

2017年CRC第15屆信用評分 研討會考察紀要

王璞玉/金融聯合徵信中心 徵信部經理
閻美晴/金融聯合徵信中心 研究部

壹、 會議目的

英國愛丁堡大學信用風險研究中心（Credit Research Centre，以下簡稱CRC），成立於1997年，主要宗旨為致力於信用風險相關理論與實務應用之研究，其中特別著重於消金信用風險與信用評分模型之建置與應用；CRC初期成立目的，原本僅是為了提供信用風險相關研發人員一個能相互探討與交流之開放式論壇，然而發展至今，儼然已是信用風險研究之重鎮，其CRC創立者Lyn C. Thomas等人所撰寫之《Credit Scoring and Its Applications》更被視為信用評分模型經典書籍。CRC除不定期於國際會議或期刊中發表相關研究成果與評論外，另固定每兩年於愛丁堡舉辦一場國際性信用評分研討會，會議舉辦之目的，主要是希冀能提供一個信用風險理論方法與實務應用交流討論的平台，聚集相關領域之產學界專業人士，並邀請相關主管機關及領

導廠商針對特定議題進行討論，所以會議內容與其他消費者信用風險領域之國際會議有本質上的差異，除有部分領導廠商藉機發表產品技術與研發外，許多國家及學術機構也趁此會議分享其研究成果。

CRC於2017年愛丁堡舉辦之第15屆信用評分研討會¹（Credit Scoring and Credit Control XV），其會議進行方式與過去歷屆相似，本屆會議共為期四天（8月29日至9月1日），首先於8月29日晚上舉辦兩場專題研討會，此兩場研討會之課程內涵為信用評分模型基本概念，主要是提供非數理背景之管理階層一個入門的課程，另8月30日至9月1日三天，CRC共安排4場主題演講及21場分組演講，共計有82場演講，在8月31日晚上亦安排一場與會者晚宴。本次研討會會議內容涵蓋主題相當多元，與會人員多數為金融機構或學術機構之資深研究人員或主管，進行演講發表的業者包

1 會議相關資料及本文參考資料皆請詳見該會議網站（<https://www.business-school.ed.ac.uk/crc-conference/>）。

含Equifax、Experian及FICO等國際型信用評分機構之代表人員，學術機構則有University of Edinburgh Business School，University of Southampton，Imperial College等，而研討內容之主題包含有：需求預測模式（Demand Modelling）、特徵變數（Features）、順序決策（Sequential Decisions）、詐欺模型（Fraud Modelling）、IFRS9、內部評等模型（IRB Modelling）、違約損失率（Loss Given Default）、違約暴險額（Exposure at Default）、模型程式（Programming）、存活分析模型（Survival Modelling）、順序決策（Sequential Decisions）、利潤評分（Profit Scoring）、房貸分析（Mortgages）和壓力測試（Stress Testing）等主要議題。本次出差赴英國參與此研討會，希冀能藉由此研討會瞭解目前信用風險管理之最新發展與趨勢，對於本中心未來信用資料與評分模型之加值服務與創新發展能有所助益。

貳、會議內容

以下分享四篇有關需求預測模式、特徵變數及順序決策之議題，其中前三篇內容涉及P2P網路借貸、替代性資料（alternative data）與普惠金融（financial inclusion）相關之研究，最後一篇內容主要是探討信用報告機構的資料與總體經濟變數的關係。

一、A Credit Scoring Model Based on Alternative Mobile Data for Financial Inclusion

(一) 摘要說明

近年來隨著互聯網經濟和消費金融的蓬勃發展，信用對中國消費者而言亦趨重要，然而，中國有超過2/3的人是沒有信用紀錄的，因為這些人從來沒有去過銀行，要對這些未曾與銀行有往來紀錄的人（unbanked people）進行所謂的信用風險評估，是有相當程度的困難。隨著大數據時代來臨，未曾與銀行有往來紀錄的人，因著許多替代性資料的產生，能有機會享受到普惠金融服務，對於信用評分模型而言，替代性資料中的手機數據（mobile phone data），普遍被認定具備良好的預測效果，因此此研究，在與中國電信業者和Cignifi合作之下，嘗試使用手機數據建構出一新的信用評分模型，研究人員相信在不久的將來，中國將會廣泛應用此類評分模型。

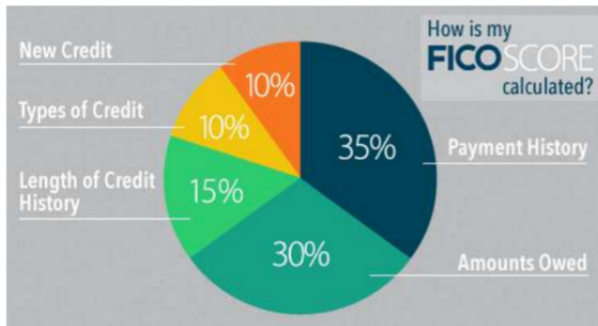
(二) 研究動機

信用評分是消費者信用風險管理極為重要的工具之一，傳統的信用評分是依循FICO模式，也就是以消費者的信用紀錄（如繳款紀錄，負債金額，信用歷史長度，信用借貸種類及新信用需求，參見圖1）所建構而成的評分模型。然而，全球有近30億的消費者是屬於信用未見者（credit invisible），特別是在新興國家市場，又由於這些國家近幾年的金融服務業急速擴展，因此在消費者替代性資料的需求是相當急迫的，替代性資料（或稱大數據）於信用評分模型之應用，亦就此成為近期持續發燒的議題之一。目前已有數家大型個人信用報告

機構，包含FICO，皆紛紛投入替代性資料於信用評分模型之相關研發，例如，2016年10月25日，FICO宣布之普惠金融計畫（Financial Inclusion Initiative），其中即包含使用替代性資料於信用評分模型之開發計畫。

目前正有數以百計的高科技企業，使用各項替代性資料發展信用評分模型，如手機資料，線上支付，電子商務資料，甚至社交網路等資料，以中國來說，大約有一百家的高科技企業，如阿里巴巴的芝麻評分，其主要是加入線上支付資訊，京東商城使用自家電子商城中，消費者之電子商務資料去發展信用評分，騰訊使用微信及QQ之社交網路資訊衡量消費者信用風險。

圖 1：FICO 評分模式



比較各類替代性資料，一般以為消費者的手機資訊，是相當適合納入信用評分模型的資料，主要原因之一是手機資訊上的還款資料，可代表該手機使用者的還款能力與還款意願，此一資訊與授信借貸的還款行為資訊，具有一

定程度的相關性，其次是考量到建模需要的資料樣本問題，通常消費者的手機資訊僅需蒐集近數月的資料，即可獲取足夠的資料樣本，樣本觀察期如有三個月的資料已視為足夠，如果能蒐集六個月的觀察期資訊則是更為完全，另一方面來看，手機資訊也常視為消費者經濟活動的一種指標。

(三) 研究背景

接下來，對中國個人信用市場進行概況說明，首先就信用報告機構來說，中國的信用報告機構，是由中國人民銀行（People's Bank of China，以下簡稱PBOC），即中國的中央銀行所建立，屬官方信用報告中心，目前成立已有10年，主要是提供信用報告與信用評分²服務，其資料涵蓋約4億消費者，其中有1億消費者是屬於無法評分者，該服務提供給核准設立的金融機構，如銀行或信用合作社（credit unions）；另中國尚有民營信用報告機構，PBOC曾於2015年1月公布首批獲得中國個人徵信牌照的8家機構，其中亦包含網路科技公司如阿里巴巴和騰訊，其後在2017年4月，PBOC基於這些機構非屬於第三方獨立單位等原因，最後拒絕核准發照，而在近三年的中國市場上，約有近500,000家的徵信公司，此為中國信用市場躍進的一種指標，同時這些徵信公司非僅僅限於提供個人信用服務，亦有部分

² 目前尚未上線。

公司對一般企業提供企業信用服務。從上述簡要的說明，呈現出目前中國整體信用報告市場之功能及其信用控管上一個混亂的現象。

在信用市場上，目前有超過2/3的消費者缺乏傳統授信借貸資料，但除此之外，消費者卻有相當多其他的資料來源與種類，如手機資料，線上支付，公用事業費用資料，電子商務資料及社交網路等資料，其中超過2/3的消費者有手機資料，然而這些手機資料多半是用於詐欺防範或是催收使用，鮮少使用於信用評分模型中；此外，因著網路借貸與個人借貸之興起，P2P網路借貸平台，最多曾有五千家，其中有三分之一目前已經倒閉，而個人借貸方面，「發薪日貸款（Payday Loan）」自上半年開始蓬勃發展，基於上述這些因素，中國目前對於信用評分領域專才有高度需求。

然而，在普惠金融、網路借貸及信用評分等議題方興未艾的發展中，如何找出一個有效的方式使用行動數據，促進普惠金融發展，仍然是中國當前的一個挑戰。此研究主題，主要是利用「手機資料」，作為建構信用評分模型中的替代性資料，在此「手機資料」是指手機通聯紀錄（Call Detail Record，CDR）和手機帳單收費資訊。

(四) 相關研究文獻

目前針對手機資料於信用評分領域之應用，在中國還尚未有類似研究，但是國際上已有部分研究探討，如Björkegren和Grissen採用手機資料上的行為變數預測違約風險，與傳

統上使用信用歷史紀錄的信用評分模型之預測力相較，其結果相近，並以加勒比海國家的借款人為樣本，使用通聯紀錄資料及貸款結果，結果顯示，前25%違約風險最高的借款人，其違約風險約是後25%的借款人的6倍；Jose San Pedro等人提出MobiScore，此模型是藉由手機資料，建置使用者財務風險模型，在過去的研究中發現，手機資料能夠傳達使用者個性與社會經濟地位；在沒有信用歷史資料時，MobiScore可以取代傳統的信用評分模型，為目前被排除於金融服務範圍的客群，提供一個取得信用的渠道，亦能用來作為一種補充性的資料，以改進傳統上基於財務資訊的評分模型，並採用拉丁美洲國家電信營運商和金融機構的資料進行實證分析，實證結果發現此方法產生的模型準確度，與傳統信用評分技術相當；另美國Cignifi公司應用手機資料於巴西的市場上進行小額信貸審核。

(五) 研究方法與樣本

由於中國手機多數都是以預付（pre-paid）卡儲值的方式，並非是採用後付款形式，因此如果只使用手機門號資訊，就授信借貸的角度，難以明確定義「違約」，為解決無法明確定義手機違約的問題，此研究採用一方法論，稱為交叉訓練（cross-training），也就是除了使用手機門號的資訊，另外從銀行取得個人信用違約的資料，試圖以手機資料衍生而成之各類行為變數為模型自變數，而信用違約紀錄與否為模型應變數，進行模型配適。

在樣本上，由中國一家商業銀行提供40多萬筆個人授信借貸資訊為樣本，所有樣本皆有一手機門號，接下來觀察樣本信用違約概況，並進行抽樣選取。

(六) 模型建置

在評分模型配適上，使用傳統處理二元分類 (binary classification) 的羅吉斯迴歸模型，其中最費時之部分，主要是從手機資料衍生相關行為變數並進行篩選。在手機變數的面向上，從以下幾個特徵進行考量：傳入手機與傳出手機的資訊、平日與周末、一天的時間分配、網內與網外、以及設備類型等，初步納入篩選的變數有超過500多個，而至最終模型變數已少於15個變數。

(七) 研究結果

另以資料基準日未來12個月為違約績效期，凡於違約績效期內出現30天以上之延遲還款紀錄者，即視為「違約者」，以資料基準日過去3個月內的表現為手機資料觀察期。最後，評估測試最佳結果可達KS值約為42。

(八) 結論與後續發展

在中國，有超過2/3的人是沒有信用紀錄的，但卻有超過2/3的人擁有手機。在信用評分模型中，加入手機資料進行信用風險評估，是能夠有助於未曾與銀行有往來紀錄的人取得金融服務，也就是解決「雞生蛋，蛋生雞」的問題。而從研究顯示目前中國於手機資料應用於信用評分上，仍有持續改善的空間，無論是在違約樣本或正常樣本，皆需要更多的訓練樣本。

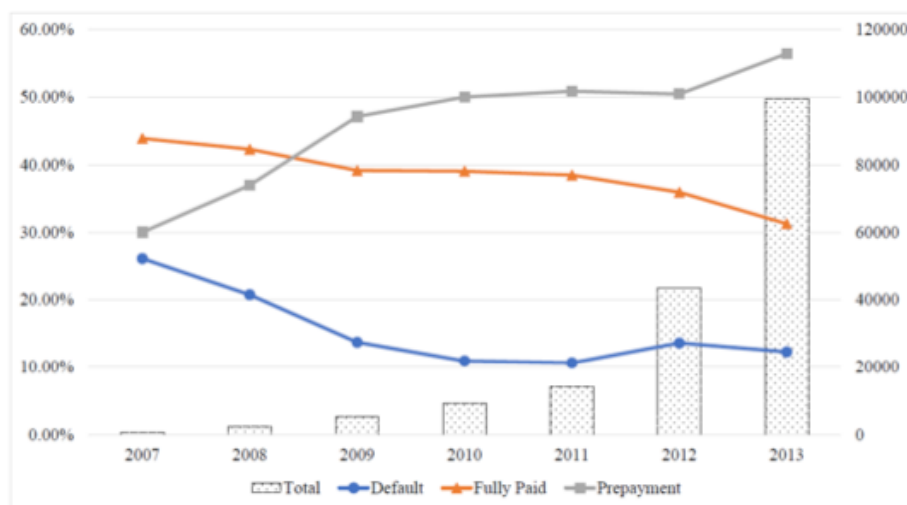
針對後續的研究方向，預期未來能將信用評分模型應用至更多消金市場上的情境分析，且能夠取得更多的替代性資料，如線上支付和公用事業費用資料，以細緻化目前模型。

二、Predicting Prepayment and Default Risks of Unsecured Consumer Loans in Online Lending

(一) 摘要說明

網路借貸 (online lending) 為借款人提供了快速融資服務，然而，平台業者和貸方會因借款人未依原本協議內容進行償還而面臨的兩個主要問題，一為提前償還 (early repayment)，一為違約未還款，此兩個因素都會影響借貸的獲利能力。對網路借貸而言，儘管借款人提前償還的比率較違約率更高，但是涉及提前償還的研究卻不是很普遍，在信用評分的文獻上，通常焦點都圍繞在違約風險的部分。此研究使用多變量羅吉斯迴歸 (multivariate logistic regression) 模型來預測潛在「提前償還」和「違約」風險的機率，並以14萬筆無擔貸款人之借貸資料，證實此兩種事件分別有其獨特的模式，同時汲取先前金融危機的教訓得知，評分模型會受到經濟情況的變化而產生不同的影響程度，因此模型亦將系統性風險納入考量，嘗試將總體經濟因素加入建模，從樣本外驗證已經顯示此方法能夠準確地預測提前償還和違約機率，最後，此研究強調面對發展迅速的網路借貸市場，提前償還相關之法規應有明確訂定之必要性。

圖 2：各年度下 P2P 網路貸款總金額與貸款結果趨勢圖



(二) P2P網路借貸

P2P網路借貸是一種藉由網際網路，投資人提供個人或企業貸款資金，借款人直接從投資人取得融資的服務管道。由於P2P網路借貸公司提供這些服務完全是以網路（online）運作，因此，相較於傳統金融機構，他們能以較低的成本營運，提供更為經濟的服務，即使P2P網路借貸公司已將借貸平台和借款人信用審查等相關費用納入考量，但往往投資人仍可取得高於銀行儲蓄和投資商品的報酬率，而借款人亦能用較低的利率借款。

圖2為各年度下中國P2P網路借貸市場的趨勢圖，其中長條圖部分為2007年至2013年每年度P2P網路借貸總金額，顯示P2P網路借貸於近幾年有明顯大幅度的一個增長，造成此成長其中一個主要的原因，乃是由於機構投資者也開始進入這個市場；進一步再觀察借貸結果，將之分為「全額繳清（fully-paid）」、

「提前償還（prepayment）」、「違約（default）」三種情況，圖2直線趨勢圖部分為2007年至2013年每年度P2P借貸結果佔比，從貸款結果中不僅發現P2P違約風險偏高，同時亦可看出P2P提前償還的比率更較違約風險高，且在近年尚有增加趨勢。

(三) 提前償還

在借貸市場上，一旦借款人提前償還，貸方則會遭受利息損失，而過去關於提前償還的研究文獻中，多數焦點是以消金市場之擔保品抵押貸款為主軸，如Ciochetti et al. (2002)，Ciochetti et al. (2003)及 Pennington-Cross (2010)有關於房屋貸款之研究，或2004年 Heitfield and Sabarwal於次級車貸上的討論。另外，欲探討提前償還之相關研究，尚涉及兩項議題，一為選擇權理論（Option Theory），二為競爭風險（Competitive Risk），以下為講者之簡要介紹說明。

在選擇權理論部分，參考Deng等(2000)³發表於Econometrica之文章，該文章主要是探究房貸市場的選擇權理論，研究發現當不動產市值超過其原有價值時，房屋持有人會執行買權，也就是提早還款，同時再進行融資抵押；當不動產市值低於其原有價值時，房屋持有人會執行賣權，選擇房貸違約，另外，該文章提出交易成本是執行買賣權決策上一個重要的影響因子。在一般商業抵押貸款時，債權人通常會針對提早還款的情形，訂定相關罰款，防止貸款人進行再融資（Steinbuks, 2015; Varliand Yildirim, 2015）。然而，在網路借貸市場上，提前償還一般並無收取額外費用，由於借款人沒有交易成本的考量，以至於提前償還的比率相對很高。

在競爭風險部分，競爭風險是在討論「造成觀察事件終止之原因」時，一種較為活潑的定義說明，也就是指所有可能會導致觀察樣本退出（the exit of an account）的事件間處於彼此互斥的狀態，更為白話解釋，是指在研究人員針對某項有興趣的研究事件時，凡是所有可能會導致觀察樣本退出觀察名單的事件發生後，會導致該觀察樣本排除於觀察名單，或是後續其他所有可能的事件皆不可能發生。競爭風險很常見於研究死亡率的議題上，在信用評分方面，違約事件，客戶流失，提前償還和結清帳戶都將導致借款人停止使用現有帳戶。

(四) 研究方法

此研究採用之統計方法為存活分析（Survival Analysis），並使用多項式羅吉斯迴歸（multinomial logistic regression）模型建模，多項式羅吉斯迴歸用於模型結果屬於間斷型選擇變數（discrete choice variable），基本模型假設為在一觀察期間內，給定觀察樣本下，事件間互相獨立：

$$\ln \left(\frac{P(d_i=j|x_i)}{P(d_i=0|x_i)} \right) = \beta_j^T x_i, j=1,2$$

$$P(d_i=0|x_i) = \frac{1}{1 + \sum_{j=1}^2 \exp(\beta_j^T x_i)}$$

$$P(d_i=j|x_i) = \frac{\exp(\beta_j^T x_i)}{1 + \sum_{j=1}^2 \exp(\beta_j^T x_i)}$$

最後，模型找出各事件預測值：

$$\hat{j} = \arg \max \{ \hat{P}(d_i = j), j = 0, 1, 2 \}$$

(五) 研究資料

在研究資料上，以P2P網路借貸下，14萬無擔貸款之個人戶為研究樣本，並依違約率由低至高依序分為A至E五個等級樣本，表1為此五個等級下三種借貸結果「全額繳清」、「提前償還」與「違約」之樣本分配表，由表1中可以看出，由A等級至E等級，違約率依序上升，然而提前償還的比率卻為依序遞減，在此意謂著，違約風險越低者，提前償還的比率有越高的傾向。

3 Deng, Y., Quigley, J.M., van Order, R., 2000. Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options. Econometrica68, 275-307。

表 1：各等級下不同借貸結果之樣本分配表

	A	B	C	D	E	總計
違約	5.32%	10.28%	15.68%	20.39%	23.59%	17,449
全額繳清	36.36%	33.26%	29.60%	27.95%	26.63%	45,074
提前償還	58.31%	56.45%	54.71%	51.66%	49.78%	78,082
總計	27,414	55,969	33,615	18,995	4,612	140,605

(六) 研究結果與發現

接下來，由三種不同的面向，即「借款人特性」、「借貸特徵」與「總體經濟變數」，分別選取出適合的模型變數，最後觀察各模型變數估計值與顯著程度，結果如表2所呈現，從表中不難發現，預測違約與預測提前償還之模型結果並不相同，部分變數於兩者預測效力沒有出現一致性，如6個月被查詢紀錄對於預測違約有用，在提前償還上則無顯著結果，又例如借款利率的增加會導致違約風險增加，但導致提前償還的可能性增加幅度更多。

表3為模型結果在訓練組與測試組下，實際結果與預測結果交叉分析表，從表3中得知，無論是訓練組或測試組，提前償還的準確度皆達87%左右，高於違約與全額繳清之準確度，另違約部分亦有不錯的預測效果，其準確度約有65%左右；表4為五個等級下預測結果準確度分析表，在表4各等級下觀察，得知提前償還之準確度在各等級下皆表現良好，最低為「A等級」的86.06%，最高「B等級」的88.12%，從表3與表4可知在模型準確度上，預測違約與預測提前償還皆有相當程度的效力，但是相對於違約，本模型對於提前償還的預測效果更為顯著。

(七) 結論

依據上述研究結果，此研究最後分析歸納以下三點結論：

- (1)在已知借款人特性、借貸特徵與總體經濟因素之條件下，使用多項式羅吉斯迴歸模型去預測授信借貸之結果，能夠準確地預測「提前償還」和「違約」兩者，其中「提前償還」在預測表現較「違約」表現好；
- (2)研究發現較高的借款利率不僅意味著會有較高的違約機會，同時也會增加提前償還的可能性，這是因為借款人並不想負擔高額利息，而在借款人特性方面，像是負債收入比（debt-to-income ratio）及FICO 評分，對於「提前償還」和「違約」兩種結果表現皆有顯著性地影響，其中FICO評分越高，借款人提前還款的機會越高；在總體經濟因素上，即「GDP成長率」、「聯邦基金利率」及「個人破產率」三項，對於「提前償還」和「違約」之發生皆有足夠的影響程度；
- (3)近幾年，「發薪日貸款」在市場上正有快速發展的趨勢，由於此類貸款有高昂的年化利率（Bhutta,2014），相較之下，目前P2P網路借貸亦屬於一種低成本短期借貸，在產品設計上多是以36期或更久的還

款期間為主，但因為P2P網路借貸對於提前償還並沒有訂定出違約金等相關罰則，在這種情況下，可能會導致借款人以P2P網路借貸作為發薪日貸款的另一種借貸選擇，並以提前償還獲取套利機會（Arbitrage Opportunity）。因此建議投資人應特別留意此類型風險，P2P網路借貸業者亦應考慮

對於「提前償還」收取違約金，以彌補在借貸組合中可能發失的損失金額。以美國P2P網路借貸著名業者LendingClub為例，該平台業者會重新包裝其借貸組合後，再行售予其他金融機構，因此為避免再發生像次貸危機這類金融巨災，合適的定價對於投資人與主管機關而言，具有相當程度之重要性。

表 2：羅吉斯迴歸模型預測結果⁴

	全額繳清vs違約			全額繳清vs提前償還		
	係數	誤差	p-value	係數	誤差	p-value
Intercept	18.043***	0.285	<.0001	0.069	0.179	0.6999
借款人特性						
負債收入比	-0.022***	0.002	<.0001	-0.005***	0.001	0.0001
工作時間	-0.001	0.005	0.8169	-0.027***	0.003	<.0001
房屋所有權-貸款	0.264***	0.061	<.0001	0.123***	0.034	0.0002
房屋所有權-租賃	0.200***	0.060	0.0008	0.013	0.034	0.7015
動用信用額度	-0.027***	0.005	<.0001	-0.030***	0.003	<.0001
總信用額度	0.035***	0.002	<.0001	0.027***	0.001	<.0001
循環金額使用率	-0.718***	0.083	<.0001	-0.515***	0.046	<.0001
法院公告紀錄	0.195	0.156	0.2108	0.042	0.091	0.6457
FICO評分	-0.025***	0.000	<.0001	0.006***	0.000	<.0001
稅收留置權	0.098	0.177	0.5799	0.064	0.107	0.5511
12個月內壞帳紀錄	0.315*	0.125	0.0116	0.449***	0.131	0.0006
破產紀錄	0.477**	0.166	0.0041	0.384***	0.097	<.0001
6個月被查詢紀錄	0.128***	0.017	<.0001	0.006	0.010	0.5283
授信特徵						
利率	0.229***	0.015	<.0001	0.298***	0.010	0.0001
還款收入比	24.200***	0.318	<.0001	-1.015***	0.215	<.0001
B等級	0.227**	0.082	0.0058	0.317***	0.041	<.0001
C等級	0.339**	0.124	0.0061	0.410***	0.066	<.0001
D等級	0.503**	0.167	0.0027	0.439***	0.092	<.0001
E等級	0.705**	0.222	0.0015	0.547***	0.126	<.0001
總體經濟因子						
GDP成長率	-1.318***	0.031	<.0001	-0.910***	0.019	<.0001
個人破產率	-2.298***	0.258	<.0001	-2.277***	0.144	<.0001
聯邦基金利率	-0.183***	0.003	<.0001	-0.163***	0.001	<.0001

4 本表*, **, ***分別代表係數顯著水準達5%, 1%, 0.1%。

表 3：訓練組與測試組於實際結果與預測結果交叉分析表

		實際結果				準確度
		違約	全額繳清	提前償還	總計	
訓練組						
預測結果	違約	7376	820	3025	11221	65.73%
	全額繳清	1062	17137	11069	29268	58.55%
	提前償還	838	5162	41849	47849	87.46%
	總計	9276	23119	55943	88338	75.12%
測試組						
預測結果	違約	3756	427	1553	5736	65.48%
	全額繳清	514	8729	5850	15093	57.83%
	提前償還	410	2636	21558	24604	87.62%
	總計	4680	11792	28961	45433	74.93%

表 4：各等級下不同借貸結果預測準確度分析表

	A	B	C	D	E	總計
違約	53.07%	63.24%	66.74%	71.32%	67.71%	17,449
全額繳清	54.86%	57.74%	58.19%	61.47%	64.78%	45,074
提前償還	86.06%	88.12%	88.05%	87.98%	86.50%	78,082
總計	72.59%	75.19%	75.32%	76.63%	75.98%	140,605

三、Personality and risk: a new chapter for credit assessment

(一) 摘要說明

針對信用市場上，缺乏信用歷史紀錄或是信用資料不足（thin-file）的客群，研究人員證明個人心理特質和信用行為之間是存在某種關聯性，並依此開發一套完整的心理量測問卷，希冀藉由測驗問答及心理量測的評分模型，即時於線上有效地評估個人信用風險高低，此一方法不僅有助於信用資料不足之客群取得信用進入信用市場，還能为金融機構開發出更多目

標客戶，幫助組織決策者更進一步瞭解客戶與其企業前景，使用更好的工具與方法對貸款申請人進行風險評估，最終以提高客戶終身價值。

(二) 研究目標

本研究目標主要是建構無擔保短期貸款的申請人評分卡（application scoring），希望信用資料不足之貸款人，能藉由此評分模型之判斷，識別出其中信用良好之客群，獲取市場競爭優勢。

圖 3：心理量測評分 - 各分數區間之違約率



此研究方法使用心理測量模型，該模型以1萬個現有客戶為其訓練組樣本，違約定義為第3個月出現1個月以上延遲還款紀錄者，並經由獨立驗證後，最終結果發現使用心理測量結果的Gini係數上升至45%，意謂著此一模型對於缺乏信用歷史紀錄或是信用資料不足之客群，可作為一有效的風險評估工具。

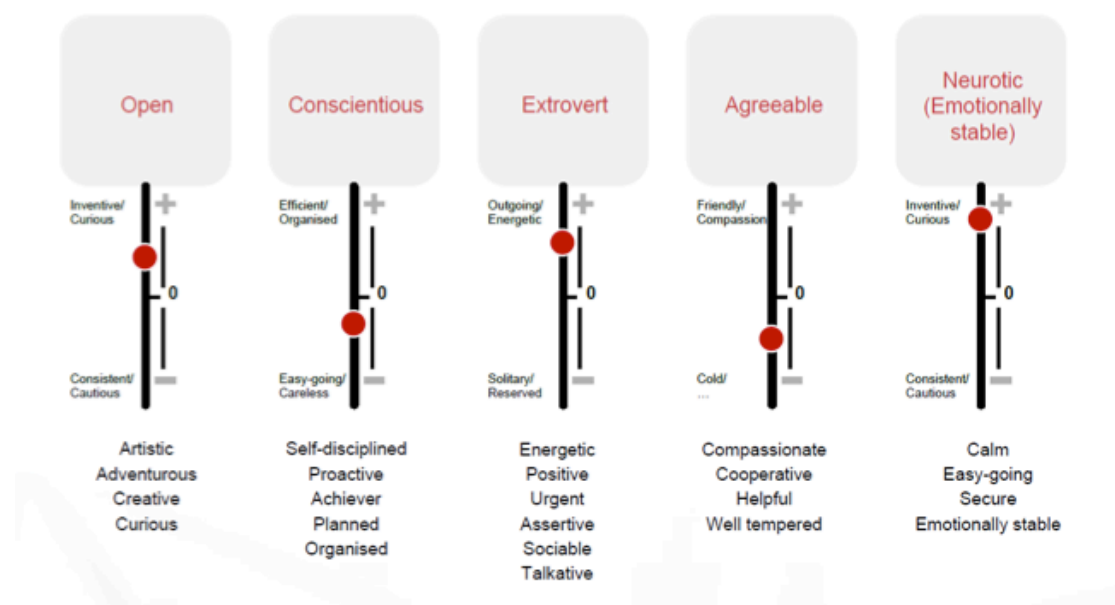
(三) 心理測量方法

由於本研究主要是使用心理測量方式建構模型，因此研究人員簡要介紹一常見的心理測量方法。一般風險評估所使用之心理測量方法，多數是使用心理學領域所稱「Big 5」人格特質理論，即「經驗開放性 (Openness to experience)」，「嚴謹性 (Conscientiousness)」，「外向性 (Extraversion)」，「親和性 (Agreeableness)」與「神經質

(Neuroticism)」，「Big 5」通常縮寫為OCEAN，每一種所代表的人格特質如下所描述，另可參見圖4：

- (1)經驗開放性：具有藝術的、愛冒險的、具創作力的、具有好奇心的；反之則為無藝術性的、實際的、興趣少的；
- (2)嚴謹性：自律的、主動積極的、追求成就的、有規劃的、有組織的；反之則為不負責的、輕率的；
- (3)外向性：有活力的、正面的、果敢的、愛好社交的、喜歡說話的；反之則為沈默話少的、冷靜的、冷淡的；
- (4)親和性：富於同情心的、樂於合作的、熱於助人的、溫和的；反之則為冷漠的、猜忌的；
- (5)神經質：冷靜的、隨和的、有安全感的、情緒穩定的；反之則為不穩定的、焦慮的、暴躁的。

圖 4：五大人格特質特徵



人格特質之相關研究與探討，至今已發展30年以上，此方法目前已能相當全面完整地評估個人特質。利用上述「Big 5」人格特質理論，研究人員嘗試找出人格特質與信用行為的關聯性，例如對於高度感覺尋求者（sensation seekers）而言，因對冒險與刺激的需求很高，因此相較於低度感覺尋求者，會傾向承擔較大的風險；個性衝動者，通常亦屬於風險較高的人，因為他們較不會去評估資訊或是衡量其他所有可能的選擇，常憑著一時興致，很快地下決定，例如衝動購物，早買早享受等；另外，現在時間（present-time）導向的人（也就是high time preference，高時間偏好）因較專注於現在，喜好當下享受與花費的感覺，較不會考量至未來，因此攸關未來之財務決定，瞭解個人於時間（Time horizon）的觀念即為相當重要的部分。

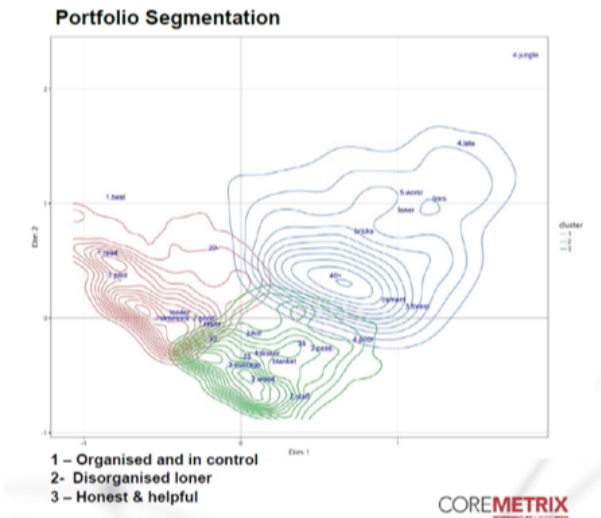
接著，針對「Big 5」人格特質設計測驗問題，受測者於完成測驗後，系統會依此進行五大人格特質分析，同時研究人員亦有發現其他幾種新的風險特徵：如金融誠信（financial integrity），樂觀偏見（optimistic bias），時間折價（temporal discounting），因此每個受測者除使用OCEAN，亦有使用其他面向進行評估。

此研究在問卷呈現上，研究人員以遊戲化的方式設計心理量測問題的介面，使受測者易於與系統互動，同時利用圖片的表達方式，使受測者能直觀且快速的回應，另針對問卷使用圖像部分特別說明，此研究認為在風險評估上使用圖像，能用來量測受測者實際的情緒穩定性，以確實瞭解申請人在不利於己身的情况下會有何種反應。

(四) 機器學習與評分結果

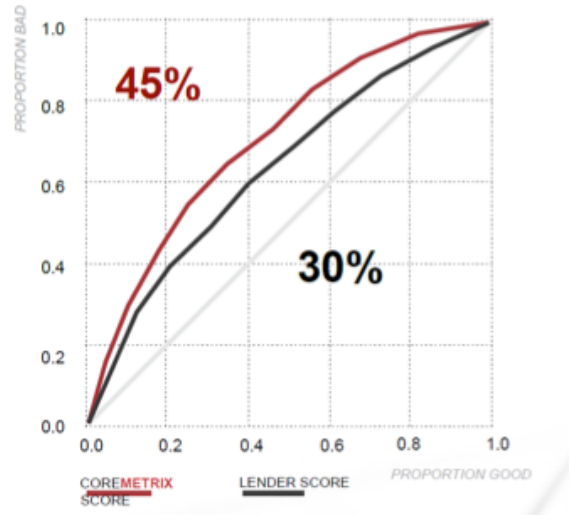
為增加模型區隔能力，找出更加有效的區隔變數與模型變數，研究人員嘗試使用目前較為新興的機器學習技術進行模型建置，如神經網絡（Neural networks），分群演算法（Clustering algorithms）及隨機森林（Random forests）。

圖 5：以機器學習方法所得之模型區隔圖



原先信用資料不足之申請人評分卡，所使用的資料為傳統信用資料，即申請人基本資料與信用報告機構的資料，其模型最佳預測力在Gini值為0.3（圖6 Lender Score），而此研究再納入心理測量資訊後，經反覆不斷地以機器學習方式進行各組合之嘗試與測試後，最後得出的模型，其預測力的表現更為穩健，相較於原有模型，在Gini值上提升50%（圖6 COREMETRIX Score）。

圖 6：傳統模型與心理量測模型之模型區隔力比較圖



講者最後以汽車保險業為實例說明，表5為傳統模型與心理量測模型理賠率交叉分析表，其中最下方灰色區塊，代表保險公司依據行內客戶評分分數之高低，區隔成三個評分客群，其平均理賠率分別為9.7%，7.9%，6.1%；在灰色上方其他顏色之區塊，代表上述三個評分客群，各自再依其上述研究人員建置之心理量測評分高低排序後，等比率分為三客群，計算其平均理賠率，藉由此一交叉分析觀察，可以發現原先理賠率最高的一群（即9.7%的客群），能更細緻化地切分為15.5%，8.1%，5.5%三群，而像是理賠率介於中間的客群（即7.9%的客群），有一部分客群於使用心理量測評分後，平均理賠率下降至4.5%，由此說明，運用心理量測評分確實能將原先風險評估模型更加以細緻化地進行區隔與分析。

四、Personality and risk: a new chapter for credit assessment

表 5：傳統模型與心理量測模型理賠率交叉分析表

Psychometric Score	Top	5.5 % 33.3%	4.5 % 33.3%	3.2 % 33.3%
	Middle	8.1 % 33.3%	6.7 % 33.3%	5.1 % 33.3%
	Bottom	15.5 % 33.3%	12.4 % 33.3%	9.9 % 33.3%
Column Average	9.7 %	7.9 %	6.1 %	
		Bottom	Middle	Top
		Insurer Score		

(一) 摘要說明

許多統計調查結果顯示，總體經濟數據於資料調整及公佈時效性問題上，如果未經妥善地處理，可能會導致總體經濟預測上有所偏誤，而對於研究人員來說，使用信用報告機構（Credit Bureau，以下簡稱CB）的信用資料，並不會涉及上述資料調整或公佈時效性等資料修正的問題。此研究主要是考量使用此類即時資料（real-time data）作為研究總體經濟數據的來源，也就是欲瞭解個人信用報告上的信用資訊，是否可納入總體經濟預測模型，提高模型準確度。具體來說，此研究將評估個人信用統計資訊（aggregate consumer credit information）對於消費支出的預測效果。

在過往的文獻研究中，曾有以消費者使用度很高的電子支付資料，作為大數據，探討該

資訊是否可作為總體經濟活動一預測指標。然而，到目前為止，以個人信用報告上的信用資訊所衍生之個人信用統計資訊，尚未被用來預測總體經濟變數。

為研究個人信用統計資訊在消費支出上的預測能力，此研究先設定一基準預測模型（baseline forecasting model），以模擬方式產生樣本外即時資料，預測下一季消費支出，接著嘗試加入CB變數於該基準模型中，最後比較兩個模型之預測能力，即比較基準預測模型與加入CB變數的模型，在均方根預測誤差（root-mean-square forecast errors，以下簡稱RMSFE）之高低。

研究初步結果證實CB資料能改善消費者支出之預測能力，然而其改進程度尚有待加強，其中延遲繳款紀錄似乎仍是預測消費支出最具指標性意義的預測變數。

(二) 研究動機

研究人員主要是欲探討在CB資料中，個人信用統計資訊對於總體經濟指數「消費者支出」是否具有預測能力，近年來雖然有相當多關於大數據的應用與研究發展，但尚未有人利用CB上的大數據資料預測總體經濟變數。

從以往研究文獻中，經濟學家在CB資料對總體經濟之關係上較無相關研究探討，相對地也較沒有使用樣本外資料進行測試（out-of-sample forecasting），雖然相關研究主題較少，但研究人員尚有些許發現，如Duarte, Rodrigues, and Rua (2017)發現ATM及POS資料對於葡萄牙消費有預測效果，或Gill, Perera, and Sunner (2012)亦使用交易資料改善澳洲消費者支出的預測結果等。

(三) 個人信用資料

個人信用資料是從個人信用報告中彙整編制而成的，並且是由三個主要信用報告機構（consumer credit reporting agencies，簡稱CRAs）所維護，在美國是指Equifax，Experian和TransUnion。個人信用報告裡的資訊一般可視為一種及時資訊，也就是說，研究人員使用CB資料庫中的這些歷史資料，皆是未被修正過的，這是資料最早進資料庫時存在的資訊。

此研究以為使用CB資料進行總體經濟之預測，具有以下五項優點：

(1)CB資料能依不同觀察單位進行歸戶計算，產生相關統計資料，如依帳戶別歸戶（account-level），依當事人歸戶

（consumer-level），或是依家庭歸戶（household-level）；

(2)在變數分析和區隔分析上，CB資料能依變數設計需求使用資料快照（snapshot）資訊及長期歷史資訊，且可依貸款人基本資料進行區隔分析；

(3)CB資料具有時效性（Timeliness），也就是說，由於一般CB資料庫每月會更新一次，當信用市場的發展趨勢有所改變時，每月報送更新的資料較能適時地反應出信用市場的變化；

(4)瞭解當地（不同地區別）信用情況，意旨CB資料可以根據消費者個人的居住地點來查看該貸款所屬的地理位置；

(5)消費者於信用使用之變動特性，如：消費者在進行與債務支付（支付順序）相關之決定時，通常與當地經濟狀況和整體國家經濟狀況有關；消費者對於未來總體經濟狀況的預期心理，常表現於還款行為與還款決策上，因此還款資訊可能有助於提高總體經濟預測力的準確度。

(四) 研究資料

此研究所使用之資料，在CB資料方面，是用TransUnion某地區下的資料，並歸戶衍生數百個信用風險特徵變數，同時亦以地區別等資料為區隔計算相關變數資訊。多數信用風險特徵變數由1992年第一季到2015年第四季皆有資料。在消費者支出方面，為取得及時資料，由費城聯邦儲備銀行（Federal Reserve Bank of Philadelphia）提供原始資料。

圖 7：信用變數與消費者支出趨勢圖

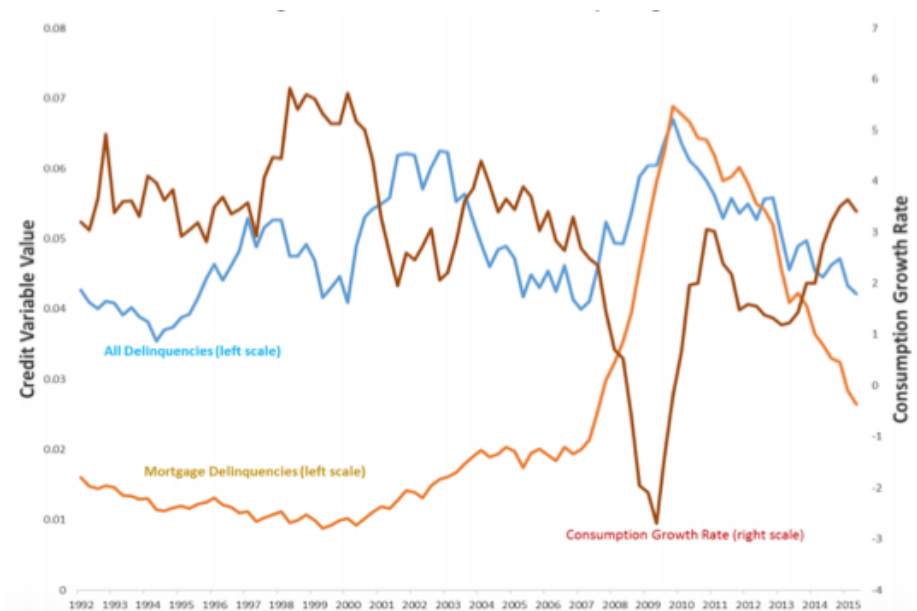
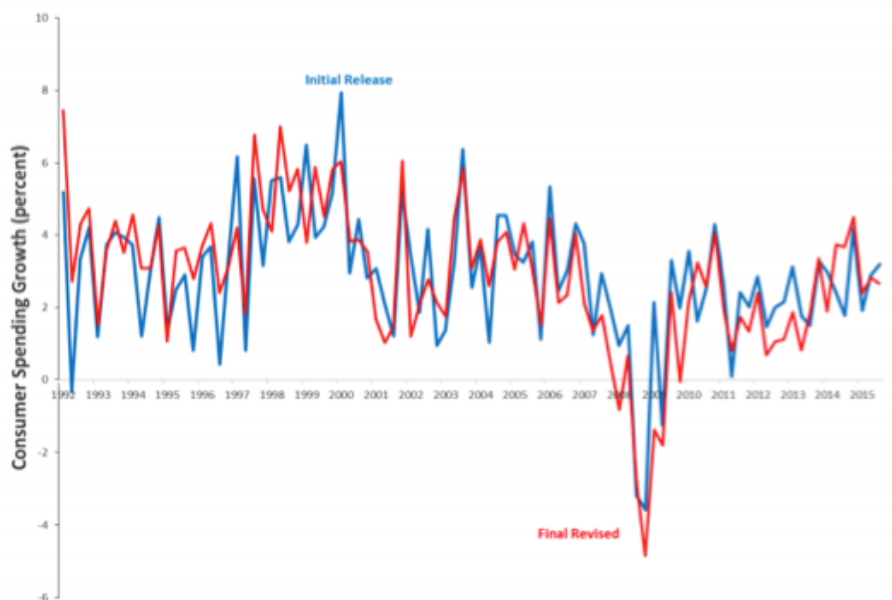


圖 8：消費者支出成長率修正前後趨勢圖



(五) 實證方法與研究結果

研究實證方法，是使用Croushore（2005）以Bram-Ludvigson（1998）模型為基礎所修正之模型，建構基準預測模型，其模型如下：

$$\Delta c_t = \alpha_0 + \sum_{i=1}^4 \alpha_1^i \Delta c_{t-i} + \sum_{i=1}^4 \alpha_2^i \Delta y_{t-i} + \sum_{i=1}^4 \alpha_3^i \Delta r_{t-i} + \sum_{i=1}^4 \alpha_4^i \Delta s_{t-i} + \varepsilon_t$$

其中， c 代表實質消費支出， y 代表實質收入（可支配收入或總收入）， r 代表3個月國庫券名目利率， s 代表實際股票價格，實值=名目價值/消費物價指數。

以上述模型，使用模擬產生的即時樣本外測試資料進行消費支出預測，另再嘗試加入1個以上的CB變數，計算基準預測模型與加入CB變數模型，最後觀察RMSFE是否有顯著下降，且分別計算出兩模型RMSFE，測試預測能力改善程度。

其中，在建模中所使用之消費者支出資料，此研究嘗試使用四種不同版本的及時資料，以觀察是否預測表現會有差異，包含初版未調整資料（I），年度調整資料（A），準基準資料（B）與最終調整後資料（F）。

而建模加入之CB變數，主要可分為被查詢紀錄，延遲還款，負債餘額，額度使用率，並再測試各CB變數之變動率，共計10個變數：

- (1)被查詢紀錄：90天內被查詢之消費者比率及其比率變動率。
- (2)延遲還款：借款人有60天以上延遲還款之比率及其比率變動率；
- (3)房貸延遲還款：房貸借款人有60天以上延遲還款之比率及其比率變動率。
- (4)負債餘額：總負債餘額與其變動率。
- (5)額度使用率：平均額度使用率與其變動率。

最後，比較各種結果之預測力改善，看是否CB資料有使預測力增加或減少，其中使用RRMSFE為觀察指數，其公式如下：

$$RRMSFE = \frac{RMSFE(\text{with credit bureau variable})}{RMSFE(\text{baseline})}$$

當RRMSFE > 1，代表加入CB變數使模型預測力更差；當RRMSFE < 1，代表加入CB變數使模型預測力更好。觀察p-value大小，當p-value < 0.05，視為該模型具備統計顯著性。

結果發現僅延遲還款類的變數能些微地改善模型預測力，但沒有任何CB變數有達到統計顯著性（即p-value皆超過0.05）。

最後，講者進行敏感度分析，嘗試改變預測期間的時間長度，同時考量是否涵蓋金融風暴的時間，以下述四種情境重新進行預測並得出結果：

- (1)縮短預測期間（從2007年第1季至2015年第3季）：模型結果雖能使RRMSE降至1以下，但不具有統計顯著性；
- (2)預測期間到2007年第4季為止（從2002年第1季至2007年第4季）：模型結果沒有任何統計顯著性，甚至有些比原本的預測力更差；
- (3)延長經濟衰退前的預測期間（從1997年第1季至2007年第4季）：模型結果發現所有CB變數預測能力表現都變差，借款人借款相關變數使得預測力顯著變差；
- (4)延長經濟衰退前的預測期間，但包括經濟衰退（從1997年第1季至2015年第3季），結果與方案3相同。

綜上所述，CB資料僅能些微改善模型預測能力，但依然不具有統計顯著性，在某些情境下，甚至還會使模型預測表現更差。

因此研究人員最後提出兩個結果：一為從分析報告證實，CB資料與消費者支出有部分關係存在，然而，在預測力上卻是無法有統計顯著性的，其次發現是延遲繳款紀錄的資料對於預測消費者支出似乎是最有預測可能性的。

參、未來建議

本次參與愛丁堡信用評分研討會，透過研討會能瞭解目前信用評分領域之最新發展趨勢與熱門議題，同時亦提供與產學界專業人士經驗分享與交流的機會，並知悉在大數據潮流與金融科技產業發展推進下，目前金融業界所提供之服務與創新之道，最重要的是，從近幾年本中心與各國信用報告機構交流或參訪之經驗，同時積極參與研討會之過程，目前本中心在未來發展與考量上，無論是針對普惠金融下信用資料不足者之議題，金融科技之於信用報告機構生存與發展，或是當事人權益之維護與提升上，與其他國家之信用報告機構，皆具有一致的方向與看法，以下茲就本次研討會彙整出相關建議，提供本中心未來營運發展參考之用。

一、致力於替代性金融之取得與應用

2011年由世界銀行與國際清算銀行（BIS）共同組成之信用報告體系專案小組所頒布之「The General Principles for Credit Reporting」，提出信用報告機構之五大指導原則，包含「資料蒐集」、「資料處理」、「治

理與風險管理」、「法規環境」及「資料跨國傳遞」，其中在「資料蒐集」部分，強調資料的正確、即時及完整性，且應有系統地蒐集各種資料來源，包括零售業、小型企業、微型金融、信用卡、保險公司、電信公司與公用事業公司等資料。從本次研討會中，不難發現普惠金融相關議題從過去幾年至今，一直是各國金融發展相當重視的主題，針對無信用紀錄者，不分學術界或產業界，皆試圖發展各種替代性資料之應用與創新，以提升信用風險評估之有效性，增加整體信用取得性，致力於達成普惠金融之目標，如中國學者對於手機資訊之應用分析，或是英國業界於金融產業使用心理量測之研究探討，不外乎都是為了普惠金融。

本中心近期曾針對過去無信用紀錄之企業與消費者，試圖由本中心資料庫中找出對信用資料不足者之風險評估方式，研究發現因缺乏信用資訊，僅有相當少數的基本資訊變數及查詢紀錄變數可供為參考指標，然而要進一步發展成如J10，J20與J21的評分模型，尚不足以建構成為有效穩定之信用風險評估工具。在國際上已有相當豐富的實證研究顯示，替代性資料對於在信用報告機構中屬於無信用資料或信用資料較少之客群，於取得信用上，有顯著之效果，世界銀行與「2017經商環境報告」之研究亦顯示，信用報告機構經由取得融資公司、交易債權人、零售業及公用事業之信用資料，可顯著提升其資料涵蓋率。

本中心於蒐集替代性資料上，提升取得信用之程度仍有相當空間可改善，基於法規法令限制，現階段除有「中小企業融資服務平台」以介接方式提供會員機構查詢企業水電資料外，尚無法蒐集各類替代性資料等非傳統信用資料；由於一般以為消費者手機資訊，即手機還款資訊與通聯紀錄，與授信借貸還款行為具備一定關聯性，同時手機資訊可視為消費者經濟活動指標之一，是眾多替代性資料中是較適合納入信用評分模型、有助於未曾與銀行有往來紀錄的人取得金融服務的資料，如美國 Cignifi 使用手機資訊為沒有信用歷史資料者提供信用評分，在相同違約風險下，其核准率提升至25%，目前亦為各新興市場提供此類信用評分服務⁵，而在筆者參訪英國信用報告機構之經驗中，亦有蒐集手機還款等資訊並考量至信用評分；同時本中心內部研究比較各種替代性資料之普及性、代表性及即時性，以電信繳款資料來看，在普及性方面，國內使用智慧型手機或平板電腦人口估計突破1,600萬人，從代表性來看，電信資料多為申請者本身以雙證件臨櫃申請、使用並繳費，資料連結度相當高，最後從即時性分析，電信繳款帳單是每月計收一次，較為即時，因此相較於水、電、瓦斯繳款資料，特別是從資料普及性來看，足以相信

電信繳款資料之蒐集對於本中心提升台灣普惠金融應有很大的助益與貢獻。

二、針對非傳統信用資料，嘗試以合作方式擴大資料蒐集範圍

經由本次研討會與英美等國之信用報告機構相關人員交流機會，可知目前在資料加值的面向上，跳脫以往傳統信用報告機構之思維，對於非傳統信用資料，並未直接向資料提供者進行蒐集並納入自身資料庫，如email、護照或是生物辨識等資料，並非由資料提供者直接報送至其資料庫，而是以合作方式，從各個合作機構中，擷取當事人相關資訊，或是由合作機構提供當事人服務。

本中心近幾年持續努力充實信用資料庫，希冀能充分協助會員機構更加瞭解自身客戶（Know Your Customer, KYC）與風險控管，目前陸續蒐集開發各類資訊，例如去（106）年12月環保署開始提供本中心企業重大污染裁罰處分資訊，或目前與財政部研議之財政資訊中心介接建置個人所得及財產資料調閱平台等，另外，政府近幾年積極推動開放資料與金融科技，已有許多開放資料可供查詢使用，像是去年7月台灣集中保管結算所建置的防制洗錢名單資料庫，或是國家發展委員會與金融監

⁵ 資料來源：「徵信與大數據」，劉新海（2016）。

督管理委員會於先前共同召開之「金融科技發展所需政府資料交換運用研商會議」，針對各業者所提出之需求（電話詐騙黑名單等），相信未來因應整個金融科技趨勢的發展下，本中心能夠有更多資料可進行加值運用，而如能在法規允許的情況下，嘗試以不同合作方法獲取非傳統信用資料，在有一定的資料品質下，如能將這些資料運用於信用評分模型，擴大評分範圍，相信對更多消費者或中小企業取得資金方面，助益匪淺，亦或是以各類網路資訊並結合詐欺黑名單開發詐欺偵測模型，亦能提升會員機構的風險控管能力。

三、對於信用資料不足之民衆，提供更多信用知識及信用取得之教育

本中心目前提供「個人線上查閱信用報告服務」並積極推廣，服務內容主要為檢視與下載信用報告，並於今（107）年1月1日開放行動裝置查閱信用報告，對於眾多低頭族而言，行動裝置服務更為方便簡易，相信後續會有更多民衆查詢信用報告與信用評分，建議本中心除了提供信用報告與信用評分相關知識與宣導外，針對信用資料不足之民衆，應思量採取較為實務之作法，幫助這些民衆取得信用，進入信用市場得以享用完整的金融服務，例如對信用資料不足者提供預約服務，以專人進行輔導諮詢等。

四、強化資料應用研究能力，提供會員機構客制化加值服務

面對金融科技的發展與提倡個資當事人權益的潮流下，各國信用報告機構於未來生存與發展皆為一個未知的變數，本中心目前雖為台灣唯一一家信用報告機構，但金融科技等網路世代的變遷相當快速，信用報告機構的功能及重要性很難不受到衝擊；本中心目前已定期提供會員機構整體授信違約資訊，中心官網開放資料專區亦有提供大數據資料庫與多項開放資料，建議後續應強化本中心資料加值運用之研究能力，協助會員機構有效利用大數據及開放資料建構指標或預測模型，或依不同會員機構之需求，提供單一客致化服務或合作方案，如結合本中心收入與負債資訊，以協助會員衡量消費者還款能力等更多決策應用。