不請自來的銀行信用評等有向下偏誤的現象嗎？— 配對方法

黃玉麗・張元・沈中華*

本文根據惠譽的銀行個別評等 (Fitch's Bank Individual Ratings, FIRs), 利用 2002 至 2005 年全球 91 國 3,881 個銀行觀察值爲樣本 (其中 3,433 個觀察值爲付費以取得之信用評等, 448 個觀察值獲得的是不請自來的信用評等), 分析兩類型銀行之信評的相對績效高低來判斷不請自來的銀行信評是否相對低於付費 (主動) 以取得之信用評等。我們應用 Rubin (1973) 與 Rosenbaum and Rubin (1983, 1985a,b) 的配對方法, 透過將兩類型銀行樣本之財務特性修正至近似, 以改善既有文獻中未處理的樣本選擇偏誤問題。我們發現在樣本配對前, 獲得不請自來信評之銀行的許多財務績效指標相對低於付費取得信評之銀行, 這解釋了爲什麼前者之信評相對較低的原因; 在樣本配對後即控制了兩類型銀行的財務特性之後, 大部分的證據顯示不請自來的銀行信評仍然相對較低。因此本文發現, 即使考慮並處理了樣本的選擇偏誤, 不請自來的銀行信評仍然存在向下偏誤的現象。

關鍵詞: 不請自來的銀行信評, 選擇偏誤, 配對方法
JEL 分類代號: C21, G10, G15, G21

1 前言

發行體評等 (issuer rating) 係針對一家企業的整體償債能力進行評比。評

*作者分別爲德明財經科技大學保險金融管理學系副教授、國立彰化師範大學商業教育學系助理教授與國立台灣大學財務金融系教授。作者感謝在 2009 年台灣經濟計量學會上台灣大學財務金融系陳聖賢教授與胡星陽教授以及參與該組會議之諸位先進的建議, 亦感謝兩位匿名評審的修正意見。文中仍有疏漏之處皆爲作者之責。

期刊資料庫

等所產生的排名和成績, 將使投資人更容易了解企業或金融機構的財務狀況是否健全; 不論是和金融機構往來, 或是購買公司股票投資, 都提供一項方便評估的管道。評等又可分爲信評機構接受發行體的委託, 向發行體收取費用, 與發行體進行訪談以取得公開及私有訊息, 繼而給予發行體或發行債務評等等級, 稱之爲主動付費取得之信用評等 (solicited rating); 與之對照者, 當信評機構主動給予發行體或發行債務評等, 亦即發行體並未向信評機構請求給予評等, 而信評機構主動給予發行體評等等級, 且依據該發行體可公開取得的資訊進行評等, 則稱爲不請自來的信用評等 (unsolicited rating)。

企業願意主動付費以獲得國際認證評等組織 (nationally recognized statistical rating organizations, NRSROs) 的評等, 目的是吸引更多國際資本市場投資人的青睞, 由於信用評等在債務市場的重要性, 且有些國家制定法規限制金融機構投資於未受評等或投機等級評等的發行體或發行債務, 使更多企業尋求獲得評等。但仍有些企業不需要國際認證評等組織的評等等級, 例如中國, 由於其管制制度與會計標準與其他國家不同, 且中國的企業多數在國內發行債券籌措資金, 因而較不需要國際認證評等組織的評等, 她們主要向中國境內的評等機構付費以獲得評等, 然而, 當一家企業需要向國際資本市場籌措資金, 或計畫到國外的證券交易所股票上市, 則會比較希望獲得國際認證評等組織給予評等, 以獲得國際的認同。

不請自來的評等一直以來受到許多爭論, 文獻發現不請自來的評等等級出現向下偏誤的現象 (Fairchild et al., 2009; Poon, 2003; Poon and Firth, 2005; Van Roy, 2006), 但巴塞爾委員會及國際清算銀行卻同意使用以不請自來的評等等級, 以評估一家銀行的資本適足比率 (Dale and Thomas, 2000); 這使得債務發行體可能因獲得向下偏誤的不請自來的評等等級, 而須支付較高的債務資金成本, 因而強烈反對信評機構主動給予發行體或發行債務信用評等 (Japan Center for International Finance, 1999)。然而, 信評機構則主張不請自來的評等可以提高信用評等在債券市場的涵蓋比率, 可以提供投資人更多投資選擇所需的資訊。此外, 不請自來的評等可以讓規模較小或成立時間較短的評等機構進入信用評等產業, 進而提高信用評等市場的競爭程度。

臺灣大學圖書館 期刊資料庫

不請自來的評等出現向下偏誤的現象, 有二個可能的原因, 第一、信評機構欲透過給予發行體一較低的不請自來的評等等級, 再透過與受評發行體進行協商, 以使發行體承諾支付評等費用, 以獲得較佳的評等等級。由於給予發行體或發行債務評等等級所收取之評等費用已是信評機構最主要的收入來源,[1] 因此信評機構希望更多的發行體支付評等費用以取得評等以增加收入, 因而有可能給予主動付費獲得評等的企業較高的評等等級。例如, 日本國際金融中心 (Japan Center for International Finance, 1999) 指出評等機構給予亞洲企業不請自來的評等, 是希望達到與企業接觸, 以使企業願意支付費用, 且同意評等機構給予評等的目的。Harington (1997) 指出部分銀行認爲穆迪投資服務所給予的不請自來的評等如同於企業收到「財務黑函」(financial blackmail)。其次, 信評機構給予發行體較低的不請自來的評等等級, 也可能是因爲評等機構採取較保守的態度, 因爲不請自來的評等, 信評機構只能透過可取得的較有限的公開訊息, 使信評機構基於保守原則而給予較差的評等。Gan (2004) 同時使用穆迪投資服務 (Moody's Investor Services) 與標準普爾信評機構 (Standard & Poor's) 的評等爲對象, 發現不請自來的評等有向下偏誤現象, 且主張係因信評機構在給予不請自來的評等時, 可獲得的訊息較有限。Fairchild et al. (2009) 分析穆迪投資服務給予日本企業的不請自來的評等, 發現不請自來的評等較低, 且認爲是信評機構所取得的資訊不完全所造成。Van Roy (2006) 也同意不請自來的評等等級較差係因爲訊息不足。

既存文獻中雖然發現不請自來的評等等級相對低於主動付費取得之評等, 但並不表示不請自來的評等出現向下偏誤的現象, 不請自來的評等較差的理由可能是發行體的自我選擇 (self-selection) 的結果。也就是說, 信用品質較佳的企業傾向選擇主動付費以獲得信用評等, 而信用品質較差的企業選擇不主動請求評等, 使信用品質較差的企業當中, 評等機構主動給予不請自來的評等的比例提高, 事實上, 兩種評等方式均能正確地反應發行體的信用品質。當獲得不請自來評等的企業認爲評等被低估, 且認爲揭露公司內部訊息可以使評等等級獲得改善時, 發行體便會請求評等機構給

[1] SEC (United States Securities and Exchange Commission) (2003) 指出穆迪投資服務與惠譽信評機構的收入中, 90% 以上來自於發行體給付之評等費用。

臺灣大學圖書館 期刊資料庫

予評等, 此時企業的評等等級也將提高。相反地, 當企業認爲不請自來的評等已正確地反應其信用品質時, 就不會請求給予評等, 因此信用品質較差的企業仍將維持相對較低的不請自來的評等等級。例如 Byoun and Shin (2002) 建立一個兩期的理論模型, 假設市場存在好的公司及不好的公司, 且公司管理者與市場參與者間存在不對稱資訊, 以探討評等機構給予發行體不請自來的評等的動機及其產生的影響, 且透過實證分析以驗證此模型的設定。首先, 此模型設定市場存在兩個個別的均衡, 其中只有好的公司會透過付費取得評等來傳遞公司訊息, 且投資人相信信評機構有辨別公司好壞的能力, 但不好的公司則選擇不傳遞公司的內部訊息。其次, 以1996至2001年間標準普爾評等機構主動給予221家發行體的評等進行檢視, 發現實證結果與模型設定一致, 多數不請自來的評等爲投機等級, 而多數付費取得之評等爲投資等級, 且好的企業會透過主動要求評等等級來傳遞企業的價值, 而不佳的企業則選擇不被評等。因此, 不請自來的評等, 一般而言獲得較差的評等等級。

另一個認爲不請自來信評不會出現向下偏誤的理由, 則主張信評機構係透過給予發行體不請自來的評等以建立其聲譽, 當信評機構想進入一個新的市場, 則可透過提供不請自來的評等, 等名聲建立後, 自然會有企業願意付費以取得評等。例如, Ramakrishnan and Thakor (1984) 和 Millon and Thakor (1985) 強調像信評機構等金融中介機構對聲譽十分重視, Cantor and Packer (1994) 也指出信評機構擔心影響其聲譽, 因此會謹慎的給予評等等級。而 Fulghieri et al. (2010) 則建構一動態理性預期模型 (dynamic rational expectations model), 檢視爲何信評機構要給予企業不請自來的評等等級, 及爲何不請自來評等平均而言低於付費取得的評等等級。此模型主張, 信評機構面對的是一種動態抵換關係 (dynamic trade-off), 雖然給予企業較佳的付費取得的評等在短期可以增加信評機構的收入, 但長期而言卻是有成本的, 因爲會提高獲得過高評等的企業, 未來無法償付債務的可能性, 反而將破壞信評機構的名聲, 繼而將使信評機構的價值下降, 也將使未來的收入下降。因此, 信評機構會在短期利益與長期利益下降間取得一平衡。此外, 此模型主張信評制度中不請自來的評等存在的意義與價值。認爲信評機構在企業已經決定是否付費取得評等之後, 會透過給予不

臺灣大學圖書館 期刊資料庫

願意付費的企業一較差不請自來的評等, 藉以威脅若企業拒絕付費取得評等, 則將獲得較差的不請自來的評等, 企圖收到更多的評等收入。其次, 透過給予不請自來的評等, 信評機構可以對投資人宣稱, 其堅持只給予較佳企業付費以取得的評等, 也可以改善信評機構的名聲。而信評機構會在評等收入增加與名聲利益之間取得一最適的評等組合。因此在均衡時, 不請自來的評等與付費取得的評等將同時存在, 且付費取得的評等均是較佳的, 而不請自來的評等較低, 但此結果並非不請自來的評等等級出現向下偏誤的現象, 而是反映了不願付費取得的評等本身品質較差的現象。

實證結果則出現不一致的現象, 如 Poon (2003) 以1998至2000年間之15國265家獲得標準普爾信評機構評等之不同產業的公司爲對象。發現付費以取得評等的公司在流動性、獲利性、槓桿程度及財務彈性上表現均優於獲得不請自來的評等的公司。他亦發現在控制了發行體的財務績效與自我選擇偏誤後, 不請自來的評等仍較主動付費以取得之評等等級爲低, 表示信評機構確實給予不請自來的發行體較差的評等。Poon and Firth (2005) 則針對82國的951家銀行的評等進行研究, 以 Heckman (1979) 兩階段方法 (two-stage method) 降低自我選擇偏誤的可能影響, 發現在控制了財務特徵和績效後, 獲得不請自來評等的企業其信用評等仍較差。Poon et al. (2009) 則將評等等級差異拆解成兩部分, 一爲在評等方式固定下財務績效所造成的差異、二是在財務績效固定下評等方式變動所造成的效果。結果發現不請自來的評等等級與主動付費以獲得之評等等級間的差異, 可以由財務績效的差異及評等方式的差異來解釋, 且部分的結果顯示評等方式的差異對銀行評等等級的解釋大於財務績效對銀行評等等級的解釋。Shimoda and Kawai (2007) 與過去看法一致, 發現不請自來的評等等級較低, 但差異的程度已逐漸減低, 平均而言少於一個評等等級的差異, 原因是公司訊息揭露程度的改善。

Butler and Rodgers (2003) 檢視1997年的360家非金融公司的債券評等, 且以發行體支付信評機構費用的高低, 判斷是付費取得信評或是不請自來的評等, 結果並未發現信評機構與債券發行體之前存在利益衝突, 亦即信評機構並未透過給予付費取得評等的發行體較高的評等等級, 以增加收入。Baker and Mansi (2002) 透過訪問債券發行體與投資人, 發現超過

臺灣大學學術期刊資料庫

半數 (58.3%) 的受訪者認爲不請自來的評等與主動付費取得之評等等級一樣精確。此外, 調查信評機構是否會爲了增加收入而給予主動付費取得評等之發行體較高的評等等級, 97.3%的受訪者認爲信評機構爲了維持其名聲, 以及考慮到未來的收入, 會維持一高品質且精確的評等等級。

由於銀行業在經濟體系中扮演重要的角色, 且銀行倒閉對一國整體經濟環境造成的影響很大, 且傳遞效果 (contagion effect) 在銀行業尤其明顯, 因此本文以惠譽信評機構 (Fitch Rating Company) 所給予全球各國的銀行個別評等 (Fitch's Bank Individual Ratings, FIRs) 等級爲研究對象。[2] 惠譽信評機構所提供的銀行個別評等服務是針對銀行的曝險程度及對風險的管理以衡量發生違約的可能性, 將銀行評等分爲 A, A/B, B, B/C, C, C/D, D, D/E, E 九個評等等級。個別評等符號之後附加一個英文字母 "s", 意指該評等雖然參酌了部份得自受評機構之資訊, 但主要仍以公開資訊做爲參考基準。[3]

過去文獻在分析不請自來的信評是否出現向下偏誤現象, 大部分的做法是比較兩組樣本評等的差異性, 也就是分析不請自來的信評與主動付費取得之信評的相對績效高低。然而, 在評估兩組樣本之評等差異時, 除了考慮評等的性質是不請自來或是主動付費取得評等的差異外, 若這兩組樣本的其他特性變數亦存在極大差異, 例如兩類型銀行樣本在規模或財務變數等特性變數上明顯不同時 (例如獲得不請自來信評之銀行的平均規模較小), 則兩組樣本的評等差異, 將無法歸因於不請自來此一因素, 也許是規模效果所造成。此外, 部分文獻在評估銀行信評的迴歸式中, 以設置虛擬變數 (不請自來者爲1, 否則爲0) 的方法欲了解不請自來評等是否有向下偏誤的現象, 但可能因該虛擬變數不具外生性 (具某些特性的銀行容易接受不請自來的信評), 使虛擬變數的估計係數產生偏誤, 即產生所謂的樣本選擇偏誤 (sample selection bias)。本文希望兩類型銀行樣本, 除了在信評

[2]惠譽 (Fitch IBCA) 的前身是成立於1913年的 Fitch Publishing Company, 於1997年與英國 IBCA 合併成爲現今的 Fitch IBCA。由於標準普爾與穆迪兩大公司全球市佔率逼近八成, 爲能獲得進一步市佔率, 分別在2000年合併 Duff and Phelps Credit Rating Co., 與在2001年合併 Thomson BankWatch, 一家國際間最大的銀行專業評等機構。

[3]個別評等旨在對一家銀行在完全獨立且無法倚賴外部支援情況下的信用狀況進行評估。個別評等乃在評估銀行的風險部位、風險偏好、及風險管理能力, 因此係代表本公司針對該銀行是否可能陷入重大經營困境, 以致必須獲得支援所表達之意見。

期刊資料庫

上區分爲不請自來與主動付費取得評等兩類型外, 其他的特性變數要相等或接近, 以達到所謂其他條件皆相同 (other things being equal) 的狀況, 方可降低或避免樣本的選擇偏誤問題。[4]

Heckman (1979) 建議使用兩階段方法 (two-stage method) 降低自我選擇偏誤。以本文來說, 第一步是利用 Probit 或 Logit 模型估計一機率函數, 即銀行是否獲得不請自來信評的決定方程式, 得到選擇偏誤調整項, 即 Inverse Mill's Ratio, 第二步則是將選擇偏誤調整項代入評估不請自來此一因素對信評影響的順序多重分類模型 (ordered probit model) 中成爲其中一個解釋變數, 如此可降低選擇偏誤的問題。然而 Hofler et al. (2004) 指出此方法的困難性。其一是必須滿足認定的要求, 即必須找到至少一個變數包括在機率模型中, 但卻不包括在績效評估的迴歸式中; 第二則是 LaLonde (1986) 的批評, 即研究中所要評估的效果 (在本文中爲不請自來此一因素對銀行信評的影響) 會因爲模型設定的不同而有極大的差異。

針對選擇偏誤, Rubin (1973) 提出配對方法 (matching method), 其基本概念是針對每個獲得不請自來評等的銀行樣本, 由所有獲得主動付費取得評等的銀行樣本中, 根據特性變數是否相同爲依據, 選擇一個或數個樣本作爲配對樣本。完成樣本配對後, 兩組樣本在特性變數上的差異將大幅減少, 甚或消失, 此時兩組樣本的相對評等高低, 將可歸因於不請自來這一因素的影響。然而, 若特性變數相當多, Rubin (1973) 的方法可能使能符合多重特性變數均需相等之標準的樣本數目大幅降低。[5]

Rosenbaum and Rubin (1983, 1985a,b) 提出傾向分數配對法 (propensity score matching method; PSM), 其基本概念是在配對過程中, 將多個特性變數對應到一個維度, 即首先根據多個特性變數估計一機率函數, 也就是銀行獲得不請自來信評的機率函數, 稱傾向分數函數 (propensity score function; PSF; 成爲試驗樣本的機率函數), 將個別樣本的特性變數代入機率函數, 將可得到該樣本的估計機率 [成爲試驗樣本的估計機率, 稱傾向分

[4]在生物與醫學研究中, 例如試驗人體用藥的治療效果, 通常是在兩群特性 (如生病時間、生病程度與持續時間、身高、體重、性別、年齡、家庭所得與工作時間) 相近的人中, 一組給予用藥 (試驗組), 另一組則無 (對照組), 爾後再透過兩組樣本病情痊癒的相對差異是否顯著來認定該藥的治療是否有效。

[5]請參考 Heckman et al. (1998, 1997)、Heckman and Robb (1985, 1986)。

臺灣大學學術期刊資料庫

數 (propensity score; PS)]。接著針對每一個獲得不請自來信評的銀行樣本, 由所有主動付費取得評等的銀行樣本中, 尋找傾向分數相同的樣本爲配對樣本。這方法將同時改善選擇偏誤與樣本配對時過多維度的問題。[6]

本文利用2002至2005年間, 來自全球91國獲得惠譽的銀行個別評等之銀行爲樣本, 共有3,881個觀察值, 其中3,433個觀察值爲付費取得之信用評等的銀行, 448個觀察值爲獲得不請自來的評等的銀行, 分析兩類型銀行信評的相對績效高低, 以判斷不請自來的銀行信評是否相對低於主動付費以取得之信用評等。同時我們應用 Rubin (1973) 以及 Rosenbaum and Rubin (1983, 1985a,b) 的配對方法, 以修正既有文獻中未處理的選擇偏誤問題。爲什麼我們不直接評估銀行主動付費以獲得評等之前與之後的評等差異性?[7] 原因之一是銀行在付費以獲取評等之前與之後的經營體質與經營環境存在差異且無法進行控制; 原因之二是文獻中方案評估 (program evaluation) 或政策效果分析 (policy impact analysis) 的做法, 皆是利用參與方案 (在此爲不請自來的評等) 與未參與方案 (付費以獲得評等) 的個體進行評估, 例如 Heckman et al. (1998) 以及一些先前在附註中提到關於配對方法的衆多應用研究, 我們的研究架構與這些研究相同。

本文共分5節, 下一節爲配對理論與傾向分數配對法的介紹, 第3節介紹計量模型, 第4節爲實證結果, 最終做成結論。

---

[6]Rubin and Thomas (1992) 證明傾向分數配對法可以有效解決選擇偏誤的問題。此方法在經濟領域的應用上相當多, 請參考 Persson (2001)、Hutchison (2004)、Glick et al. (2006)、Vega and Winkelried (2005)、Crowe (2006)、Ham et al. (2004) 以及 Shen and Chang (2008)。

[7]在448個獲得不請自來的銀行信評樣本中, 下一年度轉爲以付費方式取得評等的銀行共有25個樣本, 其中16個樣本獲得的評等等級不變, 8個樣本獲得較佳的評等等級, 付費後獲得較差的評等的銀行, 則僅有1個樣本。舉例來說, 菲律賓發展銀行 (Development Bank of the Philippines) 在2002–2003年間獲得不請自來的信評, 信評爲 C/D, 在2004–2005轉爲付費取得信評, 信評仍爲 C/D; 中國銀行 (Bank of China) 在2002–2004年間獲得不請自來的信評, 在2005即轉爲付費取得信評; 中國銀行原獲得之不請自來的信評爲 E(1分), 付費後取得之信評爲 D/E(2分); 台灣的復華銀行 (Fuhwa Bank) 在2003–2004年間獲得不請自來的信評, 在2005即轉爲付費取得信評; 復華銀行獲得不請自來的銀行信評爲 C/D, 付費取得之評等亦爲 C/D, 皆爲4分。相反地, 原以付費方式取得評等, 而撤銷評等後仍獲得不請自來的評等等級的銀行則有8個樣本, 其中6個樣本的評等等級不變, 2個樣本的評等等級被調降, 未有銀行獲得評等等級調升。

期刊資料庫

2 配對理論與傾向分數配對法

2.1 配對理論的原理

本文主旨是應用配對方法篩選樣本, 並估計銀行獲得之不請自來的信用評等, 是否低於主動付費獲得之評等, 亦即探討不請自來評等是否出現向下偏誤的現象。我們參考 Dehejia and Wahba (2002), 定義 Y_{i1} 與 Y_{i0} 爲 i 銀行獲得之不請自來評等與主動付費取得的評等。在這兩種狀態下的評等差異爲: $\Delta_i = Y_{i1} - Y_{i0}$。給定銀行未付費取得信評, 相對於其獲得主動付費評等時的評等差異則爲:[8]

$$\Delta|_{D=1} = E\left(Y_1 - Y_0|D=1\right) = E\left(Y_1|D=1\right) - E\left(Y_0|D=1\right), \quad (1)$$

其中 D 爲虛擬變數, 若銀行信評爲不請自來時等於1, 否則爲0。關於 (1) 式的估計事實上是不可行的, 因爲 (1) 式的第二項 $E(Y_0|D=1)$ 實際上無法被觀察到, 也就是說, 我們無法得到某銀行實際上未付費, 得到的卻是主動付費而取得的評等。一般的方法是使用 $E(Y_0|D=0)$ 來替換 $E(Y_0|D=1)$; 然而, 只有當 Rubin (1973) 與 Rosenbaum and Rubin (1983) 的條件獨立假設 (conditional independence assumption; CIA) 成立時, 這個替換才是適當的:

$$Y_1,\ Y_0 \perp D|X, \quad (2)$$

其中 $\perp$ 表示獨立, X 爲不受銀行信評取得方式是否爲不請自來所影響的特性變數向量。這個假設指的是給定某些特性變數 X, 實驗樣本組 (即不請自來的銀行信評樣本) 與控制樣本組 (即主動付費取得的銀行信評樣本) 可以視爲由同一個母體所抽出, 也就是樣本被指定到實驗組或控制組的過程爲隨機, 兩組樣本在特性變數 X 上, 並沒有明顯的差異性, 因此兩組樣本在參與實驗後的差異性將足以被認定爲實驗效果。當這個條件成立時, $E(Y_0|D=0)$ 成爲 $E(Y_0|D=1)$ 的不偏估計式, 因此,

[8]以下在不會產生困擾的情況下, 將省略下標 i。

臺灣大學學術期刊資料庫

$$\begin{aligned}\Delta|_{D=1} &= E\left(Y_1|D=1\right) - E\left(Y_0|D=1\right)\\ &= E_x\left[\{E\left(Y_1|X, D=1\right) - E\left(Y_0|X, D=1\right)\}|D=1\right]\\ &= E_x\left[\{E\left(Y_1|X, D=1\right) - E\left(Y_0|X, D=0\right)\}|D=1\right]\\ &= E_x\left[\Delta^d|_{D=1,X}|D=1\right]。\end{aligned} \tag{3}$$

相反地, 若上述條件不成立時, $E(Y_0|D=0)$ 將無法成爲 $E(Y_0|D=1)$ 的良好替代估計式, 兩者之間的差異即爲選擇偏誤。

根據 Heckman et al. (1998), 上述的條件獨立假設不成立的原因有二。其一爲非比較性偏誤 (noncomparability bias), 即實驗組與控制組的特性變數無法相互比較。第二則是分配偏誤 (distributional bias), 即實驗組與控制組在特性變數 X 上的分配不相同, 使兩群樣本像是從不同的母體中所抽出 (或是樣本並非是隨機地被指定到實驗組或控制組)。本文的狀況與第二個原因相近。

2.2 配對方法

本文應用配對方法, 透過將獲得不請自來評等的銀行, 與特性變數相類似但以付費取得信評的銀行樣本進行配對, 之後再進行兩類型銀行評等高低之比較。被選爲配對樣本的銀行, 可以視爲獲得不請自來信評之銀行的反事實 (counterfactual) 樣本, 即特性變數與獲得不請自來信評之銀行極爲相似的銀行 (其獲得的是主動付費取得之信評)。透過樣本配對, 兩組樣本的特性變數將趨向一致, 滿足條件獨立假設。

2.2.1 傾向分數配對

當特性變數增多時, 爲了要獲得與不請自來信評的銀行的多個特性變數皆相近之主動付費獲取信評的銀行樣本, 將變得困難。爲避免此問題, Rosenbaum and Rubin (1983) 提出傾向分數配對法, 將樣本配對時的多維度對應到一個維度。其做法是給定多個特性變數 X, 透過機率函數 (傾向分數函數) 的估計, 而得到一機率 (傾向分數):

$$P(X) = P(D=1|X) = E(D=1|X)。 \tag{4}$$

臺灣大學學術期刊資料庫

上式 $P(X)$ 表示在給定特性變數爲 X 下, 銀行獲得不請自來信評的機率。給定獲得不請自來信評與主動付費取得信評之銀行有相同之獲得不請自來信評的機率 (相同的傾向分數), 條件獨立假設將成立, 兩銀行樣本視爲由同一個母體中所抽出, 即

$$Y_1, Y_0 \perp D|X \Rightarrow Y_1, Y_0 \perp D|P(X), \tag{5}$$

而 (1) 式變爲

$$\Delta|_{D=1} = E\left(Y_1 - Y_0 \middle| D = 1\right) = E\left(Y_1 - Y_0 \middle| P(X), D = 1\right)。 \tag{6}$$

上式表示透過選取相同傾向分數之主動付費取得信評的銀行作爲獲得不請自來信評之銀行的反事實樣本, 則可將兩樣本的評等差異歸因於不請自來這一因素。

根據 Dehejia and Wahba (2002), 傾向分數配對的過程可分爲兩步驟, 步驟一是透過簡約的 Logit 或 Probit 模型估計, 以得到所有銀行的傾向分數 $P(D = 1|X)$, 而獲得不請自來信評與主動付費取得信評的銀行樣本分別以 P_i 與 P_j 來標示。步驟二則是針對每個獲得不請自來信評的銀行樣本, 在主動付費取得信評的銀行樣本中尋找傾向分數相同的樣本做爲配對樣本。透過分析配對後的兩組樣本, 將可得到 (1) 式的不偏估計。一般來說, 事實上並不容易找到傾向分數完全相同的反事實樣本, 因此通常將相同修正爲相近。配對方法的應用研究中較常使用的第一個準則是 Nearest-Neighbor Matching (往後簡稱 Nearest), 針對每個獲得不請自來信評的銀行, 從主動付費取得信評的銀行樣本中尋找傾向分數差距最小者爲配對樣本, 即

$$C\left(P_i\right) = \min_j \left|P_i - P_j\right|, \tag{7}$$

其中 $C(P_i)$ 爲經由配對過程選擇與獲得不請自來信評之銀行樣本 i 的傾向分數最相近的主動付費取得信評的銀行樣本集合, 在 Nearest 配對下, 此集合只包含一個樣本。此方法採取抽出放回, 同一控制樣本將可能會重複被選配到不同的試驗樣本。

第二個準則爲 Caliper Matching (往後簡稱 Caliper), 它要求獲得不請自來信評與主動付費取得信評的銀行樣本之間的傾向分數差距不能太大。

臺灣大學學術期刊資料庫

換句話說, 當某一獲得不請自來信評之銀行的傾向分數與主動付費取得信評的銀行之傾向分數的差異小於 η 時, 即 $|P_i - P_j| < \eta$, 將被選入 $C(P_i)$ 的集合內。根據 Shen and Chang (2008), η 通常選取爲所有樣本之傾向分數估計值標準差的1/4, 而 $C(P_i)$ 包含的樣本數則可能是1個以上。

2.2.2 以特性變數做樣本配對

除了利用傾向分數的相似性做樣本配對外, 我們也可以直接利用特性變數做配對, 本文的第三個準則爲 Mahalanobis Metric Matching (往後簡稱 Mahala)。首先定義獲得不請自來信評之銀行樣本 i 與主動付費取得信評之銀行樣本 j 之間的馬氏距離 (Mahalanobis Distance):

$$d(i, j) = \left(X_i - X_j\right)' V^{-1} \left(X_i - X_j\right), \tag{8}$$

其中 X_i 與 X_j 分別爲樣本 i 與樣本 j 的 $k \times 1$ 特性變數向量, 而 k 爲特性變數個數; V 爲特性變數之 $k \times k$ 階的變異數 — 共變異數矩陣。此方法的步驟爲, 首先計算所有樣本 i 與所有樣本 j 之間的 d, 接下來針對每個樣本 i, 從樣本 j 中挑選出 d 最小的樣本作爲配對樣本。

第四個準則爲 Mahalanobis Metric Matching with Calipers (往後簡稱 Mahala Caliper)。當樣本 i 與樣本 j 的 d 小於 δ 時, 即 $d(i, j) < \delta$, 則樣本 j 將被選爲樣本 i 的配對樣本。η 指定爲所有樣本之傾向分數估計值標準差的1/4。

2.3 驗證樣本配對的結果

當樣本配對完成後, 本文使用下列兩種方式驗證配對結果是否令人滿意, 主要在於分析樣本配對後, 兩組樣本間之特性變數差異是否仍然存在。方式一是透過各別特性變數之母體平均數的 t 檢定, 即

$$H_0\colon \left(\mu_i - \mu_j\right) = 0, \tag{9}$$

其中 μ_i 與 μ_j 分別爲獲得不請自來信評之銀行樣本 i 與主動付費取得信評之銀行樣本 j 的某個特性變數的平均數。若檢定結果無法拒絕兩組樣本間平均數的差異爲 0, 則表示在樣本配對後, 兩樣本組在該特性變數上

臺灣大學學術期刊資料庫

的平均數是相近的, 亦表示就平均數而言, 樣本配對有效降低了兩組樣本在該特性變數上的差異。[9]

第二種驗證方式是計算

$$\frac{\left(\bar{X}_{iB}-\bar{X}_{jB}\right)-\left(\bar{X}_{iA}-\bar{X}_{jA}\right)}{\left(\bar{X}_{iB}-\bar{X}_{jB}\right)}\times 100, \tag{10}$$

其中 $\bar{X}_{iB}$ 與 $\bar{X}_{jB}$ 分別表示在樣本配對前, 獲得不請自來信評之銀行樣本 i 與主動付費取得信評之銀行樣本 j 之某個特性變數的平均數, $\bar{X}_{iA}$ 與 $\bar{X}_{jA}$ 則分別代表在樣本配對後, 樣本 i 與樣本 j 之某個特性變數的平均數。這個公式計算出透過樣本配對, 某個特性變數在兩組樣本之間差異性的變動百分比。當此百分比爲正且愈大, 表示兩組樣本之某個特性變數的差異性降低, 即配對愈有效。

3 計量模型

3.1 傾向分數函數

根據所有的樣本, 考慮影響銀行獲得不請自來之信評或主動付費獲取信評的特性變數 X, 以 Probit 模型估計傾向分數函數, 即銀行獲得不請自來之信評的機率函數,

$$P(D=1)=F(\beta' X), \tag{11}$$

上式中之 β 爲影響係數。若 $F(\cdot)$ 是標準常態累積機率函數, 則爲 Probit 模型, 而:

$$X=[1, \text{COVERAGE}, \text{CAPITAL}, \text{ROA}, \text{COST}, \text{LIQUID}, \text{ASSET}], \tag{12}$$

其中 COVERAGE 爲資產品質指標, 以呆帳覆蓋率衡量, 亦即備抵呆帳佔逾期放款比率, 此比率愈高, 代表銀行應付逾期放款的能力愈強, 預期此比率愈高, 愈有利於信用評等。CAPITAL 爲資本適足性指標, 以自有資本佔風險性資產之比率來衡量, 資本適足性指標適合衡量一銀行資本的安全性, 此比率愈高表示銀行自償的能力愈強, 應付非預期債務求償的能力也

[9]既有配對理論文獻中尚未討論特性變數在二階動差以上的配對要如何進行。

臺灣大學學術期刊資料庫

愈強, 發生違約的機率也能降低, 預期此比率愈高, 銀行可獲得愈高的信用評等。ROA 爲獲利性指標, 以資產報酬率 (淨利除以總資產) 來衡量, 預期愈高的獲利性, 銀行違約的可能性愈低, 可獲得較佳的評等。COST 爲無效率性指標, 以成本佔收入比率衡量, 此比率愈高, 表示銀行營運的效率性愈低, 預期將對評等產生不利的影響。LIQUID 爲流動性指標, 以流動資產佔存款及短期資金來源之比率來衡量, 此比率愈高, 銀行握有的變現性資產愈高, 愈能應付存款的突然減少與放款的需要, 可降低發生支付不能或無法滿足客戶信用需求的窘境, 預期流動性愈高, 銀行可獲得愈高的信用等級。ASSET 爲銀行規模, 將總資產取自然對數表示, 一般認爲銀行的規模愈大, 分散資產風險的可能性也愈大, 而投資人可能相信主管當局較不願意讓大銀行倒閉, 因其倒閉所付出的成本較高。UBS Investment Bank (UBS) (2004) 指出銀行規模愈大, 可獲得愈佳的信用評等, 銀行規模與信用評等的相關性最高, 且銀行規模可以反應企業之地理位置、產品市場多角化、競爭能力、市場份額及品牌等質性資料。

信用評等機構在給予評等時通常以一較長期間的觀點來評估, 因爲理想的評等應該與景氣循環互相獨立 (rating through the cycle), 一般常採過去三到五年的平均數來衡量公司的財務績效, 本文亦依循此做法, 且以各財務比率過去3年平均數衡量。即2005年的信用評等等級的決定, 是以2004、2003及2002年的財務報表資料來解釋, 以排除景氣循環的影響。

3.2 不請自來的銀行信評向下偏誤的估計

本文以兩種方式評估銀行所獲得的不請自來的信用評等相對於主動付費取得之評等是否有向下偏誤的現象。首先是基本統計量檢定 (t 檢定), 透過計算獲得不請自來評等與主動付費取得之評等的平均評等差異, 以跋靴法 (Bootstrap Method) 建立信賴區間以檢定差異的顯著性。如果平均評等之差異顯著爲正, 表示平均來說, 不請自來的信用評等其績效相對較佳, 不請自來的評等並未出現向下偏誤的現象, 而是向上偏誤; 若平均評等之差異顯著爲負, 表示不請自來的信用評等較差, 主動付費取得之評等較佳, 亦即不請自來的評等存在向下偏誤的現象。

第二種方式爲以「獲得不請自來之信評的銀行」與「主動付費獲得評等

臺灣大學學術期刊資料庫

的銀行」爲樣本, 使用順序多重分類模型 (ordered probit model) 進行分析, 且將惠譽信評機構給予全球商業銀行的個別評等轉換成數字評等, 惠譽信評機構所提供的銀行個別評等服務是針對銀行的曝險程度及對風險的管理以衡量發生違約的可能性, 將銀行評等分爲 A, A/B, B, B/C, C, C/D, D, D/E, E 九個評等等級。A 表示受評銀行就獨立基礎而言的信用狀況極強。B 表示受評銀行就獨立基礎而言的信用狀況頗強。該銀行並無令人關切的重大問題。C 表示受評銀行就獨立基礎而言的信用狀況允當, 但卻有一或多處的問題。D 表示受評銀行面對來自內部及/或外部的劣勢。E 表示受評銀行問題嚴重, 需要或可能會需要外部支援。此外, 個別評等等級包括上列五個等級中間值: 即 A/B、B/C、C/D 及 D/E, 表示受評銀行的信用介於兩個評等等級之間。當評等爲 A, 則以數字 9 表示, 而評等 A/B 則爲 8, 評等 B 爲 7, 評等 B/C 爲 6, 評等 C 爲 5, 評等 C/D 爲 4, 評等 D 爲 3, 評等 D/E 爲 2, 評等 E 爲 1, 共分成九個等級, 數字愈大代表銀行評等愈好。[10] 順序多重分類模型假設機率分配函數爲常態機率分配, 並透過最大概似法來估計迴歸參數與數個組別的分界點 (cutoff points), 依照各樣本點所落入組界區間以決定樣本所屬組別, 因爲等級爲順序尺度 (ordinal scale), 故組別間存在優劣或順序的關係。

因變數 RATING 是惠譽信評機構針對銀行所給予可觀察到的實際信用評等等級, RATING* 是信評機構針對銀行所給予無法觀察到的連續變數以代表此銀行之信用, 且假設可觀察到的評等等級可以代表評等機構對發行體的信用所給予的無法觀察的連續變數的分類, 模型設定如下:

$$\begin{aligned} \text{RATING}_i^* = {} & \beta_0 + \beta_1 \text{UNSOLICIT_}D_i + \beta_2 \text{SOVRATING}_i \\ & + \beta_3 \text{COVERAGE}_i + \beta_4 \text{CAPITAL}_i + \beta_5 \text{ROA}_i \\ & + \beta_6 \text{COST}_i + \beta_7 \text{LIQUID}_i + \beta_8 \text{ASSET}_i + \varepsilon_i, \end{aligned} \tag{13}$$

[10] 信用評等的高低不僅能提高銀行或公司債務工具的銷售能力與談判地位, 亦有助於投資人作出其投資抉擇。不論是大型機構投資人 (如銀行、保險與投信公司), 或是散戶投資皆可利用市場上常見之信用評等公司的信用評等來評估其投資組合的風險以及其所相對應的要求報酬率。部分大型投資機構更因政府監管機構規定或內部風控機制, 禁止購買未經評等或低於某評等的投資工具。對受評機構來說, 信用評等可協助其建立企業聲譽, 用於競價投標、資金借貸、與上下游廠商議價、或作爲與其他公司進行策略聯盟談判的籌碼。作著感謝評審對此部分的補充建議。

期刊資料庫

$$\begin{aligned} \text{RATING} &= 1, \quad \text{if} \quad \text{RATING}^* < \mu_1; \\ &= 2, \quad \text{if} \quad \mu_1 \leq \text{RATING}^* < \mu_2; \\ &\cdots\cdots \\ &= 9, \quad \text{if} \quad \mu_8 \leq \text{RATING}^*。 \end{aligned}$$

樣本期間由2002年至2005年, 共有來自於全球91國的3,881個觀察值爲樣本, 爲解釋變數的係數, ε 爲標準常態隨機誤差項, 平均數爲0, 變異數爲1。$\mu_i (i = 1, \cdots, 8)$ 則爲門檻參數值, 其中 μ_1 代表評等 E 與評等 D/E 的門檻值; 而 μ_2 則代表評等 D/E 與評等 D 的門檻值, 以此類推。當信評機構主動給予銀行評等等級, 則設定虛擬變數 UNSOLICIT_D 爲1, 其他則爲0。若 β_1 顯著爲正, 表示不請自來的信用評等其績效相對較佳, 並未出現向下偏誤的現象, 而是向上偏誤; 若 β_1 顯著爲負, 表示平均而言, 不請自來的信用評等較差, 主動付費取得之評等較差較佳。不請自來的信用評等出現向下偏誤的現象。

在控制變數部分, 本文參考過去探討銀行信用評等決定的文獻 Poon and Firth (2005); Poon et al. (1999, 2009); Rojas-Suarez (2001) 納入國家主權評等、5個財務比率以及銀行規模作爲控制變數。文獻指出主權評等對發行體評等有顯著影響, Borensztein et al. (2006) 發現1997年放寬國家主權上限 (sovereign ceiling) 政策執行後, 很少銀行可獲得較國家主權評等更好的評等, 實證結果指出國家主權上限的效果仍高度顯著, 且發現國家主權評等對銀行業的影響更甚於非銀行業。而探討主權評等決定的文獻中, Cantor and Packer (1996) 發現8個變數可以解釋90%以上的穆迪及標準普爾信評機構的主權評等變異, 分別爲每人國民所得、國民生產毛額成長率、通貨膨脹率、財政平衡、經常帳餘額、負債佔出口比率、是否爲先進國家的指標, 以及自1970年以來該國政府是否曾經發生違約。Borio and Packer (2004) 則使用1996年至2003年間52個國家爲對象發現, 平均每人國民所得、通貨膨脹率及經濟成長率對主權評等的影響十分顯著。質的變數則以政治風險的代理變數及貪污指標具有顯著影響。過去的違約歷史及通貨膨脹率也可解釋主權評等。而公共債務及外債的比率則對新興市場國家的主權評等影響較爲明顯。因此本文納入主權評等以控制各國總體經濟

臺灣大學圖書館 期刊資料庫

情況的差異。[11]

國家主權評等 SOVRATING 與發行體評等相似, 將字母評等轉換成數字評等, 當主權評等爲 AAA 則以數字9表示, 而評等 AA 則爲8, 評等 A 爲7, 評等 BBB 爲6, 評等 BB 爲5, 評等 B 爲4, 評等 CCC 爲3, 評等 CC 爲2, 評等 C 爲1, 共分成九個等級, 數字愈大代表主權評等愈好。5個財務比率以過去3年的平均數表示, 以排除景氣循環的影響, 包括爲資產品質指標 COVERAGE; 資本適足性指標 CAPITAL; 資產報酬率 ROA; 無效率性指標 COST; 流動性指標 LIQUID, 此外, ASSET 爲銀行規模。其定義與前述相同 (變數的定義請參考表1)。本文中的長期信用評等、財務資訊、國家主權評等, 資料均來自惠譽 Fitch IBCA 的 Bankscope 資料庫。

4 實證結果

4.1 樣本配對前的敘述統計分析

表2報告本文的樣本資料, 共有來自91國的3,881個銀行的觀察值, 其中3,433個觀察值爲2002至2005年間付費以取得之信用評等的銀行家數, 448個觀察值爲獲得不請自來的評等的銀行家數。發現與過去文獻一致,

[11] Cantor and Packer (1996) 發現八個變數可以解釋90%以上的穆迪及標準普爾信評機構的主權評等變異, 他們分別爲每人國民所得、國民生產毛額成長率、通貨膨脹率、財政平衡、經常帳餘額、負債佔出口比率、是否爲先進國家的指標, 以及自1970年以來該國政府是否曾經發生違約。Borio and Packer (2004) 則使用1996年至2003年間52個國家爲對象發現, 平均每人國民所得、通貨膨脹率及經濟成長率對主權評等的影響十分顯著; 質的變數則以政治風險的代理變數及貪污指標具有顯著影響; 過去的違約歷史及通貨膨脹率也可解釋主權評等; 而公共債務及外債的比率則對新興市場國家的主權評等影響較爲明顯。Bissoondoyal-Bheenick (2005) 則以量化資料分析標準普爾及穆迪信評機構的主權評等決定發現, 每人實質國民所得及通貨膨脹率在不同信評機構、不同年份, 均是影響主權評等最重要的兩個經濟變數, 而其他總體經濟變數則有不同的效果。因此, 本文納入國家主權評等應該可以補捉影響銀行信評高低的國家因素。跨國信用評等決定的文獻, 也多納入國家主權評等爲控制變數, 如 Borensztein et al. (2006) 發現1997年放寬國家主權上限 (Sovereign Ceiling) 政策執行後, 很少銀行可獲得較國家主權評等更好的評等, 實證結果指出國家主權上限的效果仍高度顯著, 且發現國家主權評等對銀行業的影響更甚於非銀行業。因此現有文獻均以國家主權評等等級來控制國家因素對發行體評等的影響。例如 Poon (2003)、Poon and Firth (2005)、Poon et al. (2009), 均以主權評等控制國家的影響。

此部分作者感謝匿名評審的提醒與修正建議。

期刊資料庫

表 1: 變數定義

變數	定義
是否爲不請自來信評之虛擬變數 (UNSOLICIT_D)	當評等機構主動給予銀行評等等級, 則設定虛擬變數爲 1, 其他則爲 0。
銀行評等 (RATING)	惠譽信評機構給予全球商業銀行的長期信用評等, 當評等爲 A 則以數字 9 表示, 而評等 A/B 則爲 8, 評等 B 爲 7, 評等 B/C 爲 6, 評等 C 爲 5, 評等 C/D 爲 4, 評等 D 爲 3, 評等 D/E 爲 2, 評等 E 爲 1, 共分成九個等級, 數字愈大代表銀行評等愈好。
國家主權評等 (SOVRATING)	當主權評等爲 AAA 則以數字 9 表示, 而評等 AA 則爲 8, 評等 A 爲 7, 評等 BBB 爲 6, 評等 BB 爲 5, 評等 B 爲 4, 評等 CCC 爲 3, 評等 CC 爲 2, 評等 C 爲 1, 共分成九個等級, 數字愈大代表主權評等愈好。
覆蓋比率 (COVERAGE)	過去 3 年備抵呆帳佔逾期放款比率的平均數
資本比率 (CAPITAL)	過去 3 年自有資本佔風險性資產比率的平均數
資產報酬率 (ROA)	過去 3 年淨收入佔總資產比率的平均數
成本對收入比率 (COST)	過去 3 年成本佔收入比率的平均數
流動比率 (LIQUID)	過去 3 年流動資產佔存款及短期資金來源比率的平均數
資產總額 (ASSET)	過去 3 年總資產取自然對數值的平均數

說明: 本文中的長期信用評等、財務資訊、國家主權評等, 資料均來自惠譽 Fitch IBCA 的 Bankscope 資料庫。

獲得不請自來評等的銀行主要爲亞洲地區國家, 包括柬埔寨、中國、香港、印度、印尼、南韓、澳門、馬來西亞、巴基斯坦、菲律賓、斯里蘭卡、台灣及越南。

表 3 報告不請自來的評等等級與付費請求評等的評等等級的分布情況。發現付費請求評等等級最多出現爲 B 評等等級, 共有 886 個樣本資料獲得 B, 其次爲評等 B/C, 再其次爲 C, 此三個評等等級共佔 60% 的樣本。而獲得不請自來的評等等級的銀行, 評等等級最多出現在 D/E 以及 E, 分別有 123 及 101 個觀察值。再其次爲評等等級 D, 有 99 個觀察值。當僅由評等

臺灣大學學術期刊資料庫

表 2: 樣本銀行在各國分配

	主動付費取得信評之銀行家數	獲得不請自來信評之銀行家數	總和		主動付費取得信評之銀行家數	獲得不請自來信評之銀行家數	總和
阿爾巴尼亞	1	0	1	南韓	36	12	48
安道爾	11	0	11	科威特	22	0	22
阿根廷	15	0	15	拉脫維亞	7	0	7
澳洲	43	0	43	黎巴嫩	9	0	9
奧地利	7	0	7	立陶宛	19	0	19
亞塞拜然	4	0	4	盧森堡	9	0	9
巴林	24	0	24	澳門	6	6	12
孟加拉	2	20	22	馬來西亞	30	15	45
白俄羅斯	7	0	7	馬爾他	5	0	5
比利時	32	0	32	墨西哥	27	0	27
貝南	3	0	3	荷蘭	55	0	55
百慕達	11	0	11	紐西蘭	6	0	6
波士尼亞	1	0	1	挪威	38	0	38
巴西	47	0	47	阿曼	13	0	13
保加利亞	5	0	5	巴基斯坦	0	16	16
加拿大	25	0	25	巴拿馬	15	0	15
智利	19	0	19	祕魯	4	0	4
中國	5	59	64	菲律賓	30	27	57
哥倫比亞	8	0	8	波蘭	31	0	31
克羅埃西亞	7	0	7	葡萄牙	40	0	40
塞浦路斯	15	0	15	卡達	12	0	12
捷克共和國	12	0	12	羅馬尼亞	12	0	12
丹麥	14	0	14	俄羅斯聯邦	95	0	95
多明尼加	19	0	19	斯里蘭卡	0	20	20
厄瓜多爾	1	0	1	沙烏地阿拉伯	36	0	36
埃及	11	0	11	塞爾維亞	1	0	1
薩爾瓦多	17	0	17	新加坡	20	0	20
愛發尼亞	4	0	4	斯洛伐克	17	0	17
芬蘭	16	0	16	斯洛維尼亞	24	0	24
法國	110	0	110	南非	36	0	36
喬治亞共和國	4	0	4	西班牙	182	0	182
德國	150	0	150	超國家組織	3	0	3

續接下頁

臺灣大學學術期刊資料庫

承接上頁

	主動付費取得信評之銀行家數	獲得不請自來信評之銀行家數	總和		主動付費取得信評之銀行家數	獲得不請自來信評之銀行家數	總和
希臘	22	0	22	瑞士	26	0	26
瓜地馬拉	1	0	1	瑞典	31	0	31
香港	39	38	77	台灣	63	79	142
匈牙利	7	0	7	泰國	40	0	40
冰島	8	0	8	千里達托貝哥	14	0	14
印度	16	111	127	突尼西亞	5	0	5
印尼	33	23	56	土耳其	73	0	73
愛爾蘭	31	0	31	烏克蘭	15	0	15
以色列	12	0	12	阿拉伯聯合國	26	0	26
義大利	119	0	119	英國	218	0	218
日本	126	0	126	美國	931	2	933
約旦	9	0	9	委內瑞拉	33	0	33
哈薩克	23	0	23	越南	5	20	25
肯亞	14	0	14	合計	3,433	448	3,881

說明: 本文樣本共有來自91國的3,881家銀行, 其中3,433家爲2002至2005年間付費以取得之信用評等的銀行家數, 448家爲獲得不請自來的評等的銀行家數。

等級的分配來看, 不請自來的評等等級確實出現向下偏誤的情況, 但此結果可能是自我選擇偏誤所造成, 即財務績效較佳的銀行, 傾向付費以取得信評機構的評等, 而財務績效表現不佳的銀行傾向不請求評等, 當信評機構主動給予評等, 由於其較差的財務績效, 自然也就獲得較差的評等等級。

表4報告獲得不請自來信評與主動付費取得信評之銀行的評等與績效變數的敍述統計量。發現不請自來的銀行信評, 平均而言低於主動付費取得之信評 (分數分別爲2.7634與5.3122)。有趣的是, 部分財務績效與特性變數在兩組銀行樣本中呈現顯著的差異性, 例如, 獲得不請自來評等的銀行, 平均而言有相對較低的國家主權評等 (分數分別爲7.0926與8.5168)、覆蓋比率 (分別爲62.452%與158.2%)、資本適足率 (分別爲11.577%與15.318%)、資產報酬率 (分別爲 −0.0513%與1.1453%)、成本對收入比率 (分別爲56.179%與60.452%)、相對較高的流動比率 (分別爲33.286%與31.677%) 以及較低的資產規模 (分別爲6.7377與6.9935)。這表示獲得

臺灣大學學術期刊資料庫

表 3: 惠譽銀行個別評等統計分析

		付費取得的信評		不請自來的信評		合計	
		數目	百分比	數目	百分比	數目	百分比
A	受評銀行就獨立基礎而言的信用狀況極強。其特色可能包括具備絕佳的獲利及完整的資產負債結構、企業規模、管理、經營環境或未來前景。	98	2.85%	0	0.00%	98	2.53%
A/B		279	8.13%	3	0.67%	282	7.27%
B	受評銀行就獨立基礎而言的信用狀況頗強。該銀行並無令人關切的重大問題。其特色可能包括強健的獲利及完整的資產負債結構、企業規模、管理、經營環境或未來前景。	886	25.81%	2	0.45%	888	22.88%
B/C		644	18.76%	12	2.68%	656	16.90%
C	受評銀行就獨立基礎而言的信用狀況允當，但卻有一或多處的問題。有關問題可能包括獲利及資產負債表的結構、企業規模、管理、經營環境或未來前景。	522	15.21%	53	11.83%	575	14.82%

續接下頁

臺灣大學學術期刊資料庫

承接上頁

		付費取得的信評		不請自來的信評		合計	
		數目	百分比	數目	百分比	數目	百分比
C/D		369	10.75%	55	12.28%	424	10.93%
D	受評銀行面對來自內部及/或外部的劣勢。該行令人關切的部分可能包括獲利及資產負債表的結構、企業規模、管理、經營環境或未來前景。新興市場的銀行必然面對較多來自外部的潛在缺失。	373	10.87%	99	22.10%	472	12.16%
D/E		175	5.10%	123	27.46%	298	7.68%
E	受評銀行問題嚴重，需要或可能會需要外部支援。	87	2.53%	101	22.54%	188	4.84%
合計		3,433	100%	448	100%	3,881	100%

說明: 個別評等等級包括上列5個等級中間值: 即 A/B、B/C、C/D 及 D/E。

不請自來信評的銀行, 相對於主動付費取得信評之銀行, 其所處的國家主權評等較低, 資產品質較差、資本適足性較不足、獲利能力較差及銀行資產規模較小; 然而, 獲得不請自來評等的銀行在代表管理能力的成本對收入比率以及代表流動性風險的流動比率的表現上是相對優勢的。但目前的分析結果僅透過敍述統計量的比較, 仍沒有統計顯著性的證據。

表5報告變數間的相關係數矩陣。由第1欄之係數發現, 獲得不請自來信評的銀行普遍來說信用評等較低、國家主權評等、覆蓋比率、資本比率、資產報酬率、成本對收入比率均較低、流動比率較高而資產規模較小, 與表4的結果相似。

自我選擇偏誤觀點主張銀行獲得的不請自來信評較差, 可能是因爲財

臺灣大學學術期刊資料庫

表 4: 敍述統計量

	所有銀行信評樣本				不請自來的銀行信評樣本				主動評等的銀行信評樣本			
	平均數	標準差	最小值	最大值	平均數	標準差	最小值	最大值	平均數	標準差	最小值	最
是否爲不請自來												
信評之虛擬變數	0.0746	0.2628	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	0.0000	0.0000	0.0
銀行評等	5.1219	2.0217	1.0000	9.0000	2.7634	1.4753	1.0000	8.0000	5.3122	1.9383	1.0000	9.0
國家主權評等	8.4230	1.7338	1.0000	10.000	7.0926	1.1787	5.0000	10.000	8.5168	1.7282	1.0000	10
覆蓋比率	151.18	141.20	−24.448	953.69	62.452	73.337	1.1800	645.42	158.20	142.89	−24.448	95
資本比率	15.035	17.616	−68.805	808.00	11.577	7.8329	−68.805	38.946	15.318	18.157	−19.016	80
資產報酬率	1.0551	3.0875	−70.425	60.940	−0.0513	4.3941	−70.425	3.8700	1.1453	2.9379	−34.705	60
成本對收入比率	60.136	26.254	−323.28	307.34	56.179	35.816	−192.61	307.34	60.452	25.314	−323.28	28
流動比率	31.802	41.375	0.0100	829.72	33.286	19.009	7.2380	189.35	31.677	42.726	0.0100	82
資產總額	6.9744	0.8555	2.9075	9.1079	6.7377	0.6077	5.0610	8.7217	6.9935	0.8697	2.9075	9.1

說明: 資料期間爲2002至2005年。所有資料共3,881個銀行–年 (bank-year) 樣本點, 獲得不請自來信評之銀行共448個樣本點, 主動付費取得信評之銀行共3
個樣本點。

臺灣大學圖書館 期刊資料庫

表 5: 變數之間的相關矩陣

	是否爲不請自來信評之虛擬變數	銀行評等	國家主權評等	覆蓋比率	資本比率	資產報酬率	成本對收入比率	流動比率	資產總
是否爲不請自來信評之虛擬變數	1.0000								
銀行評等	−0.3132	1.0000							
國家主權評等	−0.2225	0.6149	1.0000						
覆蓋比率	−0.1734	0.3985	0.3013	1.0000					
資本比率	−0.0345	−0.0114	−0.0979	0.0022	1.0000				
資產報酬率	−0.1107	0.2738	0.1500	0.2696	0.0981	1.0000			
成本對收入比率	−0.1154	−0.1113	0.0269	−0.0613	−0.0405	−0.0554	1.0000		
流動比率	0.0454	−0.1549	−0.3378	−0.0718	0.2531	0.0156	−0.0624	1.0000	
資產總額	−0.0610	0.2317	0.3478	−0.1143	−0.2143	−0.0723	−0.0008	−0.2348	1.000

說明: 資料期間爲2002至2005年。所有資料共3,881個銀行–年 (bank-year) 樣本點, 獲得不請自來信評之銀行共448個樣本點, 主動付費取得信評之銀行共3,433個樣本點。

臺灣大學學術期刊資料庫

務績效表現不佳而傾向不請求評等, 而財務績效較佳的銀行, 則相對較傾向付費以取得評等。至目前爲止, 我們的證據部分地證實此說法, 也就是銀行評等屬於不請自來抑或是主動付費取得, 有可能是自我選擇所造成的。另外, 一方面由於敍述統計的結果缺乏顯著性的探討, 另一方面則是兩組樣本的特性與績效變數存在相當程度的差異, 必須先將這差異排除或降低, 才能認銀行信用評等的差異是來自於不請自來或是主動付費這一事實所導致的結果。

4.2 傾向分數函數的估計

表6報告以 Probit 模型估計傾向分數函數的實證結果, 傾向分數函數即銀行獲得不請自來信用評等的機率函數。模型設定之被解釋變數爲銀行信評是否爲不請自來的虛擬變數 (是爲1, 否則爲0), 解釋變數則包括覆蓋比率、資本比率、資產報酬率、成本對收入比率、流動比率與資產總額共六個變數。

首先, 設定 (A) 發現, 六個解釋變數中有四個之估計係數顯著爲負, 包括覆蓋比率 (−0.0066)、資本比率 (−0.0436)、成本對收入比率 (−0.0094) 以及資產總額 (−0.3823)。表示獲得不請自來信評的銀行普遍來說, 覆蓋比率較低、資本比率較低、成本對收入比率較低且資產總額較小。除了成本對收入比率較低代表管理能力較佳之外, 其餘的三個變數皆顯示獲得不請自來信評的銀行財務狀況表現是相對較差的, 此結果可部分解釋不請自來銀行信評相對較低的緣故。亦即當銀行的財務績效相對較差時, 因不希望增加大衆對他們的表現是相對較差的印象, 傾向不主動要求獲得評等, 而評等機構根據公開且事實上相對較差的財務資訊下, 給予較差的評等等級。這可以解釋爲不請自來的信用評等事實上是有內生性選擇的成份在內。其次, 由於設定 (A) 之資產報酬率與流動比率的估計係數並不顯著, 根據 Dehejia and Wahba (2002) 的簡約原則, 設定 (B) 僅納入其餘四個解釋變數。發現估計係數結果與設定 (A) 相類似, 四個變數皆顯著, 且正負號不變。以下我們將根據設定 (B) 的估計結果, 計算每家銀行的傾向分數, 並進行樣本配對。

臺灣大學學術期刊資料庫

表 6: 傾向分數函數的估計: Probit 模型

	設定 (A)	設定 (B)
變數名稱	估計係數	
常數項	2.9448 (6.74)	2.8534 (6.88)
覆蓋比率	−0.0066 (−11.0)	−0.0067 (−11.2)
資本比率	−0.0436 (−5.31)	−0.0477 (−6.39)
資產報酬率	−0.0038 (−0.22)	
成本對收入比率	−0.0094 (−5.99)	−0.0095 (−6.23)
流動比率	0.0001 (0.12)	
資產總額	−0.3823 (−7.19)	−0.3631 (−7.10)
有效觀察值數目	3,408	3,649
Pseudo R-square	0.1873	0.1862

說明: 下表是以 Probit 模型並根據所有樣本估計傾向分數函數的實證結果。被解釋變數爲銀行信評是否爲不請自來的虛擬變數 (是爲 1, 否則爲 0)。設定 (A) 的解釋變數爲覆蓋比率、資本比率、資產報酬率、成本對收入比率、流動比率與資產總額六個特性變數; 設定 (B) 的解釋變數爲覆蓋比率、資本比率、成本對收入比率與資產總額; 模型的估計以混合 (Pooled) 的方式。資料期間由 2002 至 2005 年。括號中爲估計係數之 t 值。

4.3 樣本配對

估計出傾向分數函數後, 可得到每家銀行樣本的傾向分數, 接著利用前述之四種配對演算法, 進行樣本配對。

表7 報告在樣本配對前與配對後, 四個特性變數在兩組樣本之間的平均數差異檢定結果。發現兩組樣本之特性變數, 在樣本配對前與配對後明

臺灣大學 期刊資料庫

表 7: 兩組樣本的的特性變數平均數差異: 配對前與配對後

	樣本配對前			配對演算 Nearest 配對			Caliper 配對			Mahala 配對			Mahala Caliper 配對		
	不請自來評等	主動要求評等	差距 (t 值)	不請自來評等	主動要求評等	差距 (t 值)	不請自來評等	主動要求評等	差距 (t 值)	不請自來評等	主動要求評等	差距 (t 值)	不請自來評等	主動要求評等	差 (t
覆蓋比率	56.606	160.63	−104.02*** (−10.9)	56.606	69.854	−13.248** (−2.00)	56.466	70.066	−13.600** (−2.01)	56.606	61.143	−4.537 (−0.68)	56.592	56.819	−0.2 (−0.0
資本比率	12.125	14.270	−2.1450*** (−2.70)	12.125	12.071	0.0540 (0.11)	12.438	12.467	−0.0290 (−0.07)	12.125	12.457	−0.332 (−0.69)	12.291	12.404	−0.1 (−0.2
成本對收入比率	52.902	61.243	−8.3410*** (−5.52)	52.902	55.043	−2.1410 (−0.98)	54.599	56.138	−1.5390 (−0.78)	52.902	53.352	−0.45 (−0.18)	53.541	53.613	−0.0 (−0.0
資產總額	6.8276	7.0609	−0.2333*** (−4.34)	6.8276	6.5412	0.2864*** (4.58)	6.8398	6.5439	0.2959*** (4.66)	6.8276	6.8345	−0.0069 (−0.13)	6.9330	6.9318	0.0 (0.0

說明: 本表顯示樣本配對前與配對後, 獲得不請自來信評與主動付費取得信評之銀行的4個特性變數的平均數差異檢定結果。Nearest 配對是在選擇一個與獲得不請自來信評之傾向分數最相近的主動付費銀
做爲配對樣本; Caliper 方法是以傾向分數之差異小於某個值來作爲樣本配對的準則; Mahala 方法爲選擇與獲得不請信評的銀行之 Mahalanobis 距離最小的主動付費銀行作爲配對樣本; Mahala Calipers 以是
Mahalanobis 距離小於某個值來作爲樣本配對的準則。 樣本配對前, 獲得不請自來信評之銀行共 448 個樣本點, 主動付費取得信評之銀行共 3,433 個樣本點, 經過在前四種配對法下, 主動付費取得信評之銀
樣本分別變爲 233、228、233 以及 60 個樣本點。括號中爲兩組銀行樣本特性變數差距的 t 統計量, 而 *、** 與 *** 分別表示差距在90%、95% 與 99% 的信心水準下是顯著的。

期刊資料庫

顯地改變。首先, 在樣本配對前, 兩組樣本的四個特性變數的平均數皆有顯著差異 (分別爲 −104.02、−2.145、−8.341 以及 −0.2333), 即獲得不請自來信評之銀行的覆蓋比率、資本比率、成本對收入比率以及資產總額均顯著較低; 其次, 在 Nearest 配對演算下, 平均而言, 四個特性變數只剩下兩個仍具顯著差異, 即覆蓋比率與資產總額; 第三, 在 Caliper 配對演算下, 平均而言, 覆蓋比率與資產總額仍具有顯著差異性, 與 Nearest 配對有相似的結果; 最後, 有趣的是, 在 Mahala 與 Mahala Caliper 配對演算下, 兩樣本組的四個特性變數已接近相等。先前提及, 若要將兩類型銀行的信用評等高低歸因於不請自來或是主動付費取得這一因素, 必需使兩組樣本的特性變數儘量相近, 也就是說, 若兩類型銀行的特性變數與財務績效變數都調整成相近後, 若不請自來的信評仍然相對較低, 才可推論不請自來的銀行信評有向下偏誤的現象。在 Mahala 與 Mahala Caliper 配對演算下, 皆達到此要求, 將可以降低樣本選擇偏誤的問題。

表 8 報告由樣本配對前至樣本配對後, 獲得不請自來信評與主動付費取得信評之企業的特性變數的平均數差異減少的百分比。百分比愈大, 表示樣本配對愈能降低兩組樣本在特性變數上的差異性。結果發現, Nearest 配對在差異減少的成效較差, 兩組樣本間特性變數的差異性平均降低比率爲 59.2%; Caliper、Mahala 與 Mahala-Caliper 配對的有效性依序增加, 四個特性變數差異減少百分比的平均數分別爲 60.1%、92.9% 與 98.3%; 但在 Mahala-Caliper 配對下, 配對後的主動付費取得評等之銀行樣本僅存 60 個, 犧牲相當多的自由度, 而在 Nearest、Caliper 與 Mahala 配對下, 則仍保有較多的樣本點, 分別爲 233、228 與 233 個。當出現配對有效性較高但卻損失大量樣本的情況下, 配對有效性和樣本自由度將成爲一種抵換 (Trade-off) 關係 (本文並未出現此情況)。以下將報告四種配對法下的兩組樣本之績效比較與迴歸分析結果。

4.4 不請自來 vs. 付費取得的銀行信評

4.4.1 兩組樣本的平均數 *t* 檢定

表 9 報告包括配對前與配對後的樣本中, 獲得不請自來信評之銀行與主動付費取得信評之銀行的評等平均數差異分析的結果。首先, 第 2 欄報告樣

臺灣大學學術期刊資料庫

表 8: 特性變數差異減少百分比

	配對演算			
	Nearest 配對	Caliper 配對	Mahala 配對	Mahala Caliper 配對
覆蓋比率	87.3	86.9	95.6	99.8
資本比率	97.5	98.7	84.5	94.7
成本對收入比率	74.3	81.6	94.6	99.1
資產總額	−22.8	−26.8	97.0	99.5
平均數	59.1	60.1	92.9	98.3

說明: 本表報告4個特性變數在透過樣本配對後, 其平均數在獲得不請自來信評與主動付費取得信評之銀行之間差異性的減少百分比, 同時也報告4個特性變數差異減少百分比的平均數。在某種配對法下若百分比愈高, 表示透過該方法進行樣本配對後, 兩類型銀行的特性變數差異性愈低, 配對愈可靠。

表 9: 兩組樣本的平均評等差異: 配對前與配對後

		配對演算			
	樣本配對前	Nearest 配對	Caliper 配對	Mahala 配對	Mahala Caliper 配對
銀行評等	−2.5523***	−1.3176***	−1.3245***	−1.3863***	−1.20***
	(−19.6)	(−6.95)	(−6.96)	(−9.47)	(−4.41)

說明: 本表報告在樣本配對前與配對後, 獲得不請自來信評與主動付費取得信評之銀行的平均績效差異。資料期間由2002至2005年。利用跋靴法 (Bootstrap Method) 1,000次建立信賴區間以判斷差異的顯著性。括號中爲差異估計之 Z 值, 而 *、** 與 *** 分別表示差距在90%、95%與99%的信心水準下是顯著的。

本配對前, 平均評等差異結果, 發現平均評等差異顯著爲負 (−2.5523), 表示主動付費取得之信評相對高於不請自來的信評。其次, 第3至6欄報告配對後, 平均評等是否仍有差異。結果發現在 Nearest 配對下, 平均評等差異顯著爲負 (−1.3176), 在 Caliper 配對下, 平均評等差異也顯著爲負 (−1.3245), 表示不請自來的信評在配對後仍相對較低; 雖然較樣本配對前差異的程度 (−2.5523) 有下降的趨勢, 但差異仍然存在且達統計上顯著水準。在 Mahala 與 Mahala Caliper 配對下, 亦出現相似結果。

由以上分析結果得知, 樣本配對前, 不請自來的信評相對較低, 但至此

臺灣大學學術期刊資料庫

尚不能做成不請自來信評有向下偏誤現象的結論; 但在樣本配對後, 即兩類型銀行有相近似之績效表現後, 不請自來的評等仍然相對低於付費取得之信評, 在相同的績效表現應有相同信評的假設下, 不請自來的信評相對較低, 即表示存在向下偏誤的現象。

4.4.2 Ordered-Probit 迴歸分析

接著本文利用順序多重分類模型 (ordered-probit model), 分析不請自來信評是否影響信用評等, 即檢視方程式 (13) 中的「信評是否爲不請自來的虛擬變數」(UNSOLICIT_D), 其估計係數的正負號及顯著性, 當估計係數顯著爲負, 表示在控制其他影響銀行評等等級的可能因素後, 當獲得的是不請自來之信用評等, 仍將使銀行獲得相對較低的評等等級, 也就表示不請自來的評等確實出現向下偏誤的現象。

表 10 報告配對前與配對後樣本的迴歸分析結果。[12] 第 2 欄爲配對前的結果, 發現 UNSOLICIT_D 的估計係數顯著爲負 (−0.9121), 表示當信評爲不請自來時, 相對於信評爲主動付費取得低 0.9121 個評等等級。由於這是樣本配對前的分析結果, 兩組銀行樣本的績效或特性變數仍存有差異性 (由前述結果得知, 獲得不請自來信評的銀行其績效普遍較差), 尚不能做成不請自來的銀行信評有向下偏誤的結論。控制變數的估計係數均與預期相同, 其中 SOVRATING、COVERGE、CAPITAL、ROA、LIQUID、ASSET 的估計係數均顯著爲正, COST 的估計係數顯著爲負, 表示主權評等、覆蓋比率、資本比率、資產報酬率、流動比率愈高, 以及資產規模愈大, 可獲得愈高的信用評等, 而成本對收入比率愈低的銀行, 也可獲得愈高的信用評等, 這與既有實證研究的結果相符合。

第 3 至 6 欄報告配對後樣本之實證結果。首先, 在 Nearest 配對下, UNSOLICIT_D 的估計係數顯著爲負 (−1.0726), 表示不請自來的信評, 相對低於主動付費取得之信評。而國家主權評等、覆蓋比率、資本比率與資產總額的估計係數顯著爲正, 表示國家主權評等、覆蓋比率、資本比率與資

[12] Ordered-Probit 模型的估計參數可以包括未轉換前的估計參數以及轉換後的邊際效果兩種。然而本文的分析重點, 在於觀察樣本配對前與樣本配對後, 獲得不請自來的銀行信評是否相對較低, 也就是「著重觀察是否爲不請自來信評虛擬變數之估計係數其正負顯著與否」, 因此並未就表 10 各變數的估計係數進行轉換。此部分謝謝評審的補充說明建議。

期刊資料庫

表 10: 不請自來的信用評等比較低嗎? Ordered-Probit 模型

$$\text{RATING}_i = \beta_0 + \beta_1 \text{UNSOLICIT_D}_i + \beta_2 \text{SOVRATING}_i + \beta_3 \text{COVERAGE}_i + \beta_4 \text{CAPITAL}_i + \beta_5 \text{ROA}_i + \beta_6 \text{COST}_i + \beta_7 \text{LIQUID}_i + \beta_8 \text{ASSET}_i + \varepsilon_i \circ$$

	Ordered-Probit 模型				
	樣本配對前	Nearest 配對	Caliper 配對	Mahala 配對	Mahala Caliper 配對
是否爲不請自來信評之虛擬變數	−0.9121*** (−11.5)	−1.0726*** (−8.91)	−1.0386*** (−8.52)	−1.0014*** (−9.09)	−0.9327*** (−5.58)
國家主權評等	0.4321*** (27.7)	0.2953*** (6.08)	0.3102*** (6.33)	0.3826*** (6.95)	0.4198*** (4.91)
覆蓋比率	0.0022*** (15.1)	0.0038*** (5.19)	0.0036*** (4.84)	0.0043*** (5.93)	0.0042*** (4.59)
資本比率	0.0031** (2.00)	0.0990*** (7.47)	0.1250*** (8.56)	0.0986*** (7.82)	0.1362*** (7.65)
資產報酬率	0.0907*** (9.09)	0.0175 (0.61)	0.0867** (2.53)	0.0217 (0.79)	0.0400 (1.00)
成本對收入比率	−0.0074*** (−8.92)	−0.0020 (−0.76)	−0.0019 (−0.67)	−0.0007 (−0.36)	0.0013 (0.37)
流動比率	0.0029*** (3.45)	−0.0012 (−0.49)	−0.0028 (−1.09)	0.0014 (0.45)	−0.0005 (−0.12)
資產總額	0.2404*** (9.08)	0.2758*** (2.98)	0.3102*** (3.32)	0.0393 (0.38)	−0.1842 (−1.25)
有效觀察值數目	3,304	409	404	416	260
Pseudo R-square	0.1573	0.1609	0.1732	0.1656	0.1696

說明: 本表報告利用最小平方法估計獲得不請自來信評與主動付費取得信評之銀行的信用評等差異。資料期間由 2002 至 2005 年。獲得不請自來信評之銀行在配對前與配對後皆爲 448 個樣本點, 在 Nearest、Caliper、Mahala 與 Mahal Calipers 配對下, 主動付費取得信評之銀行分別爲 233、228、233 以及 60 個樣本點。括號中爲估計係數之 t 值, 而 *、** 與 *** 分別表示估計係數在90%、95% 與 99% 的信心水準下顯著異於 0。

臺灣大學學術期刊資料庫

產總額愈高的銀行其信評愈高; 其次, 在 Caliper 配對下, UNSOLICIT_D 的估計係數仍然顯著爲負 (−1.0386), 表示不請自來的信評仍然相對較低; 第三, 在 Mahala 與 Mahala-Caliper 配對下, UNSOLICIT_D 的估計係數仍然顯著爲負, 係數分別爲 −1.0014 與 −0.9327。

由以上分析結果得知, 在使用 4 種不同的樣本配對方法進行配對, 即將兩組銀行樣本的績效與特性變數調整成近似後 (從平均數的角度來看), 發現不請自來的信評仍然相對低於付費取得之信評。若相同的銀行績效應有相同銀行信評, 但結果發現獲得不請自來信評的銀行雖與主動付費取得信評的銀行有相近似的績效表現, 但其信評卻相對較低, 因此, 不請自來的銀行信用評等確實存在向下偏誤的現象。

關於獲得不請自來的信用評等是否存有向下偏誤的既有實證文獻中, 矯正樣本選擇偏誤的方式是利用 Heckman (1979) 的兩階段估計法 (two-stage estimation procedure), 例如 Poon (2003)、Poon and Firth (2005) 即採用此方法進行估計, 因此本文亦採用此方法做爲主要分析結果的穩健性檢驗。在兩階段估計上, 第一階段, 先透過非線性的選擇方程式, 即利用最大概似法估計二元 Probit 模型, 被解釋變數爲樣本是否爲獲得不請自來銀行信評的零或一指標變數, 解釋變數則爲影響樣本是否爲獲得不請自來銀行信評的變數, 考慮前述在配對方法中, 傾向分數函數的估計上與此第一階段的含意相同, 因此我們放置銀行的覆蓋比率、資本比率、成本對收入比率與資產總額四個變數, 第一階段估計可獲得解釋變數的係數及選擇偏誤調整項。第二階段的估計上, 跟隨 Poon (2003)、Poon and Firth (2005), 除了選擇偏誤調整項外, 包括了是否爲不請自來信評之虛擬變數、國家主權評等、覆蓋比率、資本比率、資產報酬率、成本對收入比率、流動比率以及資產總額。

表 11 報告兩階段估計的分析結果。第 2 欄報告第一階段的估計結果, 發現獲得不請自來信評之銀行的覆蓋比率 (估計係數爲 −0.0062)、資本比率 (估計係數爲 −0.043)、成本對收入比率 (估計係數爲 −0.0113) 顯著較低, 資產規模顯著較小 (估計係數爲 −0.2724)。除了成本對收入比率較低代表管理能力較佳之外, 其餘的三個變數皆顯示獲得不請自來信評的銀行財務狀況表現是相對較差的。第 3 欄報告二階段的估計結果, 發現 UNSO-

臺灣大學學術期刊資料庫

表 11: 不請自來的信用評等比較低嗎? Ordered-Probit 模型
(Heckman 兩階段估計)

$$\text{UNSOLICIT_D}_i = \alpha_0 + \alpha_1 \text{COVERAGE}_i + \alpha_2 \text{CAPITAL}_i + \alpha_3 \text{COST}_i + \alpha_4 \text{ASSET}_i + \mu_i,$$

$$\text{RATING}_i = \beta_0 + \beta_1 \text{UNSOLICIT_D}_i + \beta_2 \text{SOVRATING}_i + \beta_3 \text{COVERAGE}_i + \beta_4 \text{CAPITAL}_i + \beta_5 \text{ROA}_i + \beta_6 \text{COST}_i + \beta_7 \text{LIQUID}_i + \beta_8 \text{ASSET}_i + \varepsilon_i。$$

	Heckman 兩階段估計	
	第一階段	第二階段
	估計係數	
是否爲不請自來信評之虛擬變數		1.5005
		(0.58)
國家主權評等		0.2404**
		(2.09)
覆蓋比率	−0.0062***	−0.0042
	(−10.3)	(−0.29)
資本比率	−0.0430***	0.0836
	(−5.67)	(0.95)
資產報酬率		−0.0001
		(0.00)
成本對收入比率	−0.0113***	−0.0135
	(−6.92)	(−0.58)
流動比率		−0.0013
		(−0.24)
資產總額	−0.2724***	−0.2505
	(−5.14)	(−0.41)
有效觀察值數目	3,620	3,416

說明: 本表報告利用 Heckman (1979) 之兩階段方法估計不請自來信評與主動付費取得信評之銀行的信用評等差異。資料期間由2002至2005年。所有公司共3,881個樣本點。括號中爲估計係數之 t 值, 而 *、** 與 *** 分別表示估計係數在90%、95%與99%的信心水準下顯著異於0。

LICIT_D 的估計係數不顯著, 表示不請自來的信評並未出現向下偏誤的現象。有趣的是, 在迴歸式中僅有主權評等的估計係數顯著爲正 (0.2404), 代

臺灣大學學術期刊資料庫

表銀行所處的國家主權評等愈高則其信評將會是愈高的。另外, 在表格中並未報告, 選擇偏誤調整項的估計係數爲 1.5005, Z 值爲 2.604, 爲統計顯著, 這與第一階段估計中四個解釋變數皆爲顯著的結果相呼應, 表示樣本的選擇偏誤證據是顯著存在的。

整體來說, 大部分的證據顯示不請自來的信用評等, 相對低於主動付費取得的評等, 且不論是用配對前或是配對後的樣本結果皆如此。這其中有兩層意義。首先, 在配對前樣本顯示獲得不請自來信評的銀行其績效表現是相對較差的, 因此當得到信評相對較低的結果, 僅能證實既有文獻中的自我選擇假設 (self-selection Hypothesis), 即信用品質較佳的企業傾向選擇付費以獲得信用評等, 而信用品質較差的企業選擇不主動請求評等。其次, 在樣本配對後, 兩類型銀行的績效與特性變數已調整爲相近, 仍然得到不請自來信用評等相對較低的結果, 則顯示不請自來的信評的確存在向下偏誤的現象, 可能的解釋是信評機構僅能利用未主動付費請求評等之銀行的有限公開訊息, 因而給予較差的評等等級。本文實證結果與 Poon (2003)、Poon and Firth (2005) 以及 Fairchild et al. (2009) 相一致, 他們皆發現不請自來的評等等級出現向下偏誤的現象, 而本文的貢獻在於應用晚近新發展的配對方法以降低樣本選擇偏誤, 進行不請自來之銀行信評向下偏誤的估計與認定時更爲嚴謹。

5 結論

不請自來的評等是否有向下偏誤的現象在既有實證研究中一直未有定論, 原因是部分的研究雖然發現不請自來的信評相對較低, 但獲得不請自來信評的銀行其體質卻也相對較弱, 並不能說是有向下偏誤的現象, 頂多能說是銀行自我選擇的結果 (Byoun and Shin, 2002)。部分學者則是發現在控制了銀行的績效後, 不請自來的評等仍較付費取得之評等爲低, 因此不請自來的評等應該是有向下偏誤的現象 (Fairchild et al., 2009; Poon, 2003; Poon and Firth, 2005)。

本文應用 Rubin (1973) 與 Rosenbaum and Rubin (1983, 1985a,b) 的配對方法, 透過將兩類型銀行樣本之財務績效與特性變數修正至近似, 重新評估這個議題。本文利用2002至2005年間, 來自全球91國獲得惠譽的

臺灣大學學術期刊資料庫

銀行個別評等之銀行爲樣本, 共有 3,881 個觀察值, 其中 3,433 個觀察值爲付費取得之信用評等的銀行, 448 個觀察值爲獲得不請自來的評等的銀行, 我們發現在樣本配對前, 不論是透過兩組樣本的平均評等比較或是迴歸分析, 大部分的證據顯示獲得不請自來信評之銀行的多數財務績效指標是相對地低於付費取得信評之銀行, 這解釋了爲什麼前者之信評相對較低的原因; 而在樣本配對後, 即控制兩類型銀行的財務特性之後, 大部分的證據仍然顯示不請自來的銀行信評是相對較低的。[13]

不請自來的信用評等出現向下偏誤的現象, 對於銀行甚至企業不是沒有成本的, 誠如第1節所述, 獲得較低的信評將使銀行或公司支付較高的債務資金成本, 而想要獲得較高評等則必須支付評等費用以取得評等, 形同被評等機構勒索 (Hold-up) 的現象。因此, 往後的研究我們可以探討獲得不請自來信用評等的銀行或企業, 相對於主動付費取得評等者而言, 是否付出了較高的資金成本? 是否股票報酬相對較低? 若獲得不請自來的信用評等與主動付費取得評等者有相同的財務績效, 理應獲得相同條件的資金成本與市場報酬, 卻因信評機構在有無收取信評費用的差異下, 給予不同的信評, 繼而導致其資金成本與市場報酬有所差異, 則信評機構在市場上的角色便應該受到檢討, 它甚至可能會誤導市場投資人。另外, 給定未付費而取得不請自來的信用評等是相對較低的, 有什麼辦法或有什麼因素可以減輕或矯正亦值得探討。

在配對方法的使用上是有些限制與缺點的。首先、最值得注意的就是配對變數的選擇, 在本文中我們利用推論的方法找出若干個銀行特性變數做爲樣本配對的根據, 這面臨武斷 (Arbitrary) 的風險, 而利用武斷選擇的配對變數進行樣本配對再進行相關的實證分析將容易產生不確定感。因此

[13]本文的樣本包含了開發中國家以及已開發國家, 要控制這些差異性的一種方式爲: 在一國之內進行樣本配對, 爾後的分析再控制國家變數; 更詳細地說, 透過樣本配對我們會得到每一個國家有兩組樣本, 一組是獲得不請自來信評樣本, 一組是付費取得信評之樣本, 在標記這些樣本的國家別之後, 我們將所有獲得不請自來信評的樣本編爲一組, 付費取得者編爲一組, 則可以利用本文中之迴歸分析的方式 (需對樣本設置國家虛擬變數) 來進行評估是否不請自來的信評是相對較低的。然而這樣做的缺點是有可能面對單一國家內樣本數不足以進行樣本配對的風險。因此, 未來的研究在隨著時間經過而使得樣本數夠多的情況下, 可以單獨利用某特定區域國家或某特定經濟或金融發展程度相近似的國家來進行分析此議題。此部分作著感謝評審的修正建議。

期刊資料庫

在未來採用配對方法時應該在某特定議題下嘗試找出文獻中影響樣本自我選擇的因素做爲樣本配對的變數, 爾後再進行相對應的樣本配對與分析, 才能得到較嚴謹的實證結果。第二、至目前爲止我們驗證樣本配對得好不好, 皆是使用特性變數的一階動差即平均數的相似性去做檢驗, 雖然在文獻上尚未出現, 但考慮透過檢驗二階動差之相似性而獲得證實之後應該才能陳述爲某個特性變數在兩組樣本之間的分配是相近似的。

第三、在樣本配對的第一階段即傾向分數函數的估計上, 若 Pseudo R-square 相當低 (本文約爲 0.18), 表示傾向分數函數之估計結果的配適度並不高, 也就是以目前所選擇之特性變數來進行樣本配對的完整性是不夠的, 除了有可能是傾向分數函數的設定問題之外 (例如線性或非線性設定), 在文章中描述樣本配對可以降低選擇偏誤的程度時, 將只能限縮在所選擇變數之間差異性的降低上, 而不能說已經全面地控制了兩組樣本之間所有的樣本選擇偏誤, 除非得到的傾向分數函數估計結果有相當高的 Pseudo R-square; 換另一個角度來說, 第一階段傾向分數函數估計結果的好壞, 會影響第二階段的估計結果, 這個影響, 指的是到底樣本選擇偏誤在最後 (即第二階段) 的估計結果中被控制得好還是不好, Pseudo R-square 高就代表控制得好, 反之亦然。最後, 使用配對方法將減少可分析的樣本, 當配對過程中樣本數大量減少時, 將犧牲估計時的效率性, 因此使用配對方法最好是在大樣本的情況下進行。

參考文獻

Baker, H. K. and Mansi, S. A. (2002), "Assessing credit rating agencies by bond issuers and institutional investors", *Journal of Business Finance and Accounting*, 29, 1367–1398.

Bissoondoyal-Bheenick, E. (2005), "An analysis of the determinants of sovereign ratings", *Global Finance Journal*, 15, 251–280.

Borensztein, E., Cowan, K., and Valenzuela, P. (2006), "The sovereign ceiling lite and bank credit ratings in emerging markets economies", IMF working paper.

Borio, C. and Packer, F. (2004), "Assessing new perspective on country risk", *BIS Quarterly Review*, December, 47–64.

臺灣大學學術期刊資料庫

Butler, A. W. and Rodgers, K. J. (2003), "Relationship rating: How do bond rating agencies process information?", EFA Annual Conference paper, no. 491.

Byoun, S. and Shin, Y. S. (2002), "Unsolicited credit ratings: Theory and empirical analysis", working paper, Financial Management Association Annual Meeting.

Cantor, R. and Packer, F. (1994), "The credit rating industry", *Federal Reserve Bank of New York Quarterly Review*, 19, 1–26.

——— (1996), "Determinants and impact of sovereign credit ratings", *Economic Policy Review*, 2, 37–53, Federal Reserve Bank of New York.

Crowe, C. (2006), "Testing the transparency benefits of inflation targeting: Evidence from private sector forecasters", IMF working paper.

Dale, R. S. and Thomas, S. H. (2000), "The role of credit ratings in the Basel capital adequacy proposals", ECMI short paper, no. 2.

Dehejia, R. H. and Wahba, S. (2002), "Propensity score matching methods for nonexperimental causal studies", *Review of Economics and Statistics*, 84, 151–161.

Fairchild, L., Flaherty, S., and Shin, Y. (2009), "Analysis of unsolicited credit ratings: New evidence from Moody's", *Review of Pacific Basin Financial Markets and Policies*, 12, 103–123.

Fulghieri, P., Strobl, G., and Xia, H. (2010), "The economics of unsolicited credit ratings", working paper, University of North Carolina.

Gan, Y. H. (2004), "Why do firms pay for bond ratings when they can get them for free?", working paper, Wharton School.

Glick, R., Guo, X., and Hutchison, M. (2006), "Currency crises, capital-account liberalization, and selection bias", *Review of Economics and Statistics*, 88, 698–714.

Ham, J., Li, X., and Reagan, P. (2004), "Propensity score matching: A distance-based measure of migration and the wage growth of young men", Department of Economics, Ohio State University.

Harington, H. (1997), "Not Moody – just angry", *The Banker*, February, 22–23.

Heckman, J. (1979), "Sample selection bias as a specification error", *Econometrica*, 47, 153–162.

Heckman, J., Ichimura, H., Smith, J., and Todd, P. (1998), "Characterizing selection bias using experimental data", *Econometrica*, 66, 1017–1098.

Heckman, J., Ichimura, H., and Todd, P. (1997), "Matching as an econo-

臺灣人文學術期刊資料庫

metric evaluation estimator: Evidence from evaluating a job training program", *Review of Economic Studies*, 64, 605–654.

Heckman, J. and Robb, R. (1985), "Alternative methods for evaluating the impact of interventions", in J. J. Heckman and Burton Singer (eds.), *Longitudinal Analysis of Labor Market Data, Econometric Society Monographs series, No. 10*, Cambridge: Cambridge University Press.

——— (1986), "Alternative methods for solving the problem of selection bias in evaluating the impact of treatments on outcomes", in Howard Wainer (ed.), *Drawing Inference from Self Selected Samples*, New York: Springer-Verlag.

Hofler, R., Elston, J., and Lee, J. (2004), "Dividend policy and institutional ownership: Empirical evidence using a propensity score matching estimator", discussion papers, Max Planck Institute of Economics.

Hutchison, M. M. (2004), "Selection bias and the output cost of IMF programs", discussion paper, University of California, Santa Cruz.

Japan Center for International Finance (1999), "Characteristics and appraisal of major rating companies-focusing on ratings in Japan and Asia", Tokyo: Japan Center for International Finance.

LaLonde, R. (1986), "Evaluating the econometric evaluations of training programs with experimental data", *American Economic Review*, 76, 604–620.

Millon, M. and Thakor, A. (1985), "Moral hazard and information sharing: A model of financial information gathering agencies", *Journal of Finance*, 40, 1403–1422.

Persson, T. (2001), "Currency unions and trade: How large is the treatment effect?", *Economic Policy*, 33, 435–448.

Poon, W. (2003), "Are unsolicited credit ratings biased downward?", *Journal of Banking and Finance*, 27, 593–614.

Poon, W. and Firth, M. (2005), "Are unsolicited credit ratings lower? International evidence from bank ratings", *Journal of Business Finance and Accounting*, 32, 1741–1770.

Poon, W., Firth, M., and Fung, H. G. (1999), "A multivariate analysis of the determinants of Moody's bank financial strength ratings", *Journal of International Financial Markets, Institutions and Money*, 9, 267–283.

Poon, W., Lee, J., and Gup, B. E. (2009), "Do solicitations matter in bank credit ratings? Results from a study of 72 countries", *Journal of Money, Credit and Banking*, 41, 285–314.

臺灣大學學術期刊資料庫

Ramakrishnan, R. and Thakor, A. (1984), "Information reliability and a theory of financial intermediation", *Review of Economic Studies*, 51, 415–432.

Rojas-Suarez, L. (2001), "Rating banks in emerging markets: What credit rating agencies should learn from financial indicators", Institute for International Economics, working paper, no. 01–06.

Rosenbaum, P. and Rubin, D. (1983), "The central role of the propensity score in observational studies for causal effects", *Biometrika*, 70, 41–55.

——— (1985a), "Constructing a control group using multivariate matched sampling methods that incorporate the propensity", *American Statistician*, 39, 33–38.

——— (1985b), "The bias due to incomplete matching", *Biometrics*, 41, 103–116.

Rubin, D. (1973), "Matching to remove bias in observational studies", *Biometrics*, 29, 159–183.

Rubin, D. B. and Thomas, N. (1992), "Characterizing the effect of matching using linear propensity score methods with normal distributions", *Biometrika*, 79, 797–809.

SEC (United States Securities and Exchange Commission) (2003), "Report on the role and function of credit rating agencies in the operation of the securities markets", (as required by Section 702 (b) of the Sarbanes-Oxley Act of 2002), January.

Shen, C. H. and Chang, Y. (2008), "Ambition versus conscience, does corporate social responsibility pay off — The application of matching methods", *Journal of Business Ethics*, 88, 133–153.

Shimoda, N. and Kawai, Y. (2007), "Credit rating gaps in Japan: Differences between solicited and unsolicited ratings and rating splits", Bank of Japan, Working Paper Series.

UBS Investment Bank (UBS) (2004), "The new world of credit ratings", September, 1–22.

Van Roy, P. (2006), "Is there a difference between solicited and unsolicited bank ratings and, if so, why?", National Bank of Belgium, working paper, no. 79.

Vega, M. and Winkelried, D. (2005), "Inflation targeting and inflation behavior: A successful story?", *International Journal of Central Banking*, 1, 153–175.

臺灣大學學術期刊資料庫

投稿日期: 2010年4月20日, 接受日期: 2011年5月21日

Are Unsolicited Bank Credit Ratings Biased Downward? — A Matching Approach

Yu-Li Huang

Department of Insurance and Financial Management, Takming University of Science and Technology

Yuan Chang

Department of Business Education, National Changhua University of Education

Chung-Hua Shen

Department of Finance, National Taiwan University

This paper examines whether there is downward bias in Fitch's unsolicited bank credit ratings relative to solicited ratings using a matching approach. The sample for this study consists of those banks that had Fitch's Bank Individual Ratings (FBRs) during 2002–2005. There are 3,881 observations from 91countries, including 3,433 solicited ratings and 448 unsolicited ratings. By employing matching methods developed by Rubin (1973) and Rosenbaum and Rubin (1983, 1985a,b), four matching algorithms, Nearest, Caliper, Mahala and Mahala Caliper, are used to match the financial characteristics of two groups of banks in order to correct for sample selection bias. Our evidence shows that before sample matching, banks received unsolicited ratings have weaker financial profiles and this could partially explain why they also get lower ratings. Based on after-matching samples, most of the evidence suggests that unsolicited ratings are still biased downward after controlling for differences in key financial characteristics.

Keywords: Unsolicited Bank Credit Ratings, Self-selection Bias, Matching
JEL classification: C21, G10, G15, G21

臺灣大學學術期刊資料庫