

# 不同抽樣底冊之選民母體與投票母體的輪廓 分析：以2016年總統選舉民調為例\*

莊文忠\*\*、林美榕\*\*\*、洪永泰\*\*\*\*

## 摘要

有關選舉民意調查的數據真實性或樣本代表性所引發之爭論核心在於：選舉民調的「母體」(population)定義為何？由於選舉的投票率不是百分之百，導致「選民母體」和「投票母體」不完全一致，因此利用不同抽樣底冊進行選舉民調，其樣本結構與選民母體、投票母體的一致性，便成為調查方法上值得探討的議題。本文的研究目的有三：(1)透過回顧我國選舉民調所採用之不同抽樣底冊的演進過程，探討其中所隱含的母體涵蓋率缺失；(2)釐清選舉研究中的選民母體與投票母體的內涵，並以2016年總統選舉為例進行比較分析；(3)分析TEDS針對2016年總統選舉利用不同抽樣底冊進行民調，成功樣本與選舉母體、投票母體之間的落差及加權處理的修正效果，進而探討其在理論上或實務上的意涵。

\* 本論年初稿發表於「臺灣民主參與的理論建構與實踐暨選舉研究中心三十週年學術研討會」，臺北市：國立政治大學選舉研究中心主辦（2019/5/24-5/25）。本文作者感謝「臺灣選舉與民主化調查」(TEDS)及中央選舉委員會公開釋出相關調查資料，以及研討會評論人、匿名審查委員提供寶貴的修正意見，惟一切文責由作者自負。

\*\* 世新大學行政管理學系教授。E-mail: jwj@mail.shu.edu.tw。

\*\*\* 淡江大學國際企業學系助理教授（通訊作者）。E-mail: 134660@mail.tku.edu.tw。

\*\*\*\* 臺灣大學政治學系退休教授。

本研究的分析結果發現2016年總統選舉的選民母體與投票母體的差異不大。不過，利用不同抽樣底冊進行民意調查的樣本結構與這兩種母體並不完全一致，且無論是利用何種母體結構進行事後加權調整，對投票率、得票率和催票率的修正效果有限。因此，要解決民調結果與投票結果之間的落差，有關母體的定義或許並非關鍵，對抽樣底冊選擇、問卷題目設計、調查技術精進、訪問失敗減少等環節的檢討，可能才是治本之道。

**關鍵字：**選民母體、投票母體、抽樣底冊、戶籍抽樣、地址抽樣

## 壹、前言

民意調查在臺灣早已不是新鮮事，尤其每逢選舉期間，相關的民意調查之多，常令人目不暇給。此一現象在全國最高層級的總統選舉尤為明顯，除政黨和候選人外，媒體及學術機構也會公布民調結果，而各家民調數據的差異或變化，不僅是政治人物與政論節目名嘴的討論焦點，也常是閱聽大眾茶餘飯後的話題。就過去經驗來看，民意調查在臺灣受到高度關注的原因之一在於，民進黨和國民黨這兩個主要政黨先後在黨內初選階段引進民調機制，無論是總統、縣市長或立法委員的候選人，率以民調結果作為提名或協調人選之依據；而投入選戰的候選人也經常以民調作為選戰策略擬訂與評估的重要資訊，甚至企圖利用民調數據影響選民的投票意向，也因此民調結果與選舉結果的一致性、民調數據的可信度往往成為選後檢討與評論的議題。

在過去二十多年來，選舉民調主要是以住宅電話作為「抽樣底冊」(sampling frame)，然而近年來隨著手機電話使用人口日增，住宅電話的「母體」(population)涵蓋率遭遇到重大挑戰。以2020年總統大選為例，投入民進黨黨內初選的現任總統蔡英文即主張初選民調應納入手機電話，才能讓全臺500萬「唯手機族」參與表達意見，是民主與進步的作法；而參加國民黨黨內初選的鴻海集團董事長郭台銘也持相同看法，建議應在初選全民調中納入手機調查，讓更多年輕人可以參與（蘋果日報電子報，2019年5月17日），<sup>1</sup>一時之間在政治上引爆傳統電話民調的樣本代表性爭論。

基本上，不論有關選舉民調的數據可信度或樣本代表性所引發之爭論，其問題核心即在於：選舉民調的「母體」定義為何？而民調結果能否真實反映選民的投票意向、傳統的住宅電話調查是否應納入手機電話調查等課題，則可視為是此一核心問題的延伸。近年來，更有資訊或科技業者嘗試另闢蹊徑，在選舉期間利用網路大數據分析選情走向與預測選舉結果，意欲以此一作法取代傳統的民調方法，然其背後的母體定義則似乎更為撲朔迷離。要言之，對所有抽

<sup>1</sup> 莊文忠，2019，〈「民意調查是「科學」還是「政治」？〉，蘋果日報電子報，<https://tw.appledaily.com/headline/20190517/JAQD7W5EAVUBFP3AFJLWSOQTKY/>，檢索日期：2019年5月20日。

樣調查而言，首要之務即是定義「母體」，以確認研究對象的時空範圍和資格條件，其後才能深入探究利用各種不同抽樣底冊或資料蒐集方法所獲得之調查數據，何者最能反映此一母體的真實意向或行爲。

不過，誠如洪永泰（2005, 11-12）所言，絕大多數的抽樣調查主題，因沒有人知道真正的母體參數爲何，利用樣本資料對母體數值進行推估，其誤差有多大，實屬「死無對證」的情形；而選民投票行爲乃是少數可以驗證的調查主題，不僅可以得知每項選舉的投票率，也可以看到每一政黨或候選人的得票率，選舉民調因爲有此可供驗證的母體參數數據，抽樣調查結果便必須面對嚴格的檢驗。不過，在調查實務上，利用選舉民調數據推估或預測選舉結果仍存在不小的誤差，故在方法論上實有其值得探究之處。

以2016年總統選舉爲例，「臺灣選舉與民主化調查」（Taiwan's Election and Democratization Study, 以下簡稱TEDS）在選前與選後分別以不同的抽樣底冊和調查方法進行調查（如表一所整理），在原始成本樣本（即尚未加權調整）的投票率方面，無論是選前的電話調查或是選後的面訪調查，對投票率的估計呈現不同幅度的差異：選前執行1個多月的「2016TEDS選前電訪」，係以住宅電話簿爲抽樣底冊，其調查結果與實際投票結果的差距爲4.8%；在選後6個月左右執行的「2016TEDS選後面訪」，係以戶籍資料爲抽樣底冊，投票率高估近二成；不過間隔1年半左右後進行的「2017TEDS雙軌面訪」，同時以戶籍資料和地址資料作爲抽樣底冊來源，投票率估計誤差則小於2016年的選後面訪調查，約高估6%上下。

在原始成功樣本的候選人得票率方面，選前電訪中，各候選人的估計差距約在2%至6%之間；選後6個月的面訪中，各候選人的估計亦有3%至7%的差距；但選後1年半的雙軌面訪中，各候選人的估計則是戶籍抽樣(register-based sampling)的差距約爲2%至9%，略大於選後6個月的面訪，地址抽樣(address-based sampling)的差距與選後6個月的面訪差不多，約爲2%至7%。至於在候選人的催票率方面，<sup>2</sup> 選前電訪中，各候選人的估計差距約在2%至5%之間；選

<sup>2</sup> 由於政黨或候選人的得票率計算方式爲（得票數 / 有效票數）×100%，在跨年度或跨不同性質選舉的比較中，較易受到當次選舉實際投票人數變動的影響；而政黨或候選人的催票率計算方式爲（投票人數 / 選民人數）×100%，由於選民人數的變動較小，較

後6個月的面訪中，各候選人的估計則是0%至10%的差距，但選後1年半的雙軌面訪中，各候選人的估計誤差則是戶籍抽樣略大於選後6個月的面訪，地址抽樣略小於選後6個月的面訪。

表一 2016總統選舉結果與民調結果比較：投票率與候選人得票率（未加權）

	實際 選舉 結果	2016TEDS 選前電訪 <sup>3</sup>		2016TEDS 選後面訪		2017TEDS雙軌面訪				
		住宅電話		戶籍抽樣		戶籍抽樣		地址抽樣		
		%	%	差距	%	差距	%	差距	%	差距
投票率	<b>66.3</b>	71.1	4.8	85.7	19.4	72.9	6.6	71.6	5.3	
得票率	蔡英文	<b>56.1</b>	59.3	3.2	62.6	6.5	65.4	9.3	62.8	6.7
	朱立倫	<b>31.0</b>	25.5	-5.5	27.5	-3.5	23.9	-7.1	26.3	-4.7
	宋楚瑜	<b>12.8</b>	15.2	2.4	9.8	-3.0	10.8	-2.0	10.9	-1.9
催票率	蔡英文	<b>36.7</b>	41.1	4.4	46.7	10.0	46.5	9.8	44.2	7.5
	朱立倫	<b>20.3</b>	17.7	-2.6	20.5	0.2	17.0	-3.3	18.5	-1.8
	宋楚瑜	<b>8.4</b>	10.5	2.1	7.3	-1.1	7.6	-0.8	7.7	-0.7
未投票 / 未表態		30.7		25.5		28.9		29.6		
樣本數		5,841		1,690		1,190		1,162		

註：實際選舉結果的投票率計算係以投票人數除以選舉人數；各候選人的得票率計算是以有效票數為計算分母；各候選人的催票率計算是以選舉人數為計算分母。民意調查結果的投票率計算係以所有的樣本數為計算分母；各候選人的得票率計算是以有明確回答支持對象（即扣除未投票或未表態）的樣本數為計算分母；各候選人的催票率計算是以所有樣本數為計算分母。

資料來源：黃紀（2016）。

適合用於跨年度或跨不同性質選舉的比較。

<sup>3</sup> 本項調查以2016年第14屆總統副總統及第9屆立法委員選舉為研究主軸，自2015年12月2日起，至2016年1月26日止，由政治大學選舉研究中心執行電話訪問。TEDS2016-T的問卷設計由TEDS召集人黃紀教授召集問卷草擬小組於2015年11月召開討論會議，針對研究主題與問卷內容進行討論，提交第135次TEDS委員會會議確認後定稿。選前獨立訪問自2015年12月2日開始至2016年1月14日止，共41天。以每7天一波（第一波為6天），共計完成六波獨立樣本，各有980、984、995、946、980、956個成功樣本，共計完成5,841個成功樣本。

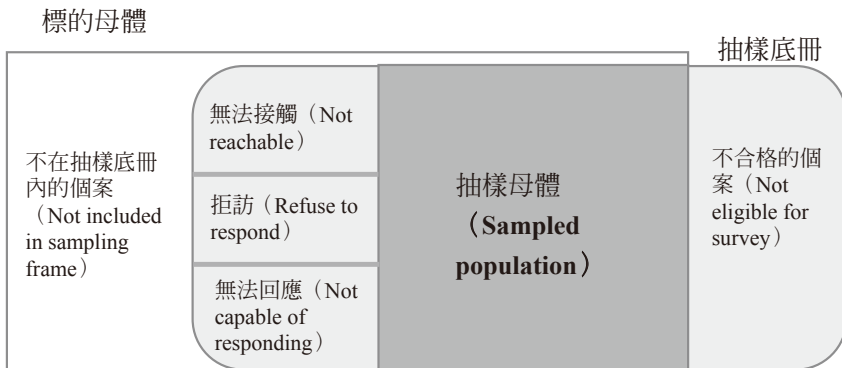
此一結果與過去研究的發現一致，即在選舉民調中最常見的是，受訪者回答有去投票的比例偏高及投票對象與實際選舉結果不盡相符的情形，選前民調普遍出現受訪者對所支持對象不表態的比例偏高，選後則是出現受訪者回答投給當選者的比例偏高。從研究方法的觀點，選舉民調對母體參數的估計偏差來源，雖然可歸咎於調查方法的選擇，或是調查工作的各個環節缺失所造成的，如問卷設計、訪員執行、資料處理等因素，此外也不乏有選民的心理認知與回答效應等，而這些誤差來源必須透過精進調查技術和訪問技巧等方式降低。但是另一個值得深入探究的問題是：不同的抽樣底冊是否也是主要的誤差來源？蓋抽樣調查的母體涵蓋率缺失極有可能是一個重要因素，此涉及抽樣調查所使用的底冊是否完整涵蓋符合資格的選民母體有關。

換言之，影響「標的母體」(target population)與「抽樣母體」(sampling population)是否一致的重要因素，應包括抽樣底冊的母體涵蓋率和調查執行的品質，雖然目前國內已有不少有關調查執行品質的研究（林宇旋等 2011；洪永泰 2003；陳光輝與劉從葦 2006；黃東益、施佳良與謝忠安 2011；鄭夙芬 2000；2006；蕭怡靖 2006；謝淑惠與李名鏞 2020），不過，探討不同母體定義和抽樣底冊的樣本代表性及推論效果的研究則相對較少（洪永泰 1996；2005）。職是之故，本研究的主要討論焦點放在標的母體與抽樣母體的一致性分析上，並以TEDS在2016年總統選舉前後利用不同抽樣底冊進行選舉民調的資料進行實證分析。

要言之，在相似的母體定義與隨機抽樣設計原則下，使用不同抽樣底冊進行調查的結果，是否描繪出不同的母體輪廓？利用「選民母體」(eligible population，如內政部的戶籍資料)和「投票母體」(voter population，如中央選舉委員會的「選舉人投票資料檔」)作為加權依據，何者的效果比較好？會不會更接近實際選舉結果？此乃是本研究所感興趣的問題。因此，本研究從方法論的觀點切入，主要目的有三：(1)透過回顧選舉調查中採用不同抽樣底冊的演進過程，探討其中所隱含的母體涵蓋率缺失；(2)釐清選舉研究中的選民母體與投票母體的內涵，並以2016年總統選舉為例進行比較；(3)分析TEDS利用不同抽樣底冊進行調查，成功樣本與選舉母體、投票母體之間的落差及加權處理的修正效果，進而探討其在理論上或實務上的意涵。

## 貳、選舉民意調查的「母體」定義

從統計的觀點，母體即是「所有研究對象的集合」，但由於社會科學的主要研究對象是人，因此，此一母體的概念經常會因為研究者使用不同抽樣底冊或調查方法等因素，而提出不同的操作性定義。Lohr(2010)將所有研究對象的集合定義為「標的母體」，透過「抽樣底冊」所建構的「抽樣母體」，即是排除了底冊中不是合格研究對象、無法接觸、無法接受訪問、及對調查題目不表示意見等個案，也就是調查中的成功樣本背後所代表的母體大小。這三者之間的關係如圖一所示。



圖一 標的母體、抽樣底冊與抽樣母體的示意關係

註：標的母體為長方形，抽樣母體為弧狀長方形，抽樣母體為深灰色正方形。

資料來源：Lohr (2010, 4)。

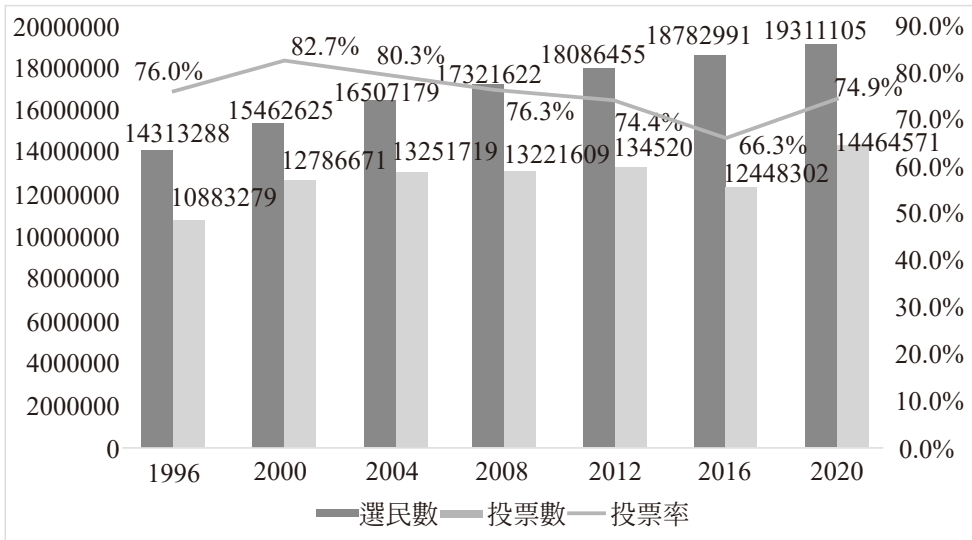
若就選舉民調而言，標的母體係指該次選舉中有投票權的全體選民，本研究將其定義為「選民母體」；不過，若是分析各個政黨或候選人的得票率，標的母體應是該次選舉中有去投票的選民，本研究將其定義為「投票母體」。當該次選舉的投票率為百分之百時，選民母體即等於投票母體，而隨著投票率越低，選民母體雖仍維持不變，但投票母體則會變得越小，造成這兩種母體在人口結構上出現明顯差異的可能性越高。至於「抽樣母體」則是指利用各種母體底冊（如戶籍資料、選舉人名冊、門牌地址、住宅電話號碼、手機電話號碼

等)所能訪問到的所有選民,實務上不論是利用何種抽樣底冊或調查方法進行調查,都可能因為底冊的不完整或訪問失敗等因素,出現母體涵蓋率的問題。因此,無論是面訪或電訪,調查成功樣本的人口結構是否等於選民母體的人口結構?而選民母體的人口結構又是否等於投票母體的人口結構?各種樣本代表性的檢定成為研究者進行統計分析之前所應該確認的工作(洪永泰 2003;黃紀與張佑宗 2003)。

根據上述定義可知,選民母體可透過我國中央選舉委員會(以下簡稱「中選會」)在選舉前所公告的資訊事先得知精確人數,並利用內政部定期公布的人口統計資料推估其人口結構,作為選舉民調的樣本代表性檢定的主要母體參數。投票母體只有在投票日當天結束後才能得知,甚至在投票結束前任何突發事件,例如2004年的「319槍擊案」、2016年的「換柱事件」和「周子瑜拿國旗事件」、2018年的「公投綁選舉亂象」、2020年的「香港反送中事件」等,都有可能導致選民改變投票意向,進而影響投票母體的規模與結構。因此,投票母體在投票前僅能透過民調數據加以推估,而其誤差往往又難以掌握,更遑論及事先得知其人口結構,以致於在選舉民調中無法作為校正樣本結構的參數來源。

選民母體與投票母體之間的關係可以圖二為例說明,我國自1996年首次舉行直接民選的總統選舉以來,選民人數呈現穩定增加的趨勢,投票人數則是有增有減,投票率除了1996年之外,呈現逐屆下降的趨勢,但在2020年又回升到七成五左右。此一結果顯示,歷次同性質選舉的選民母體通常是穩定且緩慢變化,但投票母體卻可能是變動不居的。若再進一步分析歷次選舉中投票選民的性別、年齡、地區等基本人口變數,其結構的不穩定性可能是更為明顯,值得加以探究。





圖二 1996至2020總統選舉的選民人數、投票人數與投票率

資料來源：整理自中央選舉委員會。

在人民參與各類型的政治活動中，選舉的投票參與是最典型且最受到注意的一種，不論代議民主政治的定義與內涵為何，其運作皆無法排除民眾透過選舉來參與政治的事實（莊文忠與洪永泰 2016；陳義彥 2020）。不過儘管每一位公民皆享有憲法所賦予的參政權，選舉參與更是強調「一人一票，票票等值」的平等原則，但實際上在多數民主國家的選舉中，投票行為並非法律所強制規定的公民義務，並非所有公民都會行使此一權利；且因每個人參與投票的意願和動機不同，導致擁有投票權的選民母體和實際行使投票權的投票母體之間可能會出現不一致的情形，隨著投票率越低，這兩個母體之間的選民結構就越有可能出現明顯差異，因此選民母體和投票母體的輪廓便成為眾所關注的焦點。

舉例來說，根據美國的調查與推估指出(Griffin, Teixeira, and Frey 2018, 6-7)，美國的選民結構自2020年至2036年可能出現幾個變化，並影響未來五次總統選舉，牽動民主黨和共和黨的政治板塊移動，對美國的民主政治造成衝擊。首先，符合資格的選民人口會變得更加種族多樣化，年輕、新一代的美國

人比前幾代人更具種族多樣性，2016年白人佔選民總數的69%，預計到2020年將降至67%，到2036年再降至59%；在此期間，西班牙裔人口預計將從2016年的12%增加到2036年的18%；亞洲人和其他種族從2016年的7%到增長2032年的10%。黑人選民的比例從2016年到2032年穩定上升不到1%，由12%微增至13%。其次，選民人口將呈現老化現象，65歲及以上的選民將從2016年的21%上升到2036年的27%，反觀18至64歲的選民人口將萎縮。雖然在此期間，老年人的種族多樣化差異將持續低於年輕族群，但從2020年到2036年，白人年長選民佔所有選民的比例僅上升1%，65歲及以上的選民絕大多數來自有色人種增長。第三，選民的教育程度會持續提高，此一變化在白人中尤為重要，因為有大學學歷和沒有大學學歷的人之間的政治和行為的差異往往是最大的，雖然沒有大學學歷的白人佔2016年選民的46%，但預計到了2036年將降至37%。最後，上述的全國選民結構變化在各州則是呈現不同的樣貌，有些州的變化速度較為快速，有些州則較為緩慢。

事實上，臺灣有關選民結構與投票意向的研究亦不少。舉例來說，有些研究從性別的角度切入，探討不同性別之選民在選舉時的支持對象是否存在差異，尤其是女性選民的選舉參與情形最受關注（翁秀琪與孫秀蕙 1995；黃秀端與趙湘瓊 1996；楊婉瑩 2007；楊婉瑩與林珮婷 2010；2013；楊婉瑩與劉嘉薇 2006）；有些研究是從世代的觀點，分析不同年齡層選民的政治態度對其投票意向的影響，尤其是年輕選民的政治立場與選舉參與之間的關聯性（林珮婷、鄭夙芬與王德育 2020；陳陸輝 2006；陳義彥與蔡孟熹 1997；游清鑫與蕭怡靖 2007）；有些研究則是從族群認同的視角，剖析不同省籍背景的投票意向，早期的研究較為關注本省籍（閩南人）和外省籍（大陸各省市）選民的政黨認同與投票偏好，客家族群和原住民族的選舉參與則可說是近年來頗為熱門的議題（包正豪與周應龍 2019；吳重禮與李世宏 2005；吳重禮與崔曉倩 2010；孫榮光 2020；鄭夙芬 2009）。

由此可知，選民母體的結構及變化深深地牽動政黨政治的運作和選戰策略的訂定，以致於選舉民調的樣本結構與選民母體吻合程度成為關注的焦點，實務上以選民母體作為事後加權調整的依據，自屬合理的作法。不過，從選舉研究的觀點，樣本結構與投票母體之間的一致程度，或許更為攸關選舉結果的分

析與評估，也是一個令人好奇的問題。究其原因在於，擁有選舉權的公民不一定都會去投票，因此，選民母體的結構主要是反映有投票權之公民的人口特徵輪廓，實際上有去投票之選民所構成的投票母體，才是真實反映參與選舉之公民的人口特徵輪廓。當投票率越低時，這兩種母體的輪廓不完全吻合可能性越大，甚至是造成選舉預測與實際選舉投票結果出現明顯落差的可能原因。

然而受限於民主國家選舉採取秘密投票的原則，過去大多數的研究僅能從總體層次看到選民母體的人口統計和投票率的變化，至於個人層次的投票母體，在中選會未進行選舉人投票統計相關研究之前，<sup>4</sup> 其人口結構僅能透過選前或選後的民調結果進行推估，推論誤差大小又會受到受訪者的表態意願所影響，因而成爲一個未知的謎。因此，本研究嘗試運用中選會針對2016年總統選舉委外執行「選舉人性別投票統計」研究案所建置之資料檔及TEDS針對2016年總統選舉調查所得之資料，剖析透過不同抽樣底冊所執行之選舉民調，其成功樣本結構與選民母體、投票母體之間的差異情形。

## 參、選舉民意調查的「抽樣底冊」來源

選舉民調與一般民調在母體定義方面不同之處在於，根據《公職人員選舉罷免法》第17條規定：「選舉人，除另有規定外，應於戶籍地投票所投票。」因此，除全國性的總統大選外，其他性質選舉的民調在訪問開始前必須先確認受訪者的戶籍所在地，才能判斷其是否爲合格的受訪對象，而一般民調除非有必要，否則並不需以戶籍所在地作爲樣本篩選依據，因此造成兩者的母

<sup>4</sup> 中選會爲瞭解兩性投票情形及趨勢，並配合性別主流化政策，自2008年起辦理選舉人性別投票統計作業，針對總統副總統、直轄市、縣（市）首長等選舉，以縣市別、直轄市及鄉鎮市區爲分析單位進行選舉人性別投票統計。中選會所訂定之選舉人性別投票統計資料蒐集的標準化作業流程如下：在投票日當天，各投開票所的選務工作人員於投票結束後，開始開票前，點算選舉人名冊中的不同性別選舉的領票紀錄，並完成性別投票統計表填報，之後由各鄉鎮市區公所統計填報各投開票所之性別統計表，再由各直轄市、縣（市）的選舉委員會進行資料彙整，最後交付由中選會再次統整後，即公布選舉人性別投票統計結果於中選會網站性別統計專屬網頁及行政院性別平等委員會的網頁（莊文忠與洪永泰 2016, 3），此屬於總體層次之投票母體的有限資訊。

體定義及其抽樣底冊的涵蓋率有所不同。基於研究旨趣，本節主要探討選舉民調中使用不同抽樣底冊的演進過程，作為後文實證研究設計與分析的基礎。

一般而言，抽樣調查的研究設計主要包括以下幾個過程（洪永泰 2005；龍文彬 1998）：首先，研究者應先定義母體，確認研究對象的時空範圍和資格。其次，研究者應蒐集有關調查母體的任何有用資訊，增加對母體的認識，作為抽樣設計規劃之參考，例如調查對象之分布適合分層抽樣(stratified sampling)或集群抽樣(cluster sampling)；或是分層之變數、標準及多寡；不同層之抽出比率等。第三，研究者應尋找適當的抽樣底冊，例如完整且正確的受訪者名單或查訪地址名冊，這是順利取得調查樣本的重要關鍵，也是避免訪問對象重複、遺漏的最佳手段。最後，在調查結果推估階段，研究者需要最新基本人口輪廓(profile)資料，俾據以調整原母體資料時間落後之誤差，再根據樣本所取得之數據對母體進行推論。

依照上述流程，在選舉民調中，母體的定義理應十分明確，即指在特定選舉中，擁有投票權的公民。舉例來說，TEDS自2001年啟動以來，在電訪調查中對母體的定義為：「戶籍設於臺灣地區年滿20歲以上的成年公民，受訪對象僅限於住家民衆或住商混合的民衆。」（黃紀 2016）在抽樣設計方面，若是全國性的選舉，主要是以縣市為次母體，進行分層隨機抽樣；若是縣市層級以下的選舉，則是以該選區為調查範圍，採取隨機抽樣。抽樣底冊早期主要來自於住宅電話簿，後來則是來自政治大選舉研究中心所累積的電訪資料庫，利用特定數字或隨機亂數修正抽取之電話號碼的尾數，據以製作電話樣本；為確保戶內合格樣本的中選機會，在開始訪問之前，訪員須按照「戶中抽樣」(within-household sampling)的原則，抽出應受訪對象再進行訪問。

TEDS在面訪調查中對調查母體的定義是：「以台灣地區設有戶籍，年齡在20歲以上具有選舉權的公民為調查訪問之母體。不過，並不包括軍事單位、醫院、療養院、學校、職訓中心、宿舍、看守所及監獄之居民。」（黃紀 2016）在抽樣設計方面，以開放直接民選後的歷次總統選舉民調為例，主要是採取「分層多階段PPS抽樣」原則，進行三階段抽樣，歷次總統選舉選後面訪調查的抽樣架構與設計如「附錄一」所整理。至於抽樣底冊的來源方面，歷年來隨著個人資料保護意識的升高及政府部門對公務資料使用範圍的定義改

變，而採用不同的抽樣底冊（如「附錄二」所整理）。是以，在標的母體的範圍定義明確下，整個選舉研究的過程中，影響抽樣設計與調查執行的最大變數莫過於抽樣底冊的選擇，本節以下進一步探討不同抽樣底冊的特性。

## 一、住宅電話碼簿

理論上，電話訪問的標的母體涵蓋所有可以用電話接觸到的人，至於未使用電話者則不在抽樣母體之中，而電話調查的抽樣方法主要可以分為兩大類（洪永泰與黃永政 2001, 174）：一是以住宅電話簿為主所衍生的各種方式 (directory and list based)，如等距隨機抽樣、尾數加1或減1、或是尾數隨機等方法；二是以隨機撥號產生號碼(random digit dialing, RDD)為主，利用區域碼和局碼搭配後四碼隨機的方法。近年來，隨著行動電話的普及率越來越高，有關以行動電話號碼作為抽樣底冊或採取住宅電話號碼與行動電話號碼組合的「雙底冊調查(dual frame telephone survey)」的理論與實證研究也有不少（洪永泰等 2014；張鏡文、黃東益與洪永泰 2017；許勝懋 2015；陳鴻嘉 2019），此因非本研究的討論焦點，故暫時略過不述。

基本上，無論是以住宅電話簿或RDD方法作為抽樣底冊，均是以「戶」為抽樣單位。不過由於每一家戶內的選民人數不同，再加上各個家庭可能有習慣接聽電話的人，若是在撥通電話後確認第一個接電話者為合格樣本後即進行訪問（一般稱之為「任意成人法」），可能造成同一家戶內的「人」有不等の中選機率。因此，在撥通住家電話後，應進行第二階段之戶中抽樣，依各家戶的合格人數和男女比例，以電話碼碼的尾數分配每一個人的中選機率，以符合隨機抽樣的原則，減少造成樣本偏差。

## 二、選舉人名冊

根據我國現行選舉制度，年滿20歲以上的國民係以戶籍所在地作為選舉投票權資格的認定標準，同時除了投票所工作人員得在戶籍地或工作地投票外，選民也必須在戶籍地投票。雖然不同性質的選舉有不同的設籍時間規定，也排除被褫奪公權和受禁治產宣告者，不過，實際上屬於這兩種情況的人數相對極

少，合格選民的各項統計數據和戶籍人數的差異極小，幾乎不必刻意區別而可以完全等同，在實務上可以說「合格選民就等於是戶籍資料之中年齡滿二十歲的國民」而不必擔心犯太大的錯誤（洪永泰 2005, 17）。因此，在各種選舉投票日之前，戶政機關受中央或地方選委會所託，應自戶籍登記資料中篩選出符合各項積極資格和消極資格的選民，並製作成選舉人名冊後公告閱覽。是以就選舉研究而言，選舉人名冊可說是最為理想的抽樣底冊，與標的母體完全一致，而這也曾經是臺灣選舉面訪的抽樣底冊來源。

### 三、戶籍登記簿

戶籍資料是目前國內可供抽樣設計使用之最完整的人口資料檔案之一，其內容包括全台灣近8,000個村里的重要人口特徵值，可供分析戶籍人口特性之參考，<sup>5</sup> 早期各鄉鎮之戶政事務所，有「戶籍登記簿」可供抄錄，後來隨著電子化政府的推動，已有全國連線的電子資料檔案可為抽樣及調查使用；此外，內政部也會按月公布各級行政區的人口數據，作為最終推估調查之依據（龍文彬 1998）。不過，以此資料作為民調的抽樣底冊時，其缺點是戶籍人口並不等於常住人口，透過戶籍資料進行抽樣調查，無法真實反映常住人口的結構與變遷。例如根據2010年的人口及住宅普查結果發現，臺灣民眾因為升學、就業、依親、移居等因素，超過二成的人口並非長年居住在戶籍登記地，造成了以戶籍調查與推估上的困難，其樣本代表性也不斷被質疑（朱斌好等 2015；曾憲立等 2018；龍文彬 1998）。

### 四、門牌地址資料

隨著社會大眾對保護個人資料與隱私的意識愈來愈高，我國於2012年通過

---

<sup>5</sup> 由於戶籍資料提供了完整之人口特性，所以在抽樣調查時，不但可提供各種抽樣單位之特徵變異，作為分群、分層之依據，同時也能從戶籍地址，直接抽取調查樣本。由於戶籍之動態更新隨時在進行，故隨時均可取得調查時點特定地區之人口數量資料，對於調查結果之推估作業，也非常方便，長期以來廣為各界於辦理抽樣調查時所應用（龍文彬 1998, 90）。

《個人資料保護法》，政府機關為符合法律規定，對於個人資訊的維護與管理更加嚴格，導致非公務機關申請戶籍資料用於面訪調查的難度大為提高（謝淑惠等 2018, 76）。對選舉研究而言，要利用戶籍資料作為抽樣底冊也越來越不可行，因此，TEDS自2017年起，首次啟動「戶籍資料」和「門牌地址」雙軌抽樣模式，一方面是因應未來無法取得戶籍資料檔的情況下，有備選方案可用，另一方面是檢證兩種抽樣底冊的作業流程與調查結果的異同，嘗試建立一套地址抽樣調查的標準作業程序。在2018年的縣市長選舉選後調查，首次全面採取以門牌地址為抽樣底冊的調查流程。

如前文所述，選舉研究和其他社會科學研究最大不同之處在於，無法以常住人口進行調查，在訪問前必須先確認受訪者的戶籍所在地，因此「籍在人不在」或「人在籍不在」，對選舉研究而言都是一大問題。而門牌地址抽樣基本上係以常住人口為接觸對象，再進一步透過現場戶中抽樣程序找到正確的受訪對象，在此一過程中，相較於以戶籍資料作為抽樣底冊，則明顯增加了不少調查誤差的來源，諸如難以事先排除無人設籍的門牌、難以說服保全人員登門訪問、降低接觸對象的配合意願、不易要求訪員嚴格執行抽樣程序等。

綜上所述可知，不同的抽樣底冊會影響到調查流程的設計，以選舉人名冊為抽樣底冊，是以「個人」為抽樣單位，中選樣本即為合格選民，只要能接觸到本人，即可進行訪問，無需進行戶中抽樣；以戶籍資料為底冊，同樣也是以「個人」為抽樣單位，即使接觸到本人，雖可以確定他的戶籍沒有問題，也不需要戶中抽樣，但必須確認是否具有投票權；至於以住宅電話簿或門牌地址為底冊，則是以「家戶」為抽樣單位，且事先無法確認接觸對象是否為合格選民，必須先確認戶內設籍合格人數再進行戶中抽樣。

總之，在選舉民調中，無論使用何種抽樣底冊，透過確認受訪者的戶籍所在地，主要目的即在於確保樣本結構與選民母體的一致性，以減少推論的誤差。然而由於選民母體並不同於投票母體，調查結果是否能夠如實反映選舉結果？則仍有待進一步探究，也是後文的分析與討論的焦點。

## 肆、本文的資料來源

整體而言，選民母體與該選區的整體人口結構差異不大，但投票母體可能是另一回事了。因受限於選舉採秘密投票原則，吾人只能得知投票人數和總體投票率，卻無法得知「誰去投票？」。也就是說，吾人無法取得「投票母體」的人口結構，僅能透過投票日當天的出口民調(exit poll)或選後民調，利用調查樣本推估投票人口的特徵。然而為避免干擾投票作業和選舉秩序，並非每個國家均允許出口民調，而選後民意調查又常因受訪者的無反應(nonresponse)或樂隊花車效應(bandwagon effect)，難以準確推估投票人口的母體結構。因此，為了回答本文所提出的問題，本研究利用以下幾項資料進行實證分析：

### 一、中選會「2016年選舉人性別投票統計改良之研究」

中選會在民國105年度規劃「選舉人性別投票統計改良之研究」委託研究案，該研究案以2016總統暨立法委員選舉為例，採取分層多階段集群抽樣的研究設計，<sup>6</sup>自第14任總統副總統及第9屆立法委員選舉選舉人名冊中，依「等距抽樣」抽出中選選民，建立個體層次的選民投票資料檔（莊文忠與洪永泰 2016）。此一建置完成的「2016年總統暨立法委員選舉選舉人投票資料檔」共有201,257筆選舉人資料，變數項目包含：選舉人性別、出生年月日、不同選舉種類投票、戶籍地所屬直轄市、縣（市）別、鄰等資料，其中具有總統選舉投票資格的選舉人數為201,158筆，<sup>7</sup>實際上有去投票的選舉人數為132,811筆，經加權處理後，具有全國與縣市的母體代表性，有助於瞭解不同

<sup>6</sup> 該研究案以縣市為分層變數，並選擇「選舉人數」作為集群分析的變數，將各縣市的村里依選舉人數的接近度分為二至四個集群，以等比例原則分配各集群的樣本數，第一階段抽樣是每個縣市依據「抽取率和單位大小成比例（Probability Proportional to size, PPS）」原則抽出10至12個中選村里，第二階段抽樣是自中選村里中抽出287至2,459個不等數量的選民（莊文忠與洪永泰 2016）。

<sup>7</sup> 根據《公職人員選舉罷免法》的規定，立法委員選舉的投票資格為在該選舉區居住滿4個月以上，而根據《總統副總統選舉罷免法》的規定，總統選舉的投票資格為在中華民國自由地區居住滿6個月以上，是以，極少數的選舉人僅有立法委員選舉的投票權而不具有總統選舉的投票權。



特徵之選舉人的投票行爲。此一資料雖非普查資料，不過因其調查範圍涵蓋各縣市，且樣本數超過20萬筆，應可充分反映投票母體的性別、年齡及地區等特徵，故本研究將其視爲準母體參數。<sup>8</sup>

## 二、2016TEDS選前電訪

TEDS2016選前電訪調查（TEDS2016-T）自2015年12月2日開始執行，至2016年1月14日結束，利用政治大選舉研究中心所累積的電訪資料庫，再運用隨機抽樣方式抽出電話號碼後，以隨機亂數修正抽中電話號碼的最後四碼製作電話樣本完成6波調查，共計訪問5,841個有效樣本。爲與母體結構更符合，該調查依據內政部出版之「中華民國人口統計年刊（2014年）」爲母體參數，針對樣本之「性別」、「年齡」、「教育程度」及「地理區域」分布，進行樣本代表性檢定，加權方式採用「多變數反覆加權法」（raking）。加權後，經樣本代表性檢定結果，顯示樣本結構與母體並無差異。

## 三、2016TEDS選後面訪

此一調查計畫係於2016年總統與立委選舉結束後6個月左右執行，調查母體爲台灣地區20歲以上有投票權之選民，採分層多階段抽樣設計，共計完成1,690份。爲使訪問成功之獨立樣本符合母體結構，該研究以內政部出版之「中華民國人口統計年刊（2014年）」爲母體參數，針對「性別」、「年齡」（五分類）、「教育程度」（五分類）與「地理區域」（六分類）等資料進行樣本代表性檢定與加權，加權方式採用「多變數反覆加權法」（raking），經樣本檢定與加權後，樣本結構與母體並無差異。

<sup>8</sup> 根據莊文忠與洪永泰（2016）的研究報告指出，此一資料檔係以各縣市次母體，獨立進行抽樣設計，分別抽出5,000筆至12,000筆不等的選舉人資料檔，再依縣市的人口比例加權爲全國選舉人資料檔，各縣市的樣本代表性檢定結果發現，雖有部分縣市的樣本結構與母體結構不一致，但差距均不到1%，主要是因爲樣本數過大造成統計檢定呈現顯著差異，經加權調整後，不論是個別縣市或全國，樣本結構均與母體結構一致；另一具有參考價值的數據是，無論加權前後，利用此一資料檔估計的投票率均爲66.02%，與母體投票率66.27%的差距僅有0.25%，與投票母體非常接近（莊文忠與洪永泰 2016, 71-78）。基於避免文章篇幅過長的考量，本文提供中選會研究報告網址連結（<https://www.ccc.gov.tw/central/cms/ElecAgeSexstatistics/25962>），有興趣的讀者可自行下載參照。

#### 四、2017TEDS雙軌面訪

本計畫為TEDS首次修正過往以戶籍資料作為抽樣底冊的調查方式，突破過去15年完全仰賴內政部戶籍資料庫為抽樣底冊的做法，開發創新之「門牌地址暨設籍選民抽樣設計」；同時基於交叉檢證調查結果的目的，採用戶籍資料訪問與門牌地址抽樣訪問並行的方式，並於2017年暑假進行面訪，最後各完成1,190和1,162個有效樣本。為使訪問成功之獨立樣本符合母體結構，該研究以內政部出版之「中華民國人口統計年刊（2015年）為母體參數，針對「性別」、「年齡」（五分類）、「教育程度」（五分類）與「地理區域」（六分類）等資料進行樣本代表性檢定與加權，加權方式採用「多變數反覆加權法」，經樣本檢定與加權後，樣本結構與母體並無差異。

#### 伍、分析結果與討論

本節以下分為選民母體與投票母體的比較、不同抽樣底冊之成功樣本的人口結構與母體參數的比較、不同抽樣底冊之成功樣本的投票率比較、不同抽樣底冊之成功樣本的得票率與催票率比較等四個部分進行分析與說明。

##### 一、選民母體與投票母體的比較

2016年總統選舉中，選民母體與投票母體之間的差異如表二所示。在性別方面，女性選民的比例略高於男性，參與投票的女性比例亦高於男性，二者之間的差異經卡方檢定在統計上已達顯著水準，但因樣本數高達13萬3千個左右，此不到1%的差異在實務上並無太大意義；<sup>9</sup>在年齡方面，20-29歲和30-39歲的選民雖然分別佔整體選民的17.0%和20.9%，不過，這兩個年齡層的投票

---

<sup>9</sup> Salkind與Frey（2020, 174-175）指出，統計上的顯著性是檢定假設的一項重要參考標準，但並不是唯一的標準，且不能脫離發生的背景被獨立解釋，因之，統計上的顯著差異是否在實務上具有意義，除了假設必須具有強健的理論基礎外，也在於研究者能否對此一結果提出充分且合理的解釋，否則，過於微小的差異在實務上是無足輕重的，可以加以忽略。

率均略低2%，分別為14.7%和18.4%；反觀50-59歲和60歲以上的選民，分別佔整體選民的19.2和23.7%，而這兩個年齡層的投票率則略高出2%至3%，分別為21.2%和27.0%。在地區方面，選民母體與投票母體之間的差異雖達到統計上的顯著，但其差異均不到1%，在實務上同樣無太大意義。

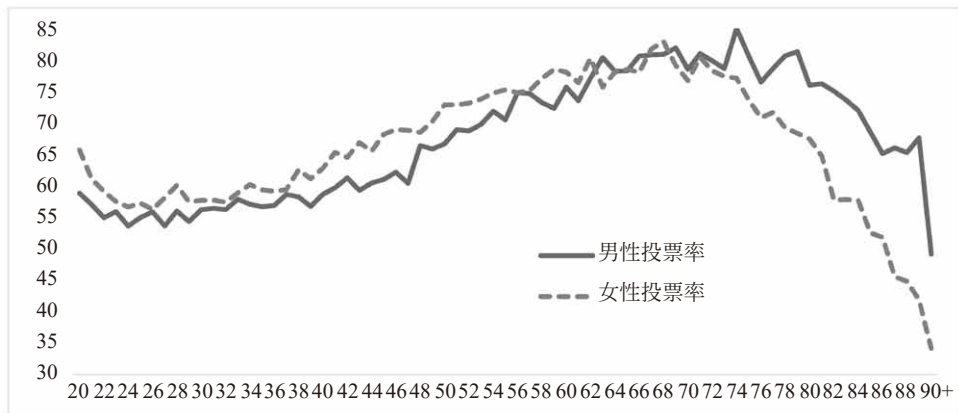
表二 2016年總統選舉的選民母體與投票母體的比例

變數	選項	選民母體	投票母體	卡方檢定
性別	男性	49.30%	48.40%	$\chi^2 = 42.677$ $P < 0.001$
	女性	50.70%	51.60%	
年齡	20-29歲	16.97%	14.69%	$\chi^2 = 1688.194$ $P < 0.001$
	30-39歲	20.88%	18.41%	
	40-49歲	19.25%	18.74%	
	50-59歲	19.19%	21.18%	
	60歲以上	23.71%	26.98%	
教育程度	國小及以下	14.84%	--	
	國、初中	12.84%	--	
	高中、職	28.06%	--	
	專科	12.34%	--	
	大學及以上	31.93%	--	
地理區域	北北基宜	32.25%	32.96%	$\chi^2 = 216.666$ $P < 0.001$
	桃竹苗	15.00%	14.98%	
	中彰投	19.04%	19.00%	
	雲嘉南	14.56%	14.26%	
	高屏澎	16.12%	16.41%	
	花東金馬	3.03%	2.39%	
總數		18,804,459	132,811	

註：選民母體係以2015年12月臺灣地區各縣市20歲以上的人口數進行計算；投票母體係以中選會「2016年選舉人性別投票統計改良之研究」所建置之201,257筆選舉人資料庫中，總統選舉有領票紀錄的132,811筆選民數進行計算，另因選舉人名冊中無選民的教育程度資料，故無法比較。

資料來源：本研究。

由選民母體和投票母體的差異可知，整體來說，女性和年長選民的投票參與程度略高，不同地區的投票參與差異並不大。不過，莊文忠與洪永泰（2016）的研究進一步指出，若將性別和年齡同時納入考慮（如圖四所示），60歲以下之各個年齡層選民中，女性選民之投票率幾乎均高於男性，在70歲以上的選民中，則是男性選民的投票率高於女性，而75歲以上的整體投票率呈現急劇下降情形，女性選民的下滑率又明顯高於男性；此外，各縣市不同的性別和年齡層的投票率亦有所差異，尤其是花東和金馬地區的投票參與較低，性別的差異亦更大。



圖四 2016年總統副總統選舉之性別與年齡投票統計

註：x軸為年齡，y軸為投票率（%）

資料來源：莊文忠與洪永泰（2016, 81）。

此一結果顯示，不同性別和年齡層組合下的選民母體與投票母體之間的差異，會比性別和年齡個別變數之間的差異來得更大，值得深入檢視選民母體和投票母體的結構。因此，本研究進一步利用性別、年齡和地理區域這三個變數，進行選民母體與投票母體的比較分析。表三詳細呈現了性別、年齡層和地區的組合下，選民母體和投票母體的結構，此數據亦可作為「事後分層加權」(post-stratification)或「多變數反覆加權」加權處理的母體參數。表四則是進一步將選民母體比例除以投票母體比例，據以反映兩者結構的差異。若比率大於

1，表示具有該特徵的選民參與投票的比例偏低；若比率小於1，表示具有該特徵的選民參與投票的比例偏高；若比率等於1，表示具有該特徵的選民參與投票的比例未偏高或偏低。由表四的數據可知，選民母體和投票母體在性別和地區的結構差異較小（除花東地區外），比率均接近1（最右邊兩欄）；年齡層的結構差異較大一些，年輕選民的投票參與比例較低，年長選民的投票參與比例較高（倒數最後一列），且59歲以下女性選民參與投票的比例均高於男性（倒數第二列和第三例）。

表三 2016年總統選舉的選民母體與投票母體的組合比例（%）

區域別	性別	20-29 歲	30-39 歲	40-49 歲	50-59 歲	60歲 以上	性別 合計	地區 合計
北北基宜	男	2.67	3.28	2.98	2.98	3.60	15.51	32.25
		2.43	2.93	2.84	3.14	4.22	15.57	32.96
	女	2.52	3.46	3.24	3.34	4.18	16.74	
		2.34	3.20	3.34	3.84	4.67	17.39	
桃竹苗	男	1.40	1.66	1.52	1.38	1.51	7.47	15.00
		1.16	1.41	1.49	1.56	1.81	7.43	14.98
	女	1.30	1.69	1.52	1.38	1.65	7.53	
		1.12	1.50	1.52	1.54	1.87	7.56	
中彰投	男	1.80	2.04	1.79	1.78	2.04	9.44	19.04
		1.49	1.71	1.64	1.89	2.35	9.08	19.00
	女	1.67	2.03	1.83	1.80	2.27	9.60	
		1.52	1.84	1.92	2.09	2.54	9.92	
雲嘉南	男	1.24	1.47	1.37	1.46	1.77	7.33	14.56
		1.08	1.30	1.26	1.50	2.00	7.15	14.26
	女	1.15	1.43	1.32	1.35	1.98	7.24	
		1.03	1.24	1.24	1.52	2.08	7.11	

區域別	性別	20-29 歲	30-39 歲	40-49 歲	50-59 歲	60歲 以上	性別 合計	地區 合計
高屏澎	男	1.39	1.64	1.56	1.54	1.87	8.00	16.12
		1.15	1.46	1.37	1.72	2.31	8.01	16.41
	女	1.29	1.62	1.55	1.57	2.09	8.11	
		1.09	1.46	1.69	1.84	2.32	8.40	
花東金馬	男	0.27	0.29	0.30	0.32	0.36	1.54	3.03
		0.14	0.16	0.21	0.27	0.39	1.17	2.39
	女	0.25	0.27	0.27	0.29	0.39	1.48	
		0.15	0.19	0.22	0.26	0.41	0.15	
性別與年齡 層合計	男	8.78	10.38	9.52	9.45	11.15	49.30	
		7.44	8.97	8.81	10.09	13.09	48.40	
	女	8.18	10.50	9.73	9.74	12.55	50.70	
		7.24	9.43	9.93	11.09	13.89	51.60	
年齡層合計		16.97	20.88	19.25	19.19	23.71	100.00	
		14.69	18.41	18.74	21.18	26.98	100.00	

註：白色底之數字為選民母體比例；灰色底之數字為投票母體比例估計。

資料來源：本研究。

表四 2016年總統選舉的選民母體與投票母體組合比例的比值

區域別	性別	20-29 歲	30-39 歲	40-49 歲	50-59 歲	60歲 以上	性別 合計	地區 合計
北北基宜	男	1.10	1.12	1.05	0.95	0.85	1.00	0.98
	女	1.08	1.08	0.97	0.87	0.90	0.96	
桃竹苗	男	1.21	1.18	1.02	0.88	0.83	1.01	1.00
	女	1.16	1.13	1.00	0.90	0.88	1.00	
中彰投	男	1.21	1.19	1.09	0.94	0.87	1.04	1.00
	女	1.10	1.10	0.95	0.86	0.89	0.97	

區域別	性別	20-29 歲	30-39 歲	40-49 歲	50-59 歲	60歲 以上	性別 合計	地區 合計
雲嘉南	男	1.15	1.13	1.09	0.97	0.89	1.03	1.02
	女	1.12	1.15	1.06	0.89	0.95	1.02	
高屏澎	男	1.21	1.12	1.14	0.90	0.81	1.00	0.98
	女	1.18	1.11	0.92	0.85	0.90	0.97	
花東金馬	男	1.93	1.81	1.43	1.19	0.92	1.32	1.27
	女	1.67	1.42	1.23	1.12	0.95	9.87	
性別與年 齡合計	男	1.18	1.16	1.08	0.94	0.85	1.02	
	女	1.13	1.11	0.98	0.88	0.90	0.98	
年齡合計		1.16	1.13	1.03	0.91	0.88	1.00	

註：灰色底之格子代表比率小於1，表示具有該特徵之投票母體所占比例大於選民母體所占比例，即投票參與的比例偏高；反之亦然。

資料來源：本研究。

## 二、不同抽樣底冊的成功樣本結構與母體參數的比較

由於調查執行過程不可避免地會受到許多因素影響而訪問失敗，可能導致樣本的代表性不佳，若直接以有偏差的成功樣本推論母體，勢必造成偏誤估計的結果，是以樣本代表性檢定之目的在於檢測樣本是否扭曲，進而影響對母體特質的推估（黃紀與張佑宗 2003, 2）。本研究以2016年總統選舉為例，分析以住宅電話、戶籍資料和地址資料作為抽樣底冊進行調查所得到的成功樣本，檢視其性別、年齡、教育及地區等變數和選民母體、投票母體的一致程度。

表五呈現未加權前的樣本結構與選民母體結構的差異。在性別方面，以戶籍資料為底冊的調查，其男性樣本所占比例偏高，甚至達到統計上顯著差異；以住宅電話和地址資料為底冊的調查，樣本結構則是較為接近選民母體。在年齡方面，以住宅電話為底冊的調查，20-29歲的樣本所占比例均偏低，40-49

歲和50-59歲的樣本所占比例偏高；而以戶籍資料和地址資料為底冊的調查，30-39歲的樣本所占比例均偏低，60歲以上的樣本所占比例均偏高；此外，由於以地址資料為抽樣底冊的調查中，20-29歲的年輕人比例也偏低，老年人的比例偏高。此是否與年輕人「籍在人不在」的比例較高有關、或是與訪員在現場進行戶中抽樣時所使用之戶中抽樣架構表的機率分配有關，值得其他研究再深入探討。在教育程度方面，以住宅電話為底冊的調查，國初中以下的樣本所占比例偏低，專科、大學及以上的樣本所占比例偏高；而在2016年和2017年的兩種底冊調查中，呈現較為不一致的結果，尚難有定論。至於地區方面，僅有以地址資料為底冊的調查，北北基宜的樣本所占比例偏低，雲嘉南的樣本所占比例略高一些，此可能與臺北市與新北市的社區型住戶較難進入訪問，或較多「籍不在人在」的外來人口較多，導致訪問失敗率較高有關；以住宅電話為底冊的調查，雖然卡方檢定達到統計上顯著水準，此可能與樣本數較大有關係，但各地區的差距均在3%以內，在95%的信心水準下，此一差異並不算大，除非研究者可以在理論與實務上提出充分且合理的解釋(Salkind and Frey 2020, 175)。

未加權前的樣本結構與投票母體的比較如表六所示，在性別方面，以住宅電話和戶籍資料為底冊的調查，其男性樣本所占比例偏高，而以地址資料為底冊的調查，樣本結構則是最為接近選民母體。在年齡方面，以地址資料為底冊的調查，20-29歲的樣本所占比例偏低甚多，60歲以上的樣本所占比例則明顯偏高；以住宅電話為底冊的調查，20-29歲和60歲以上的樣本所占比例偏低，40-49歲的樣本所占比例偏高；以戶籍資料為底冊的調查，60歲以上的樣本所占比例偏高。在地區方面，僅有以地址資料為底冊的調查，北北基宜的樣本所占比例偏低；以住宅電話為底冊的調查，雖然卡方檢定達到統計上顯著水準，但各地區的差距均在3%以內，其差異並不算大，其理由如上文所述。

無論是以何種抽樣底冊為依據進行調查，成功樣本的人口結構若依據選民母體或投票母體進行加權處理後，均會與母體結構達到一致，故無需再檢視加權後之樣本結構的代表性。然而值得討論的是，經過加權處理後的調查數據，對投票率、得票率、催票率的偏差估計是否具有顯著的修正效果，提高與母體數據的一致性，以下兩小節乃針對此一議題進行檢證。



表五 2016年總統選舉調查成功樣本結構與選民母體的比較（未加權）

變數	選項	選民母體		2016TEDS 選前電話		2016TEDS 選後面訪		2017TEDS雙軌面訪			
		住宅 電話	卡方檢定	戶籍 抽樣	卡方檢定	戶籍 抽樣	卡方檢定	地址 抽樣	卡方檢定	戶籍 抽樣	卡方檢定
性別	男性	49.3%	$\chi^2 = 0.640$	51.4%	$\chi^2 = 2.874$	53.3%	$\chi^2 = 7.534$	48.1%	$\chi^2 = 0.661$		
	女性	50.7%	$P = 0.424$	48.6%	$P = 0.090$	46.7%	$P = 0.006$	51.9%	$P = 0.416$		
年齡	20-29歲	17.0%	$\chi^2 = 310.440$	15.6%	$\chi^2 = 36.680$	15.5%	$\chi^2 = 34.161$	8.1%	$\chi^2 = 147.375$		
	30-39歲	20.9%	$P < 0.001$	16.7%	$P < 0.001$	16.1%	$P < 0.001$	16.3%	$P < 0.001$		
	40-49歲	19.3%		18.8%		19.3%		18.6%			
	50-59歲	19.2%		19.9%		19.0%		20.1%			
	60歲以上	23.7%		29.0%		30.0%		36.8%			
教育 程度	國小及以下	14.8%	$\chi^2 = 209.121$	19.2%	$\chi^2 = 39.545$	17.2%	$\chi^2 = 21.385$	21.0%	$\chi^2 = 46.513$		
	國、初中	12.8%	$P < 0.001$	9.6%	$P < 0.001$	10.3%	$P < 0.001$	10.3%	$P < 0.001$		
	高中、職	28.1%		26.7%		24.6%		23.7%			
	專科	12.3%		11.1%		12.1%		14.1%			
	大學及以上	31.9%		33.5%		35.9%		31.0%			

變數	選民母體	2016TEDS 選前電話		2016TEDS 選後面訪		2017TEDS雙軌面訪			
		住宅 電話	卡方檢定	戶籍 抽樣	卡方檢定	戶籍 抽樣	卡方檢定	地址 抽樣	卡方檢定
地理 區域	北北基宜	32.3%	$\chi^2 = 72.713$	33.0%	$\chi^2 = 6.081$	31.8%	$\chi^2 = 1.227$	27.5%	$\chi^2 = 19.313$
	桃竹苗	15.0%	$P < 0.001$	15.0%	$P = 0.298$	14.8%	$P = 0.942$	15.6%	$P = 0.002$
	中彰投	19.0%		17.6%		18.7%		19.1%	
	雲嘉南	14.6%		16.0%		15.4%		18.2%	
	高屏澎	16.1%		15.9%		16.6%		16.6%	
花東金馬	3.0%	2.1%	2.5%		2.8%		2.9%		
總數	18,804,950	5,841		1,690		1,190		1,162	

註1：表格內的調查數據為扣除掉拒答樣本數後重新計算百分比和檢定樣本代表性。

註2：選民人數為內政部2015年12年底的人口統計資料。

註3：淺色底表示樣本數據比母體數據偏低3%以上，深色底表示樣本數據比母體數據偏高3%以上。

資料來源：本研究。

表六 2016年總統選舉調查投票受訪者結構與投票母體估計的比較（未加權）

變數	選項	投票母體 (估計)		2016TEDS選前電訪		2016TEDS選後面訪		2017TEDS雙軌面訪			
		住宅 電話	卡方檢定	戶籍 抽樣	卡方檢定	戶籍 抽樣	卡方檢定	地址 抽樣	卡方檢定	卡方檢定	
性別	男性	48.4%	$\chi^2 = 21.246$ $P < 0.001$	50.3%	$\chi^2 = 2.041$ $P = 0.153$	54.5%	$\chi^2 = 12.903$ $P < 0.001$	47.5%	$\chi^2 = 0.284$ $P = 0.594$		
	女性	51.6%		49.7%		45.5%		52.5%			
年齡	20-29歲	14.7%	$\chi^2 = 128.714$ $P < 0.001$	14.0%	$\chi^2 = 22.417$ $P < 0.001$	12.2%	$\chi^2 = 17.440$ $P = 0.002$	5.8%	$\chi^2 = 93.784$ $P < 0.001$		
	30-39歲	18.4%		14.7%		16.4%		16.5%			
	40-49歲	18.7%		19.1%		20.2%		17.0%			
	50-59歲	21.2%		20.8%		19.0%		21.2%			
	60歲以上	27.0%		31.5%		32.3%		39.4%			
地理 區域	北北基直	33.0%	$\chi^2 = 38.397$ $P < 0.001$	33.5%	$\chi^2 = 5.612$ $P = 0.346$	34.4%	$\chi^2 = 4.189$ $P = 0.523$	27.4%	$\chi^2 = 21.223$ $P = 0.001$		
	桃竹苗	15.0%		14.6%		14.9%		14.7%			
	中彰投	19.0%		16.9%		16.9%		18.3%			
	雲嘉南	14.3%		15.6%		15.6%		18.5%			
	高屏澎	16.4%		16.9%		15.6%		18.3%			
	花東金馬	2.4%		2.4%		2.6%		2.9%			
總數		132,811		1,448		868		832			

註1：表格內的調查數據為扣除掉未決定、未投票和拒答的樣本數後重新計算百分比和檢定樣本代表性。

註2：投票人數為中選會2016年委託研究案資料（莊文忠與洪永泰 2016）。

註3：淺色底表示樣本數據比母體數據偏低較多，深色底表示樣本數據比母體數據偏高較多。

資料來源：本研究。

### 三、不同抽樣底冊之成功樣本的投票率比較

理論上，調查成功樣本結構與母體結構的不一致，可能會導致估計與推論的偏差，因此加權處理成爲一個兼顧理論和實務的理性選擇（包正豪與周應龍 2014；洪永泰 2003；黃紀與張佑宗 2003），實務上最常見的做法即是利用基本的人口變數作爲加權依據。然而加權處理實際上真的能夠減少估計的偏誤嗎？抑或只是讓研究者覺得心安理得的做法？另外就選舉研究而言，對投票率、候選人或政黨得票率的估計，在選民母體和投票母體不完全一致的情況下，依據何種母體人口結構進行校準，可以獲得較好的估計結果？這也是值得探究的問題。因此，本研究利用不同抽樣底冊調查所得成功樣本，以選民母體和投票母體的參數爲依據，採取「多變數反覆加權」進行處理，並利用單一樣本 $t$ 檢定(One-Sample  $t$  Test)分析加權處理前後的投票率和得票率，其結果如表七至表九所示。

表七爲加權前的投票率估計，本研究以中選會委託研究案所公開釋出的選舉人投票統計作爲母體投票率的參數，基本上無論是TEDS在選前以住宅電話爲底冊的電訪或選後以戶籍資料爲底冊的面訪，對投票率的估計普遍呈現顯著差異（除電訪中的女性和50至59歲外），且大多數的投票率估計偏高，尤其是2016TEDS選後面訪，各種樣本特徵有去投票的比例均偏高一成五至二成左右，花東金馬地區甚至偏高三成左右。但同樣是以戶籍資料爲底冊的2017年面訪，投票率估計偏高的情形就降低了不少，而以地址資料爲底冊的2017年面訪，大多數的投票率估計均未達到顯著差異。就經驗而言，選前調查的投票率估計偏高，應與社會期許效應(social desirability)有關，而選後半年的面訪相較於選後一年半的面訪，受訪者對自己有沒有投票應該是印象較深刻，但調查結果反而偏離母體投票率甚多，此是否仍受到社會期許效應或其他因素所影響，值得他文再深入分析。至於以地址資料爲底冊的調查，雖然也有部分樣本特徵的投票率估計偏高，但反而較爲接近實際的投票率，同樣值得後續研究再深究其原因。

表八是利用選民母體的基本人口變數進行加權後的投票率估計結果。此一結果顯示，無論是選前以住宅電話爲底冊的電訪或選後兩次以戶籍資料爲底

冊的面訪，加權處理的修正效果十分有限，大多數的投票率估計的檢定結果和未加權的結果差異不大；反而是以地址資料為底冊的面訪，其修正效果略佳，如在性別和地區變數上，有數個投票率估計變成統計上不顯著。表九是利用投票母體的人口特徵進行加權後的投票率估計結果，此和表八的檢定結果非常接近，同樣是以地址資料為底冊之面訪的修正效果略佳，在性別和地區變數上，有數個投票率估計變成不顯著。此一結果顯示，以選民母體和投票母體的基本人口變數進行加權處理，兩者的修正效果差不多，這應是與前文表三和表四所呈現的數據差異不大有關。不過，若未來選舉的投票率持續下降亦或是推動不在籍投票制度等，選民母體與投票母體的人口結構是否可能出現更大的差異、其加權處理的修正效果為何、是否影響選舉預測的準確度等，均仍有待檢證。

#### 四、不同抽樣底冊之成功樣本的得票率與催票率比較

最後，在選舉研究中，研究者除了關心代表政治參與程度的投票率外，政黨和候選人的得票率估計更是矚目的焦點，本研究同樣針對2016年總統選舉三位候選人的得票率和催票率進行加權處理前後的分析。比較未加權前的表一和依據選民母體的人口變數加權後的表十可以看到，加權處理的修正效果相對有限。就選前以住宅電話為底冊的電訪而言，投票率估計稍微接近母體投票率，但仍有顯著差異，蔡英文和朱立倫的得票率和催票率估計差距反而更大，宋楚瑜的得票率和催票率估計幾乎不變；以選後兩次戶籍資料為底冊的面訪而言，無論是投票率、得票率或催票率的估計，加權前後的數據均非常接近；選後以地址資料為底冊的面訪而言，投票率估計的修正效果較為明顯，和母體投票率無顯著差異，得票率估計的變化不大，對三位候選人的催票率估計則是拉近和母體催票率的差距。

表七 2016年總統選舉調查成功樣本的投票率估計比較（加權前）

變數	選項	母體投票率 (估計)		2016TEDS選前電訪		2016TEDS選後面訪		2017TEDS雙軌面訪		
		住宅 電話	檢定 <i>t</i>	戶籍 抽樣	檢定 <i>t</i>	戶籍 抽樣	檢定 <i>t</i>	地址 抽樣	檢定 <i>t</i>	檢定 <i>t</i>
性別	男性	75.8%	<i>t</i> = 13.679***	83.9%	<i>t</i> = 15.268***	74.6%	<i>t</i> = 5.668***	70.7%	<i>t</i> = 3.041**	
	女性	66.7%	<i>t</i> = -0.606	87.6%	<i>t</i> = 17.722***	71.0%	<i>t</i> = 1.996*	72.5%	<i>t</i> = 2.895**	
年齡	20-29歲	75.9%	<i>t</i> = 10.880***	76.5%	<i>t</i> = 7.428***	57.3%	<i>t</i> = 0.054	51.6%	<i>t</i> = -1.053	
	30-39歲	71.3%	<i>t</i> = 9.243***	75.5%	<i>t</i> = 6.758***	74.0%	<i>t</i> = 4.962***	72.7%	<i>t</i> = 4.449***	
	40-49歲	70.8%	<i>t</i> = 5.308***	87.1%	<i>t</i> = 12.060***	76.1%	<i>t</i> = 4.182***	65.7%	<i>t</i> = 0.438	
	50-59歲	71.3%	<i>t</i> = -1.328	89.3%	<i>t</i> = 9.743***	73.0%	<i>t</i> = 0.037	75.7%	<i>t</i> = 0.970	
	60歲以上	70.1%	<i>t</i> = -4.046***	93.1%	<i>t</i> = 15.630***	78.4%	<i>t</i> = 1.528	77.0%	<i>t</i> = 0.905	
地理 區域	北北基宜	71.9%	<i>t</i> = 4.141***	87.1%	<i>t</i> = 13.757***	79.1%	<i>t</i> = 5.540***	71.3%	<i>t</i> = 1.480	
	桃竹苗	70.7%	<i>t</i> = 2.918**	83.5%	<i>t</i> = 7.478***	73.3%	<i>t</i> = 2.181*	67.4%	<i>t</i> = 0.402	
中彰投 雲嘉南 高屏澎 花東金馬	中彰投	69.8%	<i>t</i> = 2.991**	82.5%	<i>t</i> = 7.511***	66.2%	<i>t</i> = 0.099	68.5%	<i>t</i> = 0.822	
	雲嘉南	69.4%	<i>t</i> = 3.270**	83.4%	<i>t</i> = 8.299***	73.8%	<i>t</i> = 2.812**	72.6%	<i>t</i> = 2.620**	
	高屏澎	73.9%	<i>t</i> = 4.559***	91.4%	<i>t</i> = 14.128***	68.2%	<i>t</i> = 0.296	78.8%	<i>t</i> = 3.915***	
	花東金馬	69.6%	<i>t</i> = 4.261***	81.4%	<i>t</i> = 4.895***	69.7%	<i>t</i> = 2.178*	70.6%	<i>t</i> = 2.344*	
總數		201,159	5,841	1,690		1,190		1,162		

註1：表格內的調查數據為將未決定、未投票和拒答的樣本數編碼為「未投票」後重新計算投票率。

註2：母體投票率估計為中選會2016年委託研究案的統計資料。

註3：\*表示 $P < 0.05$ ；\*\*表示 $P < 0.01$ ；\*\*\*表示 $P < 0.001$ 。

資料來源：本研究。

表八 2016年總統選舉調查成功樣本的投票率估計比較（依選民母體加權後）

變數	選項	母體投票率 (估計)		2016TEDS選前電訪		2016TEDS選後面訪		2017TEDS雙軌面訪		
		住宅 電話	檢定 t	戶籍 抽樣	檢定 t	戶籍 抽樣	檢定 t	地址 抽樣	檢定 t	
性別	男性	64.8%	75.1% t = 12.808***	82.2%	t = 13.166***	74.1%	t = 5.085***	67.0%	t = 1.109	
	女性	67.2%	t = -1.147	86.8%	t = 16.975***	70.5%	t = 1.817	69.7%	t = 1.279	
年齡	20-29歲	57.1%	t = 13.038***	75.2%	t = 7.032***	56.7%	t = -0.075	50.8%	t = -1.750	
	30-39歲	58.2%	t = 9.461***	75.4%	t = 7.456***	72.4%	t = 4.927***	70.9%	t = 4.253***	
	40-49歲	64.3%	t = 4.912***	86.8%	t = 12.063***	76.4%	t = 4.241***	64.1%	t = -0.008	
	50-59歲	72.9%	t = -1.281	89.8%	t = 10.020***	74.1%	t = 0.429	75.7%	t = 0.893	
地理 區域	60歲以上	75.1%	t = -5.527***	93.3%	t = 14.650***	78.2%	t = 1.248	76.9%	t = 0.641	
	北北基宜	67.5%	t = 4.069***	85.9%	t = 12.331***	79.0%	t = 5.506***	68.7%	t = 0.531	
區域	桃竹苗	66.0%	t = 2.518*	81.9%	t = 6.547***	70.9%	t = 1.491	61.8%	t = -1.139	
	中彰投	65.9%	t = 2.096*	81.3%	t = 7.022***	65.5%	t = -0.115	65.5%	t = -0.123	
	雲嘉南	64.6%	t = 2.226*	82.5%	t = 7.466***	73.0%	t = 2.499*	67.6%	t = 0.830	
	高屏澎	67.2%	t = 4.729***	90.8%	t = 13.348***	67.4%	t = 0.084	77.7%	t = 3.410**	
	花東金馬	52.0%	t = 3.537**	80.8%	t = 5.306***	71.4%	t = 2.256*	67.9%	t = 1.777	
總數		201,159	5,841	1,690		1,190		1,162		

註1：表格內的調查數據為將未決定、未投票和拒答的樣本數編碼為「未投票」後重新計算投票率。

註2：母體投票率估計為中選會2016年委託研究案的統計資料。

註3：\*表示 $P < 0.05$ ；\*\*表示 $P < 0.01$ ；\*\*\*表示 $P < 0.001$ 。

資料來源：本研究。

表九 2016年總統選舉調查成功樣本的投票率估計比較（依投票母體加權後）

變數	母體投票率估計	2016TEDS選前電訪		2016TEDS選後面訪		2017TEDS雙軌面訪			
		住宅電話	卡方檢定	戶籍抽樣	卡方檢定	戶籍抽樣	卡方檢定	地址抽樣	
性別									
男性	64.8%	75.9%	$t = 13.838^{***}$	83.7%	$t = 14.620^{***}$	74.6%	$t = 5.448^{***}$	68.4%	$t = 1.830$
女性	67.2%	67.1%	$t = -0.140$	87.5%	$t = 17.991^{***}$	71.0%	$t = 2.055^*$	71.5%	$t = 2.289^*$
年齡									
20-29歲	57.1%	76.0%	$t = 12.893^{***}$	77.1%	$t = 7.484^{***}$	57.1%	$t = 0.045$	53.2%	$t = -0.987$
30-39歲	58.2%	71.4%	$t = 9.524^{***}$	75.7%	$t = 7.245^{***}$	74.0%	$t = 5.290^{***}$	73.0%	$t = 4.799^{***}$
40-49歲	64.3%	70.8%	$t = 4.735^{***}$	86.7%	$t = 11.734^{***}$	75.9%	$t = 4.043^{***}$	65.7%	$t = 0.457$
50-59歲	72.9%	71.3%	$t = -1.191$	89.4%	$t = 10.214^{***}$	72.6%	$t = -0.065$	75.1%	$t = 0.775$
60歲以上	75.1%	70.2%	$t = -4.208^{***}$	93.2%	$t = 15.372^{***}$	78.2%	$t = 1.378$	76.7%	$t = 0.687$
地理區域									
北北基宜	67.5%	71.9%	$t = 4.243^{***}$	86.9%	$t = 13.476^{***}$	79.1%	$t = 5.571^{***}$	70.0%	$t = 1.084$
桃竹苗	66.0%	71.1%	$t = 3.342^{**}$	83.5%	$t = 7.541^{***}$	73.0%	$t = 2.096^*$	64.9%	$t = -0.266$
中彰投	65.9%	70.0%	$t = 2.963^{**}$	82.6%	$t = 7.806^{***}$	65.9%	$t = 0.047$	66.8%	$t = 0.282$
雲嘉南	64.6%	69.4%	$t = 2.944^{**}$	83.3%	$t = 7.763^{***}$	73.4%	$t = 2.639^{**}$	69.1%	$t = 1.214$
高屏澎	67.2%	74.4%	$t = 5.144^{***}$	91.3%	$t = 14.075^{***}$	67.7%	$t = 0.149$	79.1%	$t = 3.982^{***}$
花東金馬	52.0%	68.6%	$t = 4.282^{***}$	82.5%	$t = 4.877^{***}$	69.0%	$t = 1.950$	67.9%	$t = 1.801$
總數	201,159	5,841		1,690		1,190		1,162	

註1：表格內的調查數據為將未決定、未投票和拒答的樣本數編碼為「未投票」後重新計算投票率。

註2：母體投票率估計為中選會2016年委託研究案的統計資料。

註3：\*表示 $P < 0.05$ ；\*\*表示 $P < 0.01$ ；\*\*\*表示 $P < 0.001$ 。

資料來源：本研究。



進一步比較依選民母體和依投票母體的人口變數進行加權處理的修正效果可以發現（見表十和表十一），就選前以住宅電話為底冊的電訪而言，兩者的投票率估計差異不大，但以投票母體進行加權時，對蔡英文和朱立倫的得票率及催票率估計有較好一些的修正效果，宋楚瑜的得票率和催票率估計幾乎不變；就選後以兩次戶籍資料為底冊的面訪而言，無論是投票率、得票率或催票率的估計，兩種加權方法的數據均非常接近，差距均在1%以內；就選後以地址資料為底冊的面訪而言，除投票率估計的差距稍微大一些（超過1%），兩種加權方法的得票率和催票率估計差異不大。此一結果同樣顯示，在選民母體與投票母體的結構差異不大的情況下，利用任何一個母體數據進行加權處理，其修正效果差不多。不過，在以住宅電話為底冊的電訪中，以投票母體進行加權時，對蔡英文和朱立倫的得票率及催票率估計有較好的修正效果，其可能的原因在於電訪中普遍存在年輕人受訪比例偏低，而老人受訪比例偏高的現象，若用選民母體進行加權，雖然將樣本結構調整成和母體一致，但估計誤差反而比較大；而投票母體反映了年輕人的投票率偏低，年長者投票率偏高的事實，以此一母體進行加權處理，可以獲得較好的修正效果。然因本研究僅有一次選舉的實證資料，此一論點仍有待累積更多研究的檢證。

## 陸、結論

就母體的涵蓋率而言，理論上選舉民調的最佳抽樣底冊應該是選舉人名冊，雖然此一名冊可能因為戶籍登記資料未能及時更新、有些合格選民長期不在戶籍地、或投票日因各種因素而無法回到戶籍所在地行使投票權造成高估選民人數，但其涵蓋率還是比原始的戶籍資料檔來得好。不過在《個人資料保護法》通過後，選舉民調在實務上已不可能再以選舉人名冊作為抽樣底冊，戶籍資料檔與選民母體的一致性雖有些許落差，以其作為選舉研究的抽樣底冊，亦屬於相當不錯的選擇（尤其是在戶籍資料檔全面電腦化之後，抽樣的成本大幅降低）。然而在政府主管機關對「公務機關」從嚴認定後，TEDS便難以再以戶籍資料檔作為抽樣底冊，僅能再退而求其次，改以門牌地址資料庫作為抽樣

表十 2016年總統選舉結果與民調結果的比較\_投票率與候選人得票率(依選民母體加權後)

實際 選舉結果	2016TEDS 選前電話		2016TEDS 選後面訪		2017TEDS雙軌面訪			
	住宅電話		戶籍抽樣		戶籍抽樣		地址抽樣	
	住宅 電話	卡方檢定	戶籍 抽樣	卡方 檢定	戶籍 抽樣	卡方檢定	地址 抽樣	卡方檢定
投票率	70.6%	$t = 7.213^{***}$	84.6%	$t = 20.778^{***}$	72.3%	$t = 4.598^{***}$	68.3%	$t = 1.482$
得 票 率	66.3%	$t = 8.008^{***}$	62.5%	$t = 4.642^{***}$	64.5%	$t = 5.062^{***}$	62.6%	$t = 3.725^{***}$
	蔡英文	$t = -12.884^{***}$	27.3%	$t = -2.963^{**}$	24.2%	$t = -4.571^{***}$	25.7%	$t = -3.384^{**}$
	朱立倫	$t = 4.343^{***}$	10.3%	$t = -2.942^{**}$	11.3%	$t = -1.375$	11.7%	$t = -0.921$
催 票 率	12.8%	$t = 9.452^{***}$	46.2%	$t = 7.843^{***}$	45.4%	$t = 6.042^{***}$	42.1%	$t = 3.712^{***}$
	蔡英文	$t = -10.156^{***}$	20.2%	$t = -0.128$	17.1%	$t = -2.955^{**}$	17.3%	$t = -2.709^{**}$
	朱立倫	$t = 5.244^{***}$	7.6%	$t = -1.238$	8.0%	$t = -0.561$	7.9%	$t = -0.636$
宋楚瑜	10.5%		26.0%		29.5%		32.7%	
未投票 / 未表態	31.2%							
總數		5,841		1,690		1,190		1,162

註1：選舉結果中，投票率的計算係以投票人數除以選舉人數；各候選人的得票率計算是以有效票數為計算分母；各候選人的催票率計算是以選舉人數為計算分母。

註2：民意調查結果中，投票率的計算係以所有的樣本數為計算分母；各候選人的得票率計算是以有明確回答支持對象（即扣除未投票或未表態）的樣本數為計算分母；各候選人的催票率計算是以所有樣本數為計算分母。

註3：\*表示 $P < 0.05$ ；\*\*表示 $P < 0.01$ ；\*\*\*表示 $P < 0.001$ 。

資料來源：黃紀（2016）。

表十一 2016年總統選舉結果與民調結果的比較\_投票率與候選人得票率(依投票母體加權後)

實際選舉結果	2016TEDS 選前電話		2016TEDS 選後面訪		2017TEDS雙軌面訪			
	住宅電話	卡方檢定	戶籍抽樣	卡方檢定	戶籍抽樣	卡方檢定	地址抽樣	卡方檢定
投票率	71.4%	$t = 8.585^{***}$	85.6%	$t = 22.614^{***}$	72.7%	$t = 5.009^{***}$	69.9%	$t = 2.712^{**}$
得票率	蔡英文	$t = 4.604^{***}$	62.5%	$t = 4.670^{***}$	65.2%	$t = 5.565^{***}$	62.3%	$t = 3.589^{***}$
	朱立倫	$t = -8.445^{***}$	27.5%	$t = -2.756^{**}$	24.1%	$t = -4.705^{***}$	25.9%	$t = -3.263^{**}$
	宋楚瑜	$t = 4.114^{***}$	10.0%	$t = -3.316^{**}$	10.7%	$t = -1.973^*$	11.8%	$t = -0.872$
催票率	蔡英文	$t = 7.343^{***}$	46.6%	$t = 8.159^{***}$	46.3%	$t = 6.625^{***}$	42.8%	$t = 4.199^{***}$
	朱立倫	$t = -5.560^{***}$	20.5%	$t = 0.243$	17.1%	$t = -2.949^{**}$	17.8%	$t = -2.199^*$
	宋楚瑜	$t = 5.232^{***}$	7.5%	$t = -1.474$	7.6%	$t = -1.052$	8.1%	$t = -0.357$
未投票 / 未表態	30.6%		25.4%		29.0%		31.3%	
總數		5,841		1,690		1,190		1,162

註1：選舉結果中，投票率的計算係以投票人數除以選舉人數；各候選人的得票率計算是以有效票數為計算分母；各候選人的催票率計算是以選舉人數為計算分母。

註2：民意調查結果中，投票率的計算係以所有的樣本數為計算分母；各候選人的得票率計算是以有明確回答支持對象（即扣除未投票或未表態）的樣本數為計算分母；各候選人的催票率計算是以所有樣本數為計算分母。

註3：\*表示 $P < 0.05$ ；\*\*表示 $P < 0.01$ ；\*\*\*表示 $P < 0.001$ 。

資料來源：黃紀（2016）。

底冊，此對以戶籍所在地為合格選民認定標準的選舉民調而言，不僅無法事先完成最後一個階段的抽樣，訪員必須找中選地址後依戶中抽樣原則抽出真正的受訪者，多增加一個階段的抽樣誤差。同時也因為無法事先得知受訪者的姓名與特徵，而導致訪問失敗的比例增加。至於以住宅電話簿或RDD方法做為電訪的抽樣底冊，雖然與選民母體的差異較大，其涵蓋率也隨著行動電話的普及率提高而下降，但目前仍是主要的抽樣底冊來源，即使是採取雙底冊調查，基於成本考量，大多數還是以住宅電話為主。不過，未來也有可能改以行動電話為主，甚至發展出以行動電話為抽樣底冊的新調查模式。

理論上，無論是採用何種抽樣底冊，在投票率不是百分之百的情況下，選舉母體並不等於投票母體乃是常態，且隨著各民主國家的選舉投票率日益下滑，投票母體與選舉母體之間的差異有可能拉大；反觀若是未來推動不在籍制度，則是有可能拉近兩者之間的差異，因此何種抽樣底冊的調查成功樣本結構與標的母體最為接近，自是一個值得析探的議題。不過，值得說明的是，抽樣底冊的選擇雖然與母體的涵蓋率高度相關，但母體涵蓋率高低與調查結果的一致性並非完全相關。蓋涵蓋率主要是探討母體中的每一個個案是否都有被抽中的機率，涵蓋率越高，代表被排除在抽樣底冊之外的個案數越少，在隨機抽樣的原則下，中選樣本的結構與母體的結構越可能趨於一致；而調查結果與實際投票結果是否一致，雖會受到抽樣底冊之涵蓋率的影響，但包括調查方式、調查時間、問卷設計、訪問技巧及受訪者回答意願、訪問失敗率等，均會影響民調數據與實際投票結果的差異性，故「涵蓋率低是否必然出現估計效果差」，研究者除了分析不同抽樣底冊的樣本代表性外，亦應進一步分析不同抽樣底冊之調查數據與實際選舉結果的一致程度，並利用個體資料探究可能的解釋因素。

本研究以2016年總統選舉為例，利用未加權的調查資料，比較住宅電話、戶籍資料和地址資料這三種抽樣底冊的樣本結構，與選民母體及投票母體之間的一致程度。此外，基於以個體層次的實證調查結果進行母體推論統計分析，很重要的一個前提要件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（包正豪與周應龍 2014, 159），因此，本研究也分別利用選民母體和投票母體的基本人口變數進行加權處理，並比較分析其差異。要言之，

本研究從方法論的觀點，針對不同抽樣底冊來源之選舉民調進行比較分析，包括樣本結構的代表性；投票率、得票率、催票率估計的正確性。儘管在加權處理的調整效果等雖獲得一些初步的研究發現，兼具理論與實務之參考意涵，但仍有不少值得未來深入探究和解釋的議題。

本研究透過循證分析所獲致的幾個重要結論如下：首先，在「2016年總統暨立法委員選舉選舉人投票資料檔」中，總統選舉的投票率為六成六左右，就性別、年齡層和地區的個別變數來看，選民母體與投票母體的整體結構，僅在年齡層方面有較大差異；但若就變數的組合來看，除花東地區外，選民母體和投票母體在性別和地區的結構差異較小，且59歲以下女性選民參與投票的比例均高於男性，而60歲以上則是男性選民參與投票的比例較高，以致於投票母體的性別比例和選民母體看似較為一致。

其次，針對不同抽樣底冊進行調查的成功樣本結構來看，沒有任何一種抽樣底冊的成功樣本結構與選民母體是完全一致的。以戶籍資料為底冊的調查在地區方面的代表性較好，以住宅電話簿或地址資料為底冊的調查在性別方面的代表性較佳，而這些底冊均在年齡層和教育程度方面與選民母體的結構不一致。同樣地，沒有任何一種抽樣底冊的成功樣本結構與投票母體是完全一致的，僅有以戶籍資料為底冊的調查在地區方面的代表性較好，以地址資料為底冊的調查在性別方面的代表性較佳，兩次以戶籍資料為底冊的調查在性別方面的代表性並不穩定，而這些底冊均在年齡層方面與投票母體的結構不一致。

最後，無論是利用選民母體或投票母體的基本人口變數進行加權處理，兩者對投票率、得票率和催票率的修正效果較為有限。在投票率方面，除了以地址資料為底冊之調查的修正效果略佳外，以住宅電話簿和以戶籍資料為底冊的調查，不同人口特徵的投票率估計檢定結果和加權處理前的差異不大；在得票率和催票率方面，以住宅電話為底冊的調查，利用選民母體進行加權，蔡英文和朱立倫的得票率和催票率估計差距反而更大，利用投票母體進行加權，蔡英文和朱立倫的得票率和催票率估計差距則可以縮小一些；以戶籍資料或地址資料為底冊的調查，無論是利用選民母體或投票母體進行加權，其數據均和加權前差不多。因此，要解決民調結果與投票結果之間的落差，有關母體的定義並不是關鍵，可能還是必須回到抽樣底冊的選擇、問卷題目的設計、調查技術的

精進、訪問失敗的減少等環節，才是治本之道。

本研究的貢獻在於明確指出選民母體與投票母體之概念內涵與操作性定義的不同外，並以2016年總統大選為例，引進中選會所建置之選民投票資料檔作為投票母體，比較分析選民母體與投票母體的差異及以兩者作為母體參數進行加權處理的修正效果，提供更為客觀與深入的證據。就此言之，若選民母體和投票母體的差異不大，則選舉民調利用何者作為加權依據，其修正效果差異不大。在實務上，投票母體係屬於事後取得的數據，且每一次的母體結構可能會隨著選舉性質、競爭程度、甚至是天候因素的影響，而有較大的變動，因此以其作為選前民意調查的加權依據恐有困難。但是選後作為評估或是校正調查樣本的代表性，應有其客觀的參考價值。

本研究的限制包括：首先2016年總統選舉是否為「典型」大選的確是一個有趣且值得探討的問題，不僅是與選民的投票意願有關、也涉及政黨或候選人的得票率和催票率，除了國民黨的換柱事件、投票前爆發周子瑜國旗事件等，這些事件的確可能增加選舉民調數據與實際選舉結果之間的不一致性，然因本文主要是從方法論的觀點進行分析，並未從選民的觀點進行討論，無法評估影響受訪者的因素，此為本研究的限制，也是其他研究可再深入探討的議題；其次，因本研究只是單次選舉的調查資料，而且受限於時間和成本的考量，本研究所使用之面訪資料的樣本數各有1,000多個，若要進行更多變數的組合，可能會出現某些組合的樣本數太少而導致估計不穩定的結果；最後，由於本研究目前只處理母體涵蓋率問題，未處理訪問失敗問題對樣本結構的影響。因此，這些研究發現還需要其他選舉結果的重複驗證。

## 參考文獻

### I. 中文部分

- 包正豪、周應龍，2014，〈以電訪樣本資料庫作為解決原住民調查研究難題之可行性的再討論〉，《選舉研究》，21(2): 147-175。
- (Pao, Cheng-hao, and Ying-lung Chou. 2014. "Yi dian fang yang ben zi liao ku zuo wei jie jue yuan zhu min diao cha yan jiu nan ti zhi ke xing xing de zai tao lun" [Using Telephone Survey Sample Data as a Feasible Resolution to the Difficulty of Indigenous Survey]. *Journal of Electoral Studies* 21(2): 147-175.)
- ，2019，〈原住民選民的族群投票行為：2012年山地原住民立委選舉之個案研究〉，《調查研究——方法與應用》，42: 165-193。
- (-----, 2019. "Yuan zhu min xuan min de zu qun tou piao xing wei: 2012 nian shan di yuan zhu min li wei xuan ju zhi ge an yan jiu" [Ethnic Voting Behavior of Taiwan's Indigenous Electorate: A Case Study of the 2012 Mountain Indigenous Legislators Election]. *Survey Research-Method and Application* 42: 165-193.)
- 朱斌妤、黃東益、洪永泰、李仲彬、曾憲立，2015，《數位國家治理（2）：國情追蹤與方法整合》，台北：國家發展委員會。
- (Chu, Pin-yu, Tong-yi Huang, Yung-tai Hung, Chung-pin Lee, and Hsien-lee Tseng. 2015. "Shu wei guo jia zhi li (2): guo qing zhui zong yu fang fa zheng he" [Public Value and Electronic Governance (2): National Status and Methodological Integration]. Taipei: National Development Council.)
- 吳重禮、李世宏，2005，〈政治賦權、族群團體與政治參與：2001年縣市長選舉客家族群的政治信任與投票參與〉，《選舉研究》，12(1): 69-115。
- (Wu, Chung-li, and Shih-hung Lee. 2005. "Zheng zhi fu quan, zu qun tuan ti yu zheng zhi can yu: 2001 nian xian shi zhang xuan ju ke jia zu qun de zheng zhi xin ren yu tou piao can yu" [Political Empowerment, Ethnic Groups, and Political Participation: Political Trust and Voting Participation of the

Hakkas in the 2001 County Magistrate/City Mayoral Elections in Taiwan]. *Journal of Electoral Studies* 12(1): 69-115.)

吳重禮、崔曉倩，2010，〈族群、賦權與選舉評價——2004年與2008年總統選舉省籍差異的實證分析〉，《臺灣民主季刊》，7(4): 137-182。

(Wu, Chung-li, and Hsiao-chien Tsui. 2010. “Zu qun, fu quan yu xuan ju ping jia: 2004 nian yu 2008 nian zong tong xuan ju sheng ji cha yi de shi zheng fen xi” [Ethnicity, Empowerment, and Electoral Evaluation: Taiwanese-Mainlander Differences in the 2004 and 2008 Presidential Elections]. *Taiwan Democracy Quarterly* 7(4): 137-182.)

林宇旋、楊景閔、洪百薰、劉怡妏、吳秀英、邱淑媿，2011，〈提高社區面訪調查回應率之可行策略與成效分析〉，《台灣公共衛生雜誌》，30(3): 290-299。

(Lin, Yu-hsuan, Chin-min Yang, Baai-shyun Hurng, I-wen Liu, Shioh-ing Wu, and Shu-ti Chiou. 2011. “Ti gao she qu mian fang diao cha hui ying lu zhike xing ce lue yu cheng xiao fen xi” [Practical Strategies to Improve the Response Rate for a Household Interview Survey]. *Taiwan Journal of Public Health* 30(3): 290- 299.)

林珮婷、鄭夙芬、王德育，2020，〈臺灣認同的世代差異與投票抉擇——以2016年總統選舉為例〉，《選舉研究》，27(1): 85-124

(Lin, Pei-ting, Su-feng Cheng, and T.Y. Wang. 2005. “Tai wan ren tong de shi dai cha yi yu tou piao jue ze: Yi 2016 nian zong tong xuan ju wei li” [Generational Difference of Taiwan Identity - the Effects on Vote Choice in the 2016 Presidential Election]. *Journal of Electoral Studies* 27(1): 85-124.)

洪永泰，1996，〈抽樣調查中樣本代表性的問題〉，《調查研究》，1: 7-37。

(Hung, Yung-tai. 1996. “Chou yang diao cha zhong yang ben dai biao xing de wen ti” [The Problems of Sample Representativeness on Survey]. *Survey Research* 1: 7-37.)

-----，2003，〈原始樣本、替代樣本、與追蹤樣本的比較：「2001年台灣選舉與民主化調查研究」訪問失敗問題的探討〉，《選舉研究》，10(2):



- 37-58。
- (----- 2003. “Yuan shi yang ben, ti dai yang ben, yu zhui zong yang ben de bi jiao: "2001 nian Taiwan xuan ju yu min zhu hua diao cha yan jiu" fang wen shi bai wen ti de tan tao” [The Nonresponse Problems of the 2001 TEDS Survey]. *Journal of Electoral Studies* 10(2): 37-58.)
- , 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——方法與應用》, 18: 9-44。
- (-----, 2005. “Tai wan di qu chou yang diao cha ge zhong mu ti ding yi, chou yang di ce, he han gai lv de bi jiao” [The Sampling Frames and Coverage Rates of Different Types of Survey Populations in Taiwan]. *Survey Research- Method and Application* 18: 9-44.)
- 洪永泰、洪百薰、林宇璇、呂孟穎、許勝懋、吳淑惠、卓仲彥、徐書儀, 2014, 〈手機使用對臺灣地區電話調查涵蓋率之影響評估〉, 《調查研究——方法與應用》, 31: 7-30。
- (Hung, Yung-tai, Baai-shyun Hurng, Yu-hsuan Lin, Ming-ing Lu, Shen-mao Hsu, Shu-hui Wu, Chung-yen Cho, and Shu-yi Hsu. 2014. “Shou ji shi yong dui Taiwan di qu dian hua diao cha han gai lv zhi ying xiang ping gu” [An Evaluation of the Effect of Cellphone-only Users on Telephone Survey Coverage Rates in Taiwan]. *Survey Research- Method and Application* 31: 7-30.)
- 洪永泰、黃永政, 2001, 臺灣地區電話隨機撥號抽樣方法之研究, 《選舉研究》, 7(1): 173-184。
- (Hung, Yung-tai, and Yung-cheng Huang. 2001. “Tai wan di qu dian hua sui ji bo hao chou yang fang fa zhi yan jiu” [Telephone Sampling: Random Digit Dialing in Taiwan]. *Journal of Electoral Studies* 7(1): 174-184.)
- 孫榮光, 2020, 〈2016年總統選舉桃竹苗客家 / 非客家選民支持蔡英文的因素比較〉, 《人文社會科學研究》, 14(2): 73-93。
- (Sun, Jung-kuang. 2020. “2016 nian zong tong xuan ju tao zhu miao ke jia/fei ke jia xuan min zhi chi Cai ying-wen de yin su bi jiao” [A Comparison of Factors Leading to Hakka/Non-Hakka Voters’ Support of Tsai Ing-wen in Taoyuan,

Hsinchu and Miaoli in the 2016 Presidential Election]. *NPUST Humanities and Social Sciences Research* 14(2): 73-93.)

翁秀琪、孫秀蕙，1995，〈性別政治？——從民國八十二年台灣地區縣、市長選舉看性別、傳播與政治行爲〉，《新聞學研究》，51: 87-111。

(Weng, Shieu-chi, and Hsui-hui Sun. 1995. "Xing bie zheng zhi? - Cong min guo ba shi er nian Taiwan di qu xian/shi zhang xuan ju kan xing bie, chuan bo yu zheng zhi xing wei" [Gender Politics? -Examining the Relationship Among Gender, Communication and Political Behavior in 1993 Election in Taiwan]. *Mass Communication Research* 51: 87-111.)

張鏡文、黃東益、洪永泰，2017，〈住宅電話與手機雙底冊調查的組合估計：以2016總統選舉預測為例〉，《選舉研究》，24(2): 65-96。

(Chang, Teng-wen, Tong-yi Huang, and Yung-tai Hung. 2017. "Zhu zhai dian hua yu shou ji shuang di ce diao cha de zu he gu ji: Yi 2016 zong tong xuan ju yu ce wei li" [Post-Stratified Estimation Procedures for the Dual Frame Telephone Survey in Taiwan: The Case of the 2016 Presidential Election]. *Journal of Electoral Studies* 24(2): 65-96.)

莊文忠、洪永泰，2016，〈選舉人性別投票統計改良之研究〉，計畫編號：PG10505-0031，台北：中央選舉委員會委託研究案。

(Juang, Wen-jong, and Yung-tai Hung. 2016. "Xuan ju ren xing bie tou piao tong ji gai liang zhi yan jiu" [The Improvement of Electors Statistics on Gender]. PG10505-0031. Taipei: Central Election Commission Research Project Report.)

許勝懋，2015，〈「唯手機族」對未來電話調查的挑戰與啟發〉，《調查研究——方法與應用》，34: 33-65。

(Hsu, Shen-mao. 2015. "' Wei shou ji zu" dui wei lai dian hua diao cha de tiao zhan yu qi fa" [New Challenges and Inspirations for Future Telephone Survey Research in Taiwan: The Cell-Phone-Only Factor]. *Survey Research- Method and Application* 34: 33-65.)

陳光輝、劉從葦，2006，〈台灣選舉與民主化調查固定樣本〉(TEDS panel)之

- 代表性探討〉。《選舉研究》，13(2): 75-116。
- (Chen, Kuang-hui, and Tsung-wei Liu. 2006. "'Tai wan xuan ju yu min zhu hua diao cha gu ding yang ben"(TEDS panel) zhi dai biao xing tan tao" [The Examination of Taiwan's Election and Democratization Study Panel Data]. *Journal of Electoral Studies* 13(2): 75-116.)
- 陳陸輝，2006，〈政治信任的政治後果——以2004年立法委員選舉為例〉，〈臺灣民主季刊〉，3(2): 39-61。
- (Chen, Lu-huei. 2006. "Zheng zhi xin ren de zheng zhi hou guo: Yi 2004 nian li fa wei yuan xuan ju wei li" [Political Consequences of Political Trust: The Case of 2004 Legislative Elections in Taiwan]. *Taiwan Democracy Quarterly* 3(2): 39-61.)
- 陳義彥，2020，〈政治與政治學〉，載於《政治學》，陳義彥主編，台北：五南。
- (Chen, Yih-yan. 2020. "Zheng zhi yu zheng zhi xue" [Politics and Political Science]. In "Zheng zhi xue" [Politics], ed Chen, Yih-Yan. Taipei: Wunan.)
- 陳義彥、蔡孟熹，1997，〈新世代選民的政黨取向與投票抉擇——首屆民選總統的分析〉，〈政治學報〉，29: 63-91。
- (Chen, Yih-yan, and Meng-shi Tsai. 1997. "Xin shi dai xuan min de zheng dang qu xiang yu tou piao jue ze: Shou jie min xuan zong tong de fen xi" [Party Orientation and Voting Choice of the New Generation of Voters: An Analysis Based on the 1996 President Election]. *Chinese Political Science Review* 29: 63-91.)
- 陳鴻嘉，2019，〈時間就是金錢：訪問時間觀點下的雙底冊電話調查成本比較〉。《調查研究——方法與應用》，43: 135-167。
- (Chen, Hung-chia. 2019. Shi jian jiu shi jin qian: Fang wenshi jianguan dianxia de shuangdi cedian huadiao chacheng benbi jiao" [Time is Money: A Comparison of the Cost of a Dual-Frame Telephone Survey Based on the Length of Interview]. *Survey Research-Method and Application* 43: 135-167.)
- 曾憲立、洪永泰、朱斌妤、黃東益、謝翠娟，2018，〈多元民意調查方法的比較研究〉，〈調查研究-方法與應用〉，41: 87-117。

(Tseng, Hsien-lee, Yung-tai Hung, Pin-yu Chu, Tong-yi Huang, and Tsui-chuan Hsieh. 2018. “Duo yuan min yi diao cha fang fa de bi jiao yan jiu” [A Comparison Study of Multiple Survey Methods]. *Survey Research-Method and Application* 41: 87-117.)

游清鑫、蕭怡靖，2007，〈以新選民的政治態度論臺灣民主政治的未來〉，  
《臺灣民主季刊》，4(3): 109-151。

(Yu, Ching-hsin, and Yi-ching Hsiao. 2007. “Yi xin xuan min de zheng zhi tai du lun tai wan min zhu zheng zhi de wei lai” [New Voters’ Political Attitudes and the Future of Taiwan’s Democracy]. *Taiwan Democracy Quarterly* 4(3): 109-151.)

黃秀端、趙湘瓊，1996，〈台灣婦女近十年來政治態度的變遷——民國七十二年至八十一年〉，  
《問題與研究》，35(10): 71-95。

(Hawang, Shioh-duan, and Shiang-chong Chou. 1996. “Tai wan fu nü jin shi nian lai zheng zhi tai du de bian qian: Min guo qi shi er nian zhi ba shi yi nian” [Changes in He Political Attitudes of Taiwan Women -1983~1992]. *Issues and Studies* 35(10): 71-95.)

黃紀，2016，〈2012年至2016年「選舉與民主化調查」四年期研究規劃（4/4）：2016年總統與立法委員選舉面訪案〉，科技部補助專題研究計畫，計畫編號：MOST 101-2420-H-004-034-MY4。

(Huang, Chi. 2016. “2012 nian zhi 2016 nian "Xuan ju yu min zhu hua diao cha" si nian qi yan jiu gui hua: 2016 nian zong tong yu li fa wei yuan xuan ju mian fang an” [Taiwan’s Election and Democratization Study, 2012-2016 (IV)]. MOST 101-2420-H-004-034-MY4. Taipei: Ministry of Science and Technology Research Project.)

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為例〉，  
《選舉研究》，10(2): 1-35。

(Huang, Chi, and Yu-tzung Chang. 2003. “Yang ben dai biao xing jian ding yu zui xiao cha yi jia quan: Yi 2001 nian tai wan xuan ju yu min zhu hua diao cha wei li” [On Minimum-Discrimination-Information (MDI) Method of Weighting: an Application to the 2001 Taiwan’s Election and Democratization Study(TEDS)].

*Journal of Electoral Studies* 10(2): 1-35.)

黃東益、施佳良、謝忠安，2011，〈台灣文官調查研究資料蒐集方法與調查品質：「訪員遞送與受訪者自填問卷」方法之探討〉，《調查研究——方法與應用》，25: 141-179。

(Huang, Tong-yi, Chia-liang Shih, and Chung-an Hsieh. 2011. "Tai wan wen guan diao cha yan jiu zi liao sou ji fang fa yu diao cha pin zhi: " Fang yuan di song yu shou fang zhe zi tian wen juan " fang fa zhi tan tao" [Data Collection and Non-response in Surveying Taiwan's Bureaucrats]. *Survey Research-Method-and Application* 25: 141-179.)

楊婉瑩，2007，〈政治參與的性別差異〉，《選舉研究》，14(2): 53-94。

(Yang, Wang-ying. 2007. "Zheng zhi can yu de xing bie cha yi" [The Gender Difference in Political Participation]. *Journal of Electoral Studies* 14(2): 53-94.)

楊婉瑩、林珮婷，2010，〈她們為什麼投給馬英九？探討2008年總統大選的性別差距〉，《選舉研究》，17(1): 91-128。

(Yang, Wang-ying, and Pei-ting Lin. 2010. "Ta men wei shen me tou ji ma ying jiu? Tan tao 2008 nian zong tong da xuan de xing bie cha ju" [Why Do Women Vote for Ma? The Gender Gap in the 2008 Presidential Election]. *Journal of Electoral Studies* 17(1): 91-128.)

-----，2013，〈她們改投給蔡英文嗎？2008-2012年總統大選性別差距的變動〉，《選舉研究》，20(2): 37-71。

(-----, 2013. "Ta men gai tou ji cai ying wen ma? 2008-2012 nian zong tong da xuan xing bie cha ju de bian dong" [Do Women Transfer Their Votes to Tsai? The Change of Gender Gap from 2008 to 2012 Presidential Election]. *Journal of Electoral Studies* 20(2): 37-71.)

楊婉瑩、劉嘉薇，2006，〈探索性別差距的不同型態——以台灣選民政黨認同為例〉，《東吳政治學報》，23: 115-156。

(Yang, Wang-ying, and Jia-wei Liu. 2006. "Tan suo xing bie cha ju de bu tong xing tai: Yi tai wan xuan min zheng dang ren tong wei li" [Exploring the Different Patterns of Gender Gap-A Study of Taiwan Voters' Party Identification]. *So-*

*ochow Journal of Political Science* 23: 115-156.)

鄭夙芬，2000，〈政治類調查研究中訪員錯誤類型之研究〉，《選舉研究》，7(2): 143-191。

(Cheng, Su-feng. 2000. “Zheng zhi lei diao cha yan jiu zhong fang yuan cuo wu lei xing zhi yan jiu” [The Types of Interviewer-related Error in Political Survey]. *Journal of Electoral Studies* 7(2): 143-191.)

-----，2009，〈族群、認同與總統選舉投票抉擇〉，《選舉研究》，16(2): 23-49。

(-----, 2009. “Zu qun, ren tong yu zong tong xuan ju tou piao jue ze” [Ethnicity, Identity, and Vote Choice in Taiwan]. *Journal of Electoral Studies* 16(2): 23-49.)

蕭怡靖，2006，〈「台灣選舉與民主化調查」再測信度之分析〉，《選舉研究》，13(2): 117-144。

(Shiao, Yi-ching. 2006. “‘Tai wan xuan ju yu min zhu hua diao cha’ zai ce xin du zhi fen xi” [Analysis of Test-Retest Reliability in Taiwan’s Election and Democratization Study]. *Journal of Electoral Studies* 13(2): 117-144.)

龍文彬，1998，〈以戶口及住宅普查資料檔為抽樣母體之實務應用與探討〉，《調查研究——方法與應用》，6: 87-109。

(Lung, Wen-pin. 1998. “Yi hu kou ji zhu zhai pu cha zi liao dang wei chou yang mu ti zhi shi wu ying yong yu tan tao” [Application and Discussion of Taking Household and Population Census as Sampling Population]. *Survey Research-Method and Application* 6: 87-109.)

謝淑惠、李名鏞，2020，〈監測訪問時間的調查品質管制程序〉，《調查研究——方法與應用》，44: 7-39。

(Hsieh, Shu-hui, and Ming-yung Lee. 2020. “Jian ce fang wen shi jian de diao cha pin zhi guan zhi cheng xu” [Monitoring Interview Duration for Survey Quality Control]. *Survey Research-Method and Application* 44: 7-39.)

謝淑惠、詹大千、劉雪涵、傅仰止，2018，〈大型面訪調查的抽樣架構變革：從戶籍抽樣到地址抽樣〉，《中國統計學報》，56(4): 75-97。

(Hsieh, Shu-hui, Ta-chien Chan, Hsueh-han Liu, and Yang-chih Fu. 2018. “Da xing mian fang diao cha de chou yang jia gou bian ge: cong hu ji chou yang dao di zhi chou yang” [An Exploratory Study of the Transition from Register-Based Sampling to Address- Based Sampling in Large Face-to-Face Surveys]. *Journal of the Chinese Statistical Association* 56(4): 75-97.)

## II. 外文部分

- Cheng, Su-feng. 2006. “Survey Participation in Taiwan: Evidence from the ESC Surveys.” *Journal of Electoral Studies* 13(1): 163-195.
- Griffin, R., R. Teixeira, and W. H. Frey. 2018. “America’s Electoral Future: Demographic Shifts and the Future of the Trump Coalition.” [https://www.brookings.edu/research/americas-electoral-future\\_2018/](https://www.brookings.edu/research/americas-electoral-future_2018/) (accessed March 14, 2019)
- Lohr, Sharon L. 2010. *Sampling: Design and Analysis*. 2nd ed., Boston, MA: Brooks/ Cole.
- Salkind, N. J., and B. B. Frey. 2020. *Statistics for People Who (Think They) Hate Statistics*, 7th, California: Sage Publications, Inc.

## 附錄

附錄資訊可於當期期刊出刊後，於《選舉研究》期刊網頁查詢。<sup>10</sup>

附錄一：歷屆總統選舉調查的抽樣設計與抽樣底冊。

附錄二：歷年面訪計畫的抽樣底冊來源。

---

<sup>10</sup> 選舉研究期刊網頁網址：<http://www.jestw.com/tc/index.php>。

# **A Demographic Analysis of Eligible Population and Voter Population under Different Sampling Frames: The Case of the 2016 Presidential Election Survey**

*Wen-jong Juang<sup>\*</sup>, Mei-rong Lin<sup>\*\*</sup>, and Yung-tai Hung<sup>\*\*\*</sup>*

## **Abstract**

The controversy over the authenticity of election polls or the representativeness of samples is the definition of “population” in election polls. Since voter turnout is not absolute, there is a difference between “eligible population” and “voter population”. Therefore, the demographic consistency between eligible population and voter population with different sample frames in election polls is a topic that has become crucial in survey methodology. The purpose of this study is threefold: (1) exploring the error of population coverage through reviewing the evolution of different sampling frames used in election polls in Taiwan; (2) clarifying the implication of eligible population and voter population in election studies and analyzing the 2016 presidential election campaign in Taiwan; (3) analyzing the difference of the polls conducted by TEDS for the 2016 presidential election by using different sample frames, and examining the effect of the weighting adjustment by population parameters, and exploring its theoretical and practical implications.

---

\* Professor, Department of Public Policy and Management, Shih Hsin University.

\*\* Associate Professor, Department of International Business, Tamkang University. (corresponding author).

\*\*\* Retired Professor, Department of Political Science, National Taiwan University.



This study found that there was a minor difference between the eligible population and the voter population in the 2016 presidential election in Taiwan. However, the polls from different sampling frames were not completely consistent with the polls from these two populations. The correction effect of the weighting adjustment by the different population parameter on voter turnout, vote gain, and encouraging voter participation is limited. Therefore, to resolve the gap between the poll results and the voting results, the definition of the population might not be critical. The selection of sample frames, the design of questionnaires, the improvement of interview skills, and the reduction of interview failures are more important factors.

**Keywords:** eligible population, voter population, sampling frame, register-based sampling, address-based sampling.

