

債信評等主觀評估內部一致性 暨外部一致性之探討：以銀行及 證券實務專家之差異分析為例

The Internal and External Consistency on Credit Rating: the Difference between Bankers and Security Analysts

郭敏華 *Min-Hua Kuo*

世新大學

Shih Hsin University

90 年 11 月 23 日收稿、91 年 2 月 21 日第一次修改、91 年 4 月 10 日第二次修改、91 年 11 月 15 日接受刊登

摘 要

在債信評等的傳統研究中，為減低主觀判斷的干擾，學術界僅擷取量化因素進入評等模型，期能提高債信評估的客觀性，然而在實務上，債信評等必須考量諸多質化因素，也是一個無法迴避的事實，而在各種內在心理偏誤及外在雜訊的干擾下，主觀判斷所夾帶的偏誤使專家判斷的一致性與可信度不無疑慮，以致學術與實務之間的鴻溝始終難以弭平。本研究的主要宗旨即在正視此一議題：利用嚴謹的學術研究方法，利用內部一致性與外部一致性二種角度，分別探討主觀評斷是否存在重大偏誤。針對評估債信能力具有一定專業水準的實務專家，本研究一方面以問卷調查方式觀察不同專家對同一受評對象債信品質的共識度（外部一致性），一方面以實際財務資料檢測專家評分時實際依據的財務因素與其所宣稱之重要因素是否一致（內部一致性），俾以瞭解專業知識的主觀判斷是否具有足以信賴的一致性。分析結果顯示，量質因素的相對重要性無分軒輊，任一皆不可偏廢。在評等結果的外部一致性方面，背景不同之專業人士對受評公司的債信等級存在相當程度的歧異，但其評定等級之 Pearson 相關及三項無母數相關係數皆達統計顯著水準，應屬差強人意；內部一致性方

面，專家進行債信評估所採用的因素與其所宣稱最重要的因素亦尚稱吻合。總體而言，專業判斷的信賴度雖未臻完善，仍具有一定程度的一致性。

關鍵詞：債信評等，信用評等，專業判斷，主觀判斷，行為財務學

Abstract

To analyze the internal and external consistency in the credit rating made by professionals, this study surveys the professional opinions about the scheme of credit rating. It also regresses the credit scores made by professionals on the major financial dimensions emphasized by the professionals themselves, so as to explore the internal consistency of professional judgement. The empirical results demonstrate that quantitative and qualitative factors are equally important in rating credits. Some divergence among professionals is evidenced, but it is substantially decreased when examining homogeneous sectors, the banking or security industry. The financial factors elected by professionals are able to explain the credit score quite well. Although both internal and external consistency are supported to a certain degree, great substance is remained to improve.

Key words: credit rating, professional judgment, behavioral finance

壹、研究動機與目的

所謂債信評等，係針對舉債公司或債券本身的債信品質進行評估，並給予一個相應等級或分數的過程。該評等結果，是一個綜合了所有重要因素與環節後所得到關於信用品質的綜合性結果（郭敏華，1999）。而所謂的重要因素與環節，涵蓋層面至廣，從國家風險、產業風險到個別公司的營運風險與財務風險等等，不一而足，其中有可具體衡量的數量因素，也有難以量化的質性因素。在國內，針對大眾型債券的債信評等未臻成熟，但是此一觀念在實務上的應用已有長久歷史，特別是與授信風險關係最密切的金融業。以銀行的授信業務來說，借款者的違約風險直接關乎銀行的資產品質與獲利能力，因此在核貸時必須針對客戶的債信能力進行詳細評估，在實質上這就是一種債信評等。銀行授信評估準則中有所謂的五 P 原則，包括 People（借款人）、Purpose（資金用途）、Payment（還款能力）、Protection（債權保障）、Perspective（未來展望），在評

估借款人的債信品質時，須就此五大因素加以分析，但是五 P 只是一個方向性的指導方針，至於究竟何等因素決定了債信評質，各因素中的相對重要性又如何，迄無定論。其中只有還款能力可利用財務資料進行量化分析，其他各因素構面都需要大幅倚重專業人員的主觀判斷。換言之，債信品質是一個多元性的概念，需要從不同面向進行評估，純然的量化模型必不周延，過於倚重質化因素又可能失之主觀，因為在主觀判斷下較難系統化地避免個人因素或雜訊等之干擾，難免有很大的潛在誤差。事實上，晚近行為財務學 (Behavioral Finance) 的崛起，即是鑒於個人心理傾向與主觀判斷對財務行為具有不容忽視的影響 (郭敏華，2002)。

以全球最具權威的評等機構 S&P's 與 Moody's 為例，其所公布的評等結果具有很高的公信力，但是根據調查，在同時接受這二家機構進行債信評等的公司中，有高達近四分之一的受評公司得到不同的評等結果！從此一角度來說，所謂的專業判斷是否具有足夠的信賴度，值得進一步省思。也許就是基於此一緣故，學術界不斷投入大量專業研究，期望發展出能夠以量化因素做為客觀判斷依據的評估模型 (參郭敏華，1999、2000)。經過多年的努力，各式各樣的評估模型不斷推出，雖具有一定的準確性，卻始終無法對各種攸關因素及各因素間彼此的相對重要性達成共識。細觀渠等研究可發現，採取不同的研究方法或是針對不同時間的資料，就得到不同的模型，又難免予人以資料驅動 (data driven) 之感。量化與質化間的相對重要性如何，各項因素構面間的相對重要性又如何，迄無最後共識。更諷刺的是，在實務上主觀評估佔有十分重要的地位，但是在典型的研究中，學術界所致力，正是消除主觀判斷的成分，以致學術與實務之間的鴻溝始終難以弭平。本研究的主要宗旨即在正視此一議題：利用嚴謹的學術研究方法，探究主觀評斷是否存在重大偏誤。

一般而言，欲了解有無主觀偏誤存在，有兩種取向可資切入：(1)所有實務專家對一件事情的判斷，是否存在很大的歧異 (外部一致性)。(2)專家實際的評估程序 (實際作為) 是否符合其所聲稱「應該如何」的理性觀點 (內部一致性)。基於此，本研究為探索債信評等中主觀與客觀間的取捨、質化與量化間的謎思，乃試圖針對金融業對債信能力評估判斷具有一定專業水準的從業人員，進行債信評估意見調查，以瞭解專業知識的主觀判斷是否具有足以信賴的一致性。照理說，如果要瞭解實務專家的評估因素，固然可以利用問卷等方式，請其回答，但是由於個人的內在思考架構無法觀測，再加上人類的思考系統是一個複雜系統 (system of complexity)，個人內在評估架構較諸能夠訴諸文字語

言者要複雜許多，因此專家所聲稱的重要評估因素是否與其真正的評估架構一致（也就是所謂的內部一致性），尚待深入剖析。如果專業分析師的主觀評等結果與其所指出重要評估因素的客觀表現不對稱，表示專家意見本身即欠缺內部一致性；若再加上不同專家間對同一個評估對象所做的判斷，南轅北轍歧異過大，不具外部一致性的話，或即意味著專業主觀判斷所能提供的資訊價值有所存疑，而凡此皆代表了發展數量化模型的努力是值得的。

簡言之，本研究最主要的目的在於了解債信評等中主觀專業判斷的內部一致性與外部一致性各如何，從而對債信評等結果的信賴度有所認識。具體而言，本文的研究問題有二：

- 一、內部一致性：專業人員的主觀評等結果與其所指稱之重要因素的客觀績效是否具一致性？
- 二、外部一致性：專業人員各自做出來的債信評等結果，彼此間是否具有高度的一致性？

貳、文獻探討

自從 Beaver (1966) 與 Altman (1968, 1977) 利用財務資訊提出預測危機企業的區別模型以來，相關議題的後繼研究十分豐富，而且對於利用數量化資訊來區分公司的債信能力，已累積了相當的成果，成績斐然。然而，儘管這些利用公開的數量資料所進行的區別與預測模型，都具有不錯的區別力，但是各研究所得到的模型，其內涵卻是各色各樣，各有各的變數，各有各的相對權重，幾乎無法歸納成一個具有共識的模型。此一現象或意味著，數量分析終有其衡量限制，當總體環境、競爭態勢、甚至管理團隊發生更迭時，原先建立的預測模型可能即已不再適用 (Altman, et. al., 1977)，仍須依賴專業的主觀判斷。

然而，但凡人都有七情六慾，都難免各種環境因素、心理因素的影響，而出現諸如從眾傾向（如郭敏華、蔡世烈，2002）或是代表性偏誤（如郭敏華、黃威智，2002）等「非理性」行爲。即使是訓練有素、經驗豐富的專業人士，仍無法免於若干心理學上常見的認知誤差，包括受到特定事件的過度影響、受到印象鮮明事件之影響等等，使其在評估的過程中不知不覺設定了錯誤的權數，而造成錯誤的判斷。因此所謂的專業判斷或許有其難以取代的價值，但是另一方面也不免引人深慮：在人類的有限理性及各種身心限制干擾下，主觀的

專業判斷是否反而成為重大的干擾來源呢？

爲了瞭解專家判斷在信用評等中所扮演的角色與功能，以下從兩個不同的研究方向分別檢視現有文獻的成果。

一、評等專家對於債信等級決定因素之看法

Libby (1975) 爲瞭解銀行的授信專業人員對於借款者信用變化的因素，是否能正確掌握，曾進行一項實驗研究，邀請專業人員填寫其主觀認知的攸關因素與財務比率，同時檢視實際發生困難公司之相關財務資訊的變化。結果發現銀行授信人員在其實驗中確能選取攸關之財務比率，並做出相當不錯的破產預測。Backer & Gosman (1980) 亦對幾家著名的債信評等機構以及 24 家大型銀行專業人員進行人員訪問與調查，以瞭解相關的實務專家，在進行信用評估時所重視的因素爲何。同時觀察受評公司在發生評等降級之際，有無任何財務變數具有顯著預測能力；最後再將主觀意見與客觀觀察結果做一比較，檢視關於評等重要因素的主觀看法與實際統計結果是否具有有一致性。根據 Backer & Gosman 的調查結果，所有受訪實務人士均表示財務比率在信用評等中具有相當重要性，但是受訪者對於短、中、長期的信用評估所重視的財務比率差異頗大。而檢視樣本降級前的財務狀況，發現受訪者所強調的財務變數在信用降級前果然多呈現重大改變，顯示專業判斷的正確性。

但是，變數是否產生變化是一回事，綜合各變數以預測危機是否發生又是另一回事。Abdel-Khalik & El-Sheshai (1980) 所進行的實驗研究卻有不利於專家角色的發現：銀行授信人員所能做出的破產預測，連機械化的統計模型都不如，這些所謂的專家，並未能選出具有預測能力的財務比率。專家能力在此研究中受到了質疑。

除了瞭解專家在實際評估時所重視的準則外，欲判斷專家在債信評等中的重要性與專業判斷的品質，亦可檢視不同的專家對同一對象的看法，是否存在重大歧異。Beattie & Searle (1992) 即整理了 1990 年得到 12 家國際性評等機構公佈評等的一千餘個長期債信的評等結果，其中評等結果完全相同的，佔 44%，有 35% 的評等等級略有差異（相差一個子級，例如 A 與 A+），14% 相差二個子級，6% 則相差三個或三個以上的子級。如果限定於最大的二家 Moody's 與 S&P's 來比較的話，在同時接受兩家評等的 1398 個案例中，等級完全相同的，佔 64%，在不等的 36% 中，有 16% 是 Moody's 的評等結果高於 S&P's，有 20% 是 S&P's 高於 Moody's；但是二者的皮爾森相關係數高達 0.97，顯示二者

債信評等主觀評估內部一致性暨外部一致性之探討：以銀行及證券賣家之差異分析為例

的評等結果不盡相同，但是至少在相對風險的排序上，尙具高度共識。

Ederington (1986) 也曾針對 1975 年至 1980 年間的信用評等結果作過調查。Moody's 與 S&P's 二家評等機構對於新發行的債券，有 10~15%的評等結果不同；對於已流通在外的債券，二者評等結果不同的比率較高，約佔 25%。Moon & Stotsky (1993) 針對地方政府債券的研究，發現在 252 個地方政府公債中，有 198 個二家評等公司的評等結果一致，評等結果不同者，計 54 個，佔 21.4%；其中 S&P's 評等等級高於 Moody's 者，有 41 個，低於 Moody's 者，有 13 個。粗略看來，S&P's 傾向於給予相對 Moody's 較高的評等。在所有評等等級不同的樣本中，除了一個樣本的等級差異大於一級外，雙方差異為一級。Jewell & Livingston (1998) 就 1980-1993 年間的受評債券進行調查研究。該研究發現，在 1,277 個樣本中，雙方評等結果出現歧異者，計 221 支，佔 17.31%；其中 S&P's 評定等級較高者有 121 支，佔 55% (參表 14-1)。

表 1 1980-1993 年間 Moody's 與 S&P's 債信評等結果分歧狀況

Moody's S&P's	Aaa	Aa	A	Baa	Ba	B
AAA	36	6				
AA	12	167	59			
A		19	446	31		
BBB			40	159	10	
BB				6	40	15
B					23	208

資料來源：Jewell & Livingston (1998)

由以上幾項調查分析看來，不同專業評等機構間的評等結果，確實存在若干差異，其比率約在 10%至 36%之間。

二、評等歧異之肇因探索

一般來說，客觀正確的評等是為評等機構賴以生存的基礎，具有權威性影響力的全球性評等機構，不但擁有一流的分析師，也累積了長期以來的專屬 know-how，再加上評等結果直接影響到受評公司的籌資管道與融資成本，以及投資人的投資決策，因此，評等機構的評等程序應該是高度專業，而且至為嚴謹的。S&P's 與 Moody's 都是極負盛名的專業評等機構，擁有一流的專業人才

自不在話下，更重要的是，二者都以高度公正客觀的立場著稱，所發佈的評等等級具有很高公信力與權威性。因此，如果所謂的專家，能對信用做出正確的評等，則二家對同一個受評對象，應可獲致相同的評等結果。然而由前面的說明可知，評等結果呈現差異的比重並不低，這當中是否意味著專家判斷的正確性未必無所疑慮？還是其他因素使然？值得探討。（其實，在評等機構的實務運作中，受評個案的最後評等，必須經過評等委員會充分討論後才能決定，而評等委員會係由多位評等分析師組成，各分析師間對同一個受評個案的看法可能互有不同，必須經過充分的意見交換與討論後，再投票決定。換句話說，即使是同一個評等機構內的分析師都可能各有不同的意見，那麼二家不同的評等機構出現差異，也就不足為奇了。）

為甚麼相同的債券卻得到不同的評等結果？可能原因有三：(1)不同的評等機構對受評者的看法相同，只是各級等的門檻各有不同的標準，以致出現了不同的等級；也就是說，二家公司對各信用等級所設定的標準不同。(2)不同的評等機構的評估模型確實存在系統性差異，可能是所評估的因素不同，或是各因素所屬的權重不同；也就是說，不同的評等機構對債信的決定因素有不同的看法。例如，甲機構認為獲利性比安全性重要，乙機構卻認為安全第一，如此的差異自然造成評等結果的不同。以及(3)其他誤差。

根據 Ederington (1986) 的研究，在 1975 -1980 年間新發行的債券並同時被 S&P's 與 Moody's 評為 B 級以上的 493 個債券中，有 63 個出現不同評等，比率為 13%。這些公司即列為該研究之樣本。該研究以 S&P's 所聲稱最重要的財務變數（長期負債／總資產、利息保障倍數、總資產獲利率、現金流量對長期負債比率），加上規模、債權順位二項變數後，得到六個自變數，藉以探討 S&P's 與 Moody's 對相同債券有不同評等結果的可能原因。

研究結果顯示，兩家債信評等公司所設定的信用等級間，分界值並未呈現顯著差異；兩公司對債券屬性的權重也沒有顯著差異；也就是說，信用評等結果的不同，並非因雙方對級距定義不同所造成，也非因對重要因素的看法存在歧見。換言之，差異主要來自於誤差項，而構成誤差項的可能內涵除了隨機誤差外，也可能是因為尚有其他債券屬性未被此研究採用，且兩公司對這些屬性有不同重視程度。如果是後者的話，仍表示二家的專業分析師對於債信等級的看法呈現歧異。

Moon & Stotsky (1993) 以地方政府公債為樣本，也曾進行類似研究，實證結果更不利於專家的可信賴度：S&P's 與 Moody's 不但分級的門檻標準不同，

連各項變數的相對權重也不同。

至此，對於所謂評等專家、專業主觀判斷的可信賴度，褒貶均有，仍無定論。但是除了上述觀點之外，吾人尚可從另外一個資訊角度來思考此一問題。只要債信評等機構具有一定的公信力，專業能力獲得信任與肯定，則對投資人或是市場來說，債信評等機構就具有提供額外資訊以及降低資訊不對稱的價值。根據 Hsueh & Kidwell (1988) 針對地方政府發行的公債所作的研究，S&P's 與 Moody's 的角色似具有互補性，因為他們發現，很多地方政府會同時尋求二家進行債信評等，而且如果二家機構所評定的等級相同時，借款成本可以下降。由此看來，債信評等機構的角色可能不僅止於資訊的提供，在降低資訊不對稱、降低資訊不確定，或是資訊正確無誤的保證上，也具有一定的功能與價值。此應為專業判斷的另一項（甚至更重要的）價值。

而當二家不同的債信評等機構的評等結果不同時，市場如何看待？Billingsley (1985) 與 Liu & Moore (1987) 等認為市場按較低的等級來定價；而 Hsueh & Kidwell (1988) 與 Reiter & Ziebart (1991) 卻發現市場是按較高等級來定價。這些學者都認為，評等結果的差異並非來自隨機誤差，而是因為評等機構間對受評公司的債信有不同的看法，市場會採取其中可信賴度較高者作為定價依據。

上述研究的樣本數約在 200 個至 300 個之間，按一般研究來說，此一樣本規模已綽綽有餘，但是由於評等出現差異的債券平均佔所有受評債券的 17%，受評等級以六個來說，有五種差異的可能，換句話說，每 100 支債券中，每種錯誤情形只有 3 支 ($100 \times 0.17/5 = 3.4$)，故存在樣本過小的統計推論問題 (Jewell & Livingston, 1998)。此外，這些研究將高殖利率（垃圾）債券排除在樣本之外，但高殖利率債券的違約機率遠較投資級債券為高，應該更值得研究。於是 Jewell & Livingston 以更大的樣本 (1980-1993 年中計 1,277 個樣本) 進行研究。研究結果發現，當 Moody's 與 S&P's 對同一支債券的評等結果不同時，並沒有系統性的恆高或恆低的現象，也就是說，二家評等公司的評等結果互有高低，並沒有哪一家一貫偏高或偏低。而市場對這種債券的取價，則是介於高低之間，也就是採取二個評等結果的平均數做為定價依據。換言之，不同的評等機構所進行之債信評等，對市場都具有一定的參考價值。

由上述各研究看來，主觀判斷難免發生歧異，但是市場仍給予相當的價值肯定，並據以形成取價利率。

參、研究樣本與研究設計

本研究採取問卷調查方法蒐集專業人員的看法與債信品質的主觀評等資料，據以觀察外部一致性，同時利用財務資訊建立迴歸模型，以驗證受訪者於債信評估上的內部一致性。基於銀行徵信與授信人員對於債信評估最為熟悉，而證券分析師與基金經理人以證券分析為專業，對於上市公司的信用品質與各種財業務動態最能掌握，因此本研究即以銀行徵授信人員及證券業實際從事分析或操盤之專業人員作為調查對象。

問卷設計分為三大部份，包括對償債能力或債信品質決定因素的看法、對目前數個重要產業中各上市公司償債能力的主觀評估意見，以及個人基本資料。先請受訪者針對「質化 vs. 量化」的比重進行填答，再各自就質化因素中的可能構面進行「投票式」填答，以及量化因素百分比的填答。如此設計的原因是，在質性因素中，可能涵蓋的構面變數至多至廣，每個人的看法可能相去甚遠，如果由研究者主觀列出幾項因素的話，難免有錯失重要資訊之虞；此外，個人能否對於所有因素十分清楚而肯定地排列出其相對重要性的百分比，令人存在高度疑慮，因此，質化因素部份若採量化式強制填答，最後勉力得到的問卷結果勢將存在很大的誤差，在此情形，較適合採取開放式問卷。至於量化因素部份，由於財報分析已是一門十分成熟的分析技巧，也是專業人員必須具備的一項基本技能，所以對於其中應有的構面，已具較高的共識，故可以採取結構化的填答方式。茲將問卷內容分項說明於后。

一、償債能力之決定因素

包括三個問題：

1. 在評估償債能力時，可由財務比率衡量的因素，相對於其他質化因素的重要性如何。以 10 分為滿分，比重由 1 至 9，由填答者勾選。

2. 在財務報表以外的質性因素中，有那些因素較重要。在問卷中細列各可能質化因素，由填答者勾選，可複選。

3. 在財務比率可衡量的因素中，其相對重要性如何。在問卷中按金融業實務評估習慣，列出安全性、獲利性、流動性、效率性、資金期間配適度，及其他（如勾選此項，請填答者說明）等六項，請填答者就各構面的相對重要性填寫分數，合計 100%。為避免語意混淆，在問卷中對各構面的定義均有所說明。

二、主觀評等

選擇股市中較為活絡的八個產業製表列示於縱欄中詳列產業中所有的上市公司，橫列則分為 AAA、AA、A、BBB、BB、B、CCC、CC、C 和 D 等 10 個等級，請填答者就其對各公司償債能力的主觀認知加以勾選。各等級所代表的意義，係參照 S&P's 的定義，在題後均加以詳細說明。

八個產業類股包括水泥、食品、電機、鋼鐵、汽車、電子、百貨，與觀光等。請填答者就其所熟悉的產業中至少撰擇二個產業回答。

三、個人基本資料

包括教育程度、服務機構，與擔任職務。亦在問卷中詳列可能答案，由填答者勾選。

本研究發出問卷計 120 份，經電話確認與催請回函等動作，回收 82 份，扣除回答不清、前後不一致之問卷後，有效問卷計 77 份。其中受訪者任職於銀行者有效問卷 31 份，證券業中，屬證券公司者 27 名，投信公司 13 名，未填答者 6 名。擔任職務方面，31 位銀行行員中，21 位徵信人員，6 位授信人員，4 位未填答；證券業的受訪者中，研究員 23 名、基金經理人 9 名、管理者 3 名、其他及未填答者 11 名。教育程度方面，高中職及以下 2 名，專科畢業 2 名、大學畢業 32 名、碩士 31 名，其他及未填答者 12 名。

為回答前列研究問題(1)與(2)，亦即量質因素的相對重要性與涵蓋的重要構面，本研究採用各項敘述統計工具進行分析；研究問題(4)關於外部一致性的問題，除以相關分析、F 檢定與 t 檢定加以檢視外，亦將受訪專家的評等結果與台灣經濟新報社 (TEJ) 同一時間的評等結果做一對照比較；最後，為進一步分析專家意見的內部一致性，亦即受訪者的評分架構與其所聲稱各重要因素是否一致 (研究問題(3))，本研究將各受評公司所得之平均信用分數為應變數，將各財務量化構面的衡量變數做為自變數，利用迴歸模型加以探討。各財務構面衡量變數係根據銀行實務採操作性定義如下：

安全性：負債／資產

獲利性：稅後淨利／股東權益

流動性：流動資產／流動負債

資金配適：(長期負債+股東權益)／(長期投資+固定資產)

經營效率：營收／資產

肆、實證結果與討論

經整理所回收之回卷，針對前列各項研究問題進行分析後發現，兩組專家對於質化因素與量化因素皆認為同等重要，但是對質化因素內涵的著重點略有差異，似與二組專業人士的業務性質有關。更重要的是，本研究針對二組專家的評等結果進行外部一致性檢視，發現組間差異頗大，但組內差異則明顯小了許多，意味著主觀評等受到個人背景與思考習慣的影響頗大。最後，針對主觀靜態認知與實際評等的動態過程是否一致的問題，實證結果顯示大致尚稱一致。茲就各結果說明於后。

一、質化與量化因素的相對重要性

在影響償債能力的各種量質因素中，係以財務報表量化因素（以下簡稱量化因素）較重要，抑或以財報以外的質性判斷（以下簡稱質化因素）較重要？以滿分為 10 分的情況下，在 77 份有效問卷中，量化因素相對於質化因素的重要性，眾數與中位數均為 5 分，平均數 5.1 分，（故質化因素平均 4.9 分），標準差 1.5 分，量化因素的最高相對分數為 8，最低為 2（亦即質化因素的重要性，同樣最高 8 分，最低 2 分），故平均而言，質化因素與量化因素的相對重要性可謂平分秋色，旗鼓相當。（請參表 2）

為瞭解對質化與量化因素相對重要性的看法，是否因任職於銀行或證券業而有所不同，乃進一步將回收問卷分為銀行與證券兩組加以觀察。結果發現，量化因素的相對重要性銀行組略低於證券組。銀行組對於財報因素的重要性為 4.97 比 5.03，量化因素的最高相對分數為 7，最低為 3；眾數與中位數亦均為 5；標準差 1.33，變異係數 0.27。至於證券組，量化因素的相對重要性略高，平均為 5.47 對 4.53 之比，量化因素的最高相對分數為 8，最低為 3；眾數與中位數均為 6 分；標準差 1.5，變異係數亦為 0.27。初步看來，證券專業人士似略重視量化因素（參表 3）。但是經 t 檢定，不論是銀行組、證券組，或是所有樣本，對於量化因素與質化因素的相對重要性均無顯著差異。至於銀行與證券人員對此一看法的差異，經均數差檢定，亦未達統計顯著水準。換言之，不論是銀行組或證券組，均認為量化因素與質化因素對於債信評估同等重要。

表 2 量化因素相對於質化因素的重要性統計

量化因素相對重要程度 ¹	平均數 ²	最大值	最小值	眾數	中位數	標準差	變異係數
全部樣本(77)	5.10	8	2	5	5	1.50	0.29
銀行組(31)	4.97	7	3	5	5	1.33	0.27
證券組(40)	5.47	8	3	6	6	1.50	0.27
二組均數差	0.18	p=0.870					

說明：1.以 10 分為滿分。()內為樣本數，另有 6 份問卷未註明所屬行業。

三組均未顯著異於 5；銀行組與證券組間的均數差亦未達統計顯著水準。

二、較重要的質化構面

影響債性品質的質化因素，問卷中所列包括：產業環境、管理能力、成長潛力、管理者誠信、研發能力、市場佔有率、業界地位、退票紀錄、規模、公司知名度、黨政關係、與貴公司往來關係、自有品牌等項目。其中以產業環境所獲票數最多，獲 57 票，其次依序為管理能力與成長潛力各 46 與 44 票、管理者的誠信 41 票，研發能力 35 票，以及市場佔有率 32 票等。

分就銀行與證券組觀察，在銀行組中，除產業環境的 22 票居首外，管理者誠信與管理能力所獲票數最多，分別為 21 與 19 票，其次為研發能力 17 票，與市場佔有率 14 票。在證券組中，與銀行組所重視的因素略有出入，除產業環境 30 票同樣居第一外，獲票數最多者為成長潛力的 25 票，其次為管理能力 21 票，市場佔有率 17 票。

大體而言，受訪者對產業環境與管理能力重要性的共識度最高，除此之外，銀行業人員較重視管理者誠信，有 14%的受訪者勾選此項，但是證券組不到 9%勾選之。證券業人員較重視成長潛力，有 14%的受訪者勾選此項，但是銀行組勾選者者一樣不到 9%。同樣地，退票紀錄受到重視的比例，銀行業（8.6%）明顯高於證券業（5.2%）；而證券業對規模的重視度（8.0%）又高於銀行業（2.6%）。由此看來，雖然所有的受訪者所評估的是公司的債信，但是在思考時似乎仍然會受到受訪者所屬產業特性與習慣性思考的影響。

表 3 質化因素重要構面得票數統計

	全部樣本		銀行業組		證券業組	
	票數	百分比	票數	百分比	票數	百分比
產業環境	57	16.0%	22	14.6%	30	17.2%
管理能力	46	12.9	19	12.6	21	12.1
成長潛力	44	12.4	13	8.6	25	14.4
管理者誠信	41	11.5	21	13.9	15	8.6
研發能力	35	9.8	17	11.3	16	9.2
市場佔有率	32	9.0	14	9.3	17	9.8
業界地位	26	7.3	12	7.9	13	7.5
退票紀錄	23	6.5	13	8.6	9	5.2
規模	20	5.6	4	2.6	14	8.0
公司知名度	11	3.1	3	2.0	7	4.0
黨政關係	9	2.5	5	3.3	3	1.7
與貴公司關係	8	2.2	7	4.6	1	0.6
自有品牌	4	1.1	1	0.7	3	1.7
合計	356	100.0%	151	100.0%	174	100.0%

說明：因有部份受訪者未註明任職行業，故二組總和不等於全部樣本。

三、較重要的量化構面

一般在進行財報分析時，多分就不同的構面加以評估，本研究亦按銀行實務將分析構面分為安全性、獲利性、流動性、效率性、資金期間配適度，及其他等項目，請填答者在合計 100 分的原則下，填答其對各構面彼此間相對重要性的主觀看法。

經統計，在全部樣本中，平均而言，各構面的重要性依序為獲利性(24.4)、安全性(21.9)、流動性(19.4)，以及資金期間配適度(17.0)與效率性(15.1)，其他(2.2)。銀行業組的看法，安全性第一(24.7)，獲利性第二(21.5)，其次依序為流動性(19.9)，資金期間配適度(17.5)、效率性(14.8)，其他(1.61)。至於證券業組，則認為獲利性(26.5)高於安全性(20.2)，流動性第三(19.4)，接著為資金期間配適度(17.2)、效率性(15.4)與，其他(2.68)。經檢定，銀行業組與證券業組對各構面重要性的看法差異均未達統計顯著水準。

表 4 財務因素各構面之相對重要性

		平均數	標準差	中位數	最大值	最小值
所有樣本	獲利	24.4	10.1	25	50	10
	安全	21.9	8.8	20	50	5
	流動	19.4	7.3	20	50	5
	配適	17.0	7.5	20	30	0
	效率	15.1	6.1	15	30	0
銀行組	獲利	21.5	9.8	20	50	10
	安全	24.7	10.2	25	50	5
	流動	19.9	6.0	20	30	5
	配適	17.5	6.6	20	30	5
	效率	14.8	5.3	15	25	0
證券組	獲利	26.5	9.5	25	50	10
	安全	20.2	7.0	20	40	10
	流動	19.4	8.6	20	50	10
	配適	17.2	8.2	20	30	0
	效率	15.4	6.8	15	30	0
均差檢定		獲利	安全	流動	配適	效率
	T 值	0.038	0.046	0.785	0.878	0.690
	P 值	0.970	0.963	0.435	0.383	0.492

就此一部份可發現，受訪者在思考時確有受到其本身習慣性思考影響的脈絡可尋：銀行業重安全，證券業重獲利；但經檢定二者間差異未達統計顯著水準，整體而言，對於影響債信品質的因素，專家意見頗具共識。

四、專家評等之外部一致性

此一部份主要在觀察專家主觀評等的外部一致性程度。基於實際進行債信評估的專業人員不可能熟悉所有的產業，因此本研究並不要求全部填答，而是請求受訪者就其所熟悉的產業，針對至少二個產業作答。經統計整理，以八大產業 179 家上市公司而言，平均每家公司有 25 位受訪者填答，其中銀行業 12 位，證券業 11 位；填答最多者有 49 位（佔有效樣本之 64%），最少者有 15 位

(佔有效樣本之 20%)。不論何組，選擇汽車與水泥業者最多，其可能原因或許是這兩個產業都屬於典型的傳統產業，變化少，相對穩定，熟悉者較多；另外一種可能性屬於衡量誤差，蓋水泥業與汽車業的上市公司相對較少，填答即相對容易，若然，此一樣本偏誤是為本文一項研究限制。至於最活絡的電子資訊業，平均每家公司得到填答的人數僅 21 人，可能原因是電子資訊業中十分複雜，又可細分為若干子產業，熟悉 IC 業者可能未必熟悉被動元件等，因此平均下來人數相對較低。

至於受訪者對各家公司債信的評估，本研究對各債信等級主觀給予數值分數，最高等級 AAA 級 9 分，AA 級 8 分，依此類推，CCC 級 3 分，CC 級 2 分，C 和 D 級 1 分。同一家公司的債信等級如何，不同的專家可能有不同的意見，經統計，在 179 家受評公司中，各受訪者給予同一公司評等結果，的最高級次與最低級次間的差距，以 4 個級次最多（如 AAA 與 BB、A 與 CCC 等），計 88 個，佔 46%，其次為相差 3 個級次，佔 30% 左右；179 個受評公司中，沒有任何一家公司得到完全一致的評等共識，甚至連僅差異一個級次的也沒有。但是如果分別針對銀行業組與證券業組加以分析卻可發現，專家意見的歧異性大幅下降，評等結果最大級差由 4 級減為 3 級。在全部樣本分佈中，級差在 3 級（含）以下者，佔 36%，銀行業增至 59%，證券業更提升至七成以上（72%）。以各受評對象的得分分佈計算標準差，銀行組的標準差為平均 1.05¹，高於證券組的 0.97，但並不顯著，經 F 檢定未達統計顯著水準（ $p=0.989$ ）。銀行組與證券組級差分佈變異數經 F 檢定，亦未達顯著差異（ p 值為 0.454）。

再進一步檢視個別受評公司的評等結果，在 179 個樣本公司中，有 110 個的銀行評分均數低於證券公司，69 個樣本的評等結果銀行組評分平均高於證券組，意味著銀行組的給分較證券組「嚴格」，但分數差異未達統計顯著水準。如將平均分數四捨五入換算成符號等級（如 9 分表 AAA 等），二組評定債信等級相同者，計 106 個，佔 59%；有 49 個（28%）的銀行評定等級低於證券組，有 24 個（13%）高於證券組（詳表 6、表 7）。此一歧異程度遠高於表一 Moody's 和 S&P's 二者間歧異，其中可能原因有二：一是本研究調查對象包括證券業專業人士，債信評估是其所關注的一部份，卻未必是最主要的部分，因此受訪者

¹ 平均標準差是指 179 個受評樣本中，每個樣本獲得受訪分析師評定分數所形成分佈的標準差，全部加起來平均。在未平均之前，每個樣本的標準差代表分析師對該樣本看法的歧異程度，加以平均即可大略衡量分析師對所有樣本的平均歧異程度。證券組的一致性高於銀行組，但是二組間並未達統計顯著水準。

表 5 主觀債信評估之歧異統計

最大級差	全部樣本		銀行組		證券組	
	個	比重	個	比重	個	比重
0	0	0%	0	0%	0	0%
1	0	0	4	2	13	7
2	10	6	31	17	43	24
3	53	30	71	40	73	41
4	82	46	61	34	44	25
5	29	16	10	6	6	3
6	5	3	2	1	0	0
合計	179	100%	179	100%	179	100%

在填答問卷時雖被要求針對公司的債信做評分，卻仍可能受其證券分析所習慣評估之其他因素（如成長潛力等）所影響，致二組間差異較大。第二項可能原因是 Moody's 與 S&P's 都有其內部使用的評估架構，且在發佈評等結果之前已與內部進行意見交換與討論，而本研究之受訪者來自不同公司，缺乏一致的評估架構，致評估結果歧異較大

表 6 銀行組與證券組評定債信級等差異統計

級差	0	1	-1	-2	合計
樣本數	106	23	45	5	179
比重%	59	28	25	3	100

說明：級差係指銀行組對受評公司所評定之債信平均數減去證券組者。

表 7 銀行組與證券組債信評等結果歧異狀況

銀行 \ 證券	AAA	AA	A	BBB	BB	B
AAA						
AA	1		5			
A		7		14		
BBB		4	24		5	
BB			1	12		
B						

為使兩組間的順序相關檢定更為健全 (robust)，本研究另以四項兩母體有母數及無母數檢定工具，進行銀行與證券兩組專家間的相關性檢定。由表 8 可知，不論是所評定的平均分數或是評定順序，銀行組與證券組間的相關性皆達高度統計顯著水準，可見二組專業人員對個別公司的看法雖存在一定程度的歧異，若就價信的高低順序來說，則具有相當高的一致性。

表 8 銀行組與證券組專家評等之相關檢定

兩母體無母數檢定	統計量	p-level
Pearson Correlation	0.750	0.000
Spearman Rank Order Correlations	0.726	0.000
Kendall Tau Correlations	0.539	0.000
Gamma Correlations	0.544	0.000

最後，本研究另以台灣經濟新報社所做的信用評等結果（包括所給分數與最後等級）與本研究之專家評等結果並列比較，經以 Pearson 與 Spearman 相關分析加以檢測後發現，本研究證券組專業人員與該社評等分數之相關係數為 0.48，銀行組為 0.42，均達 1%統計顯著水準。至於此二項相關係數均遠低於本研究二組間的相關係數（0.75），其原因是可以理解的，蓋該社所採評等模型仍是以量化因素為主，並未含括較主觀的質化因素。

表 9 各組等級與 TEJ 所評等級之 Pearson 與 Spearman 相關係數

Pearson/Spearman	證券組	銀行組	TEJ 分數	TEJ 等級
證券組				
銀行組	0.75/0.73			
TEJ 分數	0.47/0.49	0.42/0.41		
TEJ 等級	0.46/0.47	0.41/0.40	0.99/0.99	

說明：1. TEJ 所分等級與本研究類似，亦為 9 級，但數字越高表等級越低，上表之相關係數已先乘 -1，使等級相關能與分數相關一致。

2. 表中所有相關係數均達 1%統計顯著水準。

五、內部一致性：財務構面對信用評分之解釋力

為剖析專家意見的內部一致性，亦即受訪專業人員對受評公司所給予的評

分是否與其所聲稱重要財務因素的客觀表現相互一致，本研究取 173 家受評公司（另有 6 家公司資料不齊，乃予刪除）所評定的平均信用分數做為待解釋變數，以問卷中五大財務構面的衡量變數做為解釋變數，蒐集問卷時最近年度（民國 86 年）之實際績效利用迴歸分析進一步剖析。結果發現，五項財務變數對受評公司平均信用分數變異之解釋力，調整後 R^2 為 0.4，截距項高度顯著，與前述質化因素重要性佔五成的發現十分一致²。

而觀察各變數之標準化係數，發現安全性與獲利性確實是最重要的二項變數，與問卷發現一致；但是根據回收問卷，流動性與資金配適度在財務構面中相對重要性達 36%，在量化模型中卻都不顯著，其可能原因至少有二：第一，意味著這兩項在理性認知中相當重要的變數，於實際進行專業評估時並未給予相對應的權重。第二、本研究所採之衡量變數存在衡量誤差，未能完全衡量流動性與資金配適度。

表 10 財務構面對信用評分之解釋力

R ² =0.419 Adjusted R ² =0.401 F(5,167)=24.104 p<.00000						
N=173	標準化係數	標準誤	係數	標準誤	t(167)	p-level
截距			6.857	0.218	31.478	0.000
安全性 ¹	-0.363	0.073	-0.019	0.004	-5.009	0.000
獲利性	0.383	0.067	0.020	0.003	5.698	0.000
流動性	-0.033	0.069	0.000	0.000	-0.470	0.639
經營效率	0.161	0.064	0.188	0.074	2.535	0.012
資金配適	0.042	0.064	0.000	0.000	0.653	0.515

說明 1：以負債比率衡量，故係數理論符號應為負，即負債比率越高，則信用分數越低。

伍、研究限制

本研究之研究對象係對於債信評等具專業能力之實務專家，涵蓋商業銀行與證券公司之專業人員。由於債信評估僅是證券分析內容的一部份，因此在受

² 由於問卷是在民國 87 年間所做，故另亦取 87 年財務資料進行迴歸，可在某種程度內檢視專家對未來財務狀況的預測力，迴歸分析結果顯示，截距項仍高度顯著，但調整後 R^2 為降為 0.34。安全性與獲利性最重要，流動性與資金配適度亦不顯著。

訪者填答問卷時，雖可清楚瞭解問卷主題在於「債信評估」，而非證券投資分析，仍可能因專業習慣而傾向於投資評估。此是為本研究限制之一。

此外，本研究在迴歸分析中，對應變數債信等級採主觀配分，如 AAA 9 分、AA 8 分，以此類推，此一作法隱含了各相鄰級距間債信品質的差異係彼此相等之假設，但實際上的信用品質差距，在不同的相鄰級別間可能並不相等。若然，則本研究於迴歸模型的係數評估可能有所偏誤，此為本另一項研究限制。

陸、結 論

根據以上實證分析，不論是問卷調查結果或是量化模型的解釋力均顯示，量質因素對於債信品質之評估，可謂同等重要。債信品質的主觀評等結果，銀行組與證券組的專家評等間存在相當的歧異，但是若單就銀行或證券業本身來看，一致性大幅上升。而利用實際財務資料以量化模型檢驗的結果顯示，專家主觀評等的債信等級與其所宣稱重要因素之實際表現尚稱吻合，雖然較次要的因素（如流動性與資金配適）在數量模型中並不顯著，但是相對重要性最高的獲利性與安全性，在量化模型中均達高度統計顯著水準。總體而言，專業判斷的信賴度雖未臻完善，仍具一定程度的一致性。

由於質化因素在債信評估中具有不容忽略的重要性，量化模型無法全然取代主觀判斷，因此專業判斷具有一定價值。從本研究各專家主觀評等結果未盡一致，以及靜態認知與動態過程仍有一定差異（較次要因素在迴歸模型中不顯著）的現象進一步思考，主觀運作的盲點和認知干擾難以完全消除，因此設法降低債信評等過程中人為判斷的偏誤仍是值得鼓勵與努力的，不論是傳統習見的以數量資訊建立區別模型的研究，或是將產業競爭分析納入系統評估架構的研究，在降低主觀偏誤上都具有一定的意義與價值。簡而言之，欲精確評估廠商債信，量化模型無法全然取代質化評估，但質化判斷易受到個人主觀意識的影響，因此建立嚴謹的評估架構以降低個人主觀干擾至屬必要。

參考文獻

- 吳俊奇，1997，信用評等模型建構之研究以我國電子資訊產業上市上櫃公司為例，國立政治大學企業管理研究所碩士論文。
- 郭敏華，2002，「理性或混沌？從財務危機談起」，貨幣觀測與信用評等，37 期：

債信評等主觀評估內部一致性暨外部一致性之探討：以銀行及證券實家之差異分析為例

7-13。

郭敏華，1999，「債信等級之決定因素(上)」，證券暨期貨管理雜誌，17 卷 11 期：1-16。

郭敏華，2000，「債信等級之決定因素(下)」，證券暨期貨管理雜誌，18 卷 1 期：1-9。

郭敏華，1999，「債信評等程序暨其在金融監理上之應用」，存款保險資訊季刊，13 卷 1 期，九月：55-76。

郭敏華、蔡世烈，2002，從行為財務學的觀點探討台灣股市之從眾傾向，第二屆跨領域管理學術暨實務研討會，東海大學主辦。

郭敏華、黃威智，2002，投資人代表性捷思傾向之探討：以類似事件資訊效果之關聯性為例，第一屆全國行為財務學研討會，世新大學主辦。

陳隆麒、郭敏華、吳政穎，與盧雲江，1997，「財務評等五力分析模型—以觀光業為例」，管理與資訊學報，2 期：77-108。

謝青曉，1997，臺灣上櫃證券商信用評等模型之研究，國立政治大學企業管理研究所碩士論文。

鍾俊文、陳勇徽，1996，「銀行信用評等-本國銀行之實證分析」，基層金融，33 期，9 月：27-53。

Abdel-Khalik, A. Rashad and Kamal M. El-Sheshai. 1980. Information choice and utilization in an experiment on default prediction. *Journal of Accounting Research*, 18(2):325-342.

Altman, Edward I. and G. G. Haldeman and P. Narayanan. 1977. ZETA Analysis: A new model to identify the bankruptcy risk of corporations. *Journal of Banking and Finance* :29-54.

Altman, Edward. 1968. Financial ratios: Discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4):589-609.

Backer, Morton and Martin L. Gosman. 1980. The use of financial ratios in credit downgrade decisions. *Financial Management*, 9(1):53-56.

Beaver, William H. 1966. Financial ratios as predictors of failure. *Journal of Accounting Research*, 71-111.

Billingsley, Randall S. 1985. Split ratings and bond reoffering yields. *Financial Management*, 14(2):59-65.

Ederington, Louis H. 1986. Why split ratings occur? *Financial Management*, 15(1):37-47.

Goldstein, Avi. 1996. What's in a rating? Agency approaches to analyzing MBS. *Journal of Fixed Income*, 5(4) March: 61-73.

Hsueh, L. Paul and David S. Kidwell. 1988. The impact of a state bond guarantee on state credit markets and individual municipalities. *National Tax Journal*, 41(2) June: 235-245.

- Jewell, Jeff. and Miles Livingston. 1998. Split ratings, bond yields and underwriter spreads. *Journal of Financial Research*, 21(2):185-204.
- Libby, Robert. 1975. Accounting ratios and the prediction of failure: Some behavioral evidence. *Journal of Accounting Research*, 150-161.
- Liu, Pu. and William T. Moore. 1987. The impact of split bond ratings on risk premia. *Financial Review*, 22(1). February:71-85.
- Moon, C-G. and J. G. Stotsky. 1993. Testing the differences between the determinant of Moody's and Standard and Poor's ratings: An application of smooth simulated maximum likelihood estimation. *Journal of Applied Econometrics*, 8(1). Jan-Mar: 81-69.
- Reiter, Sara A. and David A. Ziebart. 1991. Bond yields, ratings, and financial information: Evidence from public utility issues. *Financial Review*, 26(1). February: 45-73.

