1i})$ ，<sup>2</sup> 以矯正樣本選擇性偏誤，並可觀察兩階段間是否顯著

<sup>2</sup>  $\phi$  為常態分配之機率密度函數，而  $\Phi$  則為累積機率密度函數。

相互影響。於第二階段，則將  $\lambda_i$  視為一項額外之解釋變數，與原來的解釋變數  $X_{2i}$  一起建構第二階段的 Probit 模型，以估算  $\beta_{2i}$  及  $\lambda_i$  之係數。若是  $\lambda_i$  係數顯著異於 0，代表第二階段會受到第一階段影響，故  $\lambda_i$  可解釋方程式中的樣本選擇偏誤問題，亦即金融機構在建構財務危機預警模型時必需考量樣本選擇偏誤問題，方可較準確的預測企業財務危機之發生；因此，經由前述的兩階段過程，我們可據以估算出考慮樣本選擇偏誤下的企業違約機率：(17) 式，並進而評估是否考慮樣本選擇偏誤之企業財務危機預警模型較傳統未考慮被拒絕申請者訊息之企業財務危機預警模型有較佳的預測績效。

## 2.3 模型效力檢定方法

### 2.3.1 KS (Kolmogorov-Smirnov) 檢定

KS 檢定可用於檢定財務危機預警模型對於危機公司與正常公司之區別能力好壞。其理論基礎為，若檢定之兩樣本來自相同母體。則兩樣本的累積相對次數分配應該相當接近；反之，當兩樣本的累積相對次數分配差異越大，代表在此評分階段下的兩組樣本可能來自不同母體。由於每個評分階段下各有一個累積相對次數分配差異，故選擇一個最大的累積相對次數分配差異作為檢定統計量，當檢定統計值很大時，則拒絕兩組樣本來自相同母體之虛無假設。但由於 KS 值愈大表示模型愈具有區別力，此種判斷標準並不太客觀，故 Mays (2001) 提供了一套 KS 值標準，如下表所示。

表 2 KS 值說明

KS 檢定值 (%)	說 明
KS < 20	模型完全不具區別能力，不值得採用
20 ≤ KS < 40	模型具有一般的區別能力
40 ≤ KS < 50	模型具有好的區別力
50 ≤ KS < 60	模型具有很好的區別力
60 ≤ KS < 75	模型具有優良的區別力
75 ≤ KS	模型的區別能力好到無法置信，模型可能有問題

### 2.3.2 ROC (Receiver Operating Characteristic) 曲線

ROC 曲線可描繪出不同的截斷點下型 I 誤差和型 II 誤差的關係，此關係可以圖 1 之二維圖形呈現；其中，X 軸為型 I 誤差，<sup>3</sup> Y 軸為 (1-型 II 誤差)，<sup>4</sup> 在各個不同型一誤差下，所對應的 (1-型 II 誤差) 應該要越大越好，此表示越靠近左上角曲線之預警模型擁有愈高的解釋能力，愈能準確的預測財務危機之發生。而在 ROC 曲線下的區域被稱作 AUC (area under the curve)；AUC 比率介於 0 與 1 之間，當 AUC 愈接近於 1，表示模型預測正確性愈高。而 AUC 可解釋為在所有可能的截斷點下，驗證模型對區別正常公司和出現財務危機公司的平均能力，當 AUC 區域為 0.5 時，表示模型區別正常公司和出現財務危機公司採取隨機過程，因而此時模型毫無區別能力；當 AUC 區域為 1 時，表示模型為最佳的，此時模型具有完全的區別能力。因此，對於所有合理的預警模型，AUC 實際上介於 0.5 至 1 之間。而 Hosmer and Lemeshow (2000) 則針對各種情況下的 AUC 值進一步說明其與模型區別能力之關係，如下表所述：

表 3 AUC 值說明

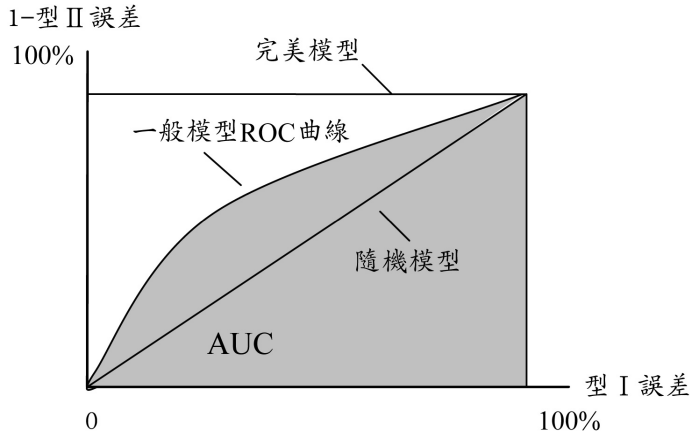
AUC 值(%)	模型區別力
< 50	模型不具區別能力
50~70	模型區別能力普通
70~80	模型區別能力佳
80~90	模型區別能力強
Over 90	模型區別能力非常強

### 2.3.3 CAP (Cumulative Accuracy Profile) 曲線

CAP 曲線係以圖形來呈現財務危機預警模型的正確區別力，透過模型計算出所有公司的信用分數並以此分數求算出模型的配適

<sup>3</sup> 型 I 誤差 (Type I Error)，公司實際上未發生財務危機，但模型卻預測其會發生財務危機。

<sup>4</sup> 型 II 誤差 (Type II Error)，公司實際上有發生財務危機，但模型卻預測不會發生財務危機。



資料來源：Engelmann et al. (2003).

圖 1 ROC 與 AUC

能力。第一步為將公司依模型所估計的違約風險由大到小排序；接著將所有公司違約風險最高的前 X% 的預測錯誤率 (alarm rate) 置於水平軸；而將所有違約公司的累積百分比:  $d(X)$  的預測正確率 (hit rate)，即：模型正確的將違約公司歸類為違約公司之數目佔所有違約公司的百分比置於縱軸，之後連結上述所有點所繪出的圖形即為 CAP 曲線，如下圖所示。當 CAP 曲線起始的斜率愈高，表示此模型配適力愈佳。而一個完美模型為可以觀察到所有正常公司具較低之違約風險，因此 CAP 曲線在初期將以接近垂直的型態急速上升，而在 Y 軸累積機率達到 1 時再以水平的方式移動，直到 X 軸累積機率達到 100% 為止；另一個極端的隨機模型為評分結果近乎隨機產生，代表著此模型不具任何配適能力。而在現實生活中，模型結果所繪出的 CAP 曲線將介於此兩個極端曲線間。另外，也可以運用單一指標來衡量 CAP 曲線對模型效力之觀察情形，稱為準確率 (accuracy ratio, AR)。圖 2 中，介於完美模型與隨機模型的面積以  $a_p$  表示；介於實際模型與隨機模型的面積則以  $a_r$  表示；AR 則被定義為兩者之比率，即： $AR = a_r / a_p$ ； $0 \leq AR \leq 1$ ，AR 愈大則表示模型的配適力愈佳。

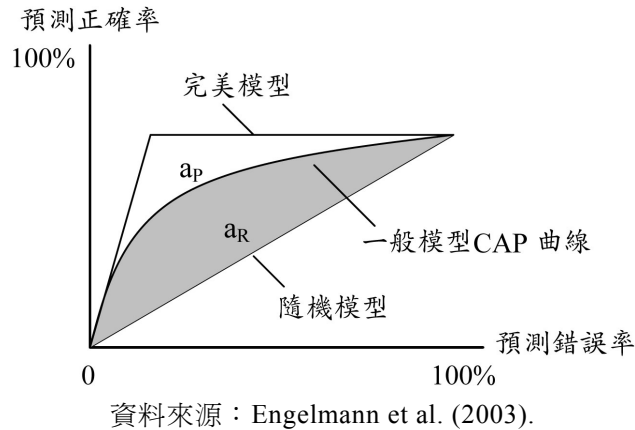


圖 2 CAP 曲線圖

### 3. 研究設計

#### 3.1 模擬樣本建立

由於本文欲探討傳統僅使用通過審核申請者樣本訊息所建立之財務危機預警模型在樣本內與樣本外之模型配適度與預測準確度，與考慮拒絕推論技術後所建立之財務危機預警模型在樣本內與樣本外之模型配適度與預測準確度差異。故本文以台灣上市公司資料為例，利用分層隨機抽樣方式選取 172 家上市公司進行分析，<sup>5</sup> 接著採 split-sample 抽樣方法，將 172 家總樣本分為兩群，其中一群為用以建構財務危機模型的訓練樣本（樣本內），另一群則為用來作為檢驗財務危機模型的測試樣本（樣本外），而訓練樣本有 132 家公司，測試樣本則有 40 家公司。

然鑑於官方對於危機企業定義過於嚴謹，以致造成危機企業樣本數過少，因此本文放寬對於財務危機企業之定義，以台灣經濟新報社 (Taiwan Economic Journal, TEJ) 所提供之財務危機與準財務

<sup>5</sup> 本文後續亦將以上市公司全樣本進行穩健性分析，觀察以全樣本或非全樣本資料建構傳統模型及修正後 Heckman 樣本選擇模型之效力關係是否具一致性。

危機定義作為本文財務危機發生標準，即只要具備下述其中一項則定義其為財務危機公司：跳票擠兌、破產倒閉、重整、杼困求援、接管、淨值為負、全額下市、財務吃緊停工、掏空挪用、暫停交易、董事長跳票、銀行緊縮、嚴重虧損、景氣不佳停工與價值減損等。透過前述財務危機定義，本文於 2000 年至 2005 年之研究期間中，所選取的 172 家模擬總樣本中，共包含了 43 家危機公司，與 129 家正常公司。

表 4 模擬樣本彙總表

修正後 Heckman 樣 本選擇模型	第一階段 (審核階段)			第二階段 (違約階段)			
	通過審核 公司	未通過審 核公司	總計	危機公司	未發生 危機公司	總計	
樣本內公司	100	32	132	樣本內公司	10	90	100
				樣本外公司	10	30	40
傳統財務危 機預警模型				樣本內公司	10	90	100
				樣本外公司	10	30	40

資料來源：台灣經濟新報社資料庫與本研究統整。

接著，本文模擬情境以進行相關分析，假設 2000 年至 2005 年間有 132 家上市公司跟某銀行申請貸款，該銀行依 132 家上市公司之台灣企業信用風險指標 (Taiwan Corporate Credit Risk Index, TCRI) 之評等等級高低作為是否核准貸款之依據，<sup>6</sup> 若公司於申貸前一年的 TCRI 信用等級小於七，則該銀行將會通過該公司的申貸案，反之，該銀行將會拒絕該公司的申貸案。2006 年時，該銀行對 132 家來申請貸款之上市公司進行調查，結果發現 132 家來跟其申請貸款之樣本內公司中的 100 家通過審核公司，在 2000 年至 2005 年中，有 10 家公司發生財務危機，有 90 家營運穩定成長；而在當時被拒絕的 32 家公司中，則有 23 家發生財務危機，有 9 家營運穩定成長；而 40 家樣本外公司中則有 10 家發生財務危機，有

<sup>6</sup> TCRI 指標係 TEJ 所發展出來的企業信用風險指標。



30 家為營運穩定成長的正常公司，故本文模擬樣本中共包含了 43 家危機公司，與 129 家正常公司，如表 4 所示。

## 3.2 研究變數選取

### 3.2.1 應變數

本文所建立的兩個模型，分別為加入拒絕推論技術後所建立之財務危機預警模型，與只以通過審核之申請者樣本所建立之傳統財務危機預警模型。在加入拒絕推論技術後所建立之財務危機預警模型中，第一階段的審核模型，其應變數為一虛擬變數，描述企業是否通過銀行之審核狀況，若是企業通過銀行審核貸款通過者以 1 表示，反之則為 0；第二階段的違約模型，應變數亦為一虛擬變數，代表公司發生財務危機的情況，若為 1 則表示企業發生財務危機，反之則為 0。而在僅以通過審核申請者樣本所建立之財務危機預警模型中，應變數亦為一虛擬變數，代表公司發生財務危機的情況，若為財務危機企業則為 1，反之則為 0。

### 3.2.2 自變數

本文對於自變數的選取，則依其屬性不同分成財務變數與非財務變數兩大類別，並以前一年資料進行財務危機預警模型建立。其中，非財務類變數之選取係根據過去相關文獻曾使用於建構財務危機預警模型之變數，<sup>7</sup> 以及金融機構在審核企業貸款時可能考慮之因素，<sup>8</sup> 包括：公司類別、<sup>9</sup> 公司年齡、公司員工人數等。而在財務變數選取方面，本文係採用過去相關文獻廣泛使用的財務變數和台灣經濟新報社所提供的財務比率共 20 個財務變數，將其依所屬特性不同歸類為五個構面，分別為獲利性：衡量公司創造利潤之能

<sup>7</sup> 如紀麗秋 (2002)、張大成 (2003)、曹曾樹 (2008) 等。

<sup>8</sup> 參考陳錦村 (2007) 第 14 頁至第 19 頁之企業適用的信用評等表。

<sup>9</sup> 本文將所有產業概分為兩大類別，一為電子產業，另一則為非電子產業。

力；流動性：衡量公司的短期債務償還能力及營運資金是否充裕、不足或是過度之現象；成長性：衡量公司在各個構面上的成長速度；效率性：衡量公司營運在諸如存貨的控制、固定資產使用效率、應收帳款品質上是否有效率；安全性：衡量公司財務槓桿的使用程度、公司舉債與還本付息的能力，如表 5 所示。

表 5 財務變數彙總表

變數構面	財務變數名稱	變數公式
獲利性	業外收支率	$\frac{\text{營業外收支淨額}}{\text{營業收入淨額}}$
	每股營業利益	$\frac{\text{營業利益}}{\text{總股數}}$
流動性	速動比率	$\frac{(\text{流動資產} - \text{存貨} - \text{預付款項} - \text{其他流動資產})}{\text{流動負債}}$
	淨營運資金佔總資產比	$\frac{\text{淨營運資金}}{\text{總資產}}$
成長性	營收成長率	$\frac{(\text{營業收入淨額} - \text{前期營業收入淨額})}{\text{前期營業收入淨額}} \times 100$
	營業毛利成長率	$\frac{\text{毛利增減額}}{\text{去年同期毛利}} \times 100$
	營業利益成長率	$\frac{(\text{本期營業利益} - \text{前期營業利益})}{\text{前期營業利益}} \times 100$
	稅後淨利成長率	$\frac{(\text{本期稅後淨利} - \text{前期稅後淨利})}{\text{前期稅後淨利}} \times 100$
效率性	淨值成長率	$(\frac{\text{淨值}}{\text{前期淨值}} - 1) \times 100$
	總資產週轉率	$\frac{\text{營業收入淨額}}{\text{平均資產總額}}$
	應收帳款週轉率	$\frac{\text{營業收入淨額}}{\text{平均}(\text{應收帳款} + \text{應收票據})}$
	存貨週轉率	$\frac{\text{營業成本}}{\text{平均存貨}}$
	每人營收	$\frac{\text{營業收入淨額}}{\text{員工人數}}$
安全性	每人營業利益	$\frac{\text{營業利益}}{\text{員工人數}}$
	負債比率	$\frac{\text{總負債}}{\text{總資產}}$
	長期資金適合率	$\frac{(\text{淨值} + \text{長期負債})}{\text{固定資產}}$
	現金流量比率	$\frac{\text{來自營業現金流量}}{\text{流動負債}}$
	利息保障倍數	$\frac{\text{所得稅及利息費用前純益}}{\text{本期利息支出}}$
	利息支出率	$\frac{\text{利息支出}}{[\text{經常利益} + \text{利息支出} \times (1 - 25\%)]}$
	淨值佔總資產比	$\frac{\text{淨值}}{\text{總資產}}$

## 4. 實證結果分析

### 4.1 基本統計量分析

本文樣本範圍包括在 2000 年到 2005 年中，發生財務危機的 43 家公司，與 129 家的正常公司。表 6 分別比較危機公司與正常公司的各項基本敘述統計分佈，由表中可知，在公司類別方面，於

2000 年到 2005 年間非電子產業發生財務危機的公司家數以百分比來看佔了 172 家總樣本的 15% 多於電子業的 10%；在公司平均年齡方面，危機公司平均年齡與正常公司則相差不大；最後，公司平均員工人數則是危機公司的平均員工人數大於正常公司。

表 6 模擬樣本分佈情形

變數		危機公司	正常公司	合計
公司類別	非電子業	15%	44%	59%
	電子業	10%	31%	41%
公司平均年齡		51%	49%	100%
公司平均員工人數		57%	43%	100%

資料來源：同表 4。

表 7 與表 8 則列出解釋變數之敘述統計量，合計共 23 個解釋變數的極小值、百分位數、平均數、極大值。為了解財務資料分佈特性，本文將上述敘述統計量區分成危機公司與正常公司兩部分來進行比較。首先，在獲利性指標變數方面，危機公司在業外收支率與每股營業利益兩指標上的表現均不及正常公司，表示危機公司的獲利能力遠不及正常公司，較易發生財務危機。在流動性指標變數方面，危機公司在速動比率與淨營運資金佔總資產比兩指標上的表現均低於正常公司，表示危機公司的短期償債能力不佳，容易造成黑字倒閉的情形發生。在成長性指標變數方面，危機公司的平均營業收入成長率雖然高於正常公司，但是在其他獲利能力成長率指標的表現多不及正常公司，表示雖然公司的營收有成長，但是若在其他成長率上出現衰退的情況，也會導致財務危機的發生。在效率性指標變數方面，由衡量經營效率的指標：存貨週轉率、應收帳款週轉率和總資產週轉率這三個變數的表現可看出，危機公司的經營效率均不及正常公司。最後在安全性指標變數方面，危機公司在各項長期償債能力指標上的表現均不及正常公司。

表 7 危機公司解釋變數之敘述統計量

變數構面	變數	最小值	5%	25%	中位數	平均數	75%	95%	最大值
非財務 變數	公司類別	1	1	1	1	1.4186	2	2	2
	公司年齡	9	10	20	28	29.3721	37.5000	51.3500	52
	員工人數	55	138.5500	280.5000	499	810.8372	837	3418.7500	5472
獲利性	業外收支率	-3.0117	-1.0918	-0.2738	-0.1372	-0.2703	-0.0665	0.0006	0.0465
	每股營業利益	-3.4461	-2.6038	-1.6819	-0.4229	-0.4518	0.2903	2.0164	3.5801
流動性	速動比率	0.1110	0.1507	0.2742	0.4639	0.7464	0.7645	2.4488	3.4017
	淨營運資金佔總 資產比	-0.2376	-0.2241	-0.0885	-0.0074	0.0515	0.1491	0.3992	0.5322
成長性	營收成長率	-0.8903	-0.8704	-0.2202	-0.0192	0.2446	0.2350	1.3377	10.2215
	營業毛利成長率	-44.8178	-11.6770	-0.8769	-0.3145	-1.5887	0.2977	1.8121	12.8188
	營業利益成長率	-491.4980	-29.0728	-2.0072	-0.7715	-13.7596	0.5319	3.3345	11.0899
	稅後淨利成長率	-133.0290	-23.5180	-3.5462	-0.6698	1.4909	0.3194	30.2218	238.7167
	淨值成長率	-0.9459	-0.7939	-0.1902	0.5475	3.0899	3.0979	13.8497	49.1799
效率性	總資產週轉率	0.1309	0.1700	0.3192	0.4475	0.5923	0.6953	1.2881	3.4715
	應收帳款週轉率	1.0936	1.3882	2.9327	5.0600	5.5127	7.7886	12.0025	13.2378
	存貨週轉率	0.5089	0.9412	2.8606	4.6203	5.8267	7.5754	15.2738	20.0529
	每人營收	1767.986	2488.329	4119.389	7583.019	9818.803	14825.590	22508.480	33794.660
	每人營業利益	-13199.1000	-3709.5400	-871.0770	-126.8570	-754.7750	134.7958	911.5700	1249.4080
安全性	負債比率	0.1592	0.3112	0.4622	0.5723	0.5366	0.6214	0.7442	0.8534
	長期資金適合率	59.7600	94.4390	144.3025	177.5100	476.5758	313.9875	948.4370	8366.0200
	現金流量比率	-0.3964	-0.2772	0.0333	0.0925	0.1114	0.1701	0.4751	1.2493
	利息保障倍數	-111.1080	-39.8915	-7.1533	-1.1261	-7.5012	0.3922	5.9080	14.8432
	利息支出率	0.0090	0.0252	0.1000	0.6179	6.8309	1.3055	13.6160	224.9053
	淨值佔總資產比	0.0220	0.0525	0.2790	0.4072	0.4052	0.5361	0.7618	0.9688

資料來源：同表 4。

表 8 正常公司解釋變數之敘述統計量

變數構面	變數	最小值	5%	25%	中位數	平均數	75%	95%	最大值
非財務 變數	公司類別	1	1	1	1	1.4186	2	2	2
	公司年齡	9	13	21	31	30.9225	38	48.1500	61
	員工人數	15	123.8000	278.7500	478	598.0698	654	1551.3000	4102
獲利性	業外收支率	-0.9761	-0.1675	-0.0426	-0.0096	-0.0286	0.0144	0.0888	0.1993
	每股營業利益	-1.8648	-0.9009	0.1530	0.8630	1.1762	1.8687	4.3975	9.3401
流動性	速動比率	0.1650	0.3279	0.5575	1.0248	1.5681	1.5535	5.3678	24.8925
	淨營運資金佔總 資產比	0.0025	0.0184	0.0799	0.1662	0.2060	0.2614	0.5728	1.1817
成長性	營收成長率	-0.8152	-0.2684	-0.0645	0.0489	0.0725	0.1960	0.4888	0.6902
	營業毛利成長率	-0.4243	0.0118	0.0794	0.2036	0.4092	0.4915	1.3436	5.3345
	營業利益成長率	-28.0547	-4.4097	-0.5568	-0.0653	-0.2025	0.4554	2.5874	16.8383
	稅後淨利成長率	-37.4487	-4.6799	-1.0054	-0.0486	0.2936	0.5770	3.5064	73.1816
	淨值成長率	-0.8471	-0.7401	-0.2011	0.0712	0.3959	0.5971	3.2713	7.0310
效率性	總資產週轉率	0.1440	0.2315	0.4555	0.6856	0.8348	1.0270	2.0014	2.7009
	應收帳款週轉率	1.0595	2.5374	3.9788	5.3672	6.6278	7.7685	15.1841	29.5093
	存貨週轉率	1.1503	1.5401	4.0585	7.1741	9.0935	11.6316	20.6273	56.3469
	每人營收	1265.963	1696.071	3453.820	5990.942	11320.660	11401.120	34153.970	143693.700
	每人營業利益	-24631.8000	-1284.5900	12.2494	246.4065	92.4349	531.1706	1787.3890	5645.4170
安全性	負債比率	0.0437	0.1495	0.2623	0.3831	0.3870	0.5102	0.6264	0.8076
	長期資金適合率	52.3000	112.3625	164.8025	256.2900	385.4021	363.2750	1394.1440	4125.6700
	現金流量比率	-0.6260	-0.2326	0.0800	0.2446	0.3234	0.4941	1.1265	2.0820
	利息保障倍數	-27.1390	-13.9000	0.1416	3.5219	41.0112	10.7454	86.6165	1865.7090
	利息支出率	-10.3815	-0.6697	-0.0271	0.0889	0.7202	0.3059	1.6637	58.2261
	淨值佔總資產比	0.0005	0.0077	0.0299	0.0923	0.1939	0.2435	0.7441	0.9370

資料來源：同表 4。

## 4.2 模型建構與分析

由於本文傳統模型與修正後 Heckman 樣本選擇模型所涵蓋之樣本訊息並不相同，故本節首先敘述加入拒絕推論技術後所建立之修正後 Heckman 樣本選擇模型之建構過程，並觀察僅考慮已通過審核之申請者樣本訊息所建立之模型，是否有樣本選擇偏誤問題存

在；接著描述以通過審核申請者樣本訊息所建立之傳統財務危機預警模型；其次則進行兩模型預測效力比較；最後則是針對本文模型效力進行穩健性分析 (robust analysis)。

#### 4.2.1 模型之建構

本文於建構模型之前，首先針對解釋變數進行篩選，此乃因解釋變數若特性重複且彼此間具高度相關時，將導致模型有過度配適情況產生，為避免此情形發生，一般以變異數膨脹因子 (variance inflation factor, VIF) 大於 10，表示該變數與其他變數間有嚴重共線性存在，為了不使模型產生過度配適，必須將變數間有高度相關者擇一逐次剔除。故本文將以全體變數 VIF 值小於 10 為標準進行檢定，而檢定結果則確認修正後 Heckman 樣本選擇模型審核階段與違約階段，以及傳統財務危機預警模型中各解釋變數彼此間並無嚴重共線性情形存在。接著本文則係以逐步迴歸分析法萃取出影響審核階段與違約階段較為重要的變數，觀察應變數 (核貸通過與否、財務危機發生與否) 與相關預測變數間的關係，來選取建構模型所需之變數，<sup>10</sup> 並利用選出的變數建構修正後 Heckman 樣本選擇模型與傳統財務危機預警模型，模型建構之意涵則詳如下述。

##### a. 修正後 Heckman 樣本選擇模型

表 9 中兩個階段的解釋變數是經由兩次逐步迴歸後所選出。<sup>11</sup> 其中第一階段的解釋變數分別為：公司類別、業外收支率、負債比率；第二階段的解釋變分別為：淨營運資金佔總資產比、存貨週轉率、 $\lambda_i$ ；由審核與違約模型所含之變數可知，其中的  $\lambda_i$  可反映在審核階段中，包括公司類別、業外收支率與負債比率資訊。而依據準判定係

<sup>10</sup> 本文所使用的選擇變數程序在不同樣本組合與樣本規模下，會選取出不同的變數組合，如本文 4.2.3 節之穩健性分析，在重複抽樣進行 1,000 次下，每次的樣本並不相同，故萃取出影響核貸通過與否與財務危機發生與否較為重要的變數亦不相同。

<sup>11</sup> 本文以 10% 的顯著水準為逐步迴歸篩選標準。

數 (pseudo coefficient of determination, pseudo  $R^2$ ) 可知，違約模型可解釋約 37.32% 的變異量。在第一階段中，由表 9 的迴歸變數估計係數之正負符號可觀察到是否獲得貸款的程度，當業外收支率該變數數值愈高，表示該公司愈容易通過審核；當公司類別為電子業時，表示該公司愈容易通過審核；當負債比率該變數數值愈高時，表示該公司的舉債程度偏高，愈不易通過審核。在第二階段中，當淨營運資金佔總資產比與存貨週轉率兩變數數值愈高時，表示該公司愈不容易發生財務危機。此外，我們可由  $\lambda_i$  來判斷第一階段審核模型與第二階段違約模型間的相關性，由表 9 中  $\lambda_i$  的檢定統計量達 5% 的顯著水準，可知兩階段模型間確實存在著顯著相關，亦即模型確實存在著樣本選擇偏誤情形，這意味著若不採取樣本選擇模型，而僅採傳統模型來估計企業財務危機情形，將會產生很大的偏誤，此實證結果也印證 Jacobson and Roszbach (2003) 與 Kim and Sohn (2007) 之審核階段與違約階段兩者間確有顯著相關存在的論點。

表 9 修正後 Heckman 樣本選擇審核模型與違約模型參數估計結果

變數名稱	係數	標準差	p 值
第一階段 (審核模型)			
截距項	1.1668	0.2672	0.0000
公司類別	3.2843	1.7839	0.0050
業外收支率	1.9254	2.1664	0.0010
負債比率	-0.2234	0.0918	0.0084
第二階段 (違約模型)			
截距項	-1.8705	1.0078	0.0635
淨營運資金佔總資產比	-5.2285	1.5087	0.0005
存貨週轉率	-0.0807	0.0042	0.0541
$\lambda_i$	0.6545	0.3090	0.0342
Pseudo $R^2 = 0.3732$			

資料來源：本研究整理。

接著我們將分別對第一階段與第二階段所選出的解釋變數意涵進行探討，首先就第一階段而言：以獲利性觀點言之，當公司的業外收支率愈高時，代表著該公司經營策略多角化的程度愈高且在本業外的獲利穩定，可推斷該公司還款能力佳，站在銀行審核貸款通過與否角度來看其通過貸款的可能性愈高；就安全性觀點言之，當公司的負債比率愈高時，代表公司財務槓桿的使用程度高，若未來公司營運不如預期情況時，公司較易發生財務危機，站在銀行審核貸款通過與否的角度來看其通過貸款的可能性愈低；就公司類別的觀點而言，由於本文樣本期間正好包含台灣電子產業盛行的年代，電子業在獲利與產業前景的表現上均優於非電子產業，故站在銀行審核貸款通過與否的角度來看，其通過貸款的可能性愈高。最後就第二階段而言：以流動性觀點言之，當公司的淨營運資金佔總資產比愈高時，代表著公司短期償債能力佳，流動性高愈不易發生週轉不靈的問題，故公司較不易發生財務危機；就效率性觀點言之，當公司的存貨週轉率愈高時，一方面代表公司商品在市場的接受度高且在銷售效率的表現上較突出，另一方面則代表企業的存貨量愈少，積壓存貨的風險相對較低，故公司較不易發生財務危機。

#### b. 傳統財務危機預警模型

表 10 傳統財務危機預警模型參數估計結果

變數名稱	係數	標準差	p 值
截距項	-1.2157	0.7697	0.1142
淨營運資金佔總資產比	-4.5669	1.5917	0.0041
存貨週轉率	-0.0853	0.0440	0.0517
負債比率	2.7263	1.4036	0.0521
Pseudo R <sup>2</sup> = 0.3660			

資料來源：本研究整理。

由表 10 的準判定係數為 0.3660 可知，傳統財務危機預警模型可解釋約 36.60% 的變異量，與前述建構之修正後 Heckman 樣本選



擇違約模型的準判定係數為 0.3732 相較可知，兩模型之準判定係數差異不大，因此本文後續將以模型配適能力與預測準確率進行進一步之分析，<sup>12</sup> 據以觀察本文修正後 Heckman 樣本選擇違約模型的配適度與預測能力是否優於傳統的財務危機預警模型。而模型之迴歸變數估計係數，則可說明各變數與公司發生財務危機間的關係，當負債比率數值愈高時，表示該公司愈容易發生財務危機；當淨營運資金佔總資產比與存貨週轉率兩變數數值愈高時，表示該公司愈不容易發生財務危機。另有關模型所選出的變數意涵而言，以流動性觀點言之，當公司淨營運資金佔總資產比愈高時，代表公司短期償債能力佳，流動性高較不易發生週轉不靈問題，故公司較不容易發生財務危機；以效率性觀點言之，當公司存貨週轉率愈高時，一方面代表公司商品在市場上的接受度高且在銷售效率的表現突出，另一方面則代表企業的存貨量愈少，積壓存貨的風險相對較低，故公司較不易發生財務危機；以安全性觀點言之，公司負債比率愈高時，代表公司財務槓桿使用程度較高，若未來公司營運不如預期狀況時，公司較容易發生財務危機。

#### 4.2.2 模型配適能力與預測準確率分析

本文首先觀察修正後 Heckman 樣本選擇違約模型與傳統財務危機預警模型對財務危機評估間是否具有統計上顯著差異，亦即兩模型之預測是否具有統計顯著差異。本文分別對兩模型之樣本內與樣本外之樣本公司所估算的違約機率值採配對樣本檢定進行分析，結果顯示樣本內與樣本外檢定之  $p$  值皆具顯著水準 10% 以上的顯著性，亦即在 10% 的顯著水準下，兩類模型於樣本內、外對財務

---

<sup>12</sup> 由於本文所建構之傳統財務危機預警模型與修正後 Heckman 樣本選擇模型，所包括的自變數並不同，故不同迴歸間的比較並不具經濟意涵。然而因本文修正後 Heckman 樣本選擇模型係加入拒絕推論技術後所建立之財務危機預警模型，而傳統財務危機預警模型則係僅以通過審核之申請者資訊所建立之模型，因此本文亦將進行模型效力比較，針對兩種涵蓋不同資訊之模型配適能力與預測能力進行相關分析。

危機衡量具顯著之差異。

其次，爲了驗證修正後 Heckman 樣本選擇違約模型與傳統財務危機預警模型，於樣本內與樣本外之配適能力與預測準確率情形，本文分別計算兩模型樣本內的 KS 值、AUC 值與 AR 值；而爲檢定兩模型對於樣本外資料之合適性，亦以樣本外資料進行驗證，檢驗模型對於樣本外資料之預測能力。當 AUC 值愈接近於 1，表示模型的預測正確性愈高；當 KS 值愈高時，代表模型本身愈能有效區別兩群體；而當模型 AR 值愈大時，則代表模型之預測準確率愈佳。

表 11 則列示兩個模型之 KS 值、AUC 值與 AR 值，其中修正後 Heckman 樣本選擇違約模型，樣本內的 KS 值、AUC 值與 AR 值分別爲 0.7357、0.8812 與 0.7623；傳統財務危機預警模型樣本內的 KS 值、AUC 值與 AR 值分別爲 0.6712、0.8817 與 0.7634；修正後 Heckman 樣本選擇違約模型，樣本外的 KS 值、AUC 值與 AR 值分別爲 0.7333、0.87 與 0.74。而傳統財務危機預警模型樣本外的 KS 值、AUC 值與 AR 值分別爲 0.6333、0.8133 與 0.6267。由於兩模型之 KS 值無論在樣本內或是樣本外均落在 60%~75% 範圍內，表示兩模型無論在樣本內或是樣本外均具有良好的區別能力，且修正後 Heckman 樣本選擇違約模型的樣本內、外 KS 值分別爲 0.7357、0.7333，明顯的優於傳統財務危機預警模型的樣本內、外 KS 值的 0.6712、0.6333，故本文實證結果可證實僅使用被接受申請者樣本建立財務危機預警模型會降低模型的預測能力。另外，兩模型之 AUC 值無論在樣本內或樣本外均落在 80%~90% 範圍內，代表兩模型無論在樣本內或是樣本外均有較強的配適能力與預測效力，且修正後 Heckman 樣本選擇違約模型在樣本外的預測效力也明顯的優於傳統財務危機預警模型。最後，兩模型在樣本內預測準確率之 AR 值表現上雖然差距不大，但 Heckman 樣本選擇違約模型在樣本外的預測準確率則明顯的優於傳統財務危機預警模型。故由模型效力檢定結果可知，本文納入未通過審核之樣本訊息所建立

之修正後 Heckman 樣本選擇違約模型，其配適力與預測力優於傳統僅考慮已通過審核之申請者樣本訊息所建立之財務危機預警模型。

表 11 部份樣本下樣本內與樣本外 KS 值、AUC 值、AR 值

		樣本內	樣本外
修正後 Heckman 樣本選擇違約模型	KS 值	0.7357	0.7333
	AUC 值	0.8812	0.8700
	AR 值	0.7623	0.7400
傳統財務危機預警模型	KS 值	0.6712	0.6333
	AUC 值	0.8817	0.8133
	AR 值	0.7634	0.6267

資料來源：本研究整理。

#### 4.2.3 穩健性分析 (robust analysis)

本小節探討在不同的樣本組合及樣本規模下，是否模型效力的檢定結果，仍呈現前文納入未通過審核樣本訊息所建立之修正後 Heckman 樣本選擇違約模型，其配適力與預測力優於傳統僅考慮已通過審核申請者樣本訊息所建立之財務危機預警模型結果。本文分別以前文上市公司部份樣本為例進行隨機抽樣，並另以上市公司全樣本資料進行與前文部份樣本資料相同步驟的模型建構過程，觀察本文修正後 Heckman 樣本選擇違約模型與傳統模型效力差異之穩定性。

本文首先針對前文之 172 家非全樣本上市公司進行隨機抽樣，將原始樣本中的樣本內、外公司隨機抽取 80% 作為另一群新的樣本內、外公司組合，重複此步驟 1,000 次，並計算兩模型的平均效力值，觀察不同樣本組合下的模型效力情形。各抽樣組合下的模型效力平均值則置於表 12 中。由表 12 可知，重複抽樣下部份樣本之樣本內與樣本外的平均 KS 值、AUC 值、AR 值亦呈現修正後 Heckman 樣本選擇違約模型，其配適力與預測力優於傳統僅考慮已通過審核申請者樣本訊息所建立之財務危機預警模型結果，且兩模

型之平均效力值，無論是 KS、AUC 或 AR 等值，皆在 1% 的顯著水準下具顯著之差異。

表 12 重複抽樣下部份樣本之樣本內與樣本外平均 KS 值、AUC 值、AR 值

		樣本內	樣本外
修正後 Heckman 樣本選擇 違約模型	平均 KS 值	0.7491	0.5322
	平均 AUC 值	0.8994	0.7673
	平均 AR 值	0.7988	0.5347
傳統財務危機預警模型	平均 KS 值	0.5276	0.4393
	平均 AUC 值	0.7627	0.7365
	平均 AR 值	0.5255	0.4731

資料來源：本研究整理。

本文接著另以上市公司全樣本資料進行與前文相同步驟之模型建構；此外，亦將全樣本資料的樣本內、外公司隨機抽取 80% 作為另一群新的樣本內、外公司組合，並進行 1,000 次的隨機抽樣，觀察在不同的樣本規模與樣本組合下，本文修正後 Heckman 樣本選擇違約模型與傳統財務危機預警模型效力差異之穩定性，模型效力檢定之結果則置於表 13 與表 14 中。實證結果亦呈現修正後 Heckman 樣本選擇違約模型，其配適力與預測力優於傳統僅考慮已通過審核之申請者樣本訊息所建立之財務危機預警模型結果，且在重複抽樣下，兩模型之平均效力值，無論是 KS、AUC 或 AR 等值，皆在 1% 的顯著水準下具顯著之差異。

表 13 全樣本下樣本內與樣本外 KS 值、AUC 值、AR 值

		樣本內	樣本外
修正後 Heckman 樣本選擇違約模型	KS 值	0.7384	0.4099
	AUC 值	0.8863	0.7016
	AR 值	0.7727	0.4033
傳統財務危機預警模型	KS 值	0.4857	0.3639
	AUC 值	0.7827	0.6803
	AR 值	0.5654	0.3607

資料來源：本研究整理。

表 14 重複抽樣下全樣本樣本內與樣本外平均 KS 值、AUC 值、AR 值

		樣本內	樣本外
修正後 Heckman 樣本選擇 違約模型	平均 KS 值	0.7441	0.5080
	平均 AUC 值	0.8945	0.7432
	平均 AR 值	0.7891	0.4844
傳統財務危機預警模型	平均 KS 值	0.6117	0.4105
	平均 AUC 值	0.8429	0.6866
	平均 AR 值	0.6857	0.3730

資料來源：本研究整理。

故由表 12 至表 14 可知，在不同的樣本組合與樣本規模下，本文考慮未通過審核申請者訊息所建構之修正後 Heckman 樣本選擇違約模型，其配適力與預測力皆有優於傳統僅考慮已通過審核之申請者樣本訊息所建立之財務危機預警模型結果，因此證實本文修正後 Heckman 樣本選擇模型與傳統財務危機預警模型效力差異是具有穩定性的。

## 5. 結論

由於過去相關文獻在建立財務危機預警模型時，皆是以已通過審核之申請者樣本來建立模型，未對未通過審核之申請者樣本進行討論，因此可能產生申請者母體分配與通過審核申請者樣本分配差異之樣本選擇偏誤問題，故由傳統模型所得到的結果不宜作為決策者之依據。本文則是以 Heckman 的樣本選擇模型為基礎建立財務危機預警模型，且亦使用僅通過審核之申請者樣本建立傳統的財務危機預警模型，並比較兩類模型在樣本內與樣本外之模型配適能力與預測準確率，期望能建立一個適用於企業貸款申請者之財務危機預警模型。

研究結果發現審核與違約兩階段模型間存在顯著相關，亦即僅考慮已通過審核之申請者樣本所建立的模型，確有樣本選擇偏誤問

題存在。而在模型的配適能力與預測準確率方面，可觀察到兩模型無論在樣本內或是樣本外均有不錯的配適能力與預測效力，但本文所建構之修正後 Heckman 樣本選擇違約模型在樣本外的預測能力明顯的優於傳統財務危機預警模型。故而未來在建立財務危機預警模型時，宜考慮模型樣本選擇偏誤問題，以拒絕推論技術建立財務危機預警模型，如此方可避免申請者母體分配與通過審核申請者樣本分配差異之樣本選擇偏誤問題，並可提高模型的預測準確率與配適能力。

而本文結果亦值得進一步擴展，就研究方法而言，由於過去使用於信用評分模型之拒絕推論技術可區分為三種類型，第一種為樣本可完全代表整個母體情況之理想技術，第二種類型為假設被接受申請者樣本分配經過某些觀察與假設後可被擴展至被拒絕申請者之分配，第三種類型為假設被接受申請者之母體分配不同於被拒絕申請者母體分配之情況。然而，由既往相關文獻可知，前兩種方法不是耗費較大成本就是無法增進模型預測效力，僅第三種方法較為可行，故本文以 Heckman (1976, 1979) 所提出的兩階段樣本選擇模型進行相關修正後，經由研究結果亦顯示，本文修正後之 Heckman (1976, 1979) 兩階段樣本選擇模型確實是可以修正傳統模型的樣本選擇偏誤問題，並增進模型之預測效力；然後續相關研究或可將前述樣本選擇偏誤的修正方法進行相關比較並進一步分析。

## 參考文獻

- 紀麗秋 Chi, Li-Chiu (2002), 「出口貿易信用風險模型之研究—以亞太地區為例」 “Study of Exporting Credit Risk Model: The Case of Asia-Pacific Basin”, 台灣金融財務季刊 *Taiwan Banking and Finance Quarterly*, 3 : 1, 81-116。 (in Chinese)
- 張大成 Chang, Ta-Cheng (2003), 「企業危機預測模型在台灣的應用與比較」 “Applying Bankruptcy Prediction Model in Taiwan’s Corporation”, 台灣銀行季刊 *Bank of Taiwan Quarterly*, 54 : 4, 147-163。 (in Chinese)
- 陳錦村 Chen, Jing-Twen (2007), 風險管理概要 *Risk Management*, 台北：新陸書局 Taipei: Shinlou。 (in Chinese)
- 曹曾樹 Tsaur, Tsen-Shu (2008), 「中小企業財務危機預警實證研究之文獻回顧」 “The Literature Review and Evidence Study of Financial Crisis Pre-Warning Model for Taiwan SMEs”, 中小企業發展季刊 *Journal of SME Development*, 9, 135-168。 (in Chinese)
- Altman, E. I. (1968), “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy,” *Journal of Finance*, 13:4, 589-609.
- Back, B., T. Laitinen, K. Sere and M. van Wezel (1996) “Choosing Bankruptcy Predictors with Discriminant Analysis, Logit Analysis and Genetic Algorithms,” in the 18th Annual Congress of the European Accounting Association, Birmingham, UK.
- Banasik, J., J. Crook and L. Thomas (2003), “Sample Selection Bias in Credit Scoring Models,” *Journal of the Operational Research Society*, 54, 822-832.

- Belkoui, A. (1980), "Industrial Bond Ratings: A New Look," *Financial Management*, 9:3, 44-51.
- Black, F. and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81:3, 637-659.
- Boyes, W., D. Hoffman and S. Low (1989), "An Econometric Analysis of the Bank Credit Scoring Problem," *Journal of Econometrics*, 40:1, 3-14.
- Charitou, A. and L. Trigeorgis (2000), "Option-Based Bankruptcy Prediction," Cyprus University and University of Columbia Working Paper.
- Chen, G. and T. Astebro (2001), "The Economic Value of Reject Inference in Credit Scoring," University of Waterloo Working Paper.
- Copas, J. B. and H. G. Li (1997), "Inference for Non-random Samples," *Journal of the Royal Statistical Society*, 59:1, 55-95.
- Crook, J. and J. Banasik (2004), "Does Reject Inference Really Improve the Performance of Application Scoring Models?" *Journal of Banking and Finance*, 28:4, 857-874.
- Ederington, L. H. (1985), "Classification Models and Bond Ratings," *Financial Review*, 20:4, 237-262.
- Engelmann, B., E. Hayden and D. Tasche (2003), "Testing Rating Accuracy," *Risk*, 16:1, 82-86.
- Farmen, T., S. Westgaard and N. van der Wijst (2004), "An Empirical Test of Option Based Default Probabilities Using Payment Behavior and Auditor Notes," Norwegian University of Science and Technology Working Paper.
- Greene, W. H. (1998), "Sample Selection in Credit-Scoring Models," *Japan and the World Economy*, 10:3, 299-316.



- Greene, W. H. (2003), *Econometric Analysis*, New Jersey: Prentice-Hall.
- Hand, D. J. (1998), "Reject Inference in Credit Operations," in E. Mays, ed., *Credit Risk Modeling: Design and Application*, New York: AMACOM.
- Heckman, J. J. (1976), "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," *The Annals of Economic and Social Measurement*, 5:4, 475-492.
- Heckman, J. J. (1979), "Sample Selection Bias as a Specification Error," *Econometrica*, 47:1, 153-161.
- Hillegeist, S., E. Kearing, D. Cram and K. Lundstedt (2004), "Assessing the Probability of Bankruptcy," *Review of Accounting Studies*, 9:1, 5-34.
- Hosmer, D. and S. Lemeshow (2000), *Applied Logistic Regression*, New York: John Wiley & Sons.
- Hsia, D. C. (1978), "Credit Scoring and the Equal Credit Opportunity Act," *The Hastings Law Journal*, 30, 371-448.
- Jacobson, T. and K. Roszbach (2003), "Bank Lending Policy, Credit Scoring and Value-at-Risk," *Journal of Banking and Finance*, 27:4, 615-633.
- Kim, Y. and S. Sohn (2007), "Technology Scoring Model Considering Rejected Applicants and Effect of Reject Inference," *Journal of the Operational Research Society*, 58, 1341-1347.
- Mays, E. (2001), *Handbook of Credit Scoring*, Chicago: Glenlake.
- Meng, C. and P. Schmidt (1985), "On the Cost of Partial Observability in the Bivariate Probit Model," *International Economic Review*, 26:1, 71-85.

- Merton, R. C. (1974), "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29:2, 449-470.
- Ohlson, J. A. (1980), "Financial Ratios and the Probabilistic Prediction of Bankruptcy," *Journal of Accounting Research*, 18:1, 109-131.
- Poirier, D. J. (1980), "Partial Observability in Bivariate Probit Model," *Journal of Econometrics*, 12:2, 209-217.
- Van de Ven, W. P. M. M. and B. M. S. van Praag (1981), "The Demand for Deductibles in Private Health Insurance: A Probit with Sample Selection," *Journal of Econometrics*, 17, 229-252.
- Vassalou, M. and Y. Xing (2004), "Default Risk in Equity Returns," *Journal of Finance*, 59:2, 831-868.
- Verstraeten, G. and D. van den Poel (2004), "The Impact of Sample Bias on Consumer Credit Scoring Performance and Profitability," *Journal of the Operational Research Society*, 56, 981-992.
- Zmijewski, M. E. (1984), "Methodological Issues Related to the Estimation of Financial Distress Prediction Models," *Journal of Accounting Research*, 22, 59-82.

# The Study of Sample Selection Bias in Corporate Financial Distress Prediction Model: An Example of Taiwan Listed Companies

Lin, Yu-Ling, Ta-Cheng Chang and  
Shin-Pin Huang

## Abstract

Traditionally, most scholars use the sample of accepted applicants in building financial distress prediction models and neglect the sample of rejected applicants. These models cannot reflect the variation of population and the interactive effects among all variables, and do not consider the problem of sample selection bias. Therefore, the fitness and prediction ability of the models would be affected. In this paper, the reject inference technology is considered in the financial distress prediction as building the modified Heckman two-stage sample selection model. Using the modified model and Taiwan's listed companies as examples, we could find that the application and the default stages are highly correlated in distress prediction. In other words, if the modified model is not used, the sample selection bias would result. After observing the fitness and prediction ability of our modified Heckman two-stage sample selection model relative to traditional financial distress prediction models, we discover that the prediction performance of the former is superior.

Keywords: Sample Selection, Reject Inference, Financial Distress  
Prediction Model, Probit Model

JEL Classification: C10, G20, G32

---

Lin, Yu-Ling, Department of Banking and Finance, Takming University of Science and Technology, No. 56, Huanshan Rd., Sec. 1, Neihu, Taipei, Taiwan, R.O.C. Tel : 886-2-26585801 ext. 5182, E-mail: [yulinglin@mail.takming.edu.tw](mailto:yulinglin@mail.takming.edu.tw). Ta-Cheng Chang, Department of International Business, Soochow University, No. 56, Kueiyang St., Sec. 1, Zhongzheng, Taipei, Taiwan, R.O.C. Tel: 886-2-23111531, E-mail: [dachen@scu.edu.tw](mailto:dachen@scu.edu.tw). Shin-Pin Huang, Yang Ming Marine Transport Corp, No. 271, Mingde 1st Rd., Cidu District, Keelung, Taiwan, R.O.C. Tel: 886-2-24559988, E-mail: [benz12354@hotmail.com](mailto:benz12354@hotmail.com).

Received 24 March 2009; revised 13 May 2009; accepted 12 October 2009.