

# 雙底冊電話調查合併組合 加權處理之比較分析\*

侯佩君\*\*

## 《本文摘要》

隨著科技進步與時代變遷，近年來臺灣面臨單底冊電話樣本的涵蓋誤差日益增大問題，不論市話調查或手機調查；而雙底冊電話調查的施行應可有效降低此一缺失。惟有關雙底冊電話調查的資料合併策略與加權處理，當下臺灣學界仍莫衷一是。本文旨在利用臺灣學界甚少觸及的變異數與偏差，以2020年雙底冊電話調查的實務資料，來評比各類型雙底冊調查的合併組合在各種加權策略下的加權效果，包括重疊設計與篩選設計；此外，本文也進一步探討非等比例樣本配置的雙底冊電話調查在資料推估上的差異，提供實務應用參考。

本文分析結果顯示：1.在雙底冊各自具有完整代表性且樣本數近乎等比例的條件下，採用一般的母體人口特徵變項來校正加權，不論是使用重疊設計或篩選設計來合併，因加權而增加的變異並不大，且僅事後分層組合估計的樣本估計值與其他合併組合有所差距。2.若雙底冊的樣本採用不

---

\* 作者非常感謝匿名審查人及編審委員的細心審閱，為本文提出寶貴建議。本文內容概由作者自行負責。

\*\* 中央研究院人文社會科學研究中心調查研究專題中心專案研發學者。

E-mail: peichun@gate.sinica.edu.tw

等比例配置，以一個完整代表性底冊搭配另一個僅具有1/3或1/2樣本數的底冊，利用篩選設計、重疊設計的直接合併及有效樣本數合併等三種方式進行比較後發現，因應加權所衍生的加權耗損不大，而彼此之間在樣本估計值上的差異也很小。

**關鍵詞：**雙底冊電話調查、加權、設計效果、加權損耗、偏差

## 壹、前言

近年來，隨著市話人口涵蓋率的逐年下降，雙底冊電話調查在臺灣學界掀起不小的討論聲量。翻閱過往歷史，早在1950-60年代美國愛荷華州立大學(Iowa State University)及普查局(Census Bureau)便已運用雙底冊調查(dual frame surveys)的設計來因應調查名單不完整的問題(Hartley 1962)。不過，當時的雙底冊調查僅是偶爾運用，並沒有成為發燒話題。直至2008年，美國國民健康調查(National Health Interview Survey, NHIS)發現近二成的家庭(20.2%)僅有手機而沒有市話(Blumberg and Luke 2009)，而2年後的調查更顯示，僅持有手機的家庭又增加了一成(31.6%)(Blumberg and Luke 2011)。此時，學界開始憂心以單一不完整的抽樣底冊進行調查，將造成估計上的偏差(bias)，故美國民意研究學會(The American Association for Public Opinion Research, AAPOR)於2010年發表的研究報告裡，即論及雙底冊電話調查的設計與合併等相關議題(The American Association for Public Opinion Research 2010)。至此，以雙底冊調查來取代單一底冊已然成為學界的共識，相關的討論與研究也如雨後春筍般地大量湧現。惟關於手機樣本如何加權，尤其當需要與市話樣本結合時，則如同美國民意研究學會所言，在當時並沒有共識存在(The American Association for Public Opinion Research 2010)。

在這波雙底冊調查的浪潮裡，臺灣算是比較後知後覺。此應是因緣於臺灣的市話用戶數一直居高不下，尤其在2005年更是達到史上最高峰，約1,362萬（如圖1所示）。惟隨著手機的持有日益普及，<sup>1</sup>近10餘年來市話用戶數開始逐步下滑；迄至2020年，用戶數共短少了287萬，平均每年約減少19萬用戶。再加上，近年來市話調查年輕樣本接觸率越來越低，而老年樣本日益偏高，市話調查的涵蓋率問題才漸次浮現而受人關注。而在市話用戶數持續下降及可觸及的年輕樣本日益偏低的態勢下，以市話接觸不到的「唯手機族」（家裡沒有市話而僅使用手機者）其佔比多寡，及其在人口結構，甚或在認知、行為

---

<sup>1</sup> 臺灣傳播資料庫2020年全國性面訪調查結果顯示，約有93.4%的民眾持有手機（張卿卿與陶振超 2021）。

或態度上是否與市話可接觸者有所差異，則對市話調查資料推估的精確性具有關鍵性的影響。在2015年之前，即有數位學者分別利用不同的面訪調查資料（2009年國民健康調查面訪案、2013年臺灣選舉與民主化面訪調查）來分析唯手機族的佔比、認知及行為差異（洪永泰等 2014；許勝懋 2015）。前述的研究發現，雖唯手機族在人口結構、政治態度或健康行為（如吃檳榔或吸煙率）上與非唯手機族有所差異，但因人數佔比低於一成（分別為9.3%、6.2%），故還不至於造成嚴重的推估偏差問題。

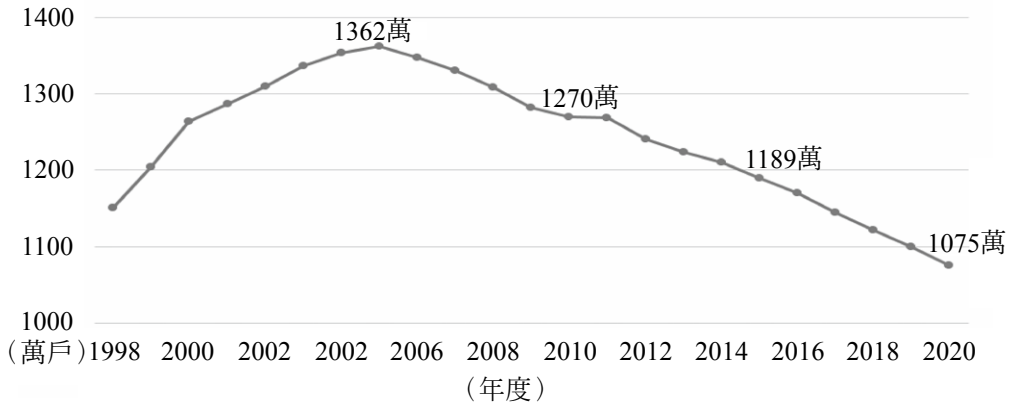


圖1 市話電話用戶數

資料來源：中華民國統計資訊網（2022）

然而，後續陳鴻嘉（2019a）於彙整中央研究院學術調查研究資料庫所釋出的調查計畫時，發現絕大多數調查所估計出的唯手機族人口佔比，已經是落在20-25%之間；少數幾個計畫的調查結果則是更高，介於26-30%，特別是2016年-2018年的調查。這代表臺灣的唯手機族不僅逐年在增加，且已經到了不容忽視的地步。尤其，在2019年的民進黨總統候選人初選民調時，雙底冊電話調查的討論更是達到了高峰，爭議點在於電話民調除了採用傳統的市話調查之外，究竟要不要納入手機調查？及雙底冊樣本要如何配置？而初選民調公布後，針對民進黨直接採用市話與手機各半的合併方式也引發學者質疑（洪永泰 2019），並認為應採用事後分層估計的母體結構比例來進行加權合併。

此時，縱然臺灣各界對於雙底冊的資料合併方式仍有爭議，但電話調查採用雙底冊的方式來進行，儼然形成了共識。而綜觀近年來臺灣學界在雙底冊電話調查的發展，雖有多位學者開始討論此類資料的合併議題，不過，絕大多數是應用前述所提及的事後分層組合估計法（洪永泰 2021；洪永泰、俞振華與高世垣 2017；張鏡文、黃東益與洪永泰 2017；蔡奇霖 2020）。

有鑑於過去10多年來，不少國外學者曾探討過多種雙底冊資料的合併途徑，而當這些途徑運用於臺灣調查時，將會呈現何種樣貌呢？為揭開這層面紗，本文旨在以貼近現實的角度及採用多元的方式，來探討當下臺灣雙底冊電話調查的合併議題。其中，本文將以臺灣學界甚少觸及的、以變異數(variance)與偏差的視角來評比各類型單底冊調查及雙底冊調查的合併組合，在各式加權策略下的加權效果。直言之，本文將從設計效果(design effect, deff)與加權損耗(weighting loss)等面向來比較各種合併組合及加權策略下的變異數差異；此外，也將利用可獲取之母體統計資料來比較偏差，嘗試理出適用於一般性調查的中庸處理之道。最後，本文也將進一步探討非等比例樣本配置的雙底冊電話調查，在實務上的可能性結果，其中，將分別以各單底冊為完整代表性的全樣本（抽樣誤差在 $\pm 3\%$ 以內），再搭配另一底冊的部分比例樣本。

## 貳、雙底冊調查之發展脈絡

Hartley (1962)是最先主張採用多底冊的估計即可達到如同以單底冊般推估精確的先驅者。他認為當一個底冊完整但抽樣成本高，而另一底冊不完整但抽樣成本卻很低時，就適合使用雙底冊的機制來互補，達到雙降的效用；亦即，除了可降低抽樣成本，還可減少樣本涵蓋率不足的問題。甚至，當所有的底冊都不完整、或進行少數人口研究（例如罕見疾病、流浪漢等）、或具有難以接觸的問題時（例如沒有樣本底冊），就相當適用多底冊調查。不過，相較於單底冊調查，採用雙底冊調查不論在研究設計或資料處理層面，都將使前述作業產生相當程度的複雜度，包括抽樣設計、加權處理、甚至是資料分析等。以下，茲針對此三個面向分別加以說明。

## 一、抽樣設計

電話調查所採用的雙底冊調查，最常見的組合是結合隨機的市話樣本與手機樣本(Random Digit Dialing sample)，亦即以市話調查為主體，同時採用手機調查來補足市話調查接觸不到的樣本。然而，當施行雙底冊電話調查時，究竟是要採取重疊設計(overlapping design)或篩選設計(screening design)？重疊設計係指執行雙底冊調查時，被抽中者全都訪問；而篩選設計則是指在執行另一底冊調查時，將重複者篩選掉不訪，例如進行市話調查時採用全樣本，而在手機調查時做篩選，僅訪問唯手機族。Kennedy (2007)曾提出要採用篩選設計的話，考量點有三個：1.估計值的變異數大小；2.是否具有無反應誤差(nonresponse error)或測量誤差(measurement error)；3.比較雙底冊調查的成本，包括篩選設計相較於重疊設計的成本差異。

首先，以變異數大小來說，最可能的影響來源是加權的處理。Kennedy認為這問題應該不存在，因為尚未看到採用篩選設計而產生較大的變異數狀況出現。惟Lohr與Brick (2014)比較了重疊設計與篩選設計在估計值的變異數表現，結果顯示在成本相同的條件下，以篩選設計而得的變異數小於重疊設計。

其次，就調查誤差來看，Brick等人(2007)認為在不同的電話使用狀態下會產生不同的回答機制，進而造成測量誤差的問題。尤其，手機調查的受訪者在回答問卷題目時，因為可隨意移動或改變場域而容易影響回答的專注力；再加上，在公共場所回答問題時，恐會因隱私問題而在回答時有所保留。Kennedy與Evertt(2007)則以實驗設計來探討兩用族在回答認知捷徑(cognitive shortcuts)類型的題目時是否具有差異，比較後發現其間並沒有差異存在。

最後，就調查成本來看，過往礙於手機的電話資費遠高於市話，以及手機調查的合作率較低，學者普遍認為雙底冊市話與手機調查的樣本配置應考量成本問題（Hu et al. 2011; Lohr and Brick 2014; The American Association for Public Opinion Research 2010; 陳鴻嘉 2019b）。然而，近十年來隨著臺灣寬頻網路的快速發展而帶動了低價網路電話的興起，再加上手機調查的合作率也不必然低於市話調查的態勢下，以重疊設計來說，當市話調查需採用戶中抽樣時，調查成本僅略高於手機調查一點，故調查成本可不再納入雙底冊樣本配置的考量範圍。至於篩選設計，Wolter等人(2015)指出在採取簡單隨機抽樣(Simple Ran-

dom Sampling, SRS)的條件下，不論是以成本固定為比較基礎，抑或是以固定變異數大小為比較基礎，篩選設計的成本均低於重疊設計，因為篩選的成本小於訪問成本；而當篩選成本高於訪問成本時，理應採用重疊途徑，例如，當唯手機族的人口佔比較低，而調查人力成本與手機費率都相當高時，因為需要耗費更多調查時間及撥打更多的手機號碼才能接觸到這群人，相對也就需要付出更多的篩選成本。惟在臺灣尚未看到採行篩選設計的實證分析結果，有待進一步驗證。

## 二、加權處理

當抽樣調查並非採取等機率抽樣、樣本的人口特徵並不符合母體結構、抑或感興趣的研究變項估計值與母體參數具有差異時，為產生不偏估計(unbiased estimates)，通常會採取加權措施用以補償抽樣設計及無反應(nonresponse)所可能產生的偏差。換言之，加權的目的，一方面是為了反應抽樣設計的不等中選機率問題(adjustment for selection probability, design weights)，另一方面則是以輔助變項來改善估計的有效性，裨益樣本能精確地反應感興趣的母體特性(adjustment for population distributions, calibration weights)。而當採行雙底冊電話調查的二群樣本都具有兩用族（同時具有市話及手機），在合併資料時則需因應抽樣底冊重疊所造成的多重中選機會問題(multiplicity)，而需在權數上進行調整(adjustment for frame overlap, composite weights)，否則將因重疊處的樣本被重複計算而造成估計上的偏差。當然，加權的結果必然是會增加誤差(error)，因為當重視偏差時，誤差就容易變高；相反地，當重視誤差時，偏差就容易變高。Little與Vartivarian (2005)曾指出，想要同時降低估計的偏差與誤差是可能的，只要加權調整時所使用的變項具有可預測調查結果及回答傾向的特性。惟此種變項不容易獲得，故加權是在偏差與誤差之間尋求一個平衡點。

## 三、資料分析

國外對於雙底冊電話調查的研究，絕大部分是集中在探討資料估計的程序，不論是想要降低誤差或偏差，而主要目的都是希冀儘量在此二者之間取得

平衡，並提高資料估計的精確度。相關的研究途徑有以下幾種，重點都是擺在如何對兩用族進行更適切的加權調整處理。綜述如下：

### (一) 雙底冊組合估計

Hartley(1962)最先倡議在雙底冊的兩用族以組合因子(composite factor)合併，用以降低變異數。此法的關鍵要點為二套樣本兩用族的組合設計權數(composite design weights)相加後應為1，公式如下：

$$y_{\text{雙底冊組合}} = y_{\text{唯市話}} + y_{\text{唯手機}} + \theta y_{\text{(市話)兩用族}} + (1 - \theta) y_{\text{(手機)兩用族}}$$

$$(0 \leq \theta \leq 1)$$

然而，Hartley理型 $\theta$ 的計算需仰賴感興趣變項的變異數與共變異數，爲了降低變異數，當感興趣的對象有所改變時就要重新計算 $\theta$ ，以實務上而言，不僅是相當不便利且具有內在不一致性的問題。Fuller與Burmeister(1972)延伸Hartley的概念，將兩用族以最大概似估計值(Maximum Likelihood Estimator)來改善Hartley的效能，惟該雙底冊調查是以採用簡單隨機抽樣爲前提，且 $y$ 值並非是線性函數(linear functions)而只是趨近線性，不適合運用於實務上所需採行的複雜抽樣調查分析。Skinner與Rao(1996)則以擬最大概似估計值(Pseudo Maximum Likelihood Estimator)來修正前者，如此一來，所有變項採用相同的加權權數，不會有內在不一致的狀況，並且適用於複雜抽樣設計的資料於統計軟體內進行變異數估計。

其後，Brick等人(2006)將 $\theta$ 以0.5帶入，稱之爲簡單組合權數(simple composite weights，以下簡稱折半合併)，此法在計算上雖然相當簡易，也可運用於複雜抽樣調查的分析，但並非是一個最有效益的選擇，尤其當雙底冊的二套樣本數比例差異大時。爲解決前述問題，Frankel等人(2007)及The American Association for Public Opinion Research(2010)提出以二套樣本兩用族之間的相對有效樣本數(the effective sample size)來估計最適 $\theta$ 值(以下簡稱有效樣本數合併)。此法在計算上並不繁複，且相較於簡單組合權數可以動態反應調查結果；亦即，當雙底冊的樣本數並非相近時，可以計算出更適切的合併權數。此外，Brick等人(2011)也從非抽樣誤差的角度來探討雙底冊調查，並發現此類調



查存有無反應誤差與測量誤差；而為能有效降低無反應誤差，Brick等人主張以外部調查資料（全國性面訪調查）的電話使用習性比例來估算 $\theta$ 。但因這項估計需具有市話樣本與手機樣本裡主要市話使用者的比例，調查實務上較不易獲取此類資訊。

### （二）單底冊途徑(Simple Frame Approach)

此研究途徑將雙底冊調查的二套樣本，視為來自單一樣本底冊的三個抽樣分層(Bankier 1986; Kalton and Anderson 1986; Meccati 2007)，不需要利用額外的母體統計值做為輔助資料。以電話調查為例，三個抽樣分層分別為：唯市話族、兩用族及唯手機族。而為了補償兩用族的多重樣本來源之中選機會問題，此研究途徑採用合併唯市話族與唯手機族的中選率來對兩用族做加權調整，裨益於有效降低估計值的偏差。惟此法比較適用於採用簡單隨機抽樣的雙底冊調查，當單一底冊的電話調查採取複雜抽樣設計，融合了抽樣分層、戶中抽樣等設計而致每個受訪單位的中選率不一致時，就無法採用此途徑。

### （三）聯合校正估計值(Joint Calibration Estimator)

Deville與Särndal (1992)為了有效改善樣本估計的精確度與正確性，於單底冊加權時以一般迴歸估計的方式(general regression (GREG) estimation)，利用已知的輔助變項來計算出一個可極小化設計權數與校正權數之間距離的最終加權權數；而此極小化距離的產生，是來自於限制「以輔助變項校正加權後的樣本總和」需符合「已知母體總和」而得。其後，Elkasabi、Heeringa與Lepkowski (2015)及Rai、Tikkiwal與Alka (2020)則進一步將Deville與Särndal此種極小化距離的概念沿用於雙底冊加權，名之為聯合校正估計值。其優點在於不需使用雙底冊加權所慣常使用的電話使用習性佔比資訊來處理加權，也就不會因為可能產生的估計偏誤而影響到整體的加權表現。但運用此法要產生趨近不偏估計效果的話，則非常仰賴所使用的校正輔助變項與研究變項之間具有高度的相關。此種要求就反而成為運用此法的關鍵門檻，畢竟與研究變項高度相關的校正輔助變項，在實務上並不容易取得。

近年來，洪永泰、俞振華與高世垣(2017)採用的事後分層組合估計來合併雙底冊電話調查。事後分層組合估計並未將各單底冊的不等機率權數納入加權考量，而僅於完成雙底冊調查後，利用性別、年齡、教育程度的交叉組合對雙

底冊樣本進行事後分層，估計出各分層於電話使用習性的比例。其後，再以此估計值做為雙底冊樣本合併的事後分層加權參數。因此，此種加權方式的加權效果完全仰賴雙底冊調查原生資料的調查品質。

從以上的文獻可得知，雙底冊調查的兩用族因具有多重來源的中選率問題而需進行資料合併處理時，採用不同的加權方式會對樣本估計值有不同影響，不論是反應在偏差或誤差的增減上；再者，尚需考量(1)加權輔助變項的可獲性，(2)加權權數是否會隨感興趣研究變項的不同而有所改變，衍生估計值不一致的問題，及(3)是否可運用於統計軟體執行複雜抽樣資料的變異數估計等。以下，將以臺灣雙底冊電話調查的實務調查結果，針對各類型的單底冊、雙底冊合併組合於各種加權處理後的效果進行比較分析。

### 參、雙底冊調查資料之分析與評比

本文分析所使用的雙底冊電話調查資料是由中央研究院人社中心下轄之調查研究專題中心於2020年底所執行的幸福溫度計調查，在市話調查與手機調查的完成數分別為1235及1255，可視為等比例樣本配置的雙底冊電話調查（張卿卿 2020）。此調查的抽樣設計、執行與變項操作說明，請參見附錄一。

本次調查所得的電話使用比率如表1所示。其中，市話調查的兩用族比手機調查多了16.8%，亦即代表唯手機族的人數多於唯市話族。若以這二個調查帶入Frankel等人(2007)的公式來推估臺灣地區18歲以上一般民衆的電話使用習性，<sup>2</sup>可估計出唯市話族(P1)的佔比為5.1%、唯手機族(P3)為22.2%，而兩用族

$$2 \quad P2 = \frac{1}{\frac{1}{\text{市話調查兩用族比率}} + \frac{1}{\text{手機調查兩用族比率}} - 1}$$

$$P1 = P2 \times \left( \frac{1}{\text{市話調查兩用族比率}} - 1 \right)$$

$$P3 = P2 \times \left( \frac{1}{\text{手機調查兩用族比率}} - 1 \right)。$$

此公式的估計結果與張鐙文、黃東益與洪永泰（2017）的三種電話使用族群估計結果相同。

(P2)為72.7%，如圖2所示。而同年度臺灣傳播資料庫同時期執行的面訪調查結果所示，唯市話族的佔比為5.9%、唯手機族為26.2%、兩用族為67.5%、都不使用者為0.4%。不同模式調查結果的比例，可說是相當接近。

表1 雙底冊調查電話使用習性

	市話調查	手機調查
唯市話(P1)	81 (6.6)	
兩用族(P2)	1154 (93.4)	960 (76.6)
唯手機(P3)		294 (23.4)
總計	1235	1254

資料來源：張卿卿（2020）。

說明：括號內為列百分比（%）。

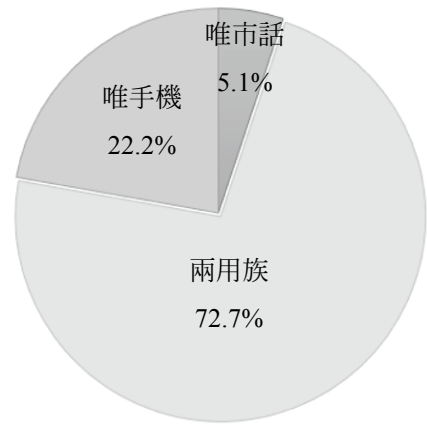


圖2 電話使用習性推估比率

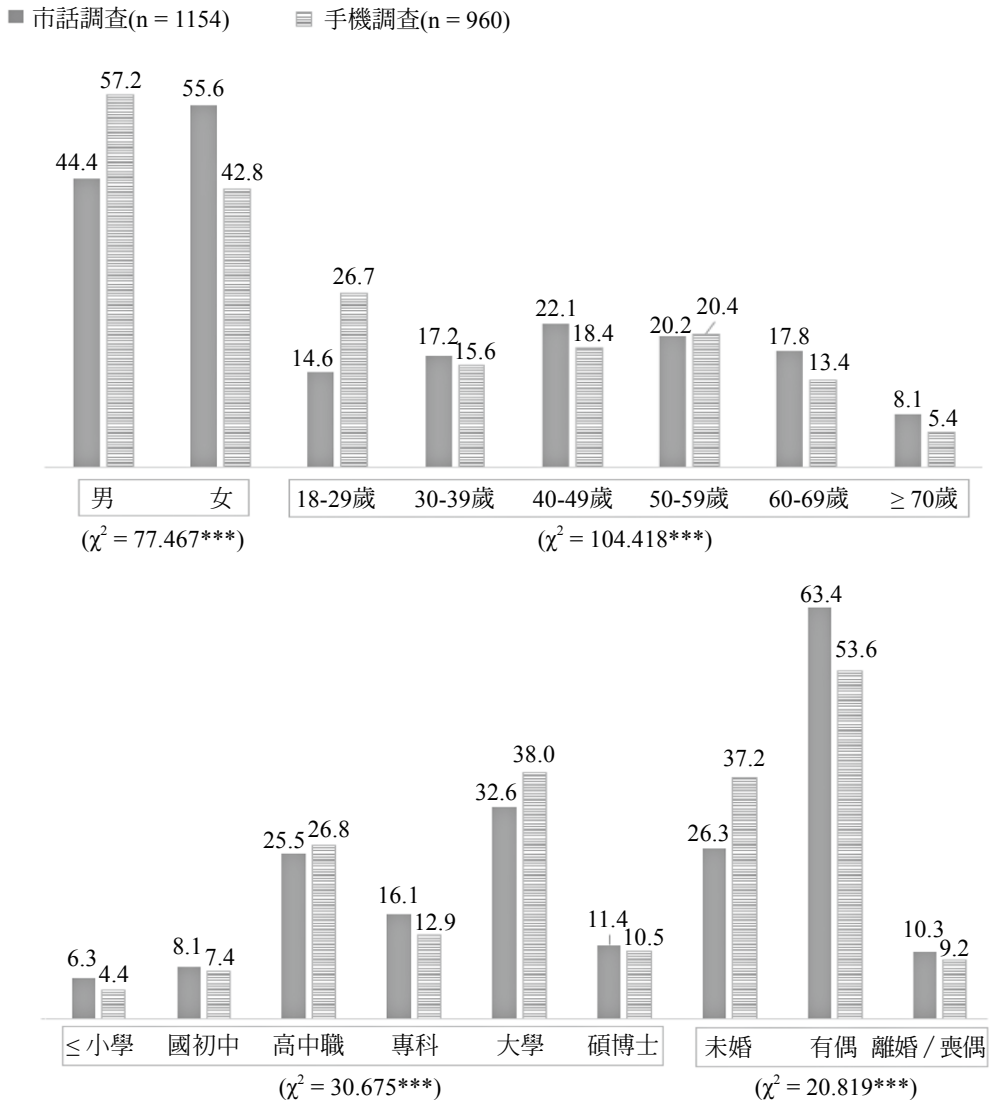
資料來源：張卿卿（2020）

就學理上而言，當雙底冊調查的抽樣底冊(sampling frame)涵蓋範圍相近時，各樣本的兩用族理當要具有相同的群像，不論是在母體人口結構，抑或是認知、行為與態度上；換言之，在雙底冊調查訪問各方的兩用族是不必要的，因為每一方的兩用族應該要能代表另一方才是。在臺灣，市話與手機市場的涵蓋率又是如何呢？以調查統計結果來看，依據2020年家庭收支調查結果所示，一般家戶持有電話機的比例為85.86%、手機的比例為96.20%；而臺灣傳播調查資料庫於2020年底針對18歲以上一般民衆所執行的面訪調查結果顯示：家裡有市話者為73%，而個人持有手機者為93.3%。最後，再以政府統計數據來看，國家通訊傳播委員會(NCC)於2020年底的電信統計，市話用戶降至1,075萬（含商用），而手機的用戶數高達2,900多萬。<sup>3</sup> 從以上數據顯示，臺灣電信市場在市話與手機的涵蓋率都不低，但民衆的手機持有率還是高於家用

<sup>3</sup> 資料取自NCC網站所提供之2020年12月份行動寬頻業務用戶數統計檔。

市話，以面訪調查來看，高約二成左右。那麼，在這樣的態勢下，採用不同電信設備所調查出來的兩用族群像是否具有差異？

從圖3所呈現的市話與手機兩用族的比較分析結果可得知，不論是在性別、年齡、教育程度、婚姻狀況、就業狀況或地區分佈等均具有統計上的顯著



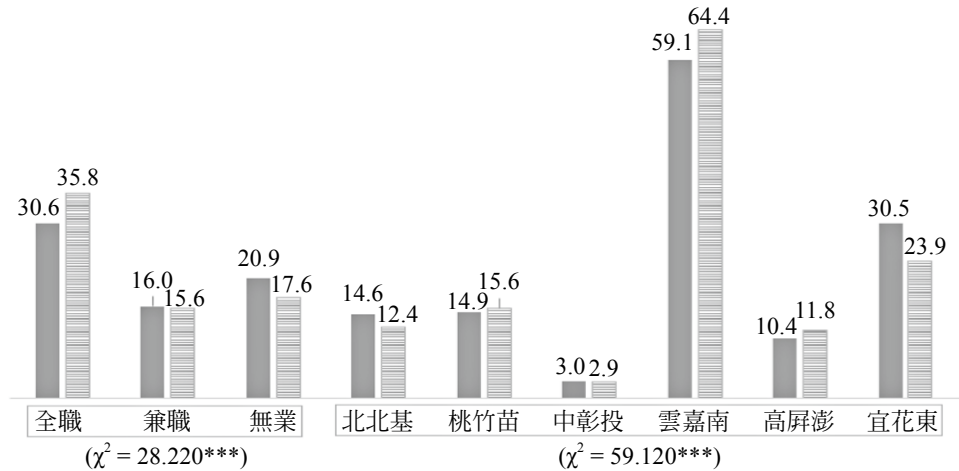


圖3 雙底冊兩用族人口特徵比較與卡方適合度檢定結果

資料來源：張卿卿（2020）。

說明：1. \*\*\*  $p < 0.001$

2. 婚姻狀況：內政部戶政司民國108年底人口數\_按年齡及婚姻狀況。因應本文分析的調查民眾年齡最低為18歲，故刪除內政部資料內15-17歲之統計值，並將調查資料內「同居」項目納入「有偶」項目。
3. 就業狀況：108年勞動情勢統計要覽：民國108年5月行政院主計總處人力資源調查。統計結果顯示15-19歲就業人口佔總就業人口不到1%，其中，15-17歲的比例更低，故忽略不納入計算。

差異 ( $p < 0.001$ )。綜觀而言，手機調查的兩用族相較於市話調查的特性是：男性較多（多12.8%）、年紀較輕（29歲以下多12.1%、60歲以上少7.1%）、教育程度略高（專科少3.2%、大學多5.4%）、未婚比例較多（多10.9%）、全職工作者較多而無業者較少（全職多5.3%、無業少6.6%）；最後，以地區來看，北部人數多而中部人數較少（北部多5.2%、中部少5.5%）。換言之，手機調查的兩用族不同於市話調查，除了男性比例高之外，還具有年輕化、高教育、高都會的樣貌。此與張鏡文等人（2017）的研究結果不同，因其發現兩用族在人口特徵上除了性別與年齡不同之外，並沒有明顯的差異。

前述的分析結果顯示，臺灣雙底冊電話調查之兩用族具有不一樣的人口

結構與地區結構，而此項結果相當有可能造成這二群人在社會認知與態度等面向上，也具有統計上顯著差異。不過，這其實已經不是什麼新鮮事了，因國內外的調查早已指出前述的研究結果（Baffour et al. 2016; Boyle, Lewis, and Tefft 2010; Kennedy 2007; 張鏡文、黃東益與洪永泰 2017）。造就兩用者群像不同的主因，不外乎是涵蓋誤差與無反應誤差所致，而此二者的影響究竟孰重孰輕，則有待進一步的實證調查來驗證。至於在測量誤差，Kennedy與Evertt (2007)則指出至少在認知類型的題目上是沒有顯著差異的。不過，這不是本文關注的焦點，但此結果反倒是指出了一個關鍵要點：即臺灣的雙底冊電話調查，不能只考慮納入唯手機族的涵蓋性問題了，亦即，不能單以「市話調查單底冊+唯手機族」等此類的篩選設計來處理，而需考慮採用重疊設計了。

既然雙底冊電話調查內兩用族的群像樣貌不同，那麼採用未加權的單底冊、雙底冊-重疊設計、雙底冊-篩選設計（含「手機+唯市話」與「市話+唯手機」二種）等狀況又會是如何呢？表2是利用六項從各類統計與政府調查所獲得的母體參數(population parameter)，包括性別、年齡、教育程度、地區、婚姻狀況、就業狀況，來觀察各單底冊與各類合併組合的表現。結果顯示，不論是執行樣本代表性檢定，或檢測母體參數是否落在該估計值的95%信賴區間（納入設計效果來計算），<sup>4</sup>各單底冊與各類合併組合的表現均不理想，惟雙底冊的合併組合終究還是略勝於單底冊一點。其中，「市話+唯手機」在性別與地區二變項有過關，算是比較優一點；其次則為「市話+手機」，在性別變項過關；再次則為「市話」，在地區變項過關。換言之，在尚未加權前，以市話調查為完整代表性的單底冊為主體來合併另一底冊，可在人口特徵變項獲致相對比較符合母體參數的估計值。表現較好的變項包括：(1)性別：其在「市話+手機」或「市話+唯手機」的優良表現，是來自於雙底冊的互補效果所致，因市話調查完成樣本裡以女性居多，而手機調查以男性居多；(2)地區：市話調查所獲得的六大地區估計值與母體參數並無統計差異存在，這樣的結果應是因緣於所採用的分層抽樣設計（參見附錄一），此設計除了為能降低空話

<sup>4</sup> 納入設計效果之樣本估計值95%信賴區間計算方式： $\hat{p} \pm 1.96 \sqrt{\frac{deff \times \hat{p}(1-p)}{n}}$

率而納入千門用戶碼使用率與撥號結果來做為分層因子之外，抽樣時也同時考量地區分佈所致。至於手機調查則因可據以做有效分層的變項過少，且絕大多數手機號碼的持有者並非單純做為個人聯絡電話，因此，完成樣本的地區分佈略遜於市話調查。

比較特別的是年齡變項，市話與手機的單底冊調查結果都顯示，在中間年齡層30-59歲的估計均表現不錯。至於年輕與年長年齡層則如預期般地呈現比例相反的態勢，且合併後的效果也不如性別變項明顯：即前述年輕與年長年齡層的母體參數仍未落入樣本估計值95%信賴區間內。而在教育程度、婚姻及就業狀況的表現則屬最差，尤其是就業狀況，在兼職與無業人口佔比差異相當大，相當有可能是政府調查無法體現社會實際的兼職人口，存在不小的黑數。

表2 未加權之比較分析

項目 % [樣本數]	市話 [n=1235]	手機 [n=1255]	唯市 話族 [n=81]	唯手 機族 [n=294]	市話 + 手機 [n=2490]	市話 + 唯手機 [n=1529]	手機 + 唯市話 [n=1336]	母體
性別								
男	43.4	58.5	29.6	62.6	51.0	47.1	56.7	49.1
女	56.6	41.5	70.4	37.4	49.0	52.9	43.3	50.9
年齡								
18-29	13.6	26.3	0.0	25.3	20.0	15.9	24.7	18.3
30-39	16.1	18.7	0.0	29.0	17.4	18.6	17.6	17.7
40-49	21.0	19.0	6.3	20.8	20.0	21.0	18.2	18.9
50-59	19.6	19.1	10.0	15.0	19.4	18.7	18.6	18.2
60-69	18.7	12.0	31.3	7.2	15.3	16.5	13.2	15.5
≥70	11.0	4.8	52.5	2.7	7.8	9.4	7.6	11.4
教育程度								
≤小學	9.7	4.1	56.8	3.4	6.9	8.5	7.3	12.2
國初中	8.8	7.9	19.8	9.2	8.4	8.9	8.6	11.7
高中職	24.7	27.3	13.6	29.0	26.0	25.6	26.4	27.7

項目 % [樣本數]	市話 [n=1235]	手機 [n=1255]	唯市 話族 [n=81]	唯手 機族 [n=294]	市話 + 手機 [n=2490]	市話 + 唯手機 [n=1529]	手機 + 唯市話 [n=1336]	母體
專科	15.4	12.0	4.9	9.2	13.7	14.2	11.6	11.8
大學	30.6	37.7	2.5	36.9	34.2	31.8	35.6	28.9
碩博士	10.8	10.9	2.5	12.3	10.9	11.1	10.4	7.8
地區								
北北基	29.6	34.8	14.8	31.4	32.2	29.9	33.6	29.9
桃竹苗	16.0	17.6	16.0	24.1	16.8	17.6	17.5	15.7
中彰投	20.8	17.5	19.8	17.2	19.1	20.1	17.6	19.2
雲嘉南	15.3	12.2	25.9	11.4	13.8	14.6	13.1	14.3
高屏澎 (金馬)	15.0	14.8	16.0	12.1	14.9	14.4	14.9	16.6
宜花東	3.3	3.1	7.4	3.8	3.2	3.4	3.4	4.3
婚姻狀況								
未婚	25.2	37.4	8.8	37.9	31.3	27.6	35.6	32.0
有偶	63.3	51.2	62.5	43.0	57.2	59.4	51.8	51.8
離婚/ 喪偶	11.5	11.5	28.7	19.1	11.5	13.0	12.5	16.3
就業狀況								
全職	56.0	65.8	12.3	70.7	61.0	58.9	62.6	55.7
兼職	10.0	11.7	4.9	11.2	10.9	10.3	11.3	2.1
無業	33.9	22.5	82.7	18.0	28.2	30.9	26.1	42.2

資料來源：張卿卿（2020）。

說明：1. 灰底標示處表示母體參數落在「納入設計效果」樣本估計值之95%信賴區間估計內。

2. 表內各項目與母體參數進行卡方適合度檢定的結果，除了「虛線框」標示者的 $p$ 值為沒有統計上的顯著差異（ $p$ 值大於0.05）、「灰線框」標示者的 $p$ 值小於0.05、「黑線框」標示者的 $p$ 值小於0.01之外，其餘項目的 $p$ 值均小於0.001。



## 肆、各類型雙底冊電話調查合併組合加權效果之比較分析

如前述所言，國外不少學者在探討雙底冊電話調查的合併組合時，是建立在採用簡單隨機抽樣的基礎上，即每位受訪樣本的中選率相等而具有相同的加權權數。但實務上卻無法這麼的理想，因執行實務調查時，爲了降低調查經費或提升資料推估的精確度，通常會運用分層、分群等多階段的複雜抽樣設計，再加上戶中抽樣機制、家戶具有多個市話號碼或個人持有多個手機號碼等多重中選等因素，必須因應每位受訪樣本的不同中選率而以加權的方式來處理。

衆所周知，加權時除了要把每個受訪樣本被選取時的不等機率還原回一致的等機率基礎之外，往往還需要處理無反應或涵蓋率等各類調查誤差對該調查資料的影響。爲解決此類問題，單純僅用一般的母體人口特徵變項（例如年齡、性別、教育程度等）來進行校正(calibration)，能夠把調查資料的內涵轉正之機率並不高。因爲這樣的做法僅是使加權後的樣本與母體的人口分佈趨近一致，但並不保證內在的差異如認知、行爲或態度等，可就此撫平。比較妥切的方式是在加權校正時，運用與感興趣的研究變項具有高度相關的輔助資料來處理，用以降低變異數及無反應偏差(Skinner 1991; Skinner and Rao 1996)。然而，在同一調查資料裡，每個研究者感興趣的變項是不同的，以實務上來說，此法不僅是相當不便利且會產生估計值不一致的問題，更遑論實務上並不易獲取這樣的數據資訊。而有鑑於調查採用簡單隨機抽樣設計及利用與感興趣變項高度相關的輔助資料來加權校正等，均過於理想性，故本文將以臺灣電話調查常態的運作模式，來分析與評比各類型單底冊與雙底冊於不同合併組合及加權處理策略的效果，嘗試導入以變異數衍生而得的設計效果、加權耗損等做爲差異評估的比較工具，同時也將比較樣本與母體之間的偏差問題。

### 一、變異數、設計效果與加權耗損

對資料分析者而言，變異數是一個相當基礎且不容輕忽的關鍵要素。這是因爲絕大多數的統計估計或假設檢定，均是以變異數爲計算基礎；其中，變異數的大小，相當程度影響到資料分析時樣本估計的精確度及檢定顯著的容易

度。誠如前述所言，爲了要兼顧資料推估的精確度、樽節調查成本及簡化調查流程等因素，實務調查鮮少會採用簡單隨機抽樣的設計，而是以分層、分群或多階段的複雜抽樣調查來取而代。而此種抽樣設計下的變異數，理當實質反應出其抽樣設計特性(design feature)的影響，包括由分層、分群或多階段抽樣等搭配起來的設計效果及加權權數等各項的總體效應，因之，絕對與簡單隨機抽樣而得者截然不同。但由於複雜抽樣設計的變異數在計算上相對複雜，爲了簡化計算，通常會採近似值(approximation)的方式來估計變異數，相關的討論可參考侯佩君（2011）一文。

爲了評估複雜抽樣調查的資料品質，Kish (1965)建構出「設計效果」此一測量指標。即以簡單隨機抽樣設計做爲比較基準，來計算二者在相同樣本數條件下的變異數比值，據此來衡量複雜抽樣調查在結合分層、分群及多階段抽樣設計及加權等總體影響的淨效應(net effect)。而在計算設計效果的同時，也衍生出有效樣本數的測量概念，用以估計採用的複雜抽樣調查所完成的樣本數，其具有的資料精確度是等同於用簡單隨機抽樣調查完成多少個樣本數。<sup>5</sup> 簡單來說，假設複雜抽樣調查的樣本數爲1,000個，設計效果爲1.5，代表此調查設計的資料精確度等同於以簡單隨機抽樣調查完成667個樣本的精確度（ $1000/1.5=666.6$ ）。由此可知，當設計效果 $>1$ ，代表複雜抽樣調查的估計精確度低於簡單隨機抽樣；設計效果 $<1$ ，則代表估計的精確度高於簡單隨機抽樣。而實務上，在計算設計效果時，大多數的統計學家會以其近似值的簡化公式來計算，即採用「 $1+$ 加權損耗（ $L_{weighting}$ ）」(Heeringa, West, and Berglund 2017)。計算公式如下。

$$\text{設計效果 (deff)} = \frac{\text{var}_{complex}(\bar{y})}{\text{var}_{SRS}(\bar{y})} \approx 1 + L_{weighting}(\bar{y})$$

[ $\bar{y}$  = 樣本平均數， $\text{var}_{complex}(\bar{y})$  = 複雜抽樣調查之變異數， $\text{var}_{SRS}(\bar{y})$  = 簡單隨機抽樣之變異數]

<sup>5</sup> 有效樣本數( $n_{eff}$ ) =  $\frac{n_{unweighted}}{deff}$  ( $n_{unweighted}$  = 未加權樣本數)。

加權損耗係指加權後估計值的變異數相較於採用簡單隨機抽樣時的增減率（計算公式如下）。例如，若計算而得的加權損耗為0.36，即代表因加權而產生的估計值其變異數比採用簡單隨機抽樣時高出36%。而若加權權數過大，則容易導致加權權數比較離散，進而造成估計不穩。那麼，究竟加權耗損在多少的範圍內是可接受的？Valliant等(2018)指出，當設計效果高於1.5，亦即加權耗損高於50%時，可能要採取一些處理措施。例如，權重削減(weight trimming)作業等。

$$L_{\text{weighting}}(\bar{y}) = \left[ \frac{\sum_{i=1}^n w_i^2}{\left( \sum_{i=1}^n w_i \right)^2} \right] \cdot n - 1 = cv^2(w) = \frac{s^2(w)}{\bar{w}^2}$$

[ $n$  = 樣本數， $w_i$  = 樣本個案的加權權數， $cv(w)$  = 加權權數的變異係數， $s(w)$  = 加權權數的標準差， $\bar{w}$  = 加權權數的平均數]

因應本文旨在評比各類型單底冊與雙底冊於不同合併組合及加權處理策略的效果，比較的項目包括加權後的變異數及樣本估計值的差異。首先，為了比較加權後在變異數上的變化差異，將應用前述提及的、以變異數概念衍生的設計效果與加權損耗概念來評比；其次，將以加權後的樣本估計值與母體統計資料做比較，從偏差的角度來評估優劣。

## 二、資料分析設計

本文在評比不同合併組合的加權處理策略時，比較的項目有二種：(1)從設計效果與加權損耗等面向來比較變異數差異，評比哪些方式容易造成估計不穩的狀況；(2)利用可獲取之母體統計資料來比較偏差，據以探討在合併與加權後，哪些方式的結果比較貼近母體而有較高的估計精確度，項目包括：婚姻狀況、就業狀況、手機電信業用戶率。<sup>6</sup> 其中，加權處理的策略有二種：(1)在

<sup>6</sup> 資料來源為國家通信傳播委員會2020年12月行動寬頻業務用戶數統計。

納入不等機率權數的基礎上，以多變項反覆加權法(raking)進行加權校正；(2)採用張鑑文、黃東益與洪永泰（2017）所運之事後分層組合估計法進行加權校正。而在進行多變項反覆加權時，除了採用臺灣學界慣常使用的母體人口特徵如年齡、性別、教育程度及六大地理區人口佔比等四變項之外（如表3第I類），也將因應雙底冊調查的資料合併處理，額外納入電話使用習性變項做為差異比較（如表3第II類）。畢竟採用雙底冊調查是源自於二種電話調查可接觸到的對象不同，而這差異的存在也是起因於電話使用的習性改變所致，故在進行雙底冊資料合併時，理應於加權時反應此特性。<sup>7</sup>至於在雙底冊的資料合併處理上，除了呈現雙底冊直接合併及篩選設計的結果之外，也將以組合設計權數的方式來合併重疊設計內的兩用族，共計有二種方式：折半合併與有效樣本數合併。單底冊、雙底冊合併組合及加權評比的各類比較組合統整如下表所示。

表3 加權評比之設計項目

比較項目		加權方式	I		II	
			不等機率加權 + 校正加權 (人口特徵變項)	不等機率加權 + 校正加權 (人口特徵變項+ 電話使用習性)	事後分層 組合估計 (電話使 用習性)	
單 底 冊	市話		I			
	手機		I			
雙 底 冊	篩選設計	市話+唯手機	I	II		II
		手機+唯市話	I	II		II

<sup>7</sup> Frankel等(2007)指出，加權時所採用母體推估變項可引用外部調查的推估資料或取自原分析樣本的推估結果，但因應Brick等(2007)擔憂手機調查恐具有無反應誤差的問題，故建議採用來自外部的全國性調查推估資料比較適切。惟本文旨在比較各種雙底冊合併途徑的差異，包括張鑑文、黃東益與洪永泰（2017）之事後分層加權合併法，故此處加權所使用的電話使用習性變項即採用該文的方式：以分析的雙底冊電話調查本身所推估之電話使用習性人口佔比做為加權時的參照，此變項的類別有三種：唯市話、唯手機及兩用族。

比較項目		加權方式	I	II
			不等機率加權 + 校正加權 (人口特徵變項)	不等機率加權 + 校正加權 (人口特徵變項+ 電話使用習性)
雙 底 冊	重 疊 訪 談 設 計	市話+手機 [直接合併]	I	II
		唯市話+唯手機+兩用族 [折半合併]	I	II
		唯市話+唯手機+兩用族 [有效樣本數合併]	I	II

資料來源：本研究

上表中各項目的加權程序與方式，說明如下：

### 1. 單底冊調查設計

本文分析所使用的單底冊調查，不論市話或手機調查，因各抽樣分層採等比例的方式來配置樣本數，故每一個電話號碼具有相同的中選率（請見附錄一說明）。惟市話調查與手機調查的抽樣底冊具有不同的電話號碼數量，故彼此之間的中選率不同；再者，市話調查的不等機率來源還包括採用戶中抽樣機制，且家戶可能具有多個市話號碼，而手機調查則來自於受訪者持有多個手機號碼。故單底冊調查在加權時，將先處理不等機率加權，還原回一致的等機率基礎之後，再運用母體人口特徵資料或電話使用習性以多變項反覆加權法進行校正加權。各單底冊的不等機率加權公式如下：

$$\text{市話調查_不等機率加權公式：} w_{set}^{LL} = \frac{1}{f_{LL}} \times \frac{1}{f_{hs}} \times \frac{1}{f_{pn}}$$

[市話電話號碼的中選率 $f_{LL}$ ，戶中抽樣中選率 $f_{hs}$ ，家中市話線數 $f_{pn}$ ]

$$\text{手機調查_不等機率加權公式：} w_{sel}^{Cell} = \frac{1}{f_{Cell}} \times \frac{1}{f_{cn}}$$

[手機電話號碼的中選率 $f_{Cell}$ ，手機門號持有數 $f_{cn}$ ]

## 2. 雙底冊調查之篩選設計

先計算單底冊的不等機率加權，接著，將單底冊與另一套的唯手機族或唯市話族採用篩選合併，再進行多變項反覆加權作業。

## 3. 雙底冊調查之重疊設計

(1)直接合併：將帶有不等機率權數的雙底冊樣本直接合併後，進行多變項反覆加權作業。

(2)折半合併：將雙底冊內「兩用族」的不等機率權數額外帶入0.5的權數後，再進行多變項反覆加權作業。

(3)有效樣本數合併：由雙底冊各自計算出不等機率權數後，再計算兩用族的有效樣本數，其後，加總雙底冊兩用族的有效樣本數做為分母、各樣本的有效樣本數當分子，分別計算出各樣本的有效樣本比率。而於雙底冊合併時，在納入原始的不等機率權數下，各樣本帶入此率做為合併時的 $\theta$ 值，最後再進行多變項反覆加權作業。

## 4. 事後分層組合估計

合併雙底冊樣本後，以性別、教育程度、年齡變項交叉組合做事後分層，<sup>8</sup> 共分為18層（2\*3\*3），分別估計出各分層內唯市話族、兩用族及唯手機族之比率做為加權的母體參數後，以此估計值對三種合併組合的各分層進行權數調整。三種合併組合包括：市話+手機、市話+唯手機、手機+唯市話。

（一）第I類比較：不等機率加權+四變項校正加權

1. 變異數比較：圖4為七項加權處理後所計算出的設計效果與加權耗損。設計效果包括二種，分別為採「不等機率加權」與採「不等機率加權+校正加權」，並計算出前二者的差異（因校正加權而增加的變異率，簡稱加權變

<sup>8</sup> 人口變項的分類項目如下：(1)性別：男、女；(2)教育程度：國中以下、高中職、大學以上；(3)年齡：18-39歲、40-59歲、60歲以上。

異率)。<sup>9</sup>

總體來看，在二階段加權的設計效果比較上，以手機的加權變異率 29.55%為最大，這代表相對於其他單底冊或雙底冊模式，手機調查的人口結構與母體差異最大，因此在以母體人口特徵變項校正時需要調整的幅度最大，其次為市話調查的13.09%，而表現最好的則是「市話+唯手機」的7.97%。

若從各階段加權後的設計效果來做細部比較，首先，在不等機率加權階段，「手機」及「手機+唯市話」的設計效果最小，這是因為相對於市話調查的戶中抽樣機制，手機調查相對要還回去的不等機率權數較少；至於雙底冊合併模式的設計效果則是差不多，介於1.12~1.20之間。其次，在校正加權後的設計效果是以「手機+唯市話」的1.17為最小，代表手機調查雖最為偏離母體的人口特徵，但併入「唯市話」後，拉近了與母體的距離；其餘則介於1.24~1.36之間。

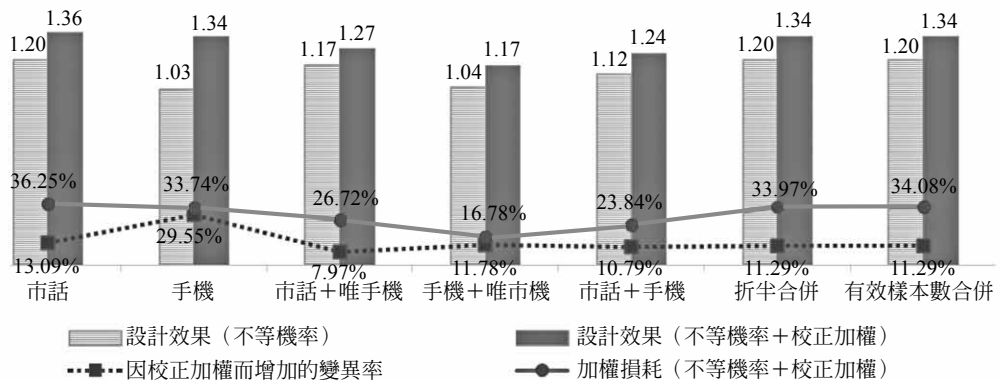


圖4 第I類四變項加權之設計效果評比

資料來源：張卿卿（2020）。

最後，再從加權損耗的面向來看，「手機+唯市話」表現最優，僅有

<sup>9</sup> 因校正加權而增加的變異率 =  $\frac{deff_{cal} - deff_{sel}}{deff_{sel}}$  (即不等機率權值設計效果( $deff_{sel}$ )與校正加權權值設計效果( $deff_{cal}$ )之差異比率)

17%，次之為「市話+手機」的24%，再次則為「市話+唯手機」的27%，至於手機、折半合併及有效樣本數合併的排名位居倒數第二，均為34%，最差則為市話的36%。不過，以上各類的加權損耗都是在可接受的範圍內，亦即，其所採用的二階段加權並未增加過多的變異率而造成估計不穩的狀況。

2. 偏差比較：此項比較是在完成校正加權後，利用可獲得的人口特徵等相關變項來比較加權後與母體參數之間的差距是否顯著，本文所觀察的變項包括婚姻狀況、就業狀況、手機電信用戶率等三項。表4所呈現的比較結果，來自於二種檢定方式：(1)以95%信賴水準（納入設計效果）來計算樣本估計值的區間估計，檢測母體參數是否落入該區間；(2)計算各變項與母體的卡方適合度檢定(Chi-Square Goodness of Fit Test)。

首先，表4結果顯示以95%信賴水準來檢定，所有模式的估計區間都僅有少數幾個有涵蓋到母體參數（請見表內灰底標示處），勉強來說，「手機」及「手機+唯市話」好一些些，但其實也僅多一個母體有落入估計區間。至於折半合併與有效樣本數合併幾乎是一致的結果，乃是因為手機與市話調查的樣本數差異很小，故二項調查的有效樣本數合併比率幾近一半一半（市話兩用族為50.77%、手機兩用族為49.23%）。其次，以卡方適合度檢定來檢測此三項變項的結果也顯示不論是單底冊調查或各式的雙底冊合併，樣本估計值與母體參數均具有統計上的顯著差異。

表4 第1類四變項加權之比較分析

項目 %	市話 [樣本數]	手機 [1255]	市話 +唯手機 [1529]	手機 +唯市話 [1336]	市話 +手機 [2490]	折半 合併 [1432]	有效樣本 數合併 [1433]	母體
婚姻狀況								
未婚	26.3	29.5	26.6	29.6	28.0	28.1	28.1	32.0
有偶	63.8	55.2	61.0	56.0	59.6	58.7	58.7	51.8
離婚/喪偶	9.9	15.3	12.5	14.4	12.4	13.2	13.2	16.3
就業狀況								
全職	57.7	59.1	57.6	58.6	58.2	57.9	57.9	55.7
兼職	10.4	11.4	10.5	10.7	10.7	10.6	10.6	2.1



項目 % [樣本數]	市話 [1235]	手機 [1255]	市話 +唯手機 [1529]	手機 +唯市話 [1336]	市話 +手機 [2490]	折半 合併 [1432]	有效樣本 數合併 [1433]	母體
無業	31.9	29.5	31.9	30.7	31.0	31.5	31.5	42.2
手機電信用戶								
中華	47.6	48.7	46.7	47.2	47.5	46.9	46.9	35.9
台灣大哥大	23.0	22.2	23.0	23.0	22.8	23.0	23.0	24.3
遠傳	20.8	20.8	21.7	21.3	21.2	21.6	21.6	24.1
台灣之星	5.3	5.6	5.4	5.8	5.5	5.6	5.6	8.7
亞太	3.3	2.7	3.1	2.6	3.0	2.9	2.9	6.9

資料來源：張卿卿（2020）。

說明：1. 灰底標示處表示母體參數落在「納入設計效果」樣本估計值之95%信賴區間估計內。

2. 表內各項目與母體參數進行卡方適合度檢定的結果，除了灰線框標示者的 $p$ 值小於0.05、黑線框標示者的 $p$ 值小於0.01之外，其餘項目的 $p$ 值均小於0.001。

## （二）第II類比較：不等機率加權 + 五變項校正加權、事後分層組合估計

校正加權的目的，在於利用母體人口特徵或感興趣的研究變項等相關參數來校正樣本，希冀降低無反應誤差，確保調查估計與已知母體人口特徵、感興趣的研究變項結果是一致的。國外有些研究單位或學者在執行雙底冊調查合併加權時，已納入電話使用習性此一輔助變項做為權重調整的參考依據(Kennedy 2007; McDougale 2016; Turakhia, Best, and Glancey 2019)，因此，本文在第II類比較的校正加權變項即加入電話使用習性此一變項來探討，而為能與張鏡文、黃東益與洪永泰（2017）的事後分層組合估計法來進行比較，此處採用的電話使用習性比率是與張鏡文等人相同、以原雙底冊本身的調查結果估計而得。

1. 變異數比較：圖5為前述七項加權處理後所計算出的設計效果與加權耗損。

首先，若比較四變項校正加權與五變項校正加權的差異，四種合併組合（篩選設計與重疊設計）在設計效果上幾乎可說是沒有什麼改變，且多採用電話使用習性來加權，在加權損耗的增加幅度也不超過1%；亦即，以此四種合

併組合來說，是否採用電話使用習性來調整權數對設計效果與加權損耗來說，幾乎是沒有影響的。

其次，以四種合併組合與事後分層組合估計來進行比較，不論是設計效果或加權損耗，事後分層組合估計的數值都比較小，此乃是因為事後分層組合估計法僅有單一階段的事後分層加權所致，其中，僅「市話+手機」組合的加權損耗略高，為11%。

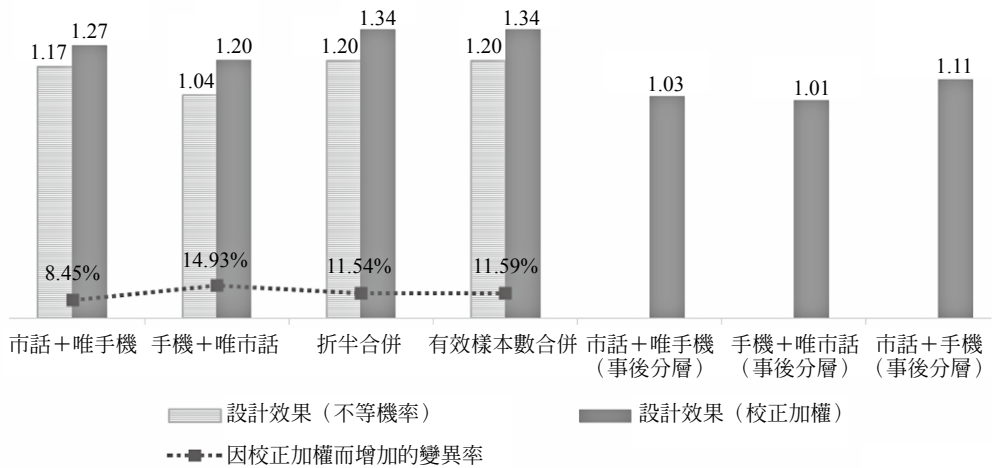


圖5 第II類五變項加權及事後分層組合估計之設計效果評比

資料來源：張卿卿（2020）。

2. 偏差比較：首先，四種合併組合的加權方式不論是採四變項校正加權與五變項校正加權（如表5所示），都同樣使用性別等四項人口特徵變項做調整，故加權後的結果都與母體一致；而若比較婚姻狀況、就業狀況與手機電信用戶率等這三項在四變項及五變項校正加權後的比例分布，則可說是幾乎沒有差異，因為同一類型合併組合之間的差距最多僅為0.2%。<sup>10</sup>

<sup>10</sup> 本文另外以外部全國性調查資料的電話使用習性佔比來做為加權的母體參數，即以臺灣傳播資料庫2020年全國性面訪調查所得來加權。結果如附錄二所示，母體參數有落入婚姻狀況、就業狀況與手機電信用戶率等項目的樣本估計值95%信賴區間者也同樣不多。

其次，就四種合併組合與事後分層組合估計相較於母體參數的表現來看，由於事後分層組合估計在加權時是以性別、年齡、教育程度進行分層後，再利用原調查本身各分層內所推估而得的電話使用習性比率做為調整變項，故加權後在人口特徵變項上，僅有「市話+唯手機」及「市話+手機」的性別變項通過樣本代表性檢定，其餘項目與母體參數都具有統計上的顯著差異；而就母體參數是否落在樣本估計值95%信賴區間的表現上，四種合併組合與事後分層組合估計都表現不佳，婚姻等三變項中僅有少數幾個項目的區間估計有包含到母體參數，其中，以事後分層的「市話+唯手機」表現最差，沒有任一項目的估計範圍涵蓋到母體參數。

表5 第II類五變項加權之比較分析

項目 %	市話+ 唯手機	手機+ 唯市話	折半 合併	有效樣本 數合併	事後分層 市話+唯 手機	事後分層 手機+唯 市話	事後分層 市話+手 機	母體
[樣本數]	[1529]	[1336]	[1432]	[1433]	[1529]	[1336]	[2490]	
性別								
男	49.1	49.1	49.1	49.1	47.2	56.8	51.0	49.1
女	50.9	50.9	50.9	50.9	52.8	43.2	49.0	50.9
年齡								
18-29	18.3	18.3	18.3	18.3	15.9	24.7	19.7	18.3
30-39	17.7	17.7	17.7	17.7	18.5	17.6	17.7	17.7
40-49	18.9	18.9	18.9	18.9	21.2	18.2	20.5	18.9
50-59	18.2	18.2	18.2	18.2	18.5	18.6	18.8	18.2
60-69	15.5	15.5	15.5	15.5	16.6	13.6	15.1	15.5
≥70	11.4	11.4	11.4	11.4	9.3	7.3	8.1	11.4
教育程度								
≤小學	12.2	12.2	12.2	12.2	8.4	7.3	7.1	12.2
國初中	11.7	11.7	11.7	11.7	8.9	8.8	8.2	11.7
高中職	27.7	27.7	27.7	27.7	25.6	26.4	26.0	27.7
專科	11.8	11.8	11.8	11.8	14.3	11.6	13.6	11.8
大學	28.9	28.9	28.9	28.9	32.0	35.6	34.3	28.9

項目 %	市話+ 唯手機	手機+ 唯市話	折半 合併	有效樣本 數合併	事後分層 市話+唯 手機	事後分層 手機+唯 市話	事後分層 市話+手 機	母體
[樣本數]	[1529]	[1336]	[1432]	[1433]	[1529]	[1336]	[2490]	
碩博士	7.8	7.8	7.8	7.8	10.9	10.4	10.9	7.8
地區								
北北基	29.9	29.9	29.9	29.9	30.3	33.9	32.2	29.9
桃竹苗	15.7	15.7	15.7	15.7	17.4	17.4	17.4	15.7
中彰投	19.2	19.2	19.2	19.2	20.0	17.5	18.9	19.2
雲嘉南	14.3	14.3	14.3	14.3	14.7	13.0	13.7	14.3
高屏澎 (金馬)	16.6	16.6	16.6	16.6	14.2	14.8	14.5	16.6
宜花東	4.3	4.3	4.3	4.3	3.4	3.3	3.3	4.3
婚姻狀況								
未婚	26.6	29.5	28.0	28.0	27.8	35.5	31.3	32.0
有偶	60.6	55.6	58.3	58.3	58.6	51.9	55.8	51.8
離婚/喪偶	12.9	14.8	13.7	13.7	13.5	12.6	12.9	16.3
就業狀況								
全職	57.7	58.9	58.0	58.0	59.0	62.7	60.9	55.7
兼職	10.7	10.9	10.8	10.8	10.1	11.4	10.7	2.1
無業	31.7	30.2	31.2	31.2	30.9	25.8	28.4	42.2
手機電信用 戶								
中華	46.8	47.8	47.1	47.1	49.6	46.5	48.3	35.9
台灣大哥大	23.0	22.7	22.8	22.8	21.6	22.7	22.1	24.3
遠傳	21.8	21.2	21.6	21.6	20.4	22.0	20.9	24.1
台灣之星	5.4	5.7	5.6	5.6	5.2	6.2	5.7	8.7
亞太	3.0	2.7	2.9	2.9	3.3	2.6	3.0	6.9

資料來源：張卿卿（2020）。

說明：1. 灰底標示處表示母體參數落在「納入設計效果」樣本估計值之95%信賴區間估計內。

2. 表內各項目與母體參數進行卡方適合度檢定的結果，除了「虛線框」標示者的 $p$ 值為沒有統計上的顯著差異（ $p$ 值大於0.05）、「灰線框」標示者的 $p$ 值小於0.05、「黑線框」標示者的 $p$ 值小於0.01之外，其餘項目的 $p$ 值均小於0.001。

最後，若比較四種合併組合與事後分層組合估計之樣本估計值是否落在彼此95%信賴區間估計內，並以「有效樣本數合併」合併組合做為比較參照組來看，事後分層組合估計的三種方式在幾項人口特徵上的估計範圍與「有效樣本數合併」估計範圍並沒有重疊。其中，以事後分層「手機+唯市話」的情況比較嚴重一點，除了男女性別差異大之外，在年齡、教育程度、地區、婚姻狀況、就業狀況也都各有數項類別有明顯差異，詳細結果請參見附錄三。蔡奇霖（2020）曾提出在使用事後分層組合估計後，應再以人口特徵變項進行校正加權，處理樣本代表性的問題，避免造成估計偏誤。

## 伍、雙底冊電話調查樣本配置比例之比較分析

前述的雙底冊電話調查分析，是建立在市話調查與手機調查這兩套樣本的樣本數幾乎是等比例配置狀態下的各種合併組合估計。然而，雙底冊調查的樣本數究竟應如何配置才比較適切呢？要回答這個問題，首先應考量的是樣本數與抽樣誤差之間的關係。當樣本數越少，抽樣誤差就越大，代表著樣本估計值的95%信賴區間範圍越廣，也就相對產生估計精確度偏低的問題。故為能獲得精確的估計，最好的方式是至少有一套樣本具有完整的代表性。Wolter等人(2019)建議人口佔比多者，該套樣本就採用完整的代表性，而人口佔比少者則採用篩選的方式。需特別注意的是，一大一小的雙底冊配置恐會產生較大的設計效果及變異數，造成加權權數具有較大的差距及產生過大的抽樣變異性(sampling variability)，故當雙底冊的樣本比例差異低於3:1時，最好採用縮小估計值(shrinkage estimator)來做合併處理。Lu等人(2013)則指出當以市話調查為主、以手機調查為輔時，若無成本考量，僅調查唯手機族是具有較優的統計估計表現，因為不需要處理樣本重疊的問題。但要在實務調查裡識別出唯手機族來訪問的挑戰性較大，因為可能要耗費較長的調查期與較多的資料蒐集成本。若有成本上的考量，在執行以市話調查為主的雙底冊電話調查（樣本數為2,000）時，使用較大樣本數的手機樣本，也可產生正確的估計結果；尤其，當手機樣本的佔比為市話樣本的30-35%時，是均方差(mean square error)最優

表現的狀態，再增加手機樣本的助益並不大。

在臺灣，電信設備的持有率，如前述所言，從2020年臺灣傳播資料庫全國性面訪調查所得之個人手機持有率是93.4%，而2020年行政院主計總處家庭收支調查估計出的家庭電話機持有率為85.9%。依照Wolter等人(2019)的說法，當雙底冊電話調查無法執行二個具完整代表性的樣本時，應該在手機調查採用完整代表性，而在市話調查採用篩選方式。但除了人口涵蓋率的考量之外，似乎應該多考量其他的影響因素，例如接觸率高低等，<sup>11</sup> 尤其，臺灣的手機號碼使用量相當大，依據NCC的統計，在2020年底的手機用戶數已超過2千9百萬，遠超過18歲以上的戶籍人口數（近2千萬人）。可想而知，手機號碼必然是被多方廣為運用，而非單純僅做為個人聯絡電話，包括：警報通知設備、純上網用、或各類型商用如股票投資捐客、廣告簡訊、詐騙電話等。再者，以本文所分析的資料結果來看（如表二），會接受手機調查訪問的人口特徵不同於市話調查，是以男性、年輕、大學程度以上的民衆居多。在此態勢下，若以手機調查為主、市話調查為輔，會是何種樣貌？

為探究非等比例配置的雙底冊調查在實務上所可能呈現的樣貌，本文將以實務調查結果來評比不同比例的雙底冊樣本配置效果；亦即，以市話與手機調查分別為單一完整代表性的全樣本（簡稱全樣本），再搭配另一底冊的部分比例樣本（簡稱比例樣本）。<sup>12</sup> 在比例樣本的樣本配置上，為能採用一般性的雙底冊合併方式，並避免產生如Wolter等人(2019)所說的過大抽樣變異，故採用的比例共有二種，一為佔全樣本樣本數的1/3比例，另一為全樣本樣本數的1/2比例。比較方式有三種，包括採重疊設計的「直接合併」、「有效樣本數合併」及篩選設計「全樣本+唯「比例樣本」（例如「市話+唯手機」）。<sup>13</sup> 分

<sup>11</sup> 以本文分析的雙底冊電話調查來看，市話調查的接觸率為30.45%，手機調查為18.65%。請見附錄一。

<sup>12</sup> 本文所使用的雙底冊電話調查在抽樣設計上是利用分層多階段RDD抽樣法（請見附錄一），事先抽好多套具有相同全國代表性的樣本，每套樣本為1000個電話號碼。正因為每套樣本具有相同的代表性，本文即利用此特性來進行比例樣本之分析。

<sup>13</sup> 此處的合併組合同樣納入不等機率權值與校正權值（性別、年齡、教育程度、地區）。另外，此分析並未納入折半合併，是因為當雙底冊配置的樣本數具有差異時，採用有效樣本數合併可以計算出更適切的合併權值，更動態反應調查結果。

析結果如下。

1. 變異數比較：從設計效果來看，首先，非等比例配置的雙底冊調查在加權後僅略差於等比例配置（如圖6，雙底冊內的單底冊樣本數相近且均大於1000）一點，差距約落在0.02-0.11之間，並以「手機+1/3市話」配置的「手機+唯市話」設計為最差，但差距僅為0.11，亦即加權損耗多增加了11%。

其次，當採行非等比例配置時，不論是搭配1/3或1/2的比例樣本，其間的差異其實不大，差距都不到0.1。其中，整體表現比較差者為「手機+1/3市話」的設計，不論是因校正加權而產生的變異率（約落在19%~34%之間）或是加權損耗（約落在28%~43%之間），都比其他的配置組合還高。這結果並不令人意外，因為以手機調查為全樣本而搭配1/3的市話調查，等同是單採手機調查的結果進行推估；尤其，只搭配唯市話的極低人口比例時更是明顯，故合併組合的成效當然最差。

總體而言，不論是以市話為全樣本或是以手機為全樣本，再搭配不同的合併組合的比例樣本，加權後所增加的變異差異並不大，也不至於過高到會產生估計不穩的狀況。其中，以「手機+1/3市話」的表現略差而已。

2. 偏差比較：首先，若與表四的雙底冊等比例配置比較，以四變項的人口特徵進行校正加權後，就婚姻狀況、就業狀況與手機電信用戶率等三個項目來看，分布比例並沒有太大差異，絕大多數的差距都在1%上下，僅「市話+1/3手機」配置的「有效樣本數合併」在「婚姻/有偶」的差異為最大(3.1%)。而若單比較非等比例配置時，差異也相當小，都落在1%上下。

其次，以樣本代表性來看（如表6所示），除了「手機+1/3市話」配置的「手機+唯市話」的婚姻變項與母體參數沒有統計上的顯著差異之外，其餘都沒通過檢定。而就樣本估計值信賴區間是否涵蓋母體參數來看，非等比例配置並沒有比等比例配置的表現還差，屬伯仲之間、不相上下。

總體來說，不論以哪個底冊當全樣本，或搭配哪種合併組合模式，以一般加權的方式處理後，彼此之間在婚姻、就業及手機電信用戶率的估計上，都沒有明顯的差異存在。

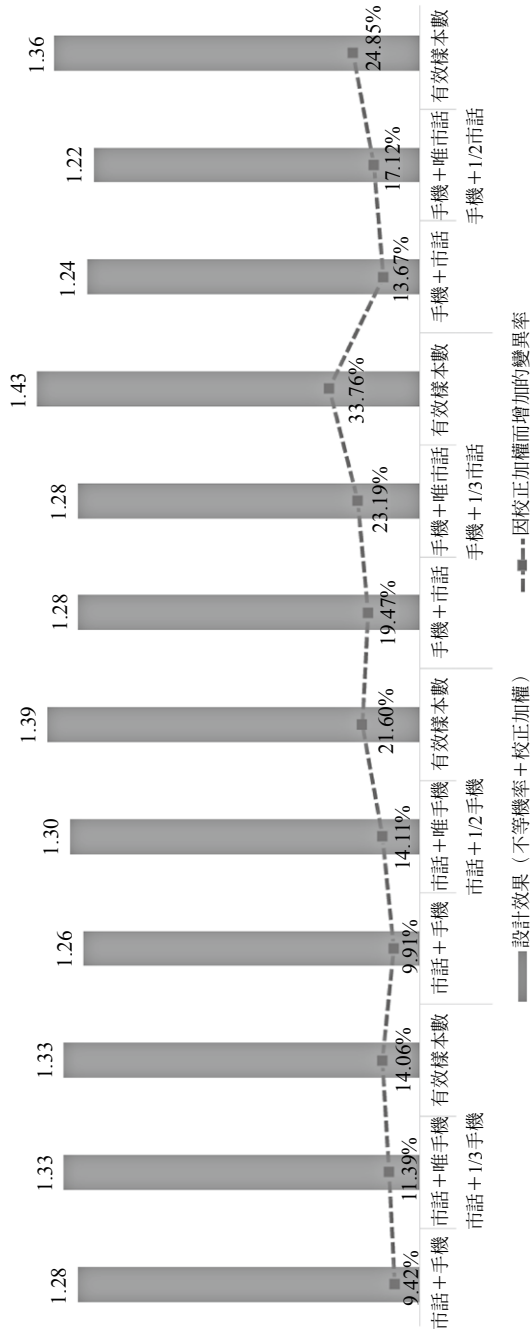


圖6 雙底冊之「全樣本」單底冊 + 「比例樣本」單底冊之加權設計效果評比

資料來源：張卿卿（2020）。



表6 雙底冊採用「全樣本」單底冊 + 「比例樣本」單底冊之四變項加權比較分析

項目 %	市話+1/3手機			市話+1/2手機			手機+1/3市話			手機+1/2市話			母體
	市話+ 手機	市話+ 唯手機	有效 樣本數 合併	市話+ 手機	市話+ 唯手機	有效 樣本數 合併	手機+ 市話	手機+ 唯市話	有效 樣本數 合併	手機+ 市話	手機+ 唯市話	有效 樣本數 合併	
[樣本數]	[1596]	[1327]	[1138]	[1936]	[1407]	[1189]	[1590]	[1276]	[1132]	[1936]	[1298]	[1180]	
<b>婚姻狀況</b>													
未婚	27.6	26.9	27.1	28.2	26.9	28.0	28.8	29.4	29.2	28.7	29.6	29.0	32.0
有偶	61.1	62.2	61.8	59.6	61.2	59.4	58.1	55.8	56.2	58.3	56.0	56.9	51.8
離婚 / 喪 偶	11.3	10.9	11.1	12.2	11.9	12.6	13.1	14.7	14.6	13.0	14.4	14.1	16.3
<b>就業狀況</b>													
全職	57.1	57.4	57.3	58.0	57.7	57.7	58.5	58.5	58.2	58.7	58.6	58.4	55.7
兼職	10.6	10.6	10.5	10.7	10.4	10.4	11.1	11.0	10.9	10.6	11.0	10.7	2.1
無業	32.3	32.0	32.2	31.3	31.9	31.9	30.4	30.5	30.8	30.7	30.4	31.0	42.2
<b>手機電信用戶</b>													
中華	48.5	47.4	47.1	48.3	47.4	47.7	48.4	48.2	47.9	48.0	47.7	47.4	35.9
台灣大哥 大	23.2	23.2	23.3	22.5	22.8	22.7	22.3	22.5	22.5	22.4	22.8	22.7	24.3
遠傳	19.5	20.6	20.6	20.5	21.0	20.9	20.9	20.9	21.1	21.1	21.1	21.4	24.1

項目	市話+1/3手機			市話+1/2手機			手機+1/3市話			手機+1/2市話			母體	
	市話+ 手機	市話+ 唯手機	有效 樣本數 合併	市話+ 手機	市話+ 唯手機	有效 樣本數 合併	手機+ 市話	手機+ 唯市話	有效 樣本數 合併	手機+ 市話	手機+ 唯市話	有效 樣本數 合併		手機+ 唯市話
[樣本數]	[1596]	[1327]	[1138]	[1936]	[1407]	[1189]	[1590]	[1276]	[1132]	[1936]	[1298]	[1180]		
台灣之星	5.4	5.7	5.8	5.5	5.6	5.6	6.0	5.7	5.8	5.7	5.7	5.8	5.7	5.8
亞太	3.3	3.2	3.2	3.2	3.1	3.1	2.5	2.7	2.6	2.9	2.7	2.8	2.7	2.8

資料來源：張卿卿（2020）。

說明：1. 灰底標示處表示母體參數落在「納入設計效果」樣本估計值之95%信賴區間估計內。

2. 表內各項目與母體參數進行卡方適合度檢定的結果，除了「虛線框」標示者的p值為沒有統計上的顯著差異（p值大於0.05）、「灰框」標示者的p值小於0.05（手機+1/3市話）、「黑框」標示者的p值小於0.01（手機+1/2市話）之外，其餘項目的p值均小於0.001。

## 陸、結語

在臺灣，不論市話與手機的持有率雖然都很高，但表2的結果顯示，在沒有使用加權前，不論是任何單底冊調查、雙底冊採重疊設計或採篩選設計等，所獲致的人口特徵結果彼此大不相同，與母體也具有統計上的顯著差異。這有部分源自於抽樣設計與無反應誤差所致，故學界往往會採取加權來補償不等機率與無反應誤差。當然，調查資料的人口特徵與母體一致，並不必然代表認知、行為與態度就與母體相符，而若能採用相對有效的輔助變項來改善估計的效益，則應該更能確保樣本得以精確地反應感興趣的母體特性。

不過，前述提及的輔助變項，在臺灣調查資源有限的環境下，是可遇而不可求的。以本文所使用的調查資料而言，並沒有相對適合可用以做為加權推估的輔助變項，而這狀況普遍存在於各調查資料。因此，本文嘗試以臺灣學界普遍使用的加權方式，從鮮少被重視的變異數與偏差視角，來評比各類型雙底冊合併組合是否存在差異，並進而探討各類型雙底冊樣本配置所呈現的效果，裨益提供實務上的應用參考。

本文的資料分析結果簡述如下。首先，雙底冊電話調查除了各底冊的唯手機族與唯市話族未被彼此涵蓋到之外，本文的資料分析結果發現，兩用族在人口結構與地區結構上具有顯著差異。不過，若進一步運用各種樣本配置的合併組合來比較加權後的估計值，例如：婚姻、就業狀況及手機電信用戶比率等三項，其間的差異則相當不顯著。這代表雖兩用族具有人口結構的差異，但在合併加權時，因具有相當程度的互補效果而弱化了原有的顯著差異問題。不過，本文在進行合併後的差異比較時，因受限於問卷題目內沒有適當的認知或態度等類型題目，而僅能以前述三項可獲取母體參照值來評比，是比較可惜之處。

其次，以加權策略來看，在執行雙底冊電話調查時，當市話調查與手機調查均為全樣本，且樣本數近乎等比例的條件下，在加權設計上不論是採四變項或五變項的校正加權、甚或是採用重疊設計或篩選合併，以加權損耗來說，僅約增加10%，代表因加權而增加的變異並不大，而在估計值的表現上差距也很小。其中，僅事後分層組合估計法因沒有納入不等機率權數及以母體人口特徵變項來做校正加權，在加權損耗的表現或許較優一點，約少了20%~30%，但

在母體人口特徵或其他如婚姻狀況等三變項上，其樣本估計值的表現就具有差異：即不僅絕大多數的樣本估計值與母體參數具有統計上的顯著差異，且與其他合併組合也具有相當程度的差距。當然，以母體人口特徵變項來進行校正加權，讓樣本的人口表徵長得與母體一致，不必然代表在估計上的絕對精確；惟將調查資料納入不等機率權數及以母體人口特徵變項進行校正加權的目的，旨在補償樣本的不等中選機率及調查的無反應誤差等，是學理上的必要。

最後，以雙底冊電話調查的樣本配置比例來看，若市話調查與手機調查的樣本數採非等比例配置，且以其中某一底冊為完整代表性的全樣本，則不論搭配的輔助調查是採用其1/3或1/2比例，在四變項校正加權後所衍生的加權耗損並不大，而彼此之間在樣本估計值上的差異也很小。惟若欲採用「手機+1/3市話」的合併組合，因為此組合等同採取純手機調查，成效略差於其他合併組合。

以前述的分析結果來看，雙底冊電話調查採用不同的加權程序實然會對樣本估計有所影響，尤其是沒有納入不等機率及考量母體人口結構的情況下。惟考量加權處理乃屬學理上的必要，其作用包括：將樣本抽取時所產生的不等機率還原回一致的等機率基礎，及因應無反應或涵蓋率等而運用母體人口結構、抑或感興趣的研究變項估計值來進行校正加權等，而為了避免加權時產生過大的加權損耗或產生過大的抽樣變異性而導致估計不穩的情況，並考量加權損耗是產自於不等機率權數及校正權數等機制下，依據本文的分析結果，建議雙底冊調查的抽樣設計採取以下策略：

1. 盡量採用不會造成不等機率權數過大的機率抽樣。
2. 為避免單底冊的抽樣誤差過大而造成估計精確度過低的問題，雙底冊內至少要有一個底冊是具有完整代表性的全樣本（抽樣誤差在 $\pm 3\%$ 以內）；而另一底冊若無法採用全樣本時，則樣本配置比例也不可差異過大，保守而言，至少不低於全樣本的1/3(Wolter et al. 2019)。
3. 當雙底冊電話調查採用重疊設計時，在考量實務調查使用簡單隨機抽樣設計，抑或是在一般調查中獲取與研究變項高度相關的校正輔助變項等可能性都極低的條件下，其實並不適合使用單底冊途徑及聯合校正估計值來合併雙底冊資料，故僅剩下雙底冊組合估計與事後分層組合估計可供選擇使

用，對此二者使用上的建議如下：

- (1)採用雙底冊組合估計時，為符合學理上的加權必要及可彈性因應調查變動的雙重考量下，相較於折半合併一律呆板地採用0.5做為組合設計權數，有效樣本數合併是比較穩健的中庸做法。因此在合併兩用族時，不僅納入了兩套樣本於抽樣設計的設計效果因子（反應單底冊的有效樣本數），又可因應雙底冊的樣本配置比例而彈性調整組合設計權數。
- (2)因事後分層組合估計是直接合併雙底冊調查，以三個人口變項進行分層後，計算推估出各分層在三種電話使用習性的佔比，並據此做事後分層校正加權，而未納入不等機率加權及非簡單隨機抽樣等相關設計效果，故此法所得的樣本估計值與其他合併組合具有比較大的差異。從學理上看，此法比較適用於採用簡單隨機抽樣及調查品質保證無虞的雙底冊調查，才比較容易產生不偏估計。

此外，就篩選設計而言，因本文的分析資料來源並非真實來自於調查實務上採行篩選式的調查，而是直接從全樣本的調查結果裡篩選出所欲分析的對象；若真實以篩選方式執行調查，不確定所需耗費的調查期及調查成本，甚或是否會產生非預期中的調查誤差。

由於本文分析所運用的雙底冊電話調查並無一般性的認知、態度與行為等類型的題目適合用來進行差異分析，故本文僅能就母體基本人口特徵變項及可獲取的婚姻、就業狀況及手機電信用戶比率等三變項的母體參數來進行評比，這是本文的研究限制所在。而單從前述這些變項來比較市話調查與手機調查的差異，明顯可得知，採用不同的電信設備（市話或手機）而可接觸到且願意接受訪問的民衆，是二群不一樣的民衆，不論是以單一完整代表性的全樣本或各自的兩用族來比較；亦即，相較於市話調查，手機調查的樣本群像是男多於女、未婚者居多，且具有年輕化、高教育及高都會的特性。故使用單一底冊調查，可能因涵蓋率、可接觸性等的問題而具有潛在的偏差問題，而此問題單靠人口基本變項加權，不必然可以扭正資料，甚至可能扭得更歪。因此，在現今缺乏單一底冊的調查結果可以全然代表臺灣母體人口的當下，施行雙底冊電話調查的必要性是存在的，而進階的研究分析，包括各種可能性合併組合的樣本配置、加權策略、誤差降低、抑或是變異數估計等，都值得吾人繼續深耕探究！

## 參考文獻

### I.中文部分

中華民國統計資訊網，2022，〈電信業者營運實績〉，中華民國統計資訊網站：<https://statdb.dgbas.gov.tw/pxweb/Dialog/varval.asp?ma=TK1101A1M&ti=%B9q%ABH%B7~%AA%CC%C0%E7%B9B%B9%EA%C1Z-%A4%EB&path=../PXfile/Transportation/&lang=9&strList=L>，檢索日期：2022年9月28日。

(National Statistics, R.O.C.(Taiwan). 2022. “Dian xin ye zhe ying yun shi ji”[Operating Performance of Telecom Operators]. National Statistics, R.O.C.(Taiwan). <https://statdb.dgbas.gov.tw/pxweb/Dialog/varval.asp?ma=TK1101A1M&ti=%B9q%ABH%B7~%AA%CC%C0%E7%B9B%B9%EA%C1Z-%A4%EB&path=../PXfile/Transportation/&lang=9&strList=L>(accessed September 28, 2022).)

侯佩君，2011，〈複雜抽樣調查的資料分析及實務應用：以SAS、STATA為例〉，《調查研究：方法與應用》，25: 91-140。

(Hou, Pei-chun, 2011. “Fu za chou yang diao cha de zi liao fen xi ji shi wu ying yong: Yi SAS, STATA wei li”[Analysis of Complex Survey Data: The Application of SAS and STATA]. *Survey Research-Method and Application* 25: 91-140.)

洪永泰，2001，〈戶中選樣之研究〉，台北：五南。

(Hung, Yung-tai. 2001. “*Hu jhong syuan yang jih yan jiu*” [Research on Within-Household Selection Methods]. Taipei: Wunan.)

-----，2019，〈2020民進黨總統初選，民調數據千萬不能當真！〉，遠見華人菁英論壇：<https://gvlf.gvm.com.tw/article/66740>，檢索日期：2022年9月28日。

(-----, 2019. “2020 min jin dang zong tong chu xuan, min tiao shu ju qian wan bu neng dang zhen” [The 2020 Democratic Progressive Party Presidential Primary

- Election, The Poll Data Must Not Be Taken Seriously]. *Global Views Leaders Forum*. <https://gvlf.gvm.com.tw/article/66740> (accessed September 28, 2022.)
- , 2021, 〈雙底冊電話調查的估計成效評估：以選前調查為例〉，《選舉研究》，28(2): 95-126。
- (-----, 2021. “Shuang di ce dian hua diao cha de gu ji cheng xiao ping gu: Yi xuan qian diao cha wei li” [Comparisons of Estimation Efficiencies among Various Dual-frame Telephone Survey Weighting Procedures: A Study of Pre-election Polls]. *Journal of Electoral Studies* 28(2): 95-126.)
- 洪永泰、俞振華、高世垣，2017，〈數位時代下地方民意探索之挑戰與回應〉，台北市政府研考會委託專題研究報告。
- (Hung, Yung-tai, Yu, Eric Chen-hua, Kao, Shih-yuan. 2017. “*Shu wei shih dai xia difang min yi tan suo zhi tiao chan yu hui ying*” [*The Challenges and Responses of Exploring Public Opinion for the Local Government under Digital Era*]. Research, Development and Evaluation Commission, Taipei City Government.)
- 洪永泰、洪百薰、林宇旋、呂孟穎、許勝懋、吳淑慧、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 lu 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.)
- 張卿卿，2020，《幸福溫度計調查》，台北：中央研究院人文社會科學研究中心。
- (Chang, C. 2020. “*Xing fu wen du ji diao cha*” [*The Happiness Thermometer Project*]. Taipei: Research Center for Humanities and Social Sciences, Academia Sinica.)
- 張卿卿、陶振超，2021，《臺灣傳播調查資料庫第二期第四次（2020年）調查：新傳播科技與生活延伸【原始數據】》，計畫編號：Most 108-2420-

H-001-005-SS2，台北：行政院科技部補助個別型計畫。

(Chang, C., and C. C. Tao. 2021. “*Tai wan chuan bo diao cha zi liao ku di er qi di si ci(2020 nian) diao cha: Xin chuan bo ke ji yu sheng huo yan shen(Yuan shi shu ju)*”[*Taiwan Communication Survey: New Communication Technologies & Life Boundary Expansion (Data Set) (Second Phase, Fourth Wave)*]. MOST 108-2420-H-001-005-SS2. Taipei: Executive Yuan Ministry of Science and Technology Research Project.)

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, 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 Procedure for the Dual Frame Telephone Survey in Taiwan: The Case of the 2016 Presidential Election]. *Journal of Electoral Studies* 24(2): 65-96.)

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: 65-96.)

陳鴻嘉，2019a，〈成本考量下雙底冊電話調查的樣本配置〉，《選舉研究》，26(1): 31-56。

(Chen, Hung-chia. 2019a. “Cheng ben kao liang xia shuang di ce dian hua diao cha de yang ben pei zhi”[Allocation of Dual-Frame Telephone Survey for Given Cost]. *Journal of Electoral Studies* 26(1): 31-56.)

-----，2019b，〈勢在必「行」：唯手機族日增趨勢下的電話調查〉，《中華傳播學刊》，35: 217-234。

(-----, 2019b. “Shi zai bi xing: Wei shou ji zu ri zeng qu shi xia de dian hua diao cha” [On Telephone Survey under the Increasing Trend of Cell Phone-only Us-



age]. *Chinese Journal of Communication Research* 35: 217-234.)

蔡奇霖，2020，〈雙底冊電話抽樣調查之事後校正〉。2020年臺灣政治學會年會「全球民主退潮？政治學的反思與應對」國際學術研討會，2020年12月5至6日，臺灣，臺北：臺灣大學。

(Tsai, Chi-lin. 2020. "Shuang di ce dian hua chou yang diao cha zhi shi hou jiao zheng" [Post-Calibration of Dual-Frame Telephone Surveys]. Presented at the 2020 International Conference on Taiwanese Political Science Association, Taipei.)

## II. 外文部分

Baffour, Bernard, Michele Haynes, Mark Western, Darren Pennay, Sebastian Misson and Arturo Martinez. 2016. "Weighting Strategies for Combining Data from Dual-frame Telephone Surveys: Emerging Evidence from Australia." *Journal of Official Statistics* 32(3): 549-578.

Bankier, M. D. 1986. "Estimators Based on Several Stratified Samples with Applications to Multiple Frame Surveys." *Journal of the American Statistical Association* 81: 1074-1079.

Blumberg, Stephen J., and Julian V. Luke. 2009. "Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July - December 2008." National Center for Health Statistics.

<https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless200905.pdf> (accessed March 22, 2021)

-----, 2011. "Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January-June 2011." National Center for Health Statistics.

<https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201112.pdf> (accessed March 22, 2021)

Brick, J. Michael, Ismael Flores Cervantes, Sunghye Lee, and Greg Norman. 2011.

- “Nonsampling Errors in Dual Frame Telephone Surveys.” *Survey Methodology* 37(1):1-12.
- Brick, J. Michael, Pat D. Brick, Sarah Dipko, Stanley Presser, Clyde Tucker, and YangYang Yuan. 2007. “Cell Phone Survey Feasibility in the U.S.: Sampling and Calling Cell Numbers versus Landline Numbers.” *Public Opinion Quarterly* 71(1):23-39.
- Brick, J. Michael, Sarah Dipko, Stanley Presser, Clyde Tucker, and Yangyang Yuan. 2006. “Nonresponse Bias in A Dual Frame Sample of Cell and Landline Numbers.” *Public Opinion Quarterly* 70: 780-793.
- Boyle, John M, Faith Lewis, and Brian Tefft. 2010. “Segmented or Overlapping Dual Frame Samples in Telephone Surveys.” *Survey Practice* 3(6).
- Deville, Jean-Claud, and Carl-Erink Särndal. 1992. “Calibration Estimators in Survey Sampling.” *Journal of the American Statistical Association* 87: 376-382.
- Elkasabi, Mahmoud A., Steven G. Heeringa, and James M. Lepkowski. 2015. “Joint Calibration Estimator for Dual Frame Surveys.” *Statistics in Transition* 16(1): 7-36.
- Frankel, Martin R., Michael P. Battaglia, Michael Link, and Ali H. Mokdad. 2007. “Integrating Cell Phone Numbers into Random Digit-Dialed (RDD) Landline Surveys.” Presented at the 2007 Joint Program Committee of the American Statistical Association.
- Fuller, WA, and LF. Burmeister. 1972. “Estimators for Samples Selected from Two Overlapping Frames.” Proceedings of the Social Statistics Section, *American Statistical Association*: 245-249.
- Hartley, H. O. 1962. “Multiple Frame Surveys.” Proceedings of the Social Statistics Section, *American Statistical Association*: 203-206.
- Heeringa, Steve G., Brady West, and Patricia A. Berglund. 2017. *Applied Survey Data Analysis*. (2nd ed.) London: Chapman & Hall.
- Hu, S.S., Lina Balluz, Michael P. Battaglia, and Martin R. Frankel. 2011. “Improving Public Health Surveillance Using a Dual-Frame Survey of Landline and Cell

- Phone Numbers.” *American Journal of Epidemiology* 173(6): 703-711.
- Kalton, G., and Dallas W. Anderson. 1986. “Sampling Rare Populations.” *Journal of the Royal Statistical Society, Series A (Statistics in Society)* 149(1): 65-82.
- Kennedy, Courtney. 2007. “Evaluating the Effects of Screening for Telephone Service in Dual Frame RDD Surveys.” *Public Opinion Quarterly* 71(5): 750-771.
- Kennedy, Courtney, and B. Stephen Evertt. 2007. “Use of Cognitive Shortcuts in Landline and Cell Phone Surveys.” *Public Opinion Quarterly* 75(2): 336-348.
- Kish, Leslie. 1965. *Survey Sampling*. New York: Wiley.
- Little, Roderick J., and Sonya Vartivarian. 2005. “Does Weighting for Nonresponse Increase the Variance of Survey Means?” *Survey Methodology* 31(2): 161-168.
- Lohr, Sharon L., and J. Michael Brick. 2014. “Allocation for Dual Frame Telephone Surveys with Nonresponse.” *Journal of Survey Statistics and Methodology* 2(4): 388-409.
- Lu, B., Timothy Sahr, Ronaldo Iachan, Matthew G. Denker, Thomas Duffy, and Daniel Weston. 2013. “Design and Analysis of Dual-frame Telephone Surveys for Health Policy Research.” *World Medical & Health Policy* 5: 217-232.
- McDougle, Robyn. 2016. *2016 Commonwealth Poll: Public Safety*. Center For Public Policy, Virginia Commonwealth University.
- Mecatti, F. 2007. “A Single Frame Multiplicity Estimator for Multiple Frame Surveys.” *Survey Methodology* 33: 151-157.
- Skinner, C. J. 1991. “On the Efficiency of Raking Ratio Estimation for Multiple Frame Surveys.” *Journal of the American Statistical Association* 86: 779-84.
- Skinner, C. J., and J. N. K. Rao. 1996. “Estimation in Dual Frame Surveys with Complex Designs.” *Journal of the American Statistical Association* 91: 349-56.
- The American Association for Public Opinion Research. 2010. *Cell Phone Task Force Report: New Considerations for Survey Researchers When Planning and Conducting RDD Telephone Surveys in The U.S. with Respondents Reached via Cell Phone Numbers*. AAPOR.
- , 2016. *Standard Definitions: Final Disposition of Case Codes and Outcome*

- Rates for Surveys*. 9th edition . AAPOR
- Turakhia, Chintan, Jonathan Best, and Sarah Glancey. 2019. *Methods Report for Pew Charitable Trusts: 2019 Philadelphia Residents Survey*. SSRS.
- Rai, Piyush Kant, G. C. Tikkiwal, and Alka. 2020. “A Joint Calibration Estimator of Population Total Under Minimum Entropy Distance Function Based on Dual Frame Surveys.” In *Statistical Methods and Applications in Forestry and Environmental Sciences*. eds. G. Chandra, R. Nautiyal, and H. Chandra. New York: Springer.
- Valliant, Richard, Jill A. Dever, and Frauke Kreuter. 2018. *Practical Tools for Designing and Weighting Survey Samples*. 2nd ed. New York: Springer.
- Wolter, Kirk M., Xian Tao, Robert Montgomery, and Philip J. Smith. 2015. “Optimum Allocation for A Dual-Frame Telephone Survey.” *Survey Methodology* 41(2): 389-401.
- Wolter, Kirk M., N. Ganesh, Kennon R. Copeland, James A. Singleton, and Meena Khare. 2019. “Estimation Tools for Reducing the Impact of Sampling and Non-response Errors in Dual-Frame RDD Telephone Surveys.” *Statistics in Medicine* 38: 4718-4732.

## 附錄

附錄資訊可於當其期刊出刊後，於《選舉研究》期刊網站查詢。<sup>14</sup>

附錄一：幸福溫度計調查簡介。

附錄二：以面訪調查估計之五變項校正加權比較分析。

附錄三：第II類五變項加權\_兩樣本估計值的95%信賴區間估計。

---

<sup>14</sup> 選舉研究期刊網頁網址：<http://www.jestw.com/tc/index.php>

# A Comparative Analysis of Weighting Approaches for Dual-Frame Telephone Surveys

*Pei-chun Hou*\*

## Abstract

With the rapid advancement of IT and the changing times, Taiwan has been facing increasing coverage error in single-frame telephone surveys (whether by landline or cell phone) in recent years. Dual-frame telephone surveys can effectively reduce this deficiency. However, there is still no consensus on how to weight dual-frame telephone samples in Taiwan.

This study uses data from a 2020 dual-frame telephone survey to compare several popular weighting approaches. The impact of these approaches on the variance and bias in both single-frame and dual-frame samples, as well as in overlap and screener designs, is discussed. Additionally, this study examines the differences between the estimates of the non-proportion dual-frame sample allocation to provide practical recommendations.

The results indicate the following: (1) if the sample allocation of the dual-frame samples is in nearly equal proportions, minor differences occur in the weighting loss between the overlap and screener designs. Only the post-stratified estimating procedure reveals some differences in the estimates with other designs. (2) If the sample allocation of the dual-frame samples is in unequal proportions, such as one single-frame national sample and the others are only 1/3 or 1/2 of the former, then after calibration, the weighting loss is only slightly increased, and these approaches

---

\* R&D Scientist, Center for Survey Research, Research Center for Humanities and Social Sciences, Academia Sinica.

produce only tiny differences in the calibrated estimates between the combinations.

**Keywords:** dual-frame telephone surveys, weighting, design effect, weighting loss, bias