政治科學論叢／第四十一期／民國 98 年 9 月／頁 151～176

以入選機率調整法修正調查推估偏差的成效評估

杜素豪*、羅婉云**、洪永泰***

摘　要

抽樣調查資料因樣本代表性失真而造成對母體推估偏差的補救辦法通常是採用加權處理，基本上以社會人口特徵爲依據。然而僅從樣本人口特徵的分布是否和母體相符來判斷樣本的代表性並不能保證樣本在認知、態度與行爲等主題變項在分布上的推論就不會有偏差。本文以總統選舉投票行爲的調查資料爲例，探討依據入選機率的次樣本分組（subclassification on the propensity score）所調整的電話調查結果在投票行爲推估方面的成效。首先利用二〇〇四年台灣地區社會變遷基本調查四期五次公民權組問卷資料，以樣本重抽法（bootstrapping）產生包含 20,000 案的擬母體（Pseudo-population），再從擬母體中以簡單隨機抽樣法抽取 200 案爲參考樣本，另外從一個典型的電話訪問調查資料中以分層隨機法抽出 800 案爲試驗樣本。兩套樣本組合成爲一套 1,000 案的新樣本。其次依據 Lee (2006) 的入選機率調整法（propensity score adjustment, PSA）進行電訪樣本「投票行爲」估計值的調整。整個流程進行 2,000 次的模擬分析。評估結果確認藉由次樣本分組所產生的入選機率調整法的確有降低推估偏差的功效。

關鍵詞：不完整資料、入選機率、入選機率調整法、投票行為

---

* 中央研究院人文社會科學研究中心調查研究專題中心副研究員，E-mail: suhao@gate.sinica.edu.tw

** 中央研究院人文社會科學研究中心調查研究專題中心研究助理，E-mail: lilaclo@gate.sinica.edu.tw

*** 臺灣大學政治學系教授，E-mail: ythung@ntu.edu.tw

收稿日期：98 年 1 月 13 日；通過日期：98 年 8 月 17 日

臺灣大學學術期刊資料庫

壹、研究背景與目的

訪問調查領域長久以來一直面臨調查資料代表性失真的問題，爲了避免造成對母體推估的偏誤，最常見的解決之道就是「加權」處理。一般作法是先就資料進行樣本代表性的檢定，針對人口特徵如地區、性別、年齡、教育程度、甚至職業、婚姻狀況等變數的分布與母體參數進行比對或統計上的適合度檢定，如果沒有通過則給予加權處理，以降低來自樣本代表性失真所產生的推估偏差。常用的加權方法有事後分層（post-stratification）、多變項反覆加權（raking）、或是前兩種方法的延伸加權法，這些方法隱含著一個理論假設，亦即所選擇進行加權的社會人口特徵變項足以透過擴大或縮小成功樣本中每一個案的權重來涵蓋未被適當納入此次調查的個案，進而得到可以代表母體的樣本（Kalton and Flores-Cervantes, 2003；Lee and Valliant, 2008）。

不過上述的加權方法無論用在面訪或電訪樣本均陸續被發現結果都不能完全解決抽樣調查推估的偏差（Kalton and Flores-Cervantes, 2003；洪永泰，2007）。其中常見的爭論是，加權的結果可能只有某些人口特徵資料接近母體參數，其他變數則不一定如此。例如劉從葦與陳光輝（2005）利用選舉研究的實證資料分析就發現在加權之後投票率與政黨得票率等行爲變項與母體參數相距較遠。李隆安與韓成業（2008）在利用多變項反覆加權的比較研究也發現加權效果事實上是有限的。這充分顯示從社會人口特徵的分布是否和母體相符來判斷樣本代表性是不夠的，還應該包括認知、態度與行爲變數的分布（例如洪永泰，2006），其主要的理由是那些訪問失敗的樣本個案或是對某些問卷題目不表示意見的受訪者很有可能是基於非抽樣因素或是意識形態上的理由而無回應，這些原因也許和人口特徵或其他因素有關，甚至存在著更加複雜的相互關聯，而複雜關聯尤其容易存在於與政治相關的訪問調查中。這種理由也可解釋沒被涵蓋在抽樣清冊中的個案很可能在社會人口特徵或與研究旨趣相關的態度上與成功樣本個案相距甚遠。

爲了突破傳統加權方法無法解決的問題，前人研究努力的方向包含從

臺灣大學學術期刊資料庫

問卷設計與考慮到認知態度或行爲的加權方法。在問卷設計方面，以投票意向調查結果的推估爲例，美國民意調查界通常使用複雜的問卷設計技巧加上有效的加權工具，不僅找出真正會去投票的受訪者也因此克服樣本代表性的缺失，例如 Traugott and Tucker (1984)；Colasanto and Mattlin (1987)；Brehm (1993)；Freedman and Goldstein (1996)；Dimock et al. (2001) 等的研究證實這些方法在美國是可行的。事實上，歷年來的美國選舉預測也都證明了美國的民意調查界確實有能力處理因爲涵蓋率和訪問失敗所造成的樣本代表性缺失。

在對認知、態度與行爲變項進行加權的研究中，Abramowitz (2005) 提出動態的政黨認同加權方法（dynamic party identification weighting），主要是利用多年且多次執行的民意調查資料（並非利用問卷設計技巧而蒐集的資料）中政黨支持比例結構所作之平均性的推估。資料分析的結果顯示此種動態加權法可以大幅度降低不同調查所隨機產生的抽樣誤差。不過，較嚴密的方法是依據具有明確之模式理論基礎的入選機率調整法（propensity scores adjustment, PSA），這個方法不僅考慮到人口特徵，也用到與推估變項有相關的認知、態度與行爲等非人口變項。此方法在早期是源於針對觀察性資料而發展的次樣本分組調整法，目的是降低觀察性資料因缺乏實驗設計中的控制組比對而產生的推估誤差，其基本原理是在有實驗組與控制組的資料中，利用模式分析觀察的特徵變項可預測某些個案是否包含於實驗組的機率值（propensity scores, PS），此機率值以配對、平均或其他方法分爲多個分群，每一分群均分別包含了一定數目的實驗組與控制組個案，透過比較實驗組與控制組在每一分群之權重的調整對實驗效果進行推估（Cochran, 1968；Rosenbaum and Rubin, 1983, 1984；Rosenbaum, 2005）。

此法的精髓在於將入選機率值可以平衡的（balancing propensity scores）分布在各分群，在推估實驗效應之時，相關共變數（預測入選機率的特徵變項）受到控制，於是每一分群中無論是實驗組或控制組的個案均具有相似的共變特徵，我們便可依此共通特性，而不是依賴依變項與共變數的關係的瞭解（亦即不需知道依變項的實際分配），作最後精確的推估（Zanutto, 2006）。此種方法已廣泛的應用在健康（羅健銘等，2007；D'Agostino, Jr.,

臺灣大學學術期刊資料庫

1998；Strümer et al., 2005）與收入（Czajka et al., 1992；Zanutto, 2006）相關的研究中。

將上述原理應用在解決調查低涵蓋率與無回應的樣本推估上的研究則是近幾年在國外熱烈討論的。稍有不同的是目前的作法是將人口變項與非人口變項一起納入統計模型之中，共同預測來自不佳的抽樣調查樣本（研究樣本）在相對於參考樣本（來自經完善抽樣結果且可產生接近母體參數的樣本）的一個特定的條件機率值（propensity score, PS），也就是本文所稱的入選機率值。應用入選機率調整法事實上需符合五個前提假設，分別是 (1) 要推估的依變項與入選機率值無關，(2) 研究單位是否被涵蓋或接受訪問不受到影響（例如電話戶中抽樣中的樣本單位的合作與否不受到其他家戶成員的影響），(3) 無論是研究樣本或參考樣本均來自非零的機率抽樣，(4) 納入模式中預測入選機率值的自變項足以代表未被納入的其他變項，以及 (5) 入選參考樣本或研究樣本的結果不會影響納入模式中的自變項（Lee and Valliant, 2008）。除了以上的前提之外，爲了能精準的推估，參考樣本的選擇也很重要。好的參考樣本需具備以下幾個條件，分別是 (1) 至少要比研究樣本還高的母體涵蓋率與回收率，(2) 比較理想的抽樣設計，(3) 透過較嚴密調查執行所回收具備高品質的調查資料，與 (4) 包含了與研究樣本相同且用模式分析入選機率的共變數（例如：社會人口特徵變項，家戶人數）（Lee and Valliant, 2008, 2009）。

國外多篇的研究論文，無論是透過模擬分析或是電訪、網路調查的比較研究，目的是在檢驗 PSA 是否可超越傳統加權方法，解決樣本代表性只考慮人口變項仍然失真的問題。Judkins et al. (2007) 的模擬分析結果顯示事後分層法、raking 與 PSA 後的估計準確度是接近的，但 Loosveldt and Sonck (2008) 比較網路調查使用一般事後分層與 PSA 的結果則發現 PSA 法效果較佳。針對電訪或（與）網路調查的分析（比較分析）結果則多建議採用 PSA 或綜合其他加權的 PSA。其中，電訪調查涵蓋率的研究指出 raking 後加上 PSA 比起只用 PSA 或 raking 在推估偏差上降低最多（Duncan and Stasny, 2001）。Lee and Valliant (2009) 在以網路調查進行的模擬研究也發現 PSA 搭配其他加權方法同時使用之後不僅可以有效的降低涵蓋率問題，也可降低 95% 的推估誤差。

臺灣大學學術期刊資料庫

在探討 PSA 如何可以消除網路調查涵蓋誤差的研究建議指出：一般只作人口或家戶特徵調整的加權方法是不足的，使用 PSA 可以納入其他非人口變項（如：態度），較能有效解決網路調查普遍的低涵蓋率問題（Dever et al., 2008）。這個建議也在針對網路調查與電訪調查的推估結果的研究中被提出，而且更加強調應該納入相關的態度變項（Schonlau et al., 2004；Schonlau et al., 2007）。該納入哪些變項則是有待未來繼續探討的，不過可以確定的是選擇納入的變項會因研究主題不同而有差異，換言之，產生入選機率的分析模式中放錯變項時推估誤差絕對會上升，不僅違背使用 PSA 的前提假設且誤導 PSA 的使用原則（Drake, 1993）。

相較於國外的相關研究，台灣地區訪問調查的樣本代表性失真現象雖然很早就有被討論（吳統雄，1984；蘇蘅，1986；隋杜卿，1986；洪永泰，1996a；黃毅志，1997），但大都僅止於觀察和比較，指出問題所在而已。大體而言，過去在國內針對樣本代表性比較有系統的研究大略有以下幾個方面：首先是從抽樣方法進行研討，多半集中在比較不同的抽樣方法或抽樣清冊所產生的樣本結構差異（林佳瑩、陳信木，1996；戴立安，1997）。其次是從戶中選樣的程序去評估樣本代表性所受到的影響（吳齊殷，1996；呂金河，1999；洪永泰，1996b）。第三是探討訪問失敗和樣本代表性的關聯（洪永泰，1989；林彩玉等，2004）。第四是全面性的樣本代表性問題（蘇建州，1998；陳肇男，2001）。第五是從時間因素分析電訪日期和時段對樣本代表性（洪永泰，2000），電話訪問調查涵蓋率和樣本代表性的影響（洪永泰，2005）。第六則是純粹比較加權方法。黃紀和張佑宗（2003）的研究結果證明以最小差異加權要比傳統的事後分層加權（post-stratification）和多變數反覆加權（raking）更好。而各種電訪加權方法的比較結果似乎差不多（洪永泰，2007）。

不過，以上來自台灣的研究多半是使用現成蒐集的資料，較少利用問卷設計技巧、實驗設計或調查方法上的比較進行樣本代表性與加權處理的分析評估，尤其是不同調查方法的比較分析。本研究另闢途徑，引進 PSA 法，以二〇〇四年總統大選投票意向電訪結果的推估爲例，將電訪樣本作爲研究樣本而社會變遷基本調查的面訪樣本爲參考樣本，首先以面訪資料爲基礎產生虛擬母體，其次透過面訪樣本與電訪樣本的合併資料，利用所

臺灣大學學術期刊資料庫

產生的次樣本之入選機率來調整電話訪問中投票行爲之估計，最後評估估計偏誤是否因此降低。

前人研究結果的確發現 PSA 用來推估網路調查結果的效果比電訪來得好（Schonlau et al., 2004），但這並不表示電訪調查就不需要採納 PSA，且文獻中針對電訪進行 PSA 的結果也顯示此法可行。台灣的電訪調查多半先採地區分層比例抽樣家戶電話號碼，在接通電話後進行戶中抽樣，所用的戶中抽樣可能來自隨機但非完全等機率（例如：Kish table）或非隨機（例如：任意成人法）。除了抽樣誤差大之外，訪問時段的限制（因找人不易，晚上找到受訪者的機率較大），有某一特定族群不在家（not-at-home）的比例較高（例如年輕人、勞工或加班者、偏遠地區居民或原住民），大哥大使用者越來越多，可透過家戶中電話接通找到人的機率越來越低，多線電話（含傳真或網路用途）以及其他影響接通困難度的因素（例如來電顯示，傳真機，答錄機篩選）越來越普遍，這些都是造成電訪涵蓋率越來越低與無回應率越來越高的來源，因此即便台灣的住宅電話使用率高達 97%，現實的電話訪問在抽樣時樣本膨脹倍數飆高已是事實（洪永泰，2007）。

有鑑於台灣的電話訪問在涵蓋率降低與無回應率與日遽增的窘境，傳統的加權方法對於母體推估的結果仍不盡然精確，因此引進 PSA 方法有其重要性，本研究先以調整電訪調查結果的估計爲例，其具體的目的分述如下：

一、比較電話訪問和虛擬母體的推估偏差。

二、藉由次樣本分組的操作，找出影響推估偏差的主要變數群。

三、試驗以 propensity score 組成的次樣本分組的修正方法降低推估偏差，並評估其功效。

貳、研究方法

一、面訪與電訪樣本的基本變項

本研究所使用的調查資料中，面訪樣本爲二〇〇四年社會變遷基本調查計畫：四期五次問卷一公民權組，調查期間爲二〇〇四年七月，以台灣

臺灣大學學術期刊資料庫

地區年滿 18 歲及以上之中華民國國民（七十四年十二月三十一日以前出生者）爲研究母體，並以台灣地區戶籍資料檔爲抽樣清冊（sampling frame），利用分層等機率三階段抽樣法（probability proportional to size, PPS）抽出受訪對象，完訪樣本數爲 1,781，刪除變項遺漏值後的樣本數爲 1,774。電訪樣本爲全國公信力民意調查公司於二〇〇八年二月二日至四日之間完成的電話訪問，以臺灣地區 20 歲及以上之民眾爲調查對象，抽樣方法以臺灣地區住宅用戶電話號碼爲底冊，採用等距方式抽出電話號碼後再將電話號碼尾兩碼改以隨機亂數代入，電話接通後以該戶電話號碼尾數進行戶中選樣，成功完訪的樣本數爲 1,083，刪除變項遺漏值後的樣本數爲 1,010。如表 1 所示，刪除變項遺漏值後，各變項比例分配與未刪除前的比例差異不大，故使用刪除變項遺漏值後的面訪與電訪樣本進行分析。資料分析採用完訪後的原始未加權資料以保持議題的單純化。

從表 1 可知，在人口變項分配上，面訪樣本中男性比女性多，電訪樣本中則是女性所佔的比例較高，在年齡層分布方面，面訪樣本中 18～29 歲、60 歲及以上者所佔比例比電訪樣本高，其他各年齡層所佔比例則是電訪樣本比面訪樣本高。教育程度在高中及以上者所佔比例是電訪樣本較高，小學及以下、國初中者所佔比例是面訪樣本較高。兩樣本在各族群比例分配方面差異不大，惟面訪樣本中台灣客家人所佔比例略高（面訪 14.37%、電訪 12.67%），而大陸各省市所佔比例略低（面訪 8.17%、電訪 12.28%）。在設籍地區方面，面訪樣本中雲嘉南、宜花東等地區所佔比例比電訪樣本高出約 5%，桃竹苗地區的比例則是比電訪樣本略高約 1%，電訪樣本中則是大台北基隆、中彰投、高屏澎等地區所佔比例比面訪樣本高。若兩樣本都跟當年的母體資料比較（內政部戶政司，2008），[1] 性別、

[1] 母體資料有兩來源。依據內政部户政司 2004 年內政統計年報，臺灣地區 20 歲及以上的人口，其中男性佔 50.43%，女性佔 49.57%；20～29 歲佔 23.26%，30～39 歲佔 22%，40-49 歲佔 22.21%，50～59 歲佔 14.82%，60 歲及以上佔 17.71%；在教育程度方面，小學及以下佔 22.43%，國、初中佔 15.24%，高中、職佔 30.12%，專科佔 14.52%，大學及以上佔 17.69%；設籍地區方面，桃竹苗佔 13.94%，中彰投佔 19.01%，雲嘉南佔 15.39%，高屏澎佔 16.84%，宜花東佔 4.69%，大台北基隆地區佔 30.13%。另依據洪永泰（1996a），台灣閩南人的比例是 75.77%，台灣客家人 10.9%，大陸各省市 12.42%，原住民 0.91%。

臺灣大學學術期刊資料庫

年齡層中 40～49 與 50～59 歲、以及教育程度中的國初中、專科與大學以上的比例是面訪樣本比較接近母體，但設籍地區與各族群中非台灣閩南人的比例則是電訪樣本比較接近於母體。

面訪樣本家中 18 歲以上的平均人口數爲 3.75 人，電訪樣本則是 2.99 人。在職業方面，軍公教、私部門職員、私部門勞工、農林漁牧人員在面訪樣本中所佔比例比電訪樣本高，私部門管理階層及專業人員、學生、家管、退休失業及其他者在電訪樣本中所佔比例比面訪樣本高。電訪樣本中政黨傾向泛藍的比例比面訪樣本高出約 21.35%，面訪樣本中「政黨傾向中立及其他者」所佔比例比電訪樣本高 14.1%。本研究所分析的投票行爲是針對二〇〇四年總統大選，在兩套樣本之間的投票行爲上呈現明顯的差異，面訪樣本以投給陳水扁、呂秀蓮者所佔比例較高（面訪 42.28%、電訪 30.99%），投給連戰、宋楚瑜者所佔比例較低（面訪 28.24%、電訪 35.35%）。[2]

二、研究步驟

㈠虛擬母體（pseudo-population data）

由於本研究沒有真正的母體資料，因此以全國性大型面訪調查—二〇〇四年台灣地區社會變遷基本調查計畫：四期五次問卷一公民權組爲基礎，利用樣本重抽法（bootstrapping）以簡單隨機抽樣且抽出不放回（without replacement）爲原則重覆的抽取面訪樣本而整合產生的一個虛擬母體。亦即從台灣地區社會變遷調查完訪的 1,774 案中進行重抽，每次簡單隨機抽取 1,000 案，20 次後得到 20 套樣本，總共產生規模爲 20,000 案的虛擬母體（pseudo-population, P^F）。

---

[2] 在此只是粗略的比較本研究所使用資料在投票行爲結果的差異，面訪的結果並非本研究的重點，因爲入選機率調整法（PSA）主要的目的是利用一個參考樣本的基本人口特徵或與投票行爲相關的態度來調整電訪中投票行爲的推估。換言之，PSA 並不會眞正用到參考樣本中的投票行爲來調整電訪的投票行爲，反而是（虛擬）母體中的投票行爲相關的人口與態度變項。其理由請見文前關於平衡分組入選機率值（PS）的說明（Zanutto, 2006）以及本文後續關於 PSA 步驟的說明。

臺灣大學學術期刊資料庫

表 1　面訪與電訪樣本變項的描述分析

	面訪樣本		電　訪　樣　本				
	N = 1,774[1]		N = 1,083			N = 1,010[2]	
	N	%	N	%	Valid %	N	%
性別							
男性	909	**51.24**	408	44.32	44.32	448	44.36
(女性)[3]	865	48.76	603	55.68	55.68	562	**55.64**
年齡							
(18～29 歲)[3]	460	**25.93**	192	17.73	18.22	184	18.22
30～39 歲	314	17.70	217	20.04	20.59	204	**20.20**
40～49 歲	347	19.56	277	25.58	26.28	271	**26.83**
50～59 歲	281	15.84	193	17.82	18.31	187	**18.51**
60 歲以上	372	**20.97**	175	16.16	16.60	164	16.24
拒答			29	2.68			
教育程度							
小學及以下	520	**29.31**	181	16.71	16.88	164	16.24
國、初中	245	**13.81**	89	8.22	8.30	86	8.51
高中、職	445	24.92	326	30.10	30.41	309	**30.59**
專科	242	13.64	167	15.42	15.58	164	**16.24**
(大學及以上)[3]	325	18.32	309	28.53	28.82	287	**28.42**
拒答			11	1.02			
省籍							
台灣閩南	1,354	**76.32**	779	71.93	74.33	749	74.16
台灣客家	255	**14.37**	132	12.19	12.60	128	12.67
台灣原住民及其他	20	**1.13**	11	1.02	1.05	9	0.89
(大陸各省市)[3]	145	8.17	126	11.63	12.02	124	**12.28**
拒答			35	3.23			
設籍地區							
桃竹苗	281	**15.84**	155	14.31	14.57	148	14.65
中彰投	256	14.43	223	20.59	20.96	216	**21.39**
雲嘉南	446	**25.14**	205	18.93	19.27	192	19.01
高屏澎	205	11.56	159	14.68	14.94	145	**14.36**
宜花東	173	**9.75**	53	4.89	4.98	51	5.05
(大台北基隆)[3]	413	23.28	269	24.84	25.28	258	**25.54**
拒答			19	1.75			
家中 18 歲以上平均人口(標準差)	**3.75** (1.708)		2.97 (1.566)			2.99 (1.552)	
職業							
軍公教人員	169	**9.53**	94	8.68	8.68	92	9.11
私部門管理階層及專業人員	137	7.72	140	12.93	12.93	135	**13.37**
私部門職員	528	**29.76**	248	22.90	22.90	236	23.37
農林漁牧	134	**7.55**	52	4.80	4.80	50	4.95
學生	100	5.64	71	6.56	6.56	65	**6.44**
家管	226	12.74	211	19.48	19.48	198	**19.60**
退休失業及其他	68	3.83	144	13.30	13.30	130	**12.87**
(私部門勞工)[3]	412	**23.22**	109	10.06	10.06	104	10.30
拒答			14	1.29	1.29		

臺灣大學學術期刊資料庫

表 1　面訪與電訪樣本變項的描述分析（續）

	面訪樣本		電　訪　樣　本				
	N = 1,774[1]		N = 1,083			N = 1,010[2]	
	N	%	N	%	Valid %	N	%
政黨傾向							
泛藍	380	21.42	449	41.46	41.46	432	**42.77**
泛綠	522	**29.43**	238	21.98	21.98	224	22.18
(中立及其他)[3]	872	**49.15**	396	36.56	36.56	354	35.05
投票行為：2004 年總統選舉投票對象							
陳呂配	750	**42.28**	328	30.29	30.29	313	30.99
連宋配	501	28.24	375	34.63	34.63	357	**35.35**
沒去投	201	11.33	167	15.42	15.42	155	**15.35**
投廢票	45	**2.54**	8	0.74	0.74	8	0.79
無反應（拒答、不知道）	201	**11.33**	124	11.45	11.45	103	10.20
當時沒投票權	76	4.28	81	7.48	7.48	74	**7.33**

註：1. 社會變遷四期五次問卷一樣本數爲 1,781，經刪除家中 18 歲以上人口數與職業變項的遺漏值後，樣本數減爲 1,774，共刪除 7 案。

2. 電訪樣本的樣本數爲 1,083，刪除年齡、教育程度、省籍、地區與職業等變項的遺漏值後，樣本數減爲 1,010，共刪除 73 案。

3. 括號內的變項組別爲入選機率邏輯迴歸模型中的對照組，各變項對照組爲女性，18～29 歲，大學及以上，大陸各省市，大台北基隆地區，私部門勞工，政黨傾向中立及其他組。

㈡合併樣本

爲了調整電訪結果之推估，需先形成進行相對於面訪之電訪入選機率預測所需的分析資料。此資料來自於將電訪樣本與面訪參考樣本合併的樣本：[3] 其中電訪樣本乃從 1,010 案中依據性別、年齡與教育程度的各分層分配比例，分層隨機抽取 800 案所形成，稱之爲 n^T。而面訪參考樣本則是從虛擬母體 P^F 中，以簡單隨機抽樣法抽取 200 案所形成，稱之爲 n^R。[4] 然後

[3] 如文前所述，選擇可靠的參考樣本是入選機率調整法的估計是否準確的重要關鍵之一。參考樣本是否可靠在於其是否來自完善抽樣設計的調查結果，而且包含足以正確的推估母體而修正未涵蓋與無回應誤差的必要的社會人口變項與非人口變項（Lee and Valliant, 2008）。由於本研究屬於初探性的介紹入選機率調整法，雖然台灣地區社會變遷基本調查不是完美的參考樣本，但在台灣已經是比較具樣本代表性的全國性大規模調查。

[4] 合併樣本所需新的兩個樣本分別抽取的案數，是可多作幾次各種組合的試驗再下定論。不過本研究完全依照 Lee (2006) 作法，且根據 Lee (2006) 文中的說明，抽取參考樣本 400 案後的入選機率調整的模擬結果與抽取 200 案模擬的結果一樣。

臺灣大學學術期刊資料庫

將抽取的面訪參考樣本與電訪樣本合併成爲一套 1,000 案的合併樣本（n^T和n^R）。

㈢入選機率調整法（propensity score adjustment, PSA）

1. 產生入選機率（propensity score, PS）

入選機率值（PS）來自於對合併樣本進行邏輯迴歸模型（logistic regression model）所推估的結果。在此邏輯迴歸模型中，依變項爲兩分類，包括面訪參考樣本與電訪樣本；其中電訪樣本設定爲代碼 1，面訪參考樣本爲 0。自變項納入人口變項與非人口變項，包括性別、年齡、教育程度、省籍、設籍地區、家中 18 歲以上人口數與職業、政黨傾向等，共八個變項。各變項的測量詳見表 1，而邏輯迴歸模型方程式爲：

$$\ln\left(\frac{P_r(T=1)}{1-P_r(T=1)}\right)=\alpha+\beta_1\,gender+\beta_2\,age+\beta_3\,educ+\beta_4\,ethnic+\beta_5\,region+\beta_6\,household+\beta_7\,occ+\beta_8\,party$$

此模型經由這七個社會人口特徵變項與一個政黨認同變項推估出來的數值即是相對於入選爲面訪參考樣本時，入選爲電訪樣本的預測機率值。

2. 產生次樣本分組的修正係數（*fc*）

將邏輯迴歸模型得出預測入選電訪樣本的機率值由小到大排序，按排序結果再將樣本平均的分爲五個次樣本分組（subclass），[5] 每一次樣本分組的案數爲 200，皆包括來自電訪與面訪參考樣本的個案。參考 Lee (2006) 的調整公式，入選機率調整公式爲以下公式(1)，若考慮電訪樣本原先權數(d)的調整公式則爲公式(2)，若原先權數未知或原調查屬於等機率時則採用公式 (3)。本研究旨在測試入選機率調整方法之成效，比較面訪參考樣

[5] 文獻中對於該分成多少次樣本分組多參考 Cochran (1968) 針對組數之研究的結論，亦即採用五個次樣本分組可降低 90% 的推估偏差，再增加次樣本分組對於降低推估偏差的效果有限，因此五個次樣本分組爲後續研究所推薦與延用。

臺灣大學學術期刊資料庫

本與電訪樣本的推估偏差，爲維持議題的單純化且初探性介紹入選機率調整法，本研究假設電訪樣本來自未知機率抽樣，而不考慮樣本原先權數，因此採用公式 (3) 計算分層修正係數（*fc*）。

$$fc = \frac{\sum_{k\in(S_c^R)} d_k^R \Big/ \sum_{k\in(S^R)} d_k^R}{\sum_{j\in(S_c^T)} d_j^T \Big/ \sum_{j\in(S^T)} d_j^T} \equiv \frac{\hat{N}_c^R \Big/ \hat{N}^R}{\hat{N}_c^T \Big/ \hat{N}^T} \tag{1}$$

$$d_j^{T.PSA} = fc\, d_j^T = \frac{\hat{N}_c^R \Big/ \hat{N}^R}{\hat{N}_c^T \Big/ \hat{N}^T} d_j^T \tag{2}$$

$$fc = \frac{n_c^R \Big/ n^R}{n_c^T \Big/ n^T} \tag{3}$$

s_c^R=次樣本分組面訪參考樣本，s_c^T=次樣本分組電訪樣本，d=樣本本身權值。

將面訪參考樣本在每一分組的個數（n_c^R）佔面訪參考樣本總數（n^R）（每套樣本爲 200 案）的比例除以電訪樣本中每一分組的個數（n_c^T）佔電訪樣本總數（n^T）（每套樣本爲 800 案）的比例即得每一分組的修正係數（*fc*）。以第一次模擬的結果爲例，在表 2 中，面訪參考樣本第一組爲 97 案，電訪樣本第一組爲 103 案，故第一組修正係數（*fc*）等於（97/200）/（103/800）= 3.76699。此修正係數即是用來調整每一次樣本分組中原先之不同投票行爲的回答人數，詳見下一步驟。

臺灣大學學術期刊資料庫

表 2　入選機率分組、分組的修正係數與調整前後的 *y* 估計值（第一次模擬結果）[1]

次樣本 分　組	面　訪 參　考 樣本數	電　訪 樣本數	電　訪　樣　本 分組的修正係數（*fc*）	*y*1 投陳呂 （人數）	*fc***y*1	*y*2 投連宋 （人數）	*fc***y*2
1	97	103	3.76699	33	124.31	24	90.41
2	55	145	1.51724	67	101.66	24	36.41
3	30	170	0.70588	63	44.47	58	40.94
4	14	186	0.30108	56	16.86	66	19.87
5	4	196	0.08163	37	3.02	114	9.31
總　計	200	800		256	290.32	286	196.94
y 估計值（第一次模擬結果）				**0.32**	**0.36**	**0.36**	**0.25**

註：1. 本表結果爲模擬抽樣第一次的結果，本研究進行 2,000 次抽樣，產生 2,000 個調整前的 *y* 估計值與調整後的 *y* 估計值。

3. 以修正係數調整電訪樣本「投票行為」估計值

接下來的步驟是根據次樣本各分組中的各修正係數調整電訪樣本依變數「投票行爲」的估計值。本研究以「2004 年總統選舉投給陳水扁、呂秀蓮」爲依變數選項一，「2004 年總統選舉投給連戰、宋楚瑜」爲依變數選項二，根據公式 (4)，電訪樣本「依變數選項一在第一次樣本分組的調整案數」是投陳呂人數 **fc*=33*3.76699=124.31，「依變數選項二第一次樣本分組的調整案數」是投連宋人數 **fc*=24*3.76699=90.41，其他分組的調整案數依此類推。如表 2 所示，最後得出電訪樣本依變數調整後的估計值分別是依變數選項一「投陳呂」（*fc*y1*）爲 0.36（調整前估計值爲 0.32），依變數選項二「投連宋」（*fc*y2*）爲 0.25（調整前估計值爲 0.36）。

$$\hat{\bar{y}}^{T.PSA}=\frac{\sum_{c}\sum_{j\in(s_c^T)} d_j^{T.PSA} y_j}{\sum_{c}\sum_{j\in(s_c^T)} d_j^{T.PSA}} \tag{4}$$

臺灣大學學術期刊資料庫

㈣模擬

以上步驟參考 Lee (2006)，同樣的重覆進行 2,000 次，最後產生 2,000 組調整後的依變數估計值，結果如表 3 所示。

參、結果與討論

爲評估入選機率調整方法的效果，根據 Lee (2006)，本研究利用以下公式 (5)～(8) 得出以面訪參考樣本爲比較基準時，電訪樣本在 PSA 調整前與調整後的估計值誤差、誤差削減比例、標準誤與標準誤增加的比例，結果如表 3 所示。

$$bias(\bar{y}^{T}) = \left[\sum_{m=1}^{M} y_m^{T} - \sum_{m=1}^{M} y_m^{R}\right] / M = \bar{y}^{T} - \bar{y}^{R} \tag{5}$$

$$p.bias(\bar{y}^{T.PSA}) = \left[\frac{\left|bias(\bar{y}^{T})\right| - \left|bias(\bar{y}^{T.PSA})\right|}{\left|bias(\bar{y}^{T})\right|}\right] \times 100 \tag{6}$$

$$se(\bar{y}^{T}) \approx \sqrt{\sum_{m=1}^{M} (y_m^{T} - \bar{y}^{T})^2 / M} \tag{7}$$

$$p.se(\bar{y}^{T.PSA}) = \left[\frac{\left|se(\bar{y}^{T.PSA})\right| - \left|se(\bar{y}^{T})\right|}{\left|se(\bar{y}^{T})\right|}\right] \times 100 \tag{8}$$

虛擬母體（P^F）在 2004 年總統大選兩組候選人的得票比例與表 3 中經抽樣 2,000 次後所得面訪參考樣本的平均估計值結果相近，虛擬母體「投票對象陳水扁、呂秀蓮」的比例爲 41.9%，「投票對象連戰、宋楚瑜」的比例爲 27.8%。若依變數選項一「投票對象陳水扁、呂秀蓮」的結果與面訪參考樣本估計值比較時，亦即未經入選機率修正係數調整時，在電訪樣

臺灣大學學術期刊資料庫

本估計誤差爲 0.10872，經過調整後所得的估計值誤差爲 0.0319861，誤差消減的比例是 70.58%。依變數選項二「投票對象連戰、宋楚瑜」在未調整前之電訪樣本估計誤差爲 0.0753969，經過調整後的估計誤差爲 0.0460872，誤差消減比例是 38.87%，從誤差消減的比例來看，經過入選機率方法調整後的電訪樣本估計值誤差比起調整前所得的誤差小，尤其是在「投票對象爲陳水扁與呂秀蓮」估計值。可見得以入選機率修正係數來調整電訪樣本估計值可明顯的降低估計誤差。

表 3　2004 年總統選舉投票行為估計的誤差與標準誤：面訪參考樣本、電訪樣本，與入選機率調整後的電訪樣本

樣　本	估計值 (*estimate*)	誤　差 (*bias*)	誤差消減比例 (*p.bias*)	標準誤 (*se*)	標準誤增加比例 (*p.se*)
投票給陳水扁、呂秀蓮（2,000 次模擬的結果）					
面訪參考	0.4186500			0.001255200	
電訪	0.3099300	0.1087200		0.000261548	
PSA 電訪	0.3866639	0.0319861	70.5794	0.000952853	264.3128604
投票給連戰、宋楚瑜（2,000 次模擬的結果）					
面訪參考	0.2778850			0.001061200	
電訪	0.3532819	0.0753969		0.000255122	
PSA 電訪	0.2317978	0.0460872	38.8739	0.000684681	168.3739544

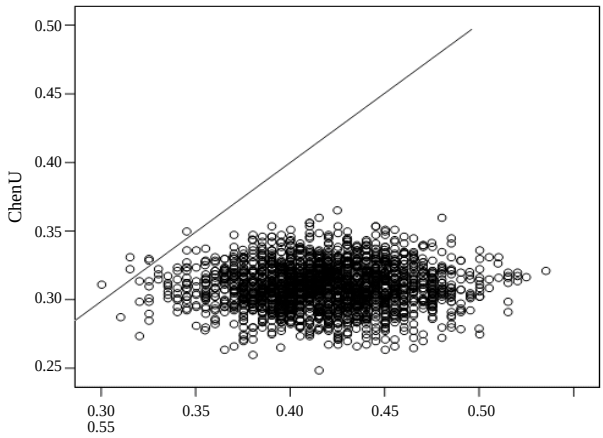
依變數選項一與選項二在未做入選機率修正係數之調整時，面訪參考樣本估計值的標準誤分別爲 0.0012552 與 0.0010612，電訪樣本的估計的標準誤則分別是 0.000261548 與 0.000255122。經過入選機率修正係數調整後的電訪樣本估計的標準誤是 0.000952853 與 0.000684681，均比未調整前電訪樣本估計的標準誤要高，但是其數值較接近於面訪參考樣本的結果。

最後，圖 1 到圖 4 分別顯示在依變數選項一與選項二中，在經過入選機率調整前與調整後的電訪樣本估計值分別與面訪參考樣本估計值相比較的散佈情形，其中的參考線標示出在經過入選機率調整方法後的電訪樣本估計值與面訪參考樣本的估計值結果相近的情況。圖 2 與圖 4 顯示，經調整後的電訪樣本估計值與面訪參考樣本估計值分別比圖 1 與圖 3 更接近，這證明入選機率調整方法對於修正電訪樣本的投票行爲估計值有其實際的

臺灣大學學術期刊資料庫

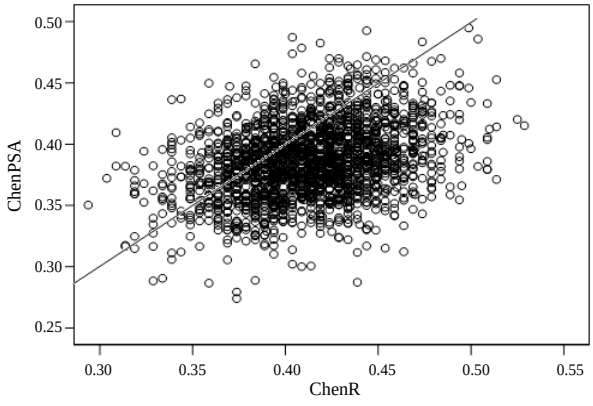
效果。再從另一角度來比較，圖 5 與圖 6 更進一步用盒形圖表示兩依變數分別在面訪參考樣本、調整前電訪樣本與調整後電訪樣本的估計值，由此圖可明顯的看到經過入選機率修正係數調整後的電訪樣本估計值的分配情況，比起未調整前的電訪樣本，更接近面訪參考樣本的估計值與其分配。

圖 1　依變數選項一（投陳呂），面訪參考樣本與未調整前電訪樣本估計值分配



ChenU：未調整前電訪樣本投陳呂估計值
ChenR：面訪參考樣本投陳呂估計值

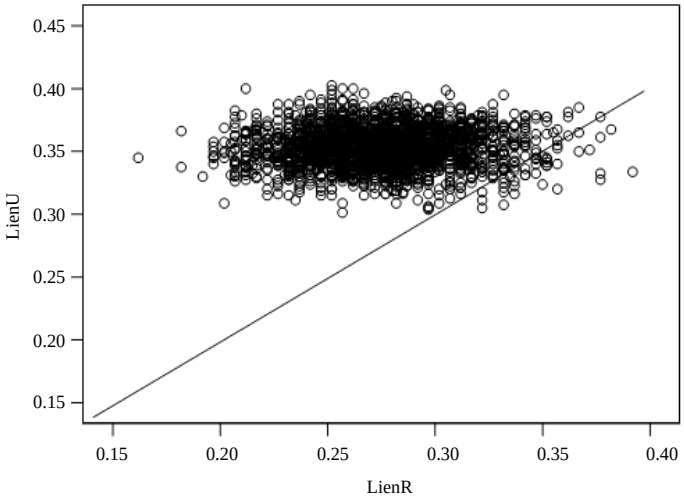
圖 2　依變數選項一（投陳呂），面訪參考樣本與 PSA 調整後電訪樣本估計值分配



ChenPSA：PSA 調整後電訪樣本投陳呂估計值
ChenR：面訪參考樣本投陳呂估計值

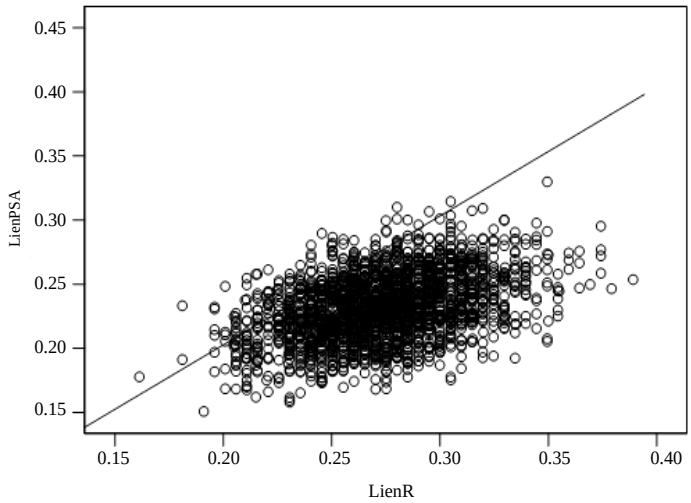
臺灣大學學術期刊資料庫

圖 3　依變數選項二（投連宋），面訪參考樣本與未調整前電訪樣本估計值分配



LienU：未調整前電訪樣本投連宋估計值
LienR：面訪參考樣本投連宋估計值

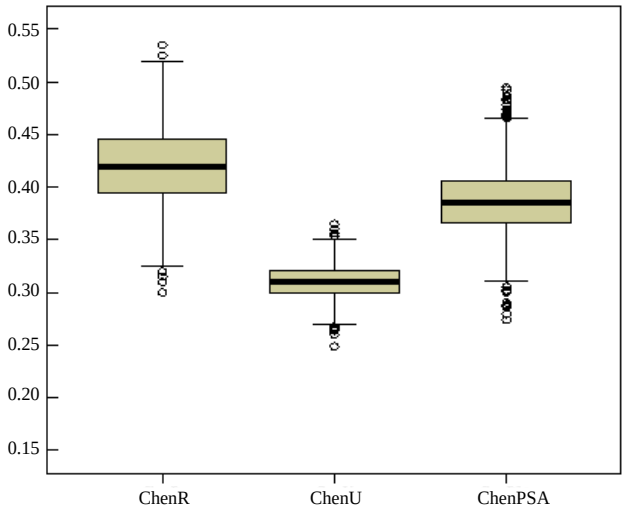
圖 4　依變數選項二（投連宋），面訪參考樣本與 PSA 調整後電訪樣本估計值分配



LienPSA：PSA 調整後電訪樣本投連宋估計值
LienR：面訪參考樣本投連宋估計值

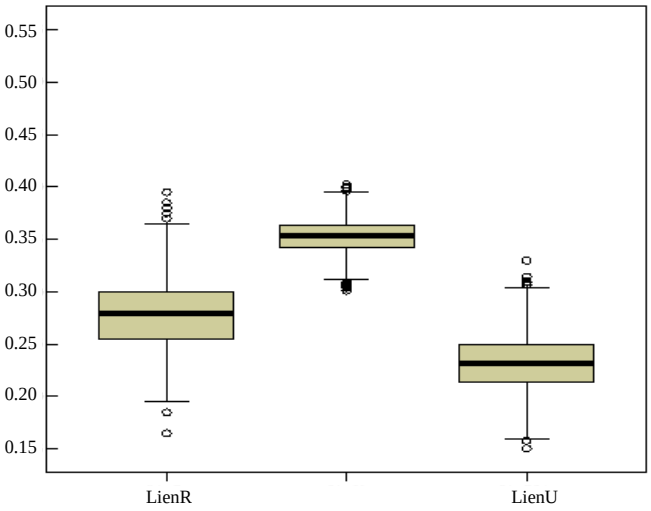
臺灣大學學術期刊資料庫

圖 5　面訪參考樣本與電訪樣本調整前後的投票行為估計值差異：依變數選項一「投票對象陳水扁、呂秀蓮」



ChenR：面訪參考樣本投陳呂估計值
ChenU：未調整前電訪樣本投陳呂估計值
ChenPSA：PSA 調整後電訪樣本投陳呂估計值

圖 6　面訪參考樣本與電訪樣本調整前後的投票行為估計值差異：依變數選項二「投票對象連戰、宋楚瑜」



LienR：面訪參考樣本投連宋估計值
LienU：未調整前電訪樣本投連宋估計值
LienPSA：PSA 調整後電訪樣本投連宋估計值

臺灣大學學術期刊資料庫

綜合以上根據估計值誤差、估計標準誤與兩種估計值分配圖的評估，我們可以肯定的是，利用入選機率調整法（PSA）來處理在主觀態度性質題目的調查推估是可推薦而有效的。當國外相關的研究已經有些歷史之際，國內這方面的研究仍有待開發。

對於未來研究的建議事實上與本研究有待改進的地方有關。首先，PSA 在國外常見的是用在調整不容易蒐集到有代表性樣本的調查資料（例如：21 世紀流行的網路調查）。這方面的調整，無論是針對電訪或網路調查，在國內至今仍無人進行研究。透過本論文引進台灣之後，希望更多學者可以從不同的抽樣、加權與統計原理比較 PSA 與其他樣本推估方法，進而證明 PSA 是否爲更具體而可行的辦法。其次，本研究僅僅利用一個沒有加權處理，無真正樣本代表性的全國性面訪調查爲參考樣本，進而產生虛擬的母體，因此在估計值與誤差消減比例的變化上也僅止於相對性的比較。有兩個建議方案可解決此權宜之計的缺憾，其一是對參考樣本或（及）其衍生的虛擬母體進行加權處理，此權數納入計算 PS 的公式之中。其二是尋找一個確切的母體資料爲比較依據。這當然更能證明 PSA 的效力。不過，比起蒐集客觀的社會人口特徵的母體資料，要蒐集到主觀態度方面的母體資料可能不是件容易的事情，這對未來研究將是一大挑戰。那麼退而求其次，如何產生更正確而接近於母體的資料變成是有待探討的問題。

這裡需提醒讀者的是應用 PSA 時首先要遵循其前提假設（Lee and Valliant, 2008），以免在不符合 PS 之應用原理之際，造成估計誤差降低的假象。其次是在分析 PSA 時，需慎重考慮要選擇多少個次樣本分組，分組內個案數太小不理想，Cochran (1968) 建議最多五個分組值得參考，但這也顯示了 PSA 的限制，有可能因此無法平衡分配 PS 的問題。再者，依變項的分布也很重要，若依變項的各項類別比例有過小案數時，似乎不見得可應用 PSA，其理由是很可能在進行次樣本分組時會沒有人被分入某一組的情形，若有，推估偏差可能會更大。最後，文獻中關於 PSA 的研究大多集中在比較估計偏差的降低情形，卻無任何研究針對標準誤上升作解釋。我們可以肯定的是產生 PS 的預測變項越多，變異量會增加，而樣本數增加時標準誤會降低，未來研究要針對此議題繼續努力時，具體的方法會是在從虛擬母體中抽取產生的參考與研究樣本時，擴大其樣本數，至於分別

臺灣大學學術期刊資料庫

該擴大到多少則有待解答。

在 PSA 的應用方面，未來值得探討的主題包括，如何利用傳統加權後 PSA 的方法，比較其與單純 PSA 的估計偏差降低情形。其次，以 PSA 降低調查涵蓋率或單位無反應（unit nonresponse）等問題的研究也亟需進行，尤其是比較電訪與網路調查方面。最後，除了本研究所探討的是投票行爲估計的調整，尚有其他值得繼續驗證的主題，至於從哪一主題開始則端視是否找得到可信賴的母體資料了。

參考書目

一、中文部份

內政部戶政司，2008，〈內政統計年報－人口年齡分配、十五歲以上人口教育程度〉，內政部統計處網頁，http://www.moi.gov.tw/stat/year. aspx，2008/04/17。

李隆安、韓成業，2008，〈The Study of Weights by the Raking Method〉，「2008 年度中國統計學社學術研討會」論文（12 月 19 日），台北：中國統計學社。

呂金河，1999，〈依號碼尾數修正的戶中選樣法〉，《中國統計學報》，37(3): 257-278。

林佳瑩、陳信木，1996，〈各種電話號碼抽樣方式之比較分析〉，《調查研究》，2: 111-141。

林彩玉、洪永泰、鄭宇庭，2004，〈調查研究中『無反應』選項問題之分析〉，《調查研究》，15: 31-60。

吳統雄，1984，《電話調查：理論與方法》，台北：聯經。

吳齊殷，1996，〈戶中抽樣與否對樣本代表性的影響：以大台北地區電話訪問爲例〉，《調查研究》，1: 39-65。

洪永泰，1989，〈抽樣調查中訪問失敗問題的處理〉，《社會科學論叢》，37: 33-52。

洪永泰，1996a，〈抽樣調查中樣本代表性的問題〉，《調查研究》，

臺灣大學學術期刊資料庫

1: 7-37。

洪永泰，1996b，《戶中選樣之研究》，台北：時英。

洪永泰，2000，〈調查時間、抽樣方法與樣本代表性的實務探討〉，「國內調查研究之抽樣設計：理論與實務研討會」論文（6 月 30 日），台北：中央研究院調查研究工作室。

洪永泰，2005，〈台灣地區抽樣調查各種母體定義、抽樣底冊、和涵蓋率的比較〉，《調查研究》，18: 9-44。

洪永泰，2006，〈台灣地區電話訪問調查資料加權處理之探討〉，「第六屆調查研究方法與應用國際學術研討會」論文(8 月 30-31 日)，台北：中央研究院人文社會科學研究中心暨調查研究專題中心。

洪永泰，2007，〈台灣地區選前電訪民意調查推估偏差的補救〉，「政治思想、政黨與選舉行爲—紀念謝延庚教授學術研討會」論文（5 月 8 日），台北：台北大學公共行政暨政策學系。

陳肇男，2001，〈大型電訪之省思〉，《調查研究》，10: 121-138。

黃　紀、張佑宗，2003，〈樣本代表性檢定與最小差異加權：以 2001 年台灣選舉與民主化調查爲例〉，《選舉研究》，10(2): 1-35。

黃毅志，1997，〈抽樣調查中訪問失敗的問題之探討：以台灣地區社會變遷調查爲例作說明〉，《調查研究》，4：113-129。

隋杜卿，1986，〈問卷調查中的樣本代表性研究〉，《思與言》，23(6): 72-86。

劉從葦、陳光輝，2005，〈Is Weighting a Routine or Something that Needs to be Justified?〉，《選舉研究》，12(2): 149-187。

戴立安，1997，〈對國內電話調查之多段隨機撥號抽樣方式的初析〉，《調查研究》，4: 131-198。

蘇建州，1998，〈穩健電話調查程序之初探〉，《調查研究》，6: 119-143。

蘇　蘅，1986，《傳播研究調查法》，台北：三民。

羅健銘、陳素秋、賴允亮、林家瑾、陳建仁，2007，〈住院癌末病患照護型態對住院醫療費用與住院天數之影響〉，《台灣衛誌》，26(4): 270-282。

臺灣大學學術期刊資料庫

二、英文部份

Abramowitz, A. I. 2005. “Just Weight! The Case for Dynamic Party Identification Weighting.” in http://www.apsanet.org. latest update 14 July 2008.

Brehm, J. 1993. *The Phantom Respondents: Opinion Surveys and Political Representation*. Ann Arbor: University of Michigan Press.

Cochran, W. G. 1968. “The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies.” *Biometrics* 24(2): 295-313.

Colasanto, Diane and Jay A. Mattlin. 1987. “Evaluation of Gallup’s Methodology for Predicting Likelihood of Voting.” Paper prepared at the 1987 Joint Statistical Meeting, Aug 17, San Francisco.

Czajka, John L., Sharon M. Hirabayashi, Roderick J. A. Little and Donald B. Rubin. 1992. “Projecting from Advance Data Using Propensity Modeling: An Application to Income and Tax Statistics.” *Journal of Business & Economic Statistics* 10(2): 117-131.

D’Agostino, Jr. Ralph B. 1998. “Tutorial in Biostatistics: Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-randomized Control Group.” *Statistics in Medicine* 17: 2265-2281.

Dever, Jill A., Ann Rafferty and Richard Valliant. 2008. “Internet Surveys: Can Statistical Adjustments Eliminate Coverage Bias?” *Survey Research Methods* 2(2): 47-62.

Dimock, Michael, Scott Keeter, Mark Schulman and Carolyn Miller. 2001. “A Voter Validation Experiment: Screening for Likely Voters in Pre-election Surveys.” Paper prepared at the 56th Annual American Association for Public Opinion Research Conference, May 17-20, Montreal, Quebec, Canada.

Drake, Christiana. 1993. “Effects of Misspecification of the Propensity

臺灣大學學術期刊資料庫

Score on Estimators of Treatment Effect." *Biometrics* 49: 1231-1236.

Duncan, Kristin B. and Elizabeth A. Stasny. 2001. "Using Propensity Scores to Control Coverages Bias in Telephone Surveys." *Survey Methodology* 27(2): 121-130.

Freedman, Paul and Ken Goldstein. 1996. "Building a Probable Electorate from Pre-election Polls: A Two-State Approach." *Public Opinion Quarterly* 60: 574-587.

Judkins, David R., David Morganstein, Paul Zador, Andrea Piesse, Brandon Barrett and Pushpal Mukhopadhyay. 2007. "Variable Selection and Raking in Propensity Scoring." *Statistics in Medicine* 26: 1022-1033.

Kalton, Graham and Ismael Flores-Cervantes. 2003. "Weighting Methods." *Journal of Official Statistics* 19(2): 81-97.

Lee, Sunghee. 2006. "Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys." *Journal of Official Statistics* 22(2): 329-349.

Lee, Sunghee and Richard Valliant. 2008. "Weighting Telephone Samples Using Propensity Scores." In *Advances in Telephone Survey Methodology*, eds. James M. Lepkowske et al. New York: John Wiley & Sons., 170-183.

Lee, Sunghee and Richard Valliant. 2009. "Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment." *Sociological Methods & Research* 37(3): 319-343.

Loosveldt, Geert and Nathalie Sonck. 2008. "An Evaluation of the Weighting Procedures for an Online Access Panel Survey." *Survey Research Methods* 2(2): 93-105.

Rosenbaum, P. R. 2005. "Observational Study." In *Volumn 3 Encyclopedia of Statistics in Behavioral Science*, eds. Brian S.

臺灣大學學術期刊資料庫

Everitt and David C. Howell. New York: John Wiley & Sons., 1451-1462.

Rosenbaum, P. R. and D. B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70(1): 41-55.

Rosenbaum, P. R. and D. B. Rubin. 1984. “Reducing Bias in Observational Studies Using Subclassification on the Propensity Score.” *Journal of the American Statistical Assocation* 79(387): 516-524.

Schonlau, Matthias, Kinga Zapert, Lisa Payne Simon, Katherine Haynes Sanstad, Sue M. Marcus, John Adams, Hongjun Kan, Rachel Turner and Sandra H. Berry. 2004. “A Comparison between Responses from a Propensity-Weighted Web Survey and Identical RDD Survey.” *Social Science Computer Review* 22: 128-138.

Schonlau, Matthias, Arthur van Soest and Arie Kapteyn. 2007. “Are ‘Webographic’ or Attitudinal Questions Useful for Adjusting Estimates from Web Surveys Using Propensity Scoring?” *Survey Research Methods* 1(3): 155-163.

Strümer, Til, Sebastian Schneeweiss, M. Alan Brookhart, Kenneth J. Rothman, Jerry Avorn and Robert J. Glynn. 2005. “Analytic Strategies to Adjust Confounding Using Exposure Propensity Scores and Disease Risk Scores: Nonsteroidal Antiinflammatory Drugs and Short-term Mortality in the Elderly.” *American Journal of Epidemiology* 161(9): 891-898.

Traugott, Michael and Clyde Tucker. 1984. “Strategies for Predicting Whether a Citizen Will Vote and Estimation of Electoral Outcomes.” *Public Opinion Quarterly* 48: 330-343.

Zanutto, Elaine L. 2006. “A Comparison of Propensity Score and Linear Regression Analysis of Complex Survey Data.” *Journal of Data Science* 4: 67-91.

臺灣大學學術期刊資料庫

A Study of Survey Nonresponse Bias Using Propensity Score Adjustment

Su-hao Tu[*], *Wan-yun Lo*[**] *and Yung-tai Hung*[***]

Abstract

It is common to use socio-demographic variables for weighting, but this does not mean that other inferences, especially on attitudes and behavior variables, will be free of bias. This article takes voting behavior in a presidential election as an example to examine the effects of adjusting the telephone survey results using sub-classifications on the propensity score as suggested by Lee (2006). We first used the Taiwan Social Change Survey (TSCS, 2004) wave 4 module for a citizenship study to produce a pseudo-population dataset based on bootstrapping sampling. A random sample of 200 cases was drawn from this dataset. In addition, a stratified random sample of 800 cases was drawn from a telephone survey. The two samples were combined into a sample of 1,000 cases. Propensity Score Adjustment (PSA) allows us to adjust and evaluate the proper estimation of voting behavior based on the sub-classification of propensity score. The procedure was repeated 2,000 times. The results showed that the PSA method does effectively reduce telephone survey estimate bias.

Keywords: Incomplete Data, Propensity Scores, Propensity Score Adjustment, Voting Behavior

---

* Associate Research Fellow of Center for Survey Research, Research Center for Humanities and Social Sciences, Academia Sinica, E-mail: suhao@gate. sinica.edu.tw.

** Research Assistant of Center for Survey Research, Research Center for Humanities and Social Sciences, Academia Sinica, E-mail: lilaclo@gate. sinica.edu.tw.

*** Professor of Department of Political Science, National Taiwan University, E-mail: ythung @ntu.edu.tw.

臺灣大學學術期刊資料庫

臺灣大學學術
期刊資料庫