選舉研究
第二十一卷第二期（103/11），pp.147-175

以電訪樣本資料庫作爲解決原住民調查研究難題之可行性的再討論 *

包正豪 **、周應龍 ***

《本文摘要》

國內以原住民爲對象之調查研究較爲缺乏，且普遍爲附屬於全國抽樣當中的小樣本分析，其主因在於台灣原住民人口數量較少，約僅占全國總人口數的 2%，輔以其地理分布相當不均勻，以及相當數量的原住民因爲生計因素而有工作／居住地與户籍所在地不同的情形，致使傳統調查方式，無論面訪或電訪，都有難以解決的抽樣困難。然而近年來，國內民意調查日益普遍，學術與民間調查機構已經累積數量極爲龐大的電話調查成功樣本，其中包含大量的漢人樣本與一定比例的原住民樣本。若能取得這些「原本不是以原住民爲母體，卻包含原住民的成功樣本」，將混雜埋藏於大量「無用」資料當中的「有價值資料」給挖掘出來，應能直接發掘篩選出所需要的原住民樣本，節約成本而解決前述抽樣困難。

有鑑於此，研究團隊與全國公信力民意調查股份有限公司合作，合併 2007 年迄今（2013 年 3 月）該公司所執行抽樣範圍爲全國的原住民電話訪問成功樣本成爲原住民電話資料庫，以爲調查之用。初步研究結果顯示，使用雙重抽樣方法所整併的電訪樣本資料庫進行電話調查，所獲得之成功樣本特徵與官方人口統計中的原住民母體特徵近似。同時，納入跨年度的原住民電話資料，能夠更有效地提升前述原住民電話資料庫的蒐集範圍，並降低樣本取得成本。

關鍵詞：民意調查、雙重抽樣、電話訪問、原住民

---

* 感謝淡江大學重點研究計畫的經費支持，使得本研究能順利進行與完成，惟本文之法律與學術倫理責任，由作者負責。

** 淡江大學全球政治經濟學系副教授。E-mail: 129456@mail.tku.edu.tw。

*** 淡江大學全球政治經濟學系助理教授（通訊作者）。E-mail: 145318@mail.tku.edu.tw。

壹、研究緣起與相關研究討論

以原住民爲研究母體的行爲調查研究，在執行調查時所面對的最重大難題是，由於台灣原住民族的人口規模極小，導致抽出合格樣本的成本過大。根據 2013 年 6 月最新的人口統計資料顯示，台灣共有原住民 530,277 人，占全國總人口數（23,344,213 人）的 2.27%（內政部統計處 2013；原住民族委員會 2013）。依此比例，如果想要有 1,068 個原住民調查樣本，以達到 3% 之抽樣容許誤差標準的話，首先必須要先成功 47,016 個全國樣本（不分原漢），才能從中篩選出足夠多的原住民樣本（1,068 個原住民樣本）。若以一般民意調查的 1/10 成功率計算，即表示必須要撥打 470,160 通電話，才能滿足前述條件。也就是說，完成一個具學術水準的原住民抽樣電話調查，所要花費的成本，是相同標準與規模之一般調查的 45 倍左右。

因此，過去鮮有以原住民爲研究母體的電話民意調查被執行過。取而代之的是，以戶籍資料爲抽樣清冊之直接到戶的面對面訪問調查。但這種直接到戶的原住民面訪調查，仍然有成本偏高的問題。以中央研究院民族學研究所與人文社會科學研究中心共同主持的「臺灣原住民社會變遷與政策評估研究計畫」爲例，費時 3 年（2006 年 3 月到 2008 年 12 月），且耗資鉅大，方完成 2,000 個樣本。[1]其間因爲經費不足，還必須由其他研究計畫的經費支應。[2]所以，無論是電話調查或是面訪調查，都是耗費長久時間與鉅額經費的龐大工程。其規模已經龐大繁複到，鮮有任何政府機關或學術機構有足夠的預算與時間，還有意願，專門針對原住民各項行爲做實證調查研究。[3] 究其根本原因，即在於「成本」二字。

檢視前述原住民行爲實證調查研究的執行成本，之所以如此駭人的原因，我們可以發現「樣本取得成本」是主因。[4]在尚未解決此一問題之前，國內從事原住民相關之實證調

[1] 「臺灣原住民社會變遷與政策評估研究計畫」的總經費未有公開資料可供查證，但以類似規模與研究時間長度的「2009 年至 2012 年『選舉與民主化調查』三年期研究規劃」作參考，其計畫經費總預算規模將超過 2,600 萬台幣。

[2] 由該研究案協同主持人林季平之國科會計畫「失業勞工的回流及連續流動及再就業：人力運用擬追蹤調查分析 (2/2)」(NSC94-2415-H-001-034) 挹注部分經費。

[3] 事實上，台灣唯一一個以原住民爲抽樣母體的專門實證行爲研究，就是中央研究院民族學研究所與人文社會科學中心共同主持的「臺灣原住民社會變遷與政策評估計畫」。另外，高德義（1996）曾經以原住民社會政治菁英爲對象，以郵寄問卷方式進行問卷調查。但高德義自承該調查在代表性上有嚴重缺陷，其研究結果僅能供參考。

[4] 面對面訪問的成本結構與電話訪問調查有所差異，前者在於每個訪問調查對象都必須有專人前往進行訪談，事後仍需透過專人進行實地複查，以確保訪談問卷不是舞弊的結果。後者則是要篩選出合格電話樣本所必須要承擔的過濾成本。兩者成本發生階段不同，但都是樣本取得過程

查研究（電話調查），譬如中華人權協會自 2007 年起所進行的「原住民人權指標年度調查」，主要採用一般調查設計，以台灣地區住宅電話號碼簿爲抽樣清冊（其意涵爲抽樣母體是台灣全體有裝設家戶電話的民衆），依系統抽樣方式抽出電話號碼。然後，將所抽出的電話號碼的後兩碼，以隨機亂數代入，使其能夠在理論上涵蓋未登錄之電話號碼。這樣的抽樣設計符合簡單隨機抽樣 (simple random sampling, SRS) 的學理要求，所抽出的樣本具有代表性，能夠推論母體的行爲與態度。但問題是，這樣的抽樣設計所針對的母體 (population) 是台灣全體民衆，而非台灣的原住民。

運用這種抽樣設計所完成的調查訪問，以「2011 年原住民人權指標調查」爲例，在 5 天內（2011 年 10 月 19 日到 23 日）完成 1,155 個成功樣本，其中原住民樣本約爲 26 個。[5] 雖然整體來說，頗具時效性。在 95% 的信賴估計區間下的抽樣誤差爲正負 2.93%，也可以接受。但若要推論次母體 (sub-population)，如原住民的行爲態度，其誤差範圍是正負 19.61%。這種推論不穩定的風險，實在高到難以令人接受，但卻是當下所謂「原住民人權指標調查報告」的重要依據。易言之，現有的原住民人權指標調查報告是建築在沙灘上的城堡，基礎極易流失。

相關研究者並非不了解這個問題，因此，所採取的補救方法是採用「德菲法」(Delphi Method)，以了解研究主題的學者專家爲訪談對象，利用一連串有系統的問卷，徵詢與研究問題有關的專家學者意見。在匿名及彼此不面對面的情況下，進行數回合的問卷調查。每次調查後分析結果，連同新問卷，再分送各專家，作爲修正先前意見的參考。如此反覆進行，直到各專家間的意見差異降至最低爲止，匯集形成一致性具體的共識。其兼具會議和傳統問卷調查法優點，是建立在「結構化的資訊流通」、「匿名的群體決定」和「專家判斷」之原理上的一種研究方法 (Murry and Hammons 1995)。

傳統的問卷調查，只要製發一次問卷就完成調查工作。德菲法則需要製發一連串密集的問卷，才能完成調查研究。同時並提供多次的回饋意見，具溝通作用。其理論基礎是由一群專家學者組成的研究小組，在數次問卷填答後，意見較不一致的專家可改變原先的意見，而趨向多數專家填答的意見，因而使大多數的專家於多次問卷作答後，意見逐漸趨向於一致，即可進行後續的分析。藉由判斷群體的集中量數與意見分析，來反應群體的共識程度和不同意見的分布情形，從而做爲實證調查經驗研究結果的補充。

德菲法是很有效，且被廣泛應用的一種研究方法，很適合在資訊欠缺的條件下使用，但同樣地，它也有很嚴重的缺點。德菲法非常容易受到與會專家學者之主觀判斷的干擾；

中，必須要支應的部分。

[5] 中華人權協會的 2011 原住民人權指標報告並未披露實際成功的原住民樣本數，此處所提的 26 個樣本是依照人口比例推估的結果。

對與會專家學者的代表性有極嚴格的要求，若代表性不平均，很容易造成偏見而扭曲最後的分析結果；對於模糊性議題而言，難以劃分優缺點，可能造成同一現象被同時解釋成優點與缺點的邏輯謬誤狀況。譬如，「2011 原住民人權指標報告」中的德菲法分析結果即承認，雖然調查結果顯示，民眾認爲政府現有措施對原住民人權的保障已經足夠，但這是「漢人眼中的原住民人權狀況」，而非「原住民本身的感受」（劉佩怡　2011, 14-15）。因此，德菲法一直都只能作爲人權指標調查的補充，而非主體。但當原住民人權指標調查的基礎就已經不穩固的時候，德菲法能夠發揮多少扭轉與輔助解釋的作用，值得存疑。

面對這樣的困境，要從事特定群體（特別是少數族群，如原住民）的行爲實證調查研究，最好是能夠掌握全體原住民的電話號碼，再以此爲抽樣清冊抽出樣本，而進行調查訪問。但橫亙於前的現實問題是，這樣的抽樣清冊是否存在？即便它存在，是否能夠取得？假設它存在，且能夠被取得，那這個抽樣清冊的品質又是如何，是否眞的涵蓋全體原住民？對於這些現實問題，現階段所有的答案都是不確定的。所以，顯然不可行。

顯然地，在從事特定群體行爲實證調查研究時，我們面臨到的困難是：「抽樣清冊建立代價極高」的問題。然而這個問題應可透過抽樣設計來解決，譬如雙重抽樣。洪永泰（2013, 124）指出，雙重抽樣 (double sampling, or two-phase sampling) 是先以低廉的代價自全體當中抽取大量的樣本，然後再從這群樣本中抽取第二次樣本。之所以採用雙重抽樣，最主要的原因在於定義母體困難，抽樣清冊無從建立或建立的代價極高。這種抽樣方式在流行病學的研究中較爲常見，因爲某些檢測太過昂貴、費時，且是侵入性，而且絕大多數民眾的檢測結果可能都是陰性。例如研究者想要瞭解老年人心臟衰竭 (heart failure) 肇因於缺血性心臟病 (ischemic heart disease) 的程度，最有效的方式是進行心導管檢查 (cardiac catheterization)，只是此項檢查正好就是昂貴、費時、侵入性，因此比較好的方式是，研究者先對患者進行第一次抽樣，進行超聲波心動圖 (echocardiogram) 的檢測，以此做一個篩檢，然後針對檢測值高的病人再進行一次抽樣，施以心導管檢查，以評估心臟衰竭的眞正原因 (Chu 1998)。

因此，爲解決特殊群體調查研究樣本取得困難的問題，中央研究院「社會意向調查小組」於 2000 年實驗性地與聯合報民意調查中心合作，使用該民調中心所累積的成功樣本資料庫，篩選出原住民電話，而以這些被篩選出的原住民電話作爲「社會意向原住民專題調查」（2000 年 9 月）的樣本來源。這種作法有效地避免了大海撈針式地過濾樣本，不必藉由傳統電話調查，用每打 1,000 通電話去累積 10 幾通原住民電話的方式來獲得所需要的原住民樣本。[6] 自然有效地降低了樣本取得成本，使得原住民行爲電話調查研究變得可

[6] 電話調查當中經常有樣本套數和樣本數不一致的問題。所謂一套樣本，係指以隨機抽樣方式抽出的一群電話號碼（實務上經常以 1,000 個電話號碼爲一套），但樣本數係指有效接通並成功訪

行。

爲確保「社會意向原住民專題調查」所使用的原住民家戶電話樣本符合隨機抽樣原則，「社會意向調查小組」僅自聯合報民意調查中心電話資料庫內，選取 1999 年 6 月到 2000 年 7 月間，以全國爲抽樣範圍的調查案之成功樣本。而後摒除其餘，只留下成功樣本當中受訪者爲原住民的家戶電話。而這些以全國爲抽樣範圍的調查案，均使用前一年度最新版的台灣地區住宅電話號碼簿爲抽樣清冊（具體年度則視調查案執行日期而異），並以系統抽樣方式抽出足夠多的電話樣本後，再以「電話號碼末兩碼以隨機亂數代入」的方式，產生一組新的電話號碼，以突破未登錄電話號碼的限制（傅仰止、簡文吟與伊慶春 2001, 110-111）。

運用上述方法，「社會意向原住民專題調查」抽出 2,406 個電話號碼，執行過程當中撥打其中的 2,399 個電話（526 筆是在 pretest 階段撥打），並成功訪問 800 位原住民受訪者（其中 135 位是在 2000 年 7 月 pretest 階段完成的樣本，而後併入同年 9 月正式訪問結果當中）。

就前述「社會意向原住民專題調查」抽樣設計的討論，可以了解「社會意向調查小組」是以正式調查開始前一年內的全國範圍調查成功樣本（約 12,000 筆）爲資料蒐集對象，然後以「族群 / 省籍」爲變數，來篩選所需要的樣本。換言之，「社會意向原住民專題調查」是從 12,000 餘筆夾雜有用與無用資訊的家戶電話當中，進行資料採礦的工作，而將「有價值的礦藏」（此處所指的，當然是家戶電話內有原住民者）給篩選出來。由於這些原住民家戶電話，在第一次被抽出時，是隨機抽樣的結果。在第二次使用時（作爲「社會意向原住民專題調查」的樣本）是雙重抽樣，依然是隨機抽樣的設計。所以，「社會意向原住民專題調查」的抽樣設計，被視爲是隨機抽樣（傅仰止、簡文吟與伊慶春 2001, 111）。

然而當我們細究這樣的抽樣設計，將會發現，「社會意向原住民專題調查」的抽樣設計，至少在理論層面上，是有瑕疵的。首先，其蒐集資料的來源是 1999 年 6 月到 2000 年 7 月，聯合報民意調查中心所執行全國範圍調查案當中的原住民成功樣本，而這些樣本是以當年度中華電信公司所出版的住宅電話號碼簿爲抽樣清冊抽出的。中華電信住宅電話號碼簿每年年底（一般來說，約爲 10 月份）會全部更新完畢，通常這時候各民調單位會改採新版的住宅電話號碼簿爲抽樣清冊。所以，「社會意向原住民專題調查」所使用的樣本，實際上應是從兩組不同的抽樣清冊當中抽出。

其次，聯合報民意調查中心平均每年約執行 160 件調查案。假設這全部都是以全國爲

問的電話號碼數量。因此，兩者之間是有差異的。以本文爲例，撥打 1,000 個電話（使用一套樣本），理論上只有會 10 餘個樣本可供分析使用。

抽樣範圍的調查案，所以「社會意向原住民專題調查」的 2,406 個樣本，來自 160 個不同的調查案。依其所言之抽樣設計，個別調查案所產生的成功樣本，確實是等機率抽出。但調查案與調查案之間，則未必是如此。如果每個調查案的完成樣本數均相異，縱使它是隨機抽取，合併成資料庫後，這資料庫內的電話號碼被抽中的機率，事實上是並不相等的。因此，從嚴謹的抽樣理論角度來看，這種嘗試以「雙重抽樣」方法利用既有資料庫，來解決特定群體機率調查研究難題的抽樣方法，似乎並不可行。

然而電話調查原本就有其限制，是基於便利性與效率，而被利用爲學術研究蒐集資料的一種方式。電話調查以住宅電話號碼簿爲抽樣清冊，其隱含之意涵爲：其母體爲台灣地區擁有家用電話者，而不是全體台灣民衆。但無論在學術研究或社會認知上，通常都用電話調查結果來推估全台灣民衆的態度。原因在於台灣地區家戶電話的高普及率。以 2013 年爲例，雖然受到手機普及率逐年提高的衝擊，台灣地區家戶電話普及率仍高達 94.8%（行政院主計總處　2012）。換言之，電話調查幾乎可以涵蓋台灣所有的人口。

就原住民部分而言，1998 年時，台灣原住民家戶當中擁有電話的比例爲 88.8%（原住民族委員會 1998, 157），而同年度，全台灣民衆家戶電話普及率爲 97.5%。其中，依可支配所得按戶數五等分位分最低者，電話普及率是 91.3%（行政院主計總處　1999）。由於原住民族委員會目前僅作經濟狀況調查（2006 和 2010），其中未有家戶電話普及率之數據，致使無官方統計數據可供參酌。然參考 1998 年的家庭收支調查結果，原住民家戶電話普及率，比依可支配所得按戶數五等分位分最低者，尚低 2.5%。而 2013 年，依可支配所得按戶數五等分位分最低者，電話普及率是 87.9%（行政院主計總處　2012）。合理推估 2013 年原住民家戶電話普及率，應該至少有 85% 以上。因此，在這樣的情形下，使用住宅電話號碼簿爲抽樣清冊而執行之電話調查的樣本代表性，應能夠涵蓋絕大多數原住民人口。

至於前述提到執行期間跨年度而造成使用不同版本住宅電話號碼簿爲抽樣清冊，是否造成涵蓋不同的問題。如果執行單位在抽出住宅電話號碼簿內的電話號碼後，便直接使用該電話號碼爲樣本，清冊不同，自然有所差異。但若執行單位並非直接使用，而是以該電話號碼爲基礎，將其後兩碼或四碼以隨機亂數代入，創造出新的樣本電話號碼。在理論上，這樣的樣本涵蓋所有有家戶電話者。因此，即便有局碼增加或減少的情形發生，該年度所抽出的樣本，依然涵蓋所有家戶電話者。使用不同版本住宅電話號碼簿爲抽樣清冊，只要每年都使用最新的版本，對個別調查案的樣本代表性，並無影響。

其次，有關於使用資料庫內樣本抽出機率不相等的問題，現象確實存在。從抽樣理論出發，雖然抽取率不同，但若不嫌煩瑣，可以在資料處理時，以加權方式，讓各樣本的抽取機率達到一致。但是否有必要依據每個調查案成功樣本的抽取率，來作加權。特別是樣

本資料來自數百個不同調查案，而這種加權所耗費的時間精力頗爲巨大時，其實是有商榷餘地的。理論上，當樣本抽取率低於 1/20 時，抽取率高低對於推論母體的影響已經幾乎可以忽略不計（洪永泰　2013, 118）。檢視「社會意向原住民專題調查」所使用的樣本，其抽取率都在 0.01% 以下，遠低於 1/20 的門檻。因此，縱使資料庫內各樣本的抽出機率有極其細微的不同，但其對於推論母體的影響微乎其微。從提升調查效率與便利性來看，使用資料庫電話資料作爲特定群體機率調查之用，從實務上來看，是可行的。

有鑑於此，在幾近不可能取得全體原住民家戶電話號碼清冊的情形下，使用電訪樣本資料庫便成爲唯一可行的辦法。而在現階段台灣社會通訊高度發展的條件下，縱使電話普及率未達 100%，但足以涵蓋絕大多數人口。同時，資料庫內樣本的抽取機率差異問題，又對推論母體影響甚微。作爲研究者，應該要注意且處理的問題是：資料庫所蒐集的原住民家戶電話是否存在系統性的偏差，因而使推論母體行爲可能產生錯誤。要避免系統性偏差，最好的辦法是盡可能地擴大資料庫，並使電話號碼來源愈廣愈好。藉由整合愈多不同來源的成功樣本，並使其數量增加愈多，我們愈可能避免掉系統性偏差的問題，理論上會愈接近母體。

貳、調查簡介與調查過程

國內民調機構衆多，但各民調機構之間，尚未有資料互換合作的經驗。事實上，民調機構普遍將調查樣本（無論成功或失敗）視爲重要機密資產，不輕易釋出。在這樣的情形下，試圖找到數個民調機構，使資料庫的資料來源多樣化，並有相當規模的樣本數量，同時還願意提供做爲學術研究之用，實屬艱鉅任務。研究團隊曾嘗試與擁有鉅量資料的政黨和大學民調中心協商資料使用或交換，但均因前述理由而無法成局。由於研究團隊成員在過去十年間，曾多次與全國公信力民意調查股份有限公司合作，建立互信，而獲該公司應允提供協助，初步解決資料取得困難的問題。其次，該公司資料庫內的樣本，並非全部自行抽樣與執行。有部分是委託其他民調機構（包括兩所大學民調中心、兩家民調公司同業）代工抽樣與執行。[7] 鑒於全國公信力民意調查股份有限公司之資料庫能符合前述「資料來源避免單一」，又能提供相當規模的樣本數量，研究團隊乃委託該公司代爲執行 2013 原住民政治態度與行爲調查案。經費來源則是由淡江大學 102 學年度重點研究計畫項目下

[7] 在樣本整理階段，所有樣本均只留下電話號碼與省籍變數，其他相關人口變數或足以辨識身分的資訊均予以刪除，故應無侵害隱私問題。其次，雖然本研究所使用資料庫內的電話來源較爲多元，不限於全國公信力民意調查股份有限公司一家，但該公司自有資料仍占本研究所使用資料的 9 成以上。至於具體委託代工之大學民調中心和民調公司同業，雖基於業務保密問題，不便披露，但應不影響研究進行。

支應，而未受其他任何政府機構、政黨、或利益團體之任何形式的資助。

2013 原住民政治態度與行爲調查，計畫本身主要目的在於，研究社會系絡 (social context) 對原住民選舉投票行爲的影響。就研究內涵而言，是嘗試援引西方政治行爲理論，透過個體層次的實證經驗研究，用以論證與解釋台灣原住民的政治行爲。既然是探討原住民選舉投票行爲，研究對象自然是有投票權人，所以在計畫成形前，即將訪問對象設定爲年滿 20 歲以上之原住民受訪者。同時因爲希望能有較大樣本規模，期盼能盡可能降低推論誤差，故將成功樣本數量的底線，設定在 1,068。經與研究團隊成員討論後，該公司考慮到過去類似 panel study 的再訪成功率，並顧及可能會有的樣本流失問題，故擴大 4.5 倍抽樣，篩選 2007 年 1 月迄今（2013 年 3 月）間，以全國爲抽樣範圍之調查中，有接受該公司（或該公司委託代工之民調機構）訪問之原住民家戶電話號碼進行訪問，共計抽出 4,873 個電話號碼。

2013 原住民政治態度與行爲調查於 2013 年 9 月 3 日（週二）正式開始執行調查，以任意選樣方式選擇合格受訪者，而至 9 月 8 日（週日）完成。期間每日晚間 18:30 至 22:00 爲調查時間，但考慮到夜間工作的可能性與具基督信仰之受訪者可能週日白天有教會活動，故於 9 月 7 日（週六）執行兩場次的調查訪問，分別是下午 13:30 到 17:00 和晚間 18:30 到 22:00。調查時間涵蓋週間晚上與週末下午和晚上。由於調查訪問時間長達 6 天，當合格受訪者不在家（訪問期間內有可能接觸到）或當時不方便接受訪問，以及無人接聽與接電話即拒訪者，均於隔兩日 / 場追蹤接觸。重複接觸次數設定爲至多 3 次，如在成功樣本數已達要求（預設值爲 1,068），而仍未成功接觸者，則視爲失敗樣本（但當場次應撥打之電話仍將全數再撥打一遍）。如此，在正式調查過程中，總計撥出 9,768 通電話，若以電話號碼（樣本）計，則是撥打了 4,873 個電話號碼。其中成功完成訪問的樣本數爲 1,102。

下表 1 是本次調查以撥出電話號碼數爲分母的所有訪問結果，其中訪問成功者占全體樣本的 22.6%。所有撥打電話之接通率爲 52.0%（電話接通 / 總電話數），接通之後的成功率爲 43.5%（訪問成功 / 電話接通數）。至於無法完成訪問之主要原因分別是：無人接聽 (21.6%) 和空號 (20.6%)，其次是無合格受訪者（該家戶內沒有原住民，占 13.7%）。其他原因者，除受訪者不在 (5.2%)、接電話即拒訪 (4.0%)、電話中 (3.0%) 外，其他原因的比例均在 3% 以下。

表 1　訪問結果表

訪問結果 [a]	接通與否	次數	百分比	重新歸類訪問結果	次數	百分比
訪問成功	電話接通	1,102	22.6	訪問成功	1,102	22.6
受訪者暫時不在或不便接聽（非當日約訪者）		254	5.2	確定有原住民受訪者，但無法在訪問期間完成訪問。	452	9.3
受訪者因臨時有事而中途拒訪（非當日約訪者）		26	0.5			
受訪者暫時不在 或 不便接聽（可當日約訪者）		15	0.3			
受訪者因臨時有事而中途拒訪（可當日約訪者）		3	0.1			
受訪者拒絕受訪（無法再訪）		16	0.3			
受訪者中途拒訪（無法再訪）		127	2.6			
受訪者訪問期間不在		11	0.2			
接電話即拒訪或接電話者有事（我很忙）		197	4.0	拒訪，且無法確認是否有原住民受訪者。	197	4.0
因語言因素無法受訪（原住民語）		33	0.7	因爲語言、生理、心理因素無法受訪，但確認有原住民受訪者。	46	0.9
因生理或心理因素無法受訪		13	0.3			
無人接聽	電話未接通	1,051	21.6	無人接聽、傳眞機、答錄機等無法確認是否有原住民受訪者。	1,138	23.4
傳眞機或其他資料傳輸設備		75	1.5			
答錄機（住宅）		12	0.2			
通話中或電話沒掛好（經立即重撥一次後仍忙線）		147	3.0	通話中或技術封閉使無法聯絡受訪者，且無法確認是否有原住民受訪者。	150	3.1
以通訊技術封閉電話		3	0.1			
停話		29	0.6	停話、空號、轉號、故障而使原樣本失效。	1,045	21.4
空號		1,006	20.6			
移機轉號		1	0.0			
電話故障		9	0.2			
答錄機（非住宅）		5	0.1	非住宅電話	74	1.5
公司、政府機關或其他非住宅	電話接通	67	1.4			
群居單位（如宿舍）		2	0.0			
戶中無合格受訪對象（或外國人）		669	13.7	無合格受訪者，但部分存疑。[b]	669	13.7
總和		4,873	100.0	總和	4,873	100.0

資料來源：2013 原住民政治態度與行爲調查。

說明：1. [a] 複數次數接觸之樣本，以最後接觸之結果爲訪問結果。

2. [b] 此類電話號碼經接通後，接電話者表示該電話所在之處所內，並無任何原住民居住。理論上，應視爲原先原住民樣本流失。但部分接電話者，經訪員辨認（或於監聽時，由督導辨認），疑爲原住民，但拒絕承認。這類電話的比例約占「戶中無合格受訪對象」的 3 成左右。但由於並非所有此類情形都會被訪員回報或被督導監聽到，故可能比例應會更高。無論如何，由於未經系統性紀錄，故以上數字僅供參考。

最初在計畫開始前，研究團隊擬仿效 2000 社會意向原住民專題調查，僅擬篩選 2012 到 2013 年間的成功樣本。但由於合作單位所屬電話樣本資料庫當中，合格成功樣本不敷所需，考慮到 2007 到 2013 年間，原住民人口數從 484,906 人，增加到 531,435 人。[8]就表 2 的數據顯示，2007 年以前增加幅度較大，特別是 2002 年的接近 3%，之後 2004 年的 2.78%。所以縱使 2001 年《原住民身分法》通過後有部分原住民恢復取得身分，但在 2007 年（含之後）的資料庫抽樣時，理論上應該多數都已經恢復並取得原住民身分，而會出現在抽樣清冊當中。縱有少數 2007 年之後陸續才回復身分者，也應屬零星情況，對本研究影響有限，不致造成系統性的差異。而同期間內，合作單位已經累積足夠多的樣本資料，因而決定將擴大資料蒐集範圍，以滿足大樣本規模的原始計畫要求。但由於電話資料庫內的樣本，係來自不同年度的調查案所累積而成，有必要檢視不同年度樣本資料的訪問結果分布狀態，以便了解使用跨年度樣本資料的影響。

表 2　歷年原住民人口數與增加情形

年度	原住民人口數	增加人數	增加比例
2001	420,300		
2002	432,516	12,216	2.91%
2003	442,388	9,872	2.28%
2004	454,682	12,294	2.78%
2005	464,961	10,279	2.26%
2006	474,919	9,958	2.14%
2007	484,174	9,255	1.95%
2008	494,107	9,933	2.05%
2009	504,531	10,424	2.11%
2010	512,701	8,170	1.62%
2011	519,984	7,283	1.42%
2012	527,250	7,266	1.40%
2013 Aug	531,435	4,185	0.79%

資料來源：作者整理自原住民族委員會之歷年原住民人口數（原住民族委員會　2013）。

[8] 原住民人口數目的增加，並不全然是自然增加。2001 年通過《原住民身分法》後，過去因爲原漢通婚冠夫姓而喪失原住民身分的個人，以及因而登記爲非原住民之子女，均可依法恢復原住民身分。

表 3 是訪問結果的跨年度比較，各年度樣本數差異頗大，主要和選舉有關。舉凡大型選舉的前一年（投票日在年初）或當年（投票日在年底），合作單位因爲調查案較多，能夠產生較多樣本進入資料庫。譬如 2008 年總統選舉，相關民調委託案都集中在 2007 年執行，所以 2007 年有 1,121 個樣本。2008 年 3 月 22 日總統選舉投票結束，其後調查案執行數目較少，所以 2008 全年，合作單位只有 345 個樣本。然而無論各年度樣本數多寡，樣本取得年度愈早者，愈可能因爲種種原因而流失，而使訪問成功的機率下降。初步檢視訪問結果爲成功的跨年度比較，發現雖然 2007 樣本訪問成功比率最低，只有 14.9%，但通盤檢視 2013 到 2007 年間的訪問成功率，也看不出來有「年度愈早，成功率愈低」的趨勢。

訪問成功率受到很多因素影響，但確定流失的樣本數量愈大，訪問成功率自然受到的影響就愈大。因此，我們再度檢視確定爲「樣本流失」的三種訪問結果（停話／空號、非住宅、戶中無合格受訪者）的跨年度比較。這三類訪問結果合併後的數據依 2007 到 2013 年排列，分別爲 33.9%、45.6%、38.0%、35.1%、37.3%、34.7%、42.3%。各年度樣本確定流失率亦沒有隨年度愈早而遞增的趨勢，而且 2007 年的確定流失率反而是最低的。如果我們只看樣本數超過 1,000 以上的三個年份（2007、2010、2011），避免樣本數較少造成推論不穩定的情形，可以發現樣本確定流失率依序爲 33.9%、35.1%、37.3%。時間愈早的流失率反而愈低。但考慮到部分「戶中無合格受訪者」可能係爲原住民，但拒絕承認（見表 1 說明 2）。將此一趨勢詮釋爲「樣本流失率穩定 35% 上下浮動」，毋寧是更恰當的解釋。

如此一來，影響各年度訪問成功率因素的解釋，就落在「無人接聽」和「通話中」兩類，無法確定到底有沒有合格原住民受訪者之上。後者比例很低，對訪問成功率的影響有限。而前者比例愈高者，確實會造成訪問成功率降低的情形。雖然 2007 和 2010 有較高比例的無人接聽（分別爲 33.3% 和 26.9%），但其他年度的無人接聽率大概都在 15% 上下 4 個百分點浮動，輔以前述「樣本流失率穩定 35% 上下浮動」，使研究團隊傾向認爲，這些無人接聽的電話有相當可能性，只是因爲單純地運氣不佳，而無法在電訪執行期間內接觸到。

表 3　訪問結果表——跨年度比較

樣本年度	訪問成功		原住民受訪者無法在訪問期間完成訪問		拒訪：無法確認有無合格受訪者		原住民受訪者因語言或生心理因素無法受訪		無人接聽／住宅答錄機／傳真機：無法確認有無合格受訪者		通話中／技術封閉：無法確認有無合格受訪者		停話／空號／轉號／故障：使原樣本失效		非住宅		戶中無合格受訪者		總和
	個案數	%	個案數	%	個案數	%	個案數	%	個案數	%	個案數	%	個案數	%	個案數	%	個案數	%	個案數
2007	167	14.9%	108	9.6%	32	2.9%	6	0.5%	373	33.3%	55	4.9%	270	24.1%	10	0.9%	100	8.9%	1,121
2008	91	26.4%	21	6.1%	12	3.5%	0	0.0%	55	15.9%	9	2.6%	62	18.0%	3	0.9%	92	26.7%	345
2009	141	29.7%	39	8.2%	28	5.9%	5	1.1%	76	16.0%	5	1.1%	79	16.6%	4	0.8%	98	20.6%	475
2010	204	18.5%	105	9.5%	47	4.3%	22	2.0%	297	26.9%	42	3.8%	256	23.2%	19	1.7%	113	10.2%	1,105
2011	462	27.0%	169	9.9%	73	4.3%	13	0.8%	321	18.7%	36	2.1%	365	21.3%	35	2.0%	239	14.0%	1,713
2012	24	34.8%	8	11.6%	3	4.3%	0	0.0%	8	11.6%	2	2.9%	6	8.7%	3	4.3%	15	21.7%	69
2013	13	28.9%	2	4.4%	2	4.4%	0	0.0%	8	17.8%	1	2.2%	7	15.6%	0	0.0%	12	26.7%	45

資料來源：2013 原住民政治態度與行爲調查。

參、調查成效評估

以個體層次的實證調查結果來做母體推論統計的分析，很重要的一個條件就是，樣本的配置應該符合母體的分布結構，否則即有推論偏差的危險。因此，要評估 2013 原住民政治態度與行爲調查的成效，最直觀的做法就是檢定成功樣本的代表性。若代表性檢定結果爲成功樣本與母體不一致時，爲避免資料分析時，造成推論的偏差，一般調查研究往往會以加權方式來調整成功樣本的結構，使其符合母體分布，然後再進行推論統計分析。就本文研究目的而言，加權與其後的資料推論分析並非重點。本文所關注的問題是，利用資料探勘與雙重抽樣方式所蒐集到的原住民電話樣本，到底有沒有結構性的偏差存在？從附錄附表 1 到附表 6 的數據來看，2013 原住民政治態度與行爲調查成功樣本，在性別、年齡、教育程度、平地或山地原住民、戶籍所在縣市，以及族群別上，都與母體的分布不一致。[9]

這樣的結果似乎暗示 2013 原住民政治態度與行爲調查的成效不彰。但電話調查的成功樣本結構，除非一開始就以配額方式設定樣本比例，實務上來看，電話調查成功樣本結構，因爲無法確知接聽電話者的身分，原本就很難有符合母體分布情形。即便採用戶中選樣方式，通常也只能夠勉強讓性別比例與母體分布相符合。其他社會人口變項的分布狀況還是很難符合。所以，這種成功樣本與母體分布不一致的結果，其實並不讓人意外。若是一般調查，這時候就會採取各類加權方式，譬如事後分層或是多變數反覆加權 (raking)，來調整樣本結構，讓它符合母體分布。如果族群和地區與母體分布差異極大，甚至產生有部分類別原本應有相當比例樣本，但實際成功個案數卻很少的狀況（反之亦然），那 2013 原住民政治態度與行爲調查，應用電話樣本資料庫來取得電話樣本的方法，就有進一步商榷的必要。倘若本調查的成功樣本結構並沒有前述情形發生，而只是些微差異存在導致代表性檢定不過關，那使用加權方式來調整成功樣本結構，以便進行推論統計分析，應該可以被視爲是合理的做法。由於台灣原住民族的社會人口分布有其特殊性，所以，即便樣本代表性檢定結果未若盡如人意，我們還是希望能夠進一步觀察成功樣本的族群與地區（戶籍所在縣市）分布是否有結構性偏差的狀況發生。

台灣地區原住民族的社會人口分布狀況迥異於一般漢人。首先，雖然統稱爲原住民，但實際上，台灣原住民至少包含 14 個不同的族群。[10] 不過，族群之間的人口規模差異甚

[9] 族群人口數的部分，由於原住民族委員會並未有官方的分齡統計，所以本文是以 2013 年 6 月間的全體族群人口數爲母體分布資料，包含 20 歲以下者，作爲檢定基準。

[10] 2013 年官方承認爲「原住民族」的族群有 14 族，但仍有其他「原住民族」正努力爭取獲得官方證明承認。譬如漢化程度已深的平埔族群，如凱達格蘭族、巴宰族、噶哈巫族、馬卡道族、洪雅族、巴布拉族、道卡斯族等族群，正發起正名運動，要求政府承認其原住民身分。

大。以 2013 年 8 月之原住民人口統計爲例（詳見下表 4），台灣地區現具原住民身分且登記所屬族群者有 514,546 人，其中阿美族便占總數的 38.3%，有近 20 萬人。其次是排灣（18.4%，94,573 人）、泰雅（16.4%，84,364 人）、布農（10.7%，55,073 人）。但這 4 族就占全體原住民的 83.7%，其他 10 族合計才只占 16.3%。這 10 族裡面，人口數在 1 萬以下的原住民族有 7 族，其中人口比例不到全體 1% 的，就有 4 族（達悟、噶瑪蘭、撒奇萊雅、邵），而撒奇萊雅和邵族的人口數更只有不到 1,000 人。

表 4　調查樣本與母體的族群別分布

樣本	次數	百分比	母體	人口數	百分比	差距
阿美	355	32.5	阿美	196,863	38.3	-5.8
排灣	210	19.2	排灣	94,573	18.4	0.8
泰雅	173	15.9	泰雅	84,364	16.4	-0.5
布農	138	12.6	布農	55,073	10.7	1.9
太魯閣	53	4.9	太魯閣	28,864	5.6	-0.7
卑南	41	3.8	卑南	13,041	2.5	1.3
魯凱	42	3.8	魯凱	12,632	2.5	1.3
賽德克	26	2.4	賽德克	8,582	1.7	0.7
鄒	31	2.8	鄒	7,063	1.4	1.4
賽夏	9	0.8	賽夏	6,303	1.2	-0.4
達悟	3	0.3	達悟	4,365	0.8	-0.5
噶瑪蘭	5	0.5	噶瑪蘭	1,334	0.3	0.2
撒奇萊雅	3	0.3	撒奇萊雅	747	0.1	0.2
邵	2	0.2	邵	742	0.1	0.1

資料來源：樣本部分來自於 2013 原住民政治態度與行爲調查，母體部分則是引用原住民族委員會 2013 年 8 月份之原住民人口統計。

其次，台灣地區原住民族雖普遍聚居於若干偏遠鄉鎮，政府甚至因此劃定原住民鄉鎮，給予特殊地位。但由於原住民鄉鎮的經濟發展程度，遠不若台灣其他地區。生活壓力迫使相當數量的原住民離開原鄉，另謀生計，而主要在基隆市、台北市、新北市、桃園縣等都會區或都會區外圍衛星城市落腳，而形成新的原住民聚落（包正豪與謝顥音　2013, 84；傅仰止　1985）。當然也有散居於其他地區的外移原住民。

在這樣的社會人口分布情形下，要針對僅占總人口數 2.27% 的特殊族群做機率抽樣調查，往往會基於便利而將調查範圍侷限在原住民鄉鎮（張茂桂、顏愛靜與傅君 1998），或選擇移居地的原住民爲研究對象（瞿海源與傅仰止　1986，轉引自傅仰止、簡

文吟與伊慶春　2001, 115）。[11] 此次 2013 原住民政治態度與行爲調查，嘗試以電話調查方式將觸角延伸到全台灣各地，是否會因爲原住民地理分布極不平均的關係，而導致樣本分布的偏差（譬如成功樣本高度集中在花東兩縣），即成爲本研究進行前最大的擔憂之一。同時，14 族原住民之間的人口規模差異過大，前四大族便占全體原住民的 83.7%。使用電話調查進行研究，能否找到人口規模較小族群的受訪者，還是有效成功樣本將高度集中在大族受訪者？就算能找到人口規模較小族群的受訪者，但其占成功樣本的比例是否能相當程度符合母體的比例，也是本研究所高度關心的問題。

爲評估 2013 原住民政治態度與行爲調查之成效，檢驗本研究是否能夠成功克服族群與居住地區分布極度不平均的困難，本研究將首先比較成功樣本與母體之間，族群與居住地區分布狀況的差異。然後再與傅仰止、簡文吟與伊慶春（2001）的研究發現做比較討論。

下表 5 是調查樣本與母體之間，平地與山地原住民的分布比例比較。雖然平地與山地原住民的行政區分是否有必要，學界與社會多有爭議（包正豪　2009），但迄今並沒有任何相關研究討論，甚至於證明，平地原住民和山地原住民的身分差異對原住民政治態度或行爲有無影響。因此，2013 原住民政治態度與行爲調查擬將其列爲自變數之一，故而有必要於此觀察調查樣本與母體之間，平地與山地原住民的分布比例。表 5 的數據顯示，調查樣本當中有 41.1% 是平地原住民，但母體的平地原住民比例卻是 47.0%。兩者之間雖存在 5.9% 的差距，但還算不上過分扭曲，應還能接受。

表 5　調查樣本與母體之平地與山地原住民比例分布

樣本	次數	百分比	母體	人口數	百分比	差距
平地原住民	453	41.1	平地原住民	249,918	47.0	-5.9
山地原住民	649	58.9	山地原住民	281,517	53.0	5.9

資料來源：樣本部分來自於 2013 原住民政治態度與行爲調查，母體部分則是引用原住民族委員會 2013 年 8 月份之原住民人口統計。

調查樣本低估平地原住民比例，可能直接導因於調查過程中找到較少阿美族受訪者。根據表 4 的數據，調查樣本與母體的族群別分布，阿美族成功樣本的比例較母體略低 5.8%。而同時根據原住民族委員會 2013 年 8 月最新原住民人口統計，全國 249,918 位平地原住民當中，阿美族占 78.3%（195,788 人），其次是排灣族的 7.8%（19,412 人）和卑南族的 5.2%（12,958 人）。這三族就占了平地原住民的 91.3%。但在調查樣本當中，排灣與卑南受訪者的比例都較母體比例略高 0.8% 和 1.3%，因而可以合理推論平地原住民受訪

[11] 瞿海源與傅仰止（1986）之著作係轉引自傅仰止、簡文吟與伊慶春（2001, 115）。

者比例較少，是因爲較少成功訪問到阿美族受訪者所造成的。

不過，整體觀察調查樣本與母體的族群別分布時，可以發現調查樣本的族群別分布，其實密切呼應母體的族群別分布。表 5 母體部分的各族群別是依人口規模由大到小排序，對照表 4 左側的調查樣本來看，幾乎是完美對應。說是「幾乎」，是因爲母體當中達悟族人口規模排序爲第 11 位，噶瑪蘭族則是第 12 位，但在調查樣本當中兩族排序倒置。考慮到達悟（4,365 人）和噶瑪蘭（1,334 人）兩族人口數不及 5,000，占全體原住民人口的比例亦不足 1%(0.8% vs. 0.3%) 的情形，樣本數量上 1 個個案的差異就足以使兩族排序變更。達悟與噶瑪蘭排序，未能呼應母體分布排序的問題，在調查實務上，幾可忽略不計。

進一步比較各族群別的樣本與母體比例，可以發現 2013 原住民政治態度與行爲調查成功樣本之族群分布，除阿美族低估 5.8% 和布農族高估 1.9% 外，其他所有族群的樣本和母體比例差距都在正負 1.5% 以下。如果我們考慮到調查對象爲年滿 20 歲以上的原住民，但母體資料卻是引用原住民族委員會 2013 年 8 月最新人口統計數，而其中還包括 0 歲到 19 歲的原住民時，如此微小差距，其實某種程度是逆向反證電話調查的功效極高。[12]

而與 2000 年社會意向原住民專題調查的結果（詳見表 6）相比較，該調查之樣本分布高估布農 (4.3%) 和賽夏 (1.6%)，而低估排灣 (2.3%) 和阿美 (1.8%)。由於該調查執行期間，官方所承認的原住民族還是 9 族，所以表 6 當中的泰雅，其實還包括太魯閣和賽德克兩族，鄒族的數據也包含了邵族，至於撒奇萊雅與噶瑪蘭則被納入阿美族。在 9 族分類的情形下，社會意向原住民專題調查樣本的族群排序分布「相對」不完美。以父親族別爲樣本族群別來看，前五大族群的樣本與母體排序一致，但母體比例當中的第六、七、八位應該爲卑南、鄒、賽夏，但實際調查果的第六、七、八位卻是賽夏、卑南、鄒。

與 2013 原住民政治態度與行爲調查相類似，2000 社會意向原住民專題調查的調查對象與當時母體亦有差異。原住民專題調查的對象是 20 到 60 歲的原住民，而母體資料則是使用 1997 年底，包含 20 歲以下和 60 歲以上原住民的數據。因此，雖然出現些微差異，卻仍能以 800 個樣本規模的調查結果，得出與當時將近 40 萬原住民人口相當接近的族群比例分布，亦再度反映出使用成功電話樣本資料庫的功效與可行性。而從 2000 社會意向原住民專題調查和 2013 原住民政治態度與行爲調查的結果比較，則可觀察到大樣本規模的好處。前者樣本數爲 800，而後者是 1,102。後者在分類更細緻（從 9 族到 14 族）的條件限制下，樣本的族群比例，比前者更加貼近母體。[13]

[12] 根據原住民族委員會 2013 年 8 月最新原住民人口統計數據，全國 531,435 位原住民當中，未滿 20 歲者有 163,297 人，占全體原住民人數的 30.7%。

[13] 由於原住民族委員會所公開之原住民族人口統計資料當中無法切割族群與年齡，因此，本研究

表 6　2000 社會意向原住民專題調查樣本與母體的族群別分布

樣本	父親族別	母親族別	母體	百分比	差距
阿美	35.8	35.7	阿美	37.6	-1.8
泰雅	23.0	22.8	泰雅	23.4	-0.4
排灣	15.5	16.1	排灣	17.8	-2.3
布農	14.8	15.5	布農	10.5	4.3
魯凱	2.8	2.6	魯凱	3.2	-0.4
卑南	3.0	2.4	卑南	2.7	0.3
鄒	1.1	1.4	鄒	1.9	-0.8
賽夏	3.4	2.8	賽夏	1.8	1.6
達悟	0.6	0.6	達悟	1.0	-0.4

資料來源：轉引自傅仰止、簡文吟與伊慶春（2001, 114）。

再從地理分布來衡量。台灣地區原住民有近 1/3(32.0%) 居住在花蓮與台東兩縣，而屏東縣主要是排灣與魯凱兩族的傳統領域，亦有 10.8% 的原住民分布。在調查伊始之際，研究團隊對於調查結果的主要憂慮在於，調查樣本的地理分布是否會過度集中在上述三縣？雖然新北市 (9.9%) 與桃園縣 (12.0%) 亦分布相當數量之原住民，但相較於傳統聚居地，都會區原住民的成功受訪比例是否會因爲工作的關係而較爲偏低？但調查樣本與母體的地理分布比較（詳見表 7），證實了研究團隊僅是杞人憂天，所擔心預想的情形或是沒有出現，或是並不嚴重。事實上，調查結果的樣本地理分布確實略微集中在原住民傳統聚居地區，而都會區縣市的成功樣本比例也略爲偏低。不過整體而言，這樣的偏差並不嚴重。

與 2000 社會意向原住民專題調查均只能使用全體原住民的母體資料來和樣本作比較評估。而其背後隱含之假設是：「20 歲以下原住民的族群及地理分布與 20 歲以上者無異」，這是囿於資料取得困難下，實務上唯一可行解決方法。

表 7 調查樣本與母體的地理分布——以縣市為單位

縣市	樣本次數	樣本百分比	母體人口數	母體百分比	差距
花蓮縣	209	19.1	91,035	17.1	2.0
臺東縣	181	16.6	79,303	14.9	1.6
桃園縣	82	7.5	63,820	12.0	-4.5
屏東縣	158	14.5	57,510	10.8	3.6
新北市	56	5.1	52,638	9.9	-4.8
高雄市	65	5.9	31,595	5.9	0.0
臺中市	52	4.8	30,260	5.7	-0.9
南投縣	86	7.9	28,556	5.4	2.5
新竹縣	55	5.0	20,523	3.9	1.2
宜蘭縣	32	2.9	16,027	3.0	-0.1
臺北市	20	1.8	15,058	2.8	-1.0
苗栗縣	25	2.3	10,926	2.1	0.2
基隆市	15	1.4	8,902	1.7	-0.3
臺南市	8	0.7	6,664	1.3	-0.5
嘉義縣	26	2.4	5,695	1.1	1.3
彰化縣	12	1.1	5,176	1.0	0.1
新竹市	6	0.5	3,459	0.7	-0.1
雲林縣	1	0.1	2,085	0.4	-0.3
嘉義市	2	0.2	916	0.2	0.0
金門縣	0	0.0	747	0.1	-0.1
澎湖縣	0	0.0	375	0.1	-0.1
連江縣	2	0.2	165	0.0	0.2
總和	1,093	100.0	531,435	100.0	

資料來源：樣本部分來自於 2013 原住民政治態度與行爲調查，母體部分則是引用原住民族委員會 2013 年 8 月份之原住民人口統計。

2013 原住民政治態度與行爲調查的樣本地理分布，在全國 22 縣市中，除金門與澎湖兩縣未有成功樣本外，其他 20 個縣市都至少有 1 個成功樣本。就算該縣市之原住民人口數不及全體原住民的 1% 也是一樣。金門縣與澎湖縣未有成功樣本的原因其實很明顯，金門有 747 位成年原住民，而澎湖只有 375 位。其比例原本就低到不易出現成功樣本。但連江縣能成功兩個個案，也非單純幸運可以解釋。因爲連江縣雖然只有 165 位成年原住民，但連江縣的設籍人口數爲 12,041 人，比例爲 1.37%，其實是遠高過金門縣的 0.6% 和澎湖縣的 0.4%。

整體來看，花蓮、台東、屏東、南投、新竹、嘉義等縣，係傳統原住民聚居地區，樣

本分布比例均略高於母體應有之比例（1.2% 到 3.6% 之間）。台北市、新北市、台中市、桃園縣等都會區原住民聚集的縣市則普遍有低估樣本比例的情形，特別是新北市（52,638 人）與桃園縣（63,820 人）分別低估 4.8% 和 4.5%。台北市和台中市則是約低估 1%（台中市爲 0.9%）。至於其他縣市，樣本與母體之間的差距，均少於 0.5%。除樣本分布比例與母體應有比例的差距有限外，這些差距對縣市排序的影響，除南投縣依母體比例應排第八位，因爲樣本成功數較多而排序第四位外，其他縣市的影響多在上下一位左右。換言之，2013 原住民政治態度與行爲調查之樣本的地理分布，大致與母體相似。

至於 2000 年社會意向原住民專題調查樣本與母體的地理分布情形（詳見表 8），大致與 2013 原住民政治態度與行爲調查的結果相似。傳統原住民聚居地區，除屏東縣低估 4.7% 和新竹縣低估 1.5% 外，其他花蓮、台東、嘉義等三縣均高估 0.9% 到 3.3% 之間。台北市、新北市、桃園縣等三大都會原住民聚集縣市，則均低估 1.0% 到 1.3% 之間。而這些差距對縣市排序的影響，也頗類本研究所得。也是除了南投縣依母體比例應排第六位，因爲樣本成功數較多而排序第三位外，對其他縣市的影響也都是在上下一位左右。

因此，兩相比較之下，2013 原住民政治態度與行爲調查的地理範圍較 2000 社會意向原住民專題調查來得廣，擴及金門與連江兩縣，但調查結果當中，澎湖和金門兩縣是沒有任何成功樣本的。後者研究範圍僅包括台灣本島各縣市與澎湖縣，只有未改制前的台南市沒有成功樣本。其次，兩個調查研究案的樣本地理分布情形，都可以觀察到「原住民傳統聚居地區縣市樣本略爲高估，而新興都會原住民聚居地區縣市樣本略爲低估」的趨勢。只是在 2013 原住民政治態度與行爲調查當中，這樣的趨勢更明顯，幅度也較 2000 社會意向原住民專題調查略高。這種趨勢有可能肇因於原鄉原住民移居台北桃園都會區比例增高，而因爲工作或其他因素，未能成功訪問。但現階段無從驗證。不過，就以上比較來看，兩個調查案之成功樣本的地理分布，並無明顯不同，也都於相當程度反映台灣原住民母體地理分布的趨勢。

表 8　2000 社會意向原住民專題調查樣本與母體的地理分布——以縣市為單位

縣市	樣本比例	母體比例	差距	縣市	樣本比例	母體比例	差距
花蓮縣	21.4	20.5	0.9	苗栗縣	1.5	2.1	-0.6
台東縣	22.0	18.7	3.3	基隆市	0.8	1.6	-0.8
桃園縣	7.9	9.2	-1.3	台南市 [a]	0.3	0.7	-0.4
屏東縣	7.8	12.5	-4.7	嘉義縣	2.4	1.1	1.3
新北市	6.6	7.6	-1.0	彰化縣	1.1	0.7	0.4
高雄市	5.6	5.1	0.5	新竹市	1.0	0.4	0.6
台中市	5.0	4.3	0.7	雲林縣	0.3	0.2	0.1
南投縣	8.5	6.0	2.5	嘉義市	0.1	0.1	0.0
新竹縣	2.4	3.9	-1.5	金門縣	-	-	-
宜蘭縣	4.4	3.0	1.4	澎湖縣	0.3	0.0	0.3
台北市	1.0	2.1	-1.1	連江縣	-	-	-

資料來源：轉引自傅仰止、簡文吟與伊慶春（2001, 116），並為配合比較 2013 原住民政治態度與行為調查而有適度修改。

說明：1. 2000 社會意向原住民專題調查執行時，縣市尚未合併改制，台灣地區僅有北高兩個直轄市。為方便比較起見，本文將以 2013 年為準，以縣市合併改制後的行政區劃來合併社會意向原住民專題調查的資料。另，社會意向原住民專題調查僅限於台灣地區，而排除金門、連江兩縣，但 2013 原住民政治態度與行為調查則有包括這兩個離島縣。

2. [a] 2000 社會意向原住民專題調查原始結果當中，未改制前的台南市沒有成功樣本。此處資料是，含未改制前的台南縣的成功樣本比例。

肆、結論與建議

台灣原住民由於人口規模小，長期以來，未受政治學界重視。縱使有部分研究關注，但普遍屬於質性訪談的小規模研究，或是附屬於全國性調查研究之下，使用零星樣本作分析，其所得結論難以推論全體原住民。但台灣的政治制度設計卻高度重視原住民的政治參與，不僅在各級選舉設立原住民專屬選區，而其保障之席次比例亦遠高於其人口規模比例。同時，自總體資料分析結果來看，原住民的政治行為顯然與漢人不相當。深究原住民政治行為不受重視的原因，個體層次實證經驗調查研究的成本過高，應為主要原因之一。為解決此一問題，本調查計畫乃思考使用「電訪樣本資料庫」與「雙重抽樣」作為解決方案，透過電訪樣本資料庫的再利用，有效減輕篩選原住民樣本的成本，使原住民機率調查研究成為可行。

原住民抽樣調查的困難度，即在抽樣清冊不易建立，因此適合用雙重抽樣的方式來解決。當然，本研究所使用的抽樣方式並不完全符合雙重抽樣的定義，首先，我們並非一次

性的從母體抽出大量樣本，詢問是否爲原住民後，再針對原住民進行抽樣，而是合併多筆資料進行篩選；其次，我們在第二階段並未抽樣，也就是說原則上只要在第一階段被篩選出是原住民的樣本，都會嘗試去訪問。雖然這二項作法並不完全符合雙重抽樣的定義，不過在第一階段，精神上我們的確是以省籍作爲篩選變數，至於第二階段全查的原因，這是在考量拒訪率之後，爲了提高最後訪問成功樣本數，減少抽樣誤差所做的決定。

從研究結果來看，觀察台灣原住民的「族群分布」與「地理分布」兩項指標，2013 原住民政治態度與行爲調查之成功樣本與母體分布，就樣本代表性的統計檢定結果來看，樣本與母體並不一致。但這並不意外，一般電訪調查樣本都有這個問題。實務上，多數以加權方式解決。但以原住民爲對象的電訪調查必須要特別注意族群和地理分布的問題。本研究結果顯示，在族群和地理分布上，樣本與母體的代表性檢定並未過關，但直接觀察次數分配，我們可以發現族群和地理分布其實並未有太離譜的偏差。事實上，研究結果顯示使用電訪樣本資料庫所做出的調查樣本分布，和母體的比例其實差異並不算大。使用加權方式來「矯正」樣本結構，應仍爲可行。與 13 年前所執行的 2000 社會意向原住民專題調查結果相比，雖然本調查案使用的樣本來源時間橫跨 6 年，但其結果卻未出現明顯偏差。甚至因爲樣本規模較大，而能更貼近母體分布。因此，電訪樣本資料庫可以被視爲是一項珍貴的研究資源，有助於研究者鎖定特定群體，依其特徵，將這些樣本自龐大資料庫當中挖掘出來，而予以利用。對於解決特定群體機率抽樣調查的難題，有相當大的幫助。

誠如傅仰止、簡文吟與伊慶春（2001, 117-118）在結論中所提建議「調查研究機構與民意調查中心在執行大型機率調查研究之餘，若能長期累積類似的資料庫，再注意時效的問題，將對其他特定的抽樣調查有莫大助益」。但是，透過本文的討論，我們雖依然贊同傅仰止、簡文吟與伊慶春（2001）對於調查研究機構累積資料庫的建議，但是擴大資料庫內資料來源異質性和樣本規模，應該會是比時效問題更重要的問題。本文的研究結果發現，如果特定群體的母體規模變化較小，以橫跨較多年度（如本研究是 6 年）和橫跨較少年度（社會意向原住民專題調查是 1 年，但跨兩個日曆年，使用兩種不同清冊）相比較，並無明顯差異。如果橫跨較多年度能夠取得更多有效樣本，調查結果的分布會更貼近母體的分布。

*　　*　　*

投稿日期：2013.12.25；修改日期：2014.04.28；接受日期：2014.06.30

附錄、2013 原住民政治態度與行為調查樣本代表性檢定

附表 1　訪問成功樣本之代表性檢定：平原 / 山原（加權前）

身分	樣本		母體	檢定結果
	人數	百分比	百分比	
平地原住民	453	41.1%	47.0%	df=1
山地原住民	649	58.9%	53.0%	卡方值 =23.1335
合計	1,102	100.0%	100.0%	$p<0.05$ 樣本與母體不一致

附表 2　訪問成功樣本之代表性檢定：性別（加權前）

性別	樣本		母體	檢定結果
	人數	百分比	百分比	
男性	414	37.6%	47.8%	df=1
女性	688	62.4%	52.2%	卡方值 =46.0427
合計	1,102	100.0%	100.0%	$p<0.05$ 樣本與母體不一致

附表 3　訪問成功樣本之代表性檢定：年齡（加權前）

年齡	樣本		母體	檢定結果
	人數	百分比	百分比	
20 至 29 歲	174	16.1%	23.8%	
30 至 39 歲	211	19.5%	23.0%	df=4
40 至 49 歲	240	22.2%	21.0%	卡方值 =67.1630
50 至 59 歲	246	22.8%	17.7%	$p<0.05$
60 歲及以上	210	19.4%	14.5%	樣本與母體不一致
合計	1,081	100.0%	100.0%	

附表 4　訪問成功樣本之代表性檢定：教育程度（加權前）

教育程度	樣本		母體	檢定結果
	人數	百分比	百分比	
不識字	20	1.8%	0.5%	
識字但未入學	16	1.5%	0.2%	
小學肄業	38	3.5%	2.4%	
小學畢業	156	14.2%	18.8%	
國、初中肄業	25	2.3%	3.5%	
國、初中畢業	133	12.1%	18.8%	df=12
高中、職肄業	39	3.5%	8.1%	卡方值 =388.8297
高中、職畢業	349	31.7%	26.2%	$p<0.05$
專科肄業	9	0.8%	2.5%	樣本與母體不一致
專科畢業	98	8.9%	5.7%	
大學肄業（含在學中）	29	2.6%	5.4%	
大學畢業	164	14.9%	6.8%	
研究所	24	2.2%	1.1%	
合計	1,100	100.0%	100.0%	

附表 5　訪問成功樣本之代表性檢定：地區（加權前）

地區	樣本		母體	檢定結果
	人數	百分比	百分比	
臺北市	20	1.8%	2.8%	
新北市	56	5.1%	9.4%	
臺中市	52	4.8%	5.1%	
臺南市	8	0.7%	1.1%	
高雄市	65	5.9%	5.7%	
基隆市	15	1.4%	1.7%	
新竹市	6	0.5%	0.6%	
嘉義市	2	0.2%	0.2%	
宜蘭縣	32	2.9%	2.8%	df=19
桃園縣	82	7.5%	11.1%	卡方值 =87.8678
新竹縣	55	5.0%	3.6%	$p<0.05$
苗栗縣	25	2.3%	2.0%	樣本與母體不一致
彰化縣	12	1.1%	0.9%	
南投縣	86	7.9%	5.3%	
雲林縣	1	0.1%	0.3%	
嘉義縣	26	2.4%	1.1%	
屏東縣	158	14.5%	11.6%	
臺東縣	181	16.6%	16.3%	
花蓮縣	209	19.1%	18.2%	
外島地區	2	0.2%	0.2%	
合計	1,093	100.0%	100.0%	

附表 6　訪問成功樣本之代表性檢定：族群別（加權前）

族群別	樣本		母體	檢定結果
	人數	百分比	百分比	
阿美	355	32.5%	38.3%	df=13 卡方值 =59.117 p<0.05 樣本與母體不一致
泰雅	173	15.9%	16.4%	
排灣	210	19.2%	18.4%	
布農	138	12.6%	10.7%	
卑南	41	3.8%	2.5%	
魯凱	42	3.8%	2.5%	
鄒	31	2.8%	1.4%	
賽夏	9	0.8%	1.2%	
達悟	3	0.3%	0.8%	
邵族	2	0.2%	0.1%	
噶瑪蘭	5	0.5%	0.3%	
太魯閣	53	4.9%	5.6%	
撒奇萊雅	3	0.3%	0.1%	
賽德克	26	2.4%	1.7%	
合計	1,091	100.0%	100.0%	

參考文獻

I. 中文部分

包正豪，2009，〈原住民籍立法委員的代表取向與問政行爲：1993-2008 之法律提案內容分析〉，《選舉研究》，16(1): 95-129。

(Pao, Cheng-hao. 2009. "Representative Orientations and Representative Behaviors of Taiwan's Indigenous Legislators: A Content Analysis of the Bills Proposed between 1993-2008." *Journal of Electoral Studies* 16(1): 95-129.)

包正豪、謝顥音，2013，〈國民黨的原住民選舉動員：一個社會文化互動途徑的初探分析〉，《台灣原住民族研究季刊》，6(2): 83-137。

(Pao, Cheng-hao, and Hao-yin Hsieh. 2013. "Indigenous Electoral Mobilization of the Kuomintang: A Preliminary Analysis of the Socio-Cultural Approach." *Taiwan Journal of Indigenous Studies* 6(2): 83-137.)

內政部統計處，2013，〈土地面積、村里鄰、戶數暨現住人口〉，內政統計查詢網：http://sowf.moi.gov.tw/stat/month/m1-01.xls，檢索日期：2013 年 10 月 10 日。

(Department of Statistics, Ministry of the Interior. 2013. "Number of Villages, Neighborhoods, Households and Resident Population." Ministry of the Interior, Land and Population Statistics Database. http://sowf.moi.gov.tw/stat/month/m1-01.xls [accessed October 10, 2013].)

行政院主計總處，1999，〈88 年家庭收支調查報告〉，臺灣地區家庭收支調查網站：http://win.dgbas.gov.tw/fies/a11.asp?year=88，檢索日期：2013 年 10 月 10 日。

(Directorate-General of Budget, Accounting and Statistics of Executive Yuan. 1999. "Report on the Survey of Family Income and Expenditure, 1999." Survey of Family Income and Expenditure Database. http://win.dgbas.gov.tw/fies/a11.asp?year=88 [accessed October 10, 2013].)

------，2012，〈101 年家庭收支調查報告〉，臺灣地區家庭收支調查網站：http://win.dgbas.gov.tw/fies/a11.asp?year=101，檢索日期：2013 年 10 月 10 日。

(------. 2012. "Report on the Survey of Family Income and Expenditure, 2012." Survey of Family Income and Expenditure Database. http://win.dgbas.gov.tw/fies/a11.asp?year=101 [accessed October 10, 2013].)

洪永泰，2013，〈抽樣〉，載於《民意調查新論》，陳陸輝主編，台北：五南。

(Hung, Yung-tai. 2013. "Sampling." In *Public Opinion Survey*, ed. Lu-huei Chen. Taipei:

Wunan.)

原住民族委員會，1998，《中華民國八十七年台灣原住民生活狀況調查報告》，台北：原住民族委員會。

(Council of Indigenous Peoples. 1998. *Living Status Survey of Indigenous People, 2009*. Taipei: Council of Indigenous Peoples.)

------，2013，〈原住民人口數統計資料〉，原住民族委員會網站：http://www.apc.gov.tw/portal/docList.html?CID=940F9579765AC6A0，檢索日期：2013 年 10 月 10 日。

(------. 2013. "Indigenous Peoples Population Statistics." Council of Indigenous Peoples Website. http://www.apc.gov.tw/portal/docList.html?CID=940F9579765AC6A0 [accessed October 10, 2013].)

高德義，1996，〈原住民與選舉政治：原住民精英選舉行爲調查分析〉，《山海文化》，11: 34-50。

(Kao, Teh-i. 1996. "Indigenous and Elections Politics: A Survey Analysis for Indigenous Elite Election Behavior." *Taiwan Indigenous Voice Bimonthly* 11: 34-50.)

張茂桂、顏愛靜、傅君，1998，《台灣原住民對保留地政策之態度：一般意見調查及原住民菁英調查》，台北：原住民族委員會。

(Chang, Mau-kuei, Ai-ching Yen, and Chun Fu. 1998. "Attitude of Taiwan Indigenous People toward Lands Reserved Policy: General Comments and Indigenous Peoples Elites Survey." Taipei: Council of Indigenous Peoples.)

傅仰止，1985，〈都市山胞研究的回顧與前瞻〉，《思與言》，23(2): 177-193。

(Fu, Yang-chih. 1985. "Review and Prospective of Urban Aboriginal Studies." *Thought and Words* 23(2): 177-193.)

傅仰止、簡文吟、伊慶春，2001，〈電訪樣本資料庫的再利用：以社會意向原住民專題調查爲例〉，《調查研究：方法與應用》，9: 109-118。

(Fu, Yang-chih, Wen-yin Chien, and Chin-chun Yi. 2001. "The Reuse of Telephone Interview Sample Database: The Case of Social Image Indigenous Special Survey." *Survey Research - Method and Application* 9: 109-118.)

劉佩怡，2011，《2011 原住民人權指標調查報告》，台北：中華人權協會。

(Liu, Pei-yi. 2011. *Report on the Survey of Indigenous Peoples Human Rights Indicators*. Taipei: Chinese Association for Human Rights.)

II. 外文部分

Chu, Nain-feng. 1998. "Two-Phase Sampling Design in Survey and Case-Control Study." *Chinese Journal of Occupational Medicine* 5(4): 219-228.

Murry, John W. Jr., and James O. Hammons. 1995. "Delphi: A Versatile Methodology for Conducting Qualitative Research." *The Review of Higher Education* 18(4): 423-436.

Using Telephone Survey Sample Data as a Feasible Resolution to the Difficulty of Indigenous Survey

Cheng-hao Pao* · Ying-lung Chou**

Abstract

There is a lack of indigenous object research inside the nation, and the research is usually attached to national sample as one of the small sample analysis. The main reason is that with only about 2% of the country's total population, indigenous population is rather few in Taiwan. Besides, uneven geographic distribution and the factor that a considerable number of workplace/residence and domicile difference due to livelihood are other reasons which cause difficulty to traditional method of investigation. Both face-to-face interview or telephone interview have handicap on sampling. However, public opinion polls had become more universal in recent years. Academic and private investigation agency all has accumulated a large number of successful phone survey samples, which contained numerous Han samples and certain proportion of indigenous samples. If there is a possibility to get these samples which were "not originally using indigenous population as statistic population, but contain successful indigenous samples," and apply data mining technology to dig the "valuable mineral" out from a large amount of "useless" information, then it will be probable to directly filter out the indigenous samples that were needed, and solve the problems mentioned above in the context of saving cost.

In this view, the research team and Taiwan Real Survey Co. had cooperated and merged the successful indigenous population phone interview samples, which are from national samplings that were executed by the company since 2007 so far (March 2013), as indigenous telephone database.

Preliminary research result shows, by using telephone interview sample database that was consolidated by data mining and double sampling, the characteristics of successful samples which got from the telephone interview are similar to the indigenous statistic population in the official population statistics. Other than that, including multi-year indigenous population phone data could enhance the collection range of the indigenous telephone database, and reduces the cost of sample acquisition.

Keywords: poll, double sampling, telephone interview, indigenous people

* Associate Professor, Department of Global Political Economy at Tamkang University.

** Assistant Professor, Department of Global Political Economy at Tamkang University (corresponding author).