

調查研究之有效空間抽樣分析 —以 2010、2012 TEDS 資料為例

徐永明* 吳怡慧**

- 一、前言
- 二、文獻回顧
- 三、研究方法
- 四、研究發現
- 五、討論與結語
- 六、參考書目

選舉研究常使用電訪或面訪等調查方法蒐集資料，並且考量樣本須具有足夠的代表性，抽樣過程運用「抽取率與抽出單位大小成比例」（probabilities proportional to size，簡稱 PPS）的原則抽出受訪者，這些受訪者所在的位置，受到抽樣成本與便利性的限制，在空間中分布並不均質；但是選舉研究中常探測的參數，像是投票行為、政黨認同、派系等，卻具有地域性的差異，因此受訪者所在的位置，對於後續研究進行空間效果

* 本文研究內容感謝國科會計畫編號 102-2410-H-031-044-MY2 經費的支持，該計畫名稱為「空間抽樣與政治研究」，促使本研究得以順利完成。

* 東吳大學政治系教授，E-mail: polymh@scu.edu.tw

** 國立臺灣大學生態學與演化生物學博士候選人。E-mail: d98b44006@ntu.edu.tw

投稿日期：2016 年 12 月 11 日；接受刊登日期：2017 年 4 月 27 日。

東吳政治學報/2017/第三十五卷第一期/頁 143-177。

分析時，是重要且有影響的。

本文為了解決空間特性對於選舉行為的影響，引用並發展新的抽樣方法：1. 在已完成的調查資料中（本文使用 TEDS2012 總統與 TEDS2010 直轄市長的選舉資料），加入空間結構（人口數）考量，重新加權計算資料的數值；2. 考慮選舉研究變項的空間自相關性（spatial autocorrelation）與空間異質性（spatial heterogeneity），提出新的空間抽樣概念，同樣在既有的抽樣調查資料上，進行樣本數縮小與樣本分布特性的修正。本文以「投票率」、「支持率」來進行相關方法的分析比較。

結果發現，方法一以人口數的空間結構差異來加權，重新統計調查資料的投票率與支持率，雖然校正結果較原本 TEDS 的結果更接近真實資料，但是方法二以空間特性為考量，依照空間物理性的距離為再抽樣的依據，其樣本的推估值與真實選舉資訊更為接近，顯然，驗證變項行為與實際政治行為之間的相關性，二維的抽樣設計方式更能有效的推估選舉行為所探測的參數。

關鍵詞：空間抽樣、選舉研究、總統選舉、直轄市長選舉、支持率、投票率

一、前言

電話訪問或面談訪問，是廣泛在選舉前用來推估選舉結果的調查資料，民調的結果與真實的選舉結果之間往往存在落差，造成評估不準確的原因有可能是資料缺失（missing data）、抽樣偏差（sampling bias）、或是受訪者不願意回答，以及回答不真實造成的錯誤報告（misreport）。

傳統的抽樣方法，抽樣的前提是假設每一個樣本獲選的機率一致，且樣本與樣本之間是獨立的，任意更換兩組樣本並不會影響統計結果。但若是從空間的角度來檢視這些已完成的調查資料，會發現，要符合樣本與樣本之間是獨立的，任意更改兩組樣本的位置，而不影響統計結果，必須考慮到樣本的分布特性；受到抽樣設計的限制，或者聚落發展、人口聚集的影響，樣本的空間分布多半較為集中，但是，相鄰的樣本，受到鄰近效應的影響，應用在空間分析時，必須捨棄部分無法區分為獨立的樣本，才能滿足樣本與樣本間獨立性的考量。過去許多選舉研究資料，並不以空間的角度來篩選適當的樣本，而是將已完成的調查資料，以加權、加入具有空間特性的參數、或提出新的統計方法來修正推估結果。

二、文獻回顧

前人既以空間異質性與空間相依性來論述空間效應對於投票行為的影響（賴進貴等，2007），提到：「受到距離的影響，越鄰近的家戶所表現出來的選舉行為會越相近（葉高華，2006）。」那麼

樣本的「空間」分布對於母體的估算則甚為重要。樣本分布的均不均質，是進行空間分析的前提之一，如林昌平等人（2015）證實：對於同年度的選舉調查，電訪資料由於樣本分布較面訪資料更為分散且均勻，其估算結果較接近真實情況（林昌平等，2015）；並且林昌平等人根據空間抽樣假設，將原本無資料的區域補上假設性樣本後，推估效果明顯改善，更能接近選舉的真實結果（林昌平等，2015）。

近年來，預測市場逐漸成為常見的選舉結果預測機制，在地方選舉方面，童振源等（2011）根據未來事件交易所的交易資料，分析預測市場對 2009 年臺灣縣市長的選舉結果，研究顯示，預測市場對當選人預測的能力，在選舉前 20 天之後，便高於民調機構，預測準確度會隨著合約到期日的接近而逐漸增加。

此外，有許多學者利用各種統計方法來建構選舉預測模型，認為選民投票給某特定候選人的機率是介於 0 與 1 之間，因此若預測選民投給某一個候選人的機率高，則他投給該位候選人的可能性就大（楊佳芳，2002；盛杏媛，1998）；用來建構選舉預測模型的參數包括：選民的政治態度及背景（盛杏媛，1994）或人口、政治態度及議題立場等（盛治仁，2001）。另有學者以 2012 年總統選舉選後的調查資料，透過多重插補法（multiple imputation）補足缺失的變數資料，此方法有調整投票偏好比例的功能，但調整誤差的幅度仍十分有限（Liu and Su, 2015）。以上研究以選舉變數推估未表態的資料，繼而預測選舉結果，是嘗試從「資料缺失」的角度來校正調查資料的準確度。

本文認為，每一份調查研究，根據其所關心的議題面向，有其適用的抽樣設計方式，而母體的「空間分布」特性，對於抽樣設計

而言，是非常重要的。樣本的空間分布特性，對於調查研究而言確實為重要參數，樣本之間不具有空間相關性，則後續將調查資料作空間分析時，才能滿足樣本在空間中的「位置」是符合獨立。對於既有的抽樣樣本，所存在的空間均質性問題，的確不乏提出修正方法的文獻，面對舊有資料的侷限性，本文嘗試透過「抽樣誤差」的修正，重新估算民調資料，比較不同的抽樣方法對於預測選舉結果的成效。

三、研究方法

2012 總統選舉前後，臺灣選舉與民主化調查 (Taiwan's Election and Democratization Study, 簡稱 TEDS) 透過電訪及面訪蒐集問卷，估計各候選人的支持率（得票數/選舉人數）及整體投票率，然而，與選後，根據各投開票所統計，得到各候選人的支持率及整體投票率來比較，TEDS 的問卷資料，在馬英九與吳敦義這組候選人，所用來推估得票情形的支持率，是高估的，相對的在蔡英文及蘇嘉全這組候選人的支持率方面，則是低估，而整體投票率，面訪與電訪兩種調查方式所獲得的資料皆是高估了（表一）。

全國性的選舉資料如此，地方性的選舉資料如何？本文查看 2010 年的直轄市長選舉（中央選舉委員會，2010；黃秀端，2010），無論當選候選人組為國民黨或民進黨，調查資料的推估結果，對於當選候選人組皆為高估，對於競爭組皆為低估。而在投票率方面，則為高估（表二）。高雄市選情較為特殊，因此將楊秋興與國民黨候選人的得票數相加，作為另一組指標比較。

表一 2012 年總統選舉結果與 TEDS 訪查資料支持率比較

| 抽樣方法 | % | | | |
|---------------------------|------|------|------|-------|
| | 國民黨 | 民進黨 | 投票率 | 樣本數 |
| 2012 年總統選舉結果 ¹ | 38.1 | 33.7 | 74.4 | |
| TEDS 電訪資料 ² | 44.8 | 25.2 | 77.2 | 2,032 |
| TEDS 面訪資料 ³ | 46.0 | 30.1 | 89.2 | 1,510 |

1. 資料來源為中央選舉委員會（2012）。
2. 資料來源為黃紀（2012）。
3. 資料來源為朱雲漢（2012）。

從以上資訊得到：選舉民調或許很難與真實投票表現一致，除了受訪者本身的回答意願及是否在選前改變心意以外，調查方法的設計也是影響總統支持率及投票率估算的因素之一。

TEDS 的抽樣設計，無論是電訪在最後一階段為求涵蓋範圍的完整性，所進行電話號碼後 2 碼或 4 碼的修正，以及面訪資料從一開始便以地理區進行不同階段的抽樣設計，無疑是以樣本能盡量在「空間」中分布平均作為抽樣考量，但是，以電話號碼、或選舉名冊隨機抽樣，仍屬於傳統抽樣方式中，將每個樣本視為一個編碼，每個編碼被選中的機率相同，因此，任意更換兩組編碼並不會改變抽樣結果。並且，考量調查便利性與抽樣成本，最後樣本的分布仍然是主要城市較密、其餘較疏或無分布（圖二、左上；圖三、左上）。然而，若是依前人研究的結論：「投票行為具有空間相依性」，那麼主要城市所抽取的樣本數多便是無效的。因此，本文認為應該考慮到投票行為，其空間自相關（spatial autocorrelation）與空間異質性（spatial heterogeneity）的特徵，發展合適的抽樣方法來進行調查。

表二 2010年直轄市長選舉結果與 TEDS 訪查資料支持率比較

| 選舉資料與調查資料 | % | | | |
|----------------------------|------|------|-------|-------|
| | 國民黨 | 民進黨 | 投票率 | 樣本數 |
| 2010 台北市長選舉結果 ¹ | 39.0 | 30.7 | 70.74 | |
| 台北市 TEDS 資料 ² | 48.3 | 27.3 | 86.2 | 1,131 |
| 2010 中市長選舉結果 ¹ | 37.0 | 35.3 | 75.0 | |
| 台中市 TEDS 資料 ² | 41.2 | 29.5 | 87.2 | 1,168 |
| 2010 高市長選舉結果 ¹ | 14.7 | 37.9 | 75.23 | |
| 高雄市 TEDS 資料 ² | 12.8 | 42.2 | 87.5 | 1,177 |
| 楊秋興+國民黨 | | | | |
| 2010 高雄市長選舉結果 ³ | 33.9 | | | |
| 高雄市 TEDS 資料 ⁴ | 32.2 | | | 1,177 |

1. 資料來源為中央選舉委員會（2010）。
2. 資料來源為黃秀端（2010）。
3. 資料來源為作者根據 1 的資料作楊秋興與國民黨候選人支持率相加處理。
4. 資料來源為作者根據 2 的資料作楊秋興與國民黨候選人支持率相加處理。

在自然科學領域，例如：生態、環境、土壤、水資源、森林以及人類疾病等調查，樣本的空間布局對於研究結果影響很大，因此發展出各式空間抽樣方法，方法的核心是以空間中的點位為抽樣單位，順應各種條件而有不同的設計，除了點位為抽樣單位以外，也包含簡單隨機抽樣、系統抽樣、分層抽樣，以及其他抽樣方法設計（Earl, 1998）。「空間抽樣」技術雖然是以考慮空間特性為特徵的

一種抽樣方法，但是它的應用很廣且多元，例如在進行環境資源或社會經濟活動等調查時，能有較高的效率，即用較少的樣本數而獲得精確度較高的區域估計（Wang et al., 2013）。在文中，Wang et al.（2013）提出三種典型的空間抽樣應用：

1. 在未知、未做過調查的地區，建立樣點或採樣系統。即在空間中進行座標點位的抽樣，爾後調查員再依據衛星定位儀，到達獲選樣點進行資料蒐集，或是將研究範圍內的空間平均分成若干網格系統，在獲選網格內進行調查，此種抽樣的前提為空間中的每一個點位或網格皆是均質的（姜成晟等，2009）。
2. 在已知的監測系統（如氣象站）或資料庫中，加入考量空間特性參數的統計方式。
3. 在已完成的調查資料中（如市政滿意度報告），依據資料屬性、抽樣方法修正其統計方法，讓結果更精確。

過去我們認為空間抽樣技術只能應用在未知的目標領域，但其實，對於舊有的、已知的調查資料，也可應用空間抽樣方法，加入空間結構的特性，修正其統計方式，讓資料結果更具可信度。

TEDS 所調查的 2012 年總統選舉資料或是 2010 年直轄市長選舉資料皆屬於已完成的抽樣調查資料，因此，本文參考 Wang et al.（2013）所提出的第三個空間抽樣應用方法，提出：

- A. 空間結構加權分析方法：**乃將各縣市的空間結構以人口數多寡為加權參數，重新計算 2012 年總統選舉 TEDS 的電訪及面訪的資料。

由於空間效應會影響到投票行為，本文認為，單純以空間結構來修正舊有的調查資料，仍無法有效解決樣本間空間相依性的問題，再者，抽樣的維度除了一維以外，應該考慮到二維的空間母體，若

是在抽樣時，能以樣本與樣本之間的物理距離關係為參數，則可避免獲選樣本之間因具有空間相依性而產生的統計誤差，以下提出兩種三式，以樣本空間分布型態作依據的空間再抽樣方法：

B. 空間再抽樣方法：根據圖一，空間抽樣的調查對象，除了未知的調查區域，例如自然環境調查以外，對於已知、具有樣本的對象，仍可提出最佳的統計推論方式。順勢 TEDS 的調查方法（PPS 之分層抽樣），則「樣本分布」為「分層」，整體抽樣層級的分佈，即「目標領域」為「空間分層」，採樣結果因為考慮空間相依性的特徵，故「統計值」歸類於「空間分層」，與不考慮空間特徵的「分層」有所區別；遂此空間再抽樣方法的思考邏輯及流程為（圖一，B、1）：

母體已知 → 分層 → 空間分層 → 採就樣本資料分析 → 空間分層

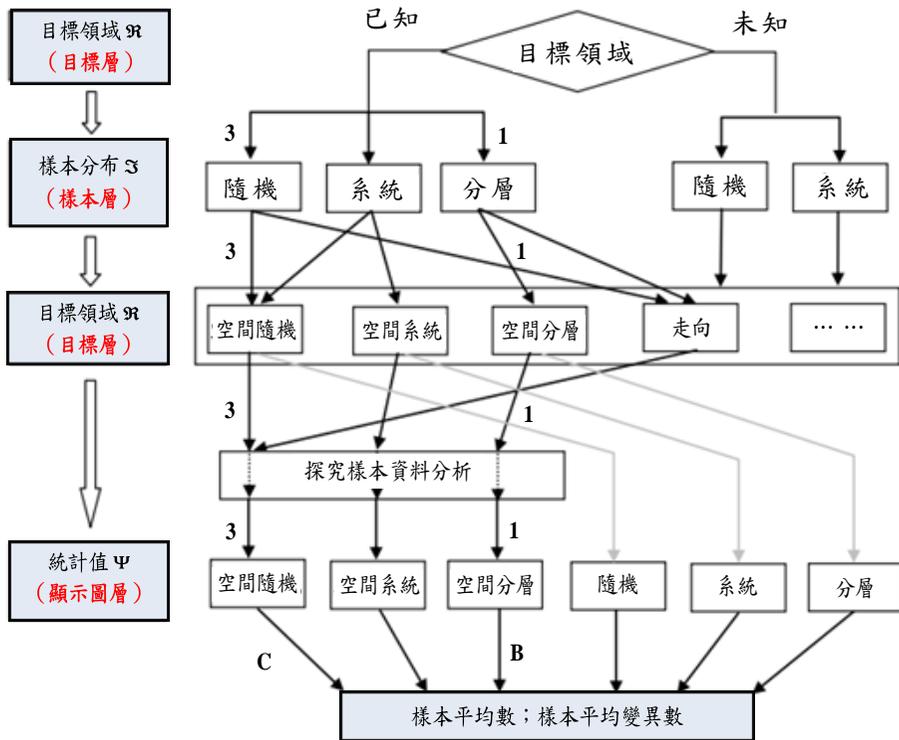
再者，作者考慮到若任意兩個樣本落在相鄰的行政區界上，仍無法避免空間相依性的存在，因此，在方法 B 的架構中，再依照「空間單元的界定」提出兩種子方法。

C. 空間抽樣方法：方法 B 是從舊有的資料再抽樣，是將原本樣本空間分布的型態作修正，因此沒有取樣的地方則不在抽樣設計考慮範圍，而方法 C 則是不受到原有樣本分布的侷限性，將全臺灣 358 個鄉鎮市視為抽樣對象，讓「樣本分布」為「隨機」，而這些對象所屬的「目標領域」為「空間隨機」，採樣結果因為考慮空間相依性的特性，故「統計值」歸類於「空間隨機」，與不考慮空間特徵的「隨機」有所區別；遂此空間抽樣方法的思考邏輯及流程為（圖一，C、3）：

母體已知 → 隨機 → 空間隨機 → 採就樣本資料分析 → 空間隨機

方法 B 與 C 在設計的考量上，有兩個著重點：第一、是希望選

取不具有空間相依性的樣本，讓樣本與樣本彼此的距離獨立，不受到鄰近效應影響。第二、是降低抽樣成本，減少樣本數；當樣本之間的距離獨立時，每一個樣本的代表性便不受到鄰近效應的影響。



圖一 空間抽樣流程圖

資料來源：Wang, Stein, Gao & Yong (2012).

(一) 以 2012 年總統選舉為例

2012 年總統選舉 TEDS 的電訪資料是由政治大學選舉研究中心執行訪問的（黃紀，2012），整個計畫是以 2012 年第 13 屆總統副

總統及第 8 屆立法委員選舉為研究主軸，從 2011 年 12 月 10 日起，至 2012 年 1 月 18 日止，由 TEDS 召集人黃紀教授召開討論會議，擬定問卷的內容包含基本資料、政治參與、選舉認知、議題立場、政黨偏好等五大區塊，以戶籍設於臺灣地區（不含金門、馬祖）年滿二十歲以上的成年公民為調查訪問的母群，並且受訪民眾僅限於住家民眾或住商混合的民眾。在執行上分成兩個主要部分：選前獨立與選後定群追蹤，選前共計完成五波獨立樣本，各有 967、976、951、958、954 筆成功樣本，而選後定群追蹤則以選前成功的 4806 筆成功樣本為母體進行訪問，成功的樣本數為 2,032 筆。選舉後的追蹤樣本資料，在抽樣方法上有兩種，一部分是以「中華電信住宅部 99-100 年版電話號碼簿」母體清冊來抽取，依據各縣市電話簿所刊電話數佔台灣地區所刊電話總數比例，決定各縣市抽出的電話數比例，接著採用等距抽樣法抽出各縣市電話樣本後，為求涵蓋範圍的完整性，則再以隨機亂數修正電話號碼的最後二碼或四碼，以求接觸到未登錄電話的住宅戶。另一部分的樣本則是來自政治大選選舉研究中心所累積的電訪資料庫，以隨機亂數修正電話號碼的最後四碼來製作電話樣本。共在全台 170 個不同的村里進行電訪。在本文後續研究中，皆以選舉後的定群追蹤樣本為分析資料，雖然選舉後的追蹤資料，可能發生受訪者根據選舉結果回答的風險，但也可能發生選舉當日跑票的情形，因此在斟酌資料使用時，本文仍以選舉後的追蹤樣本為分析資料。¹

1. 在文章中所放的 TEDS2012 電訪總樣本分布圖，是以選前成功的 4,806 筆樣本為依據，這些樣本數為選後定群追蹤的母體，因此作者以 4,806 筆樣本的所在地與樣本數來繪圖；根據「臺灣選舉與民主化調查 2012 總統與立法委員選舉電話訪問之研究設計」報告書提到，選後定群追蹤的成功樣本數為 2,032 筆，因此文章中所用來進行分析的電訪樣本數皆為 2,032 筆成功樣本數。

TEDS 的面訪資料為「2009 年至 2012 年『選舉與民主化調查』三年期研究規劃」之研究案，在三年期的計畫規畫中，第一年為「民國九十八年縣市長選舉面訪案」，針對縣市長、鄉鎮市長與縣市議員三合一選舉所進行的面訪調查；第二年為「民國九十九年直轄市市長選舉面訪案」，選取臺北市、大台中市（含舊台中縣、市）、大高雄市（含舊高雄縣、市）進行面訪；第三年則是「民國一百零一年總統與立法委員選舉面訪案」（朱雲漢，2012），希望藉由探知選舉中民意取向，讓相關研究者能夠對於選民的政治態度與投票行為等議題有進一步的探索與了解；本文在第三章第一節的總統選舉探討中，使用第三年期的研究成果，在第三章第二節的直轄市長選舉探討中，則使用第二年期的研究成果。

面訪調查是運用「抽取率與抽出單位大小成比例」（probabilities proportional to size，簡稱 PPS）的原則進行獨立樣本的抽樣，獨立樣本實際進行訪問的對象以設籍於臺灣地區且年齡在二十歲以上具有選舉權的公民為訪問母群，第一階段先依照六大地理區抽出選區，第二階段再由選區中抽出村里。確立中選的村里，以及村里的代碼等資訊之後，第三階段再交由政大選舉研究中心申請資料，並抽取正選樣本與預備樣本，並在正式訪談結束後，針對成功的獨立樣本進行再測信度訪問。共計完成 1,826 份選前的獨立樣本，以及 1,510 份選後的追蹤樣本，面訪案在全台 315 個不同的村里進行（其中包含 5 個重複村里），涵蓋選區高達 60 個，占全臺灣選區的 82%。²

I. 電訪資料

TEDS 的電訪資料為已知的調查資料，本文就已知樣本分布特

2. 文章中所放的 TEDS2012 面訪總樣本分布圖，是以 1,826 份成功的獨立樣本為依據，而後續用來進行分析的樣本則是選後的 1,510 筆追蹤樣本。

性重新進行估算，提出：

A. 空間結構加權分析：

依據國土資訊系統、社會經濟資料庫於民國 102 年所作的全臺灣以縣市為單位的人口統計資料，分成四個層級，再將 TEDS 所作的各縣市抽樣結果，包括投票率與總統支持率的平均值與之加權，得到重新分析過後的平均值。³

B. 空間再抽樣方法 1：

由於電訪資料共抽取 170 個村里，為了能盡量保有原村里數，空間再抽樣方法 1 設定再抽樣的資料量大於 170 個樣本，約為原資料量的 5~10%，在此以 5% 為抽樣比例；接著，以縣市為單位作分層抽樣，重新抽取各縣市 5% 樣本，累計為 239 筆；決定好抽樣比例後，將原資料以空間分布型態展示（圖二），依照每一筆資料的座標系統，計算兩兩之間的距離，以隨機抽樣的方式，選擇 5% 的樣本數，且這些再抽樣的樣本分布是均質的。⁴ 以上再抽樣過程採十重複，以十次結果的平均值為結果（表三）。⁵

資料所使用的座標系統為 TWD97；方法中所指的樣本在空間中的兩兩距離，為某一個樣本與其他所有樣本之間

-
3. 四個層級分別是縣市人口數達兩百萬以上、一百萬以上不足兩百萬、五十萬以上不足一百萬、以及五十萬以下。在加權方式上，0~50 萬人口的縣市投票率或候選人得票率×1（得一分）、50~100 萬人口的縣市投票率或候選人得票率 ×2（得兩分）、100~200 萬人口的縣市投票率或候選人得票率 ×3（得三分）、200 萬人口以上的縣市投票率或候選人得票率×4（得四分），重新計算過後的投票率或候選人得票率加總後，再除以所有縣市加總的得分，得到新的投票率或候選人得票率的結果。
 4. 檢定樣本的分布型態時，主要是計算某一個點，與其他點的最大距離(b)與最小距離(a)， $f(x)=1/b-a$ ， $f(x)$ 會落在 0~1 之間，再判斷 $f(x)$ 值的均質程度。
 5. 以上操作使用 ArcGIS9.3 版軟體。

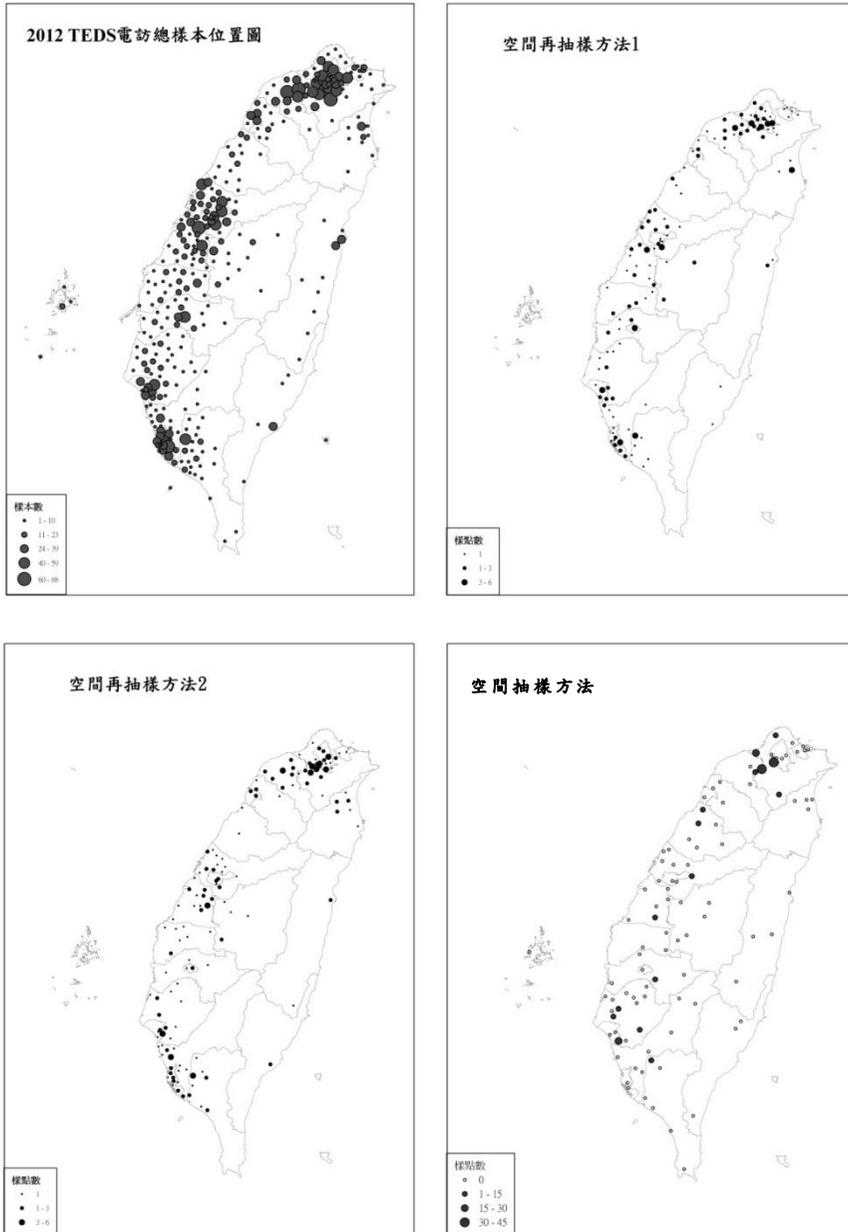
的直線距離，抽樣時，第一個樣本為隨機抽樣，接著，根據與該樣點的直線距離計算空間相關性，在沒有空間相關性的樣點中再隨機抽取一個樣本，為第二個樣本，而第三個樣本同樣也是與第一及第二個樣本沒有空間相關性的範圍中抽取出，以此類推，最後再檢定獲選樣本的分布是否為隨機分布。

空間再抽樣方法 2：

比照空間再抽樣方法 1，同樣以原資料的 5% 為抽樣比例，但是不根據縣市分層，而是直接以所有樣本在空間中的兩兩距離為依據，讓再抽樣後樣本在空間中分布是均質且隨機的（圖二）。以上再抽樣過程採十重複，以十次結果的平均值為結果（表三）

C. 空間抽樣方法：

以 358 個鄉鎮市為空間單元，進行空間隨機抽樣，抽取到的鄉鎮市從 TEDS 資料中選取樣本，若是 TEDS 沒有調查到的區域，則留白，由於 TEDS 的調查地點較為集中，空間抽樣後的獲選鄉鎮市較為分散，為了讓樣本數與前兩種抽樣方式相仿，故在進行 358 個鄉鎮市空間隨機抽樣時，以 100 個鄉鎮市為抽樣數（圖二）。操作方法使用 ArcGIS9.3 版軟體介面，將 358 個鄉鎮市的中心點標示出來，隨機選取 100 個在空間分布型態上，符合隨機分布的樣本。



圖二

(左上) 2012 總統選舉 TEDS 電訪資料之樣本數與空間分布圖；(右上) 空間再抽樣方法 1 是根據 TEDS 資料所進行的再抽樣方式，其樣本是從 TEDS 資料(左上)取樣而得，樣本數與空間分布如圖所示。(左下) 空間再抽樣方法 2 是根據 TEDS 資料所進行的再抽樣方式，其樣本是從 TEDS 資料(左上)取樣而得，樣本數與空間分布如圖所示；(右下) 空間抽樣方法是以 358 個鄉鎮市為母體進行抽樣，共抽取 100 個鄉鎮市，獲選的鄉鎮市若 TEDS 有受訪資料，則以實心表示，空心則表示缺值。

資料來源：作者以 ArcGIS9.3 軟體自繪。(左上圖) 2012 電訪總樣本位置圖乃參考黃紀(2012)。

II. 面訪資料

同樣的，TEDS 的面訪資料也為已知的調查資料，依照樣本分布特性重新進行估算，提出：

A. 空間結構加權分析：

依據國土資訊系統、社會經濟資料庫於民國 102 年所作的全臺灣以縣市為單位的人口統計資料，分成四個層級，再將 TEDS 所作的各縣市(新竹市、基隆市沒有獲選樣本，除外)抽樣結果，包括投票率與總統支持率的平均值與之加權，得到重新分析過後的平均值。

B. 空間再抽樣方法 1：

根據 TEDS 資料的空間資訊，共抽取 315 個村里，因此再抽樣時，為了能盡量保有原村里數，認定再抽樣資料量要大於 315 個樣本，約為原資料量的 20%，接著，以縣市為單位作分層抽樣，重新抽取各縣市 20% 樣本，累計為 364 筆；決定好抽樣比例後，將原資料以空間分布型態展示(圖三)，依照每一筆資料的座標系統，計算兩兩之間的距離，以隨機

抽樣的方式，選擇 20% 的樣本數，且這些再抽樣的樣本分布是均質的。以上再抽樣過程採十重複，以十次結果的平均值為結果（表三）。

空間再抽樣方法 2：

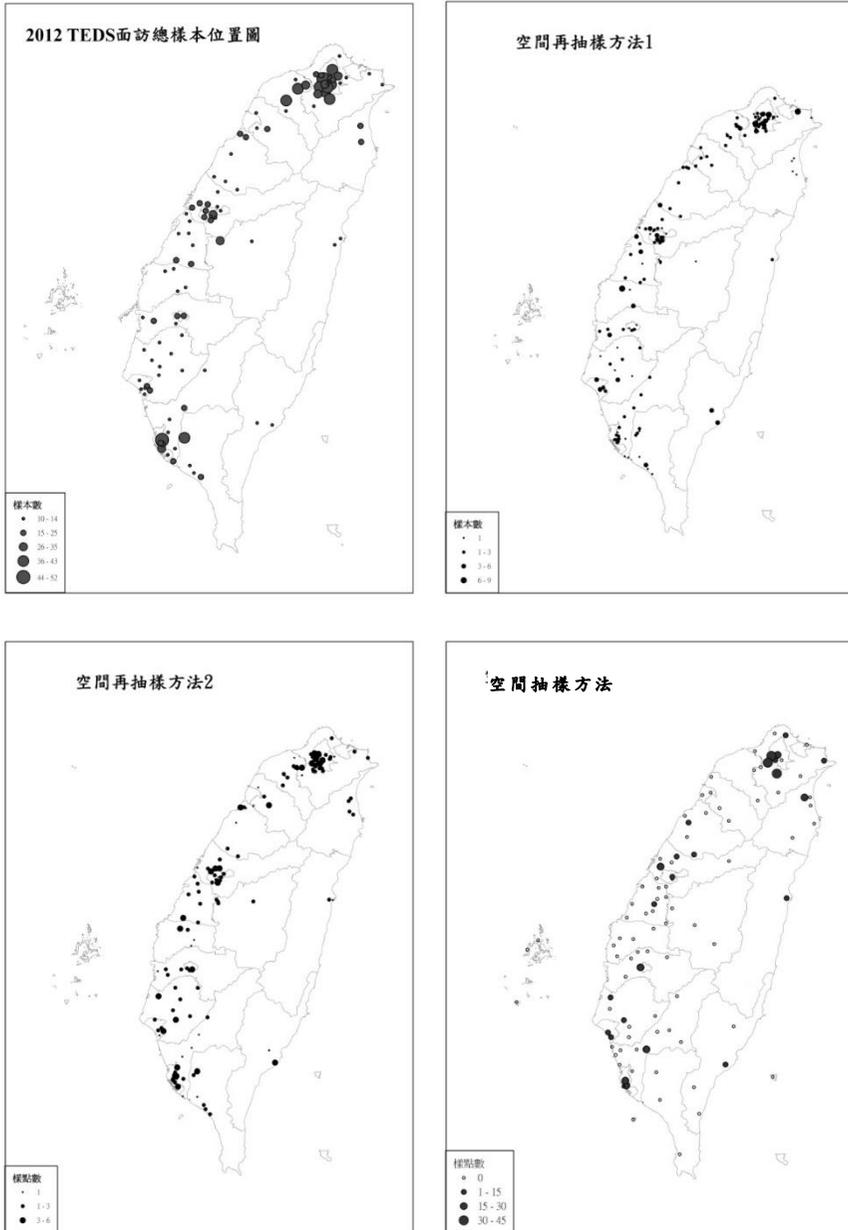
比照空間再抽樣方法 1，同樣以原資料的 20% 為抽樣比例，但是不根據縣市分層，而是直接以所有樣本在空間中的兩兩距離為依據，讓再抽樣後樣本在空間中分布是均質且隨機的（圖三）。以上再抽樣過程採十重複，以十次結果的平均值為結果（表三）。

C. 空間抽樣方法：

以 358 個鄉鎮市為空間單元，進行空間隨機抽樣，抽取到的鄉鎮市從 TEDS 資料中選取樣本，若是 TEDS 沒有調查到的區域，則留白，由於 TEDS 的調查地點較為集中，空間抽樣後的獲選鄉鎮市較為分散，為了讓樣本數與前兩種抽樣方式相仿，故在進行 358 個鄉鎮市空間隨機抽樣時，以 100 個鄉鎮市為抽樣數（圖三）。

從問卷題庫中，選擇「這次的選舉是否有去投票？」以及「請問這次選舉是投給哪一位候選人」兩個問題，進行三種空間再抽樣方法的推估，並將「馬英九與吳敦義支持率」、「蔡英文與蘇嘉全支持率」、以及整體「投票率」的推估結果，整理為表三，與 2012 總統選舉的結果及 TEDS 的訪查資料作比較。⁶

6. 本文使用 TEDS 釋出的原始資料，並且將無反應項（包括還沒有決定、不會去投、投廢票、看情形、不知道與拒答）納入計算百分比之分母。



圖三

(左上) 2012 總統選舉 TEDS 面訪資料之樣本數與空間分布圖；(右上) 空間再抽樣方法 1 是根據 TEDS 資料所進行的再抽樣方式，其樣本是從 TEDS 資料(左上)取樣而得，樣本數與空間分布如圖所示；(左下) 空間再抽樣方法 2 是根據 TEDS 資料所進行的再抽樣方式，其樣本是從 TEDS 資料(左上)取樣而得，樣本數與空間分布如圖所示；(右下) 空間抽樣方法是以 358 個鄉鎮市為母體進行抽樣，共抽取 100 個鄉鎮市，獲選的鄉鎮市若 TEDS 有受訪資料，則以實心表示，空心則表示缺值。

資料來源：作者以 ArcGIS9.3 軟體自繪。(左上圖) 2012 面訪總樣本位置圖乃參考朱雲漢(2012)。

(二) 以 2010 直轄市長選舉為例

從上一節中，我們得到空間抽樣技術在估算全國性調查資料時，是能降低成本且更有效率的，若是以台灣呈南北狹長的地理結構也能夠進行以「物理距離」為抽樣的依據，那麼地方性的選舉，縣市的幅員更為方正來看，也許可以估算的更準確，因此本文接著以直轄市長選舉為例，試算空間抽樣方法是否能應用在地方性選舉上。

TEDS 於 2010 年在台北市、大台中市(舊台中縣、市)、大高雄市(舊高雄縣、市)作了直轄市長選舉調查資料，同樣依照「抽取率與抽出單位大小成比例」的原則，但因為都市化程度不同，台北市採用「兩階段抽樣法」，即第一階段由每一個行政區(鄉鎮市層級)抽出里、第二階段則從村里抽出受訪者。台中市與高雄市由於縣市合併後，幅員廣大，且各行政區(鄉鎮市層級)變異相當大，因此先以十四項人文區位變數進行分層，⁷ 共分得五層，再採用「三

7. 人文區位的變數包括：人口密度、大專畢業以上人口比例、農漁戶數比例、工廠登記家數、自來水普及率、稅課(決算)收入、歲計(決算)支出、教科文支出/人口、人口/醫生、

階段抽樣法」，即第一階段從人文區位分層中抽出區、第二階段抽出里、第三階段再抽出受訪者（黃秀端，2010）。

這些資料為已知的調查資料，根據其抽樣方式，本文就已知樣本分布特性重新進行估算，提出：

A. 空間再抽樣方法 1：

根據 TEDS 資料的空間資訊，為了降低再抽樣樣本數，又要大於原抽樣的村里數，故以「區」作分層、區內樣本的 10% 為抽樣比例，進行分層抽樣（圖四、右上、以高雄市為例），共抽取 113 筆樣本。以上再抽樣過程採十重複，以十次結果的平均值為結果（表四）。

B. 空間再抽樣方法 2：

比照空間再抽樣方法 1 的樣本數，取整數 100 個樣本為抽樣群體，與方法 1 不同的是，方法 2 是以所有樣本本身的物理距離為參考，採取簡單隨機抽樣所獲得（圖四、左下、以高雄市為例）。以上再抽樣過程採十重複，以十次結果的平均值為結果（表四）。

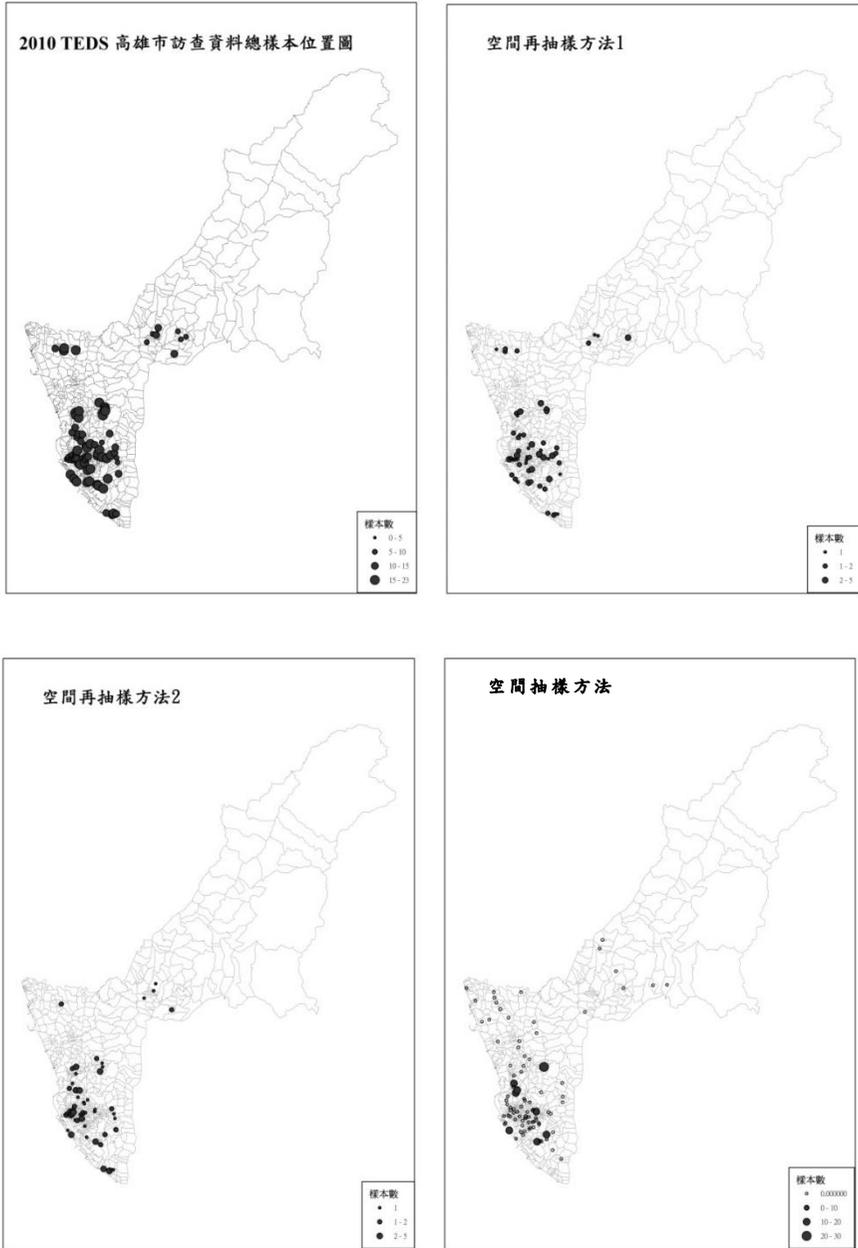
C. 空間抽樣方法：

以市內所有的里為抽樣單位，抽取整體 10% 的里為訪查對象，獲選的里若有 TEDS 的調查資料則進行計算，沒有者則是為缺值（圖四、右下、以高雄市為例）。

從問卷題庫中，選擇「這次的選舉是否有去投票？」以及「請問這次選舉是投給哪一位候選人」等問題，進行三種空間再抽樣方法的推估，並將「國民黨候選人支持率」、「民進黨候選人支持率」、以及整體「投票率」的推估結果，整理為表四（1、2），與 2010 年直轄市長選舉的結果及 TEDS 的訪查資料作比較。⁸

六十五歲以上人口比例、遷入人口比例、遷出人口比例、人口流動比例、以及青壯年（二十至三十九歲）人口比例等十四項。

8. 由於高雄市的選情特殊，因此將楊秋興的支持率考慮進來，新增一項與國民黨合併的資料。



圖四

(左上) 2010 年直轄市長選舉，TEDS 高雄市長資料之樣本數與空間分布圖；(右上) 空間再抽樣方法 1 是根據 TEDS 資料所進行的再抽樣方式，其樣本是從 TEDS 資料(左上)取樣而得，樣本數與空間分布如圖所示。(左下)空間再抽樣方法 2 是根據 TEDS 資料所進行的再抽樣方式，其樣本是從 TEDS 資料(左上)取樣而得，樣本數與空間分布如圖所示；(右下)空間抽樣方法是以里為空間單元進行抽樣，共抽取母體 10%里數，獲選的里若 TEDS 有受訪資料，則以實心表示，空心則表示缺值。

資料來源：作者以 ArcGIS9.3 軟體自繪。(左上圖) 2010TEDS 高雄市訪查資料總樣本位置圖乃參考黃秀端(2010)。

四、研究發現

(一) 以 2012 年總統選舉為例

從表三的結果得到：已知的調查資料，透過族群量差異的空間結構加權調整後(A)，支持率及投票率與原本的資料並無明顯的差異，仍不夠貼近真實的結果；作者將空間再抽樣方法或是空間抽樣方法，所估算的結果，與 TEDS 的數值作顯著性差異檢定，並無顯著性差異，但是其所估算的結果卻比 TEDS 的資料更接近真實結果，且樣本數皆只有 TEDS 資料的 20%、5%、甚至更低，表示在進行投票行為調查時，考慮取樣的空間效應是必要且能更有效的推估，成本低也較為經濟。

表三 各種空間抽樣方式所獲得的總統支持率與投票率之估計
值比較（以 2012 年）

| 抽樣方法 | % | | | |
|------------------------------|------|------|------|-------|
| | 國民黨 | 民進黨 | 投票率 | 樣本數 |
| 2012 總統選舉結果 ¹ | 38.1 | 33.7 | 74.4 | |
| TEDS 電訪資料² | 44.8 | 25.2 | 77.2 | 2,032 |
| A 空間結構加權分析 | 45.4 | 24.3 | 76.6 | 2,032 |
| B 空間再抽樣 1 | 46.1 | 25.9 | 79.9 | 239 |
| — 分層抽樣法，以縣市為單位 | | | | |
| 空間再抽樣 2 | 43.9 | 28.9 | 78.2 | 239 |
| — 以樣本間物理距離隨機抽樣 | | | | |
| C 空間抽樣 | 39.6 | 27.8 | 67.5 | 212 |
| — 以鄉鎮市間物理距離隨機抽樣 | | | | |
| TEDS 面訪資料³ | 46.0 | 30.1 | 89.2 | 1,510 |
| A 空間結構加權分析 | 45.4 | 29.2 | 88.5 | 1,510 |
| B 空間再抽樣 1 | 45.9 | 26.9 | 86.0 | 364 |
| — 分層抽樣法，以縣市為單位 | | | | |
| 空間再抽樣 2 | 43.7 | 30.5 | 89.0 | 364 |
| — 以樣本間物理距離隨機抽樣 | | | | |
| C 空間抽樣 | 45.7 | 31.1 | 89.3 | 457 |
| — 以鄉鎮市間物理距離隨機抽樣 | | | | |

1. 資料來源為中央選舉委員會（2012）。

2. 資料來源為黃紀（2012）。

3. 資料來源為朱雲漢（2012）。

以電訪資料來看，空間抽樣方法（C）所推估的國民黨支持率及投票率，皆是最接近真實得票率與投票率，民進黨支持率的推估值的確也能夠提高 TEDS 的推估效果，更貼近真實的得票情形；因此，以估算效果來看，空間抽樣方法（C）的推估最佳，能同時在兩黨的支持率與投票率的推估上，獲得顯著的效果。

就面訪資料來看，則是以空間再抽樣方法 2 與空間抽樣方法（C）的推估效果較好，能夠在國民黨與民進黨的支持率上，有效的將 TEDS 的原始資料趨近選舉結果，投票率的估算也能降低 TEDS 的估算結果與真實投票情形的誤差，因此在兩種三式的估算效果比較上是較為理想的。

面訪與電訪資料在重新作空間抽樣時，推估效果較佳的抽樣方法並不相同，以電訪而言，空間抽樣方法（C）較佳；以面訪而言，空間再抽樣方法 2 與空間抽樣方法（C）較為理想。

空間再抽樣方法 2 與空間抽樣方法（C）在面訪資料的推估結果不相上下，會有這樣的差異性，主要是受到原始資料分布的問題，面訪資料因為樣本分布較為侷限，許多空間抽樣方法（C）選擇到的鄉鎮市為缺值，因此在估計上不如面訪資料來的推估有效。即使能夠有效用來估算兩種訪查資料的空間抽樣方法不同，但相同的是，這兩種方法顯示出沒有分層的抽樣結果較有分層的更接近真實值，這是因為空間抽樣的前提是避免樣本之間的空間自相關，當分層抽樣是以行政區（縣市或鄉鎮市）為單元時，在不同分層的獲選樣本可能坐落在相鄰邊界上，仍會有樣本間空間自相關的問題；因此，在進行空間抽樣時，分層的抽樣可能不能單純以地理位置作分層，而是以其他人文社經條件來區隔，會更加理想。

(二) 以 2010 直轄市長選舉為例

這三個直轄市的抽樣結果略有不同，以鄉鎮市層級來看，台北市是每一個區內皆有抽樣樣本，台中市與高雄市則是部分區域內採樣，因此，空間再抽樣方法 1 與 2 (A、B) 對於重新估算台北市的候選人支持率及投票率是較空間抽樣方法 (C) 來的接近 TEDS 的抽樣數據，原因是空間抽樣方法 (C) 所抽取的樣點雖然分布上最為隨機，但是是以「里」為抽樣單元，礙於 TEDS 的資料限制（每個區內僅約有 4 個里有樣本分布），獲選樣本有部分缺值，雖然無法估算完整，但是能減少過於密集的樣本之間的鄰近關係，反而能夠有效的推估真實選情。而台中市與高雄市並沒有每個區內皆進行取樣，當進行空間再抽樣時，空間再抽樣方法 1 及 2 (A、B)，與空間抽樣方法 (C)，皆能調整 TEDS 的推估結果，但較無法看出哪一個方法最有效。

三種空間再抽樣方法的樣本數約略為 100 個樣本，是 TEDS 樣本數的十分之一，但是所估算的結果，受限於 TEDS 本身資料與真實情形的差距，雖然不能準確的推估，卻比 TEDS 的資料更接近於真實情形，顯示，利用空間抽樣技術，的確可以降低取樣的成本，且獲得更精確的估算。

表四(1) 各種空間抽樣方式所獲得的市長支持率與投票率之
估計值比較

| 抽樣方法 | % | | | 樣本數 |
|-------------------------------|------|------|------|-------|
| | 國民黨 | 民進黨 | 投票率 | |
| 2010 台北市長選舉結果 ¹ | 39.0 | 30.7 | 70.7 | |
| 台北市 TEDS 資料 ² | 48.3 | 27.3 | 86.2 | 1,131 |
| A 空間再抽樣方法 1 — 分層抽樣法，以區為單位 | 50.4 | 23.0 | 89.4 | 113 |
| B 空間再抽樣方法 2 — 以樣本間物理距離隨機抽樣 | 44.0 | 34.0 | 87.0 | 100 |
| C 空間抽樣方法 — 以區間物理距離隨機抽樣 | 43.7 | 28.2 | 77.5 | 71 |
| 2010 台中市長選舉結果 ¹ | 37.0 | 35.3 | 75.0 | |
| 台中市 TEDS 資料 ² | 41.2 | 29.5 | 87.2 | 1,168 |
| A 空間再抽樣方法 1 — 分層抽樣法，以區為單位 | 40.0 | 30.0 | 94.0 | 100 |
| B 空間再抽樣方法 2 — 以樣本間物理距離隨機抽樣 | 35.6 | 29.7 | 82.2 | 118 |
| C 空間抽樣方法 — 以區間物理距離隨機抽樣 | 43.3 | 29.1 | 91.5 | 141 |

1. 資料來源為中央選舉委員會（2010）。

2. 資料來源為黃秀端（2010）。

表四(2) 各種空間抽樣方式所獲得的市長支持率與投票率之估計值比較

| 抽樣方法 | % | | | | 樣本數 |
|-------------------------------|------|------|--------------------|------|-------|
| | 國民黨 | 民進黨 | 國+楊秋興 ³ | 投票率 | |
| 2010 高雄市長選舉結果 ¹ | 14.7 | 37.9 | 33.9 | 75.2 | |
| 高雄市 TEDS 資料 ² | 12.8 | 42.2 | 32.2 | 87.5 | 1,177 |
| A 空間再抽樣方法 1 — 分層抽樣法，以區為單位 | 15.3 | 43.2 | 29.7 | 89.8 | 118 |
| B 空間再抽樣方法 2 — 以樣本間物理距離隨機抽樣 | 17.0 | 35.0 | 37.0 | 85.0 | 100 |
| C 空間抽樣方法 — 以區間物理距離隨機抽樣 | 17.1 | 43.1 | 37.4 | 91.1 | 123 |

1. 資料來源為中央選舉委員會（2010）。
2. 資料來源為黃秀端（2010）。
3. 資料來源：於 2010 高市長選舉結果欄，為作者根據 1 的資料，作楊秋興與國民黨候選人支持率相加處理；於高雄市 TEDS 資料欄，為作者根據 2 的資料，作楊秋興與國民黨候選人支持率相加處理。

過去對於調查資料的估算，總是樣本數越多，越能趨近於母體，然而，從上一章的結果來看，無論是總統選舉、或是直轄市長選舉，基於調查成本考量與實際操作的可行性，一般是選擇抽樣調查，無法做到普查；因此，從樣本試算而來的估計值，與母體的實際值之間是有差距的，這部分可以用資料缺失值（missing data）或是抽樣偏差（sampling bias）來校正，本文便是使用後者來進行已知樣本的再估算。結果發現，無論是將全國性的調查資料、或是地方性的調查資料重新估算，我們認為，將投票行為的空間相依性考慮進抽

樣設計後，不但可以減少取樣數，也能得到較精確的估計值；尤其空間抽樣方法（C）的估算結果在總統選舉以及直轄市長選舉中皆有較佳的表現，目前，受限於 TEDS 的調查資訊，空間抽樣方法（C）無法有完整的有效資料，若是未來可先以空間抽樣技術抽取調查樣本，再進行調查，相信更能提高調查的準確性。

五、討論與結語

傳統抽樣方法考量城鄉人口的比例差距，會先對母群體進行「分層」，包括「人口」、「族群」或「社經背景」等人文因子，讓每一層級皆有代表樣本，達到各層級個體的抽樣機率相等；但是，考慮抽樣便利與抽樣成本，這些在選舉時期所進行的抽樣調查，樣本的地理位置會有聚集性，然而，以考量抽樣便利與抽樣成本的操作方式，往往沒有顧慮到這些樣本之間是不是存在鄰近效應，且樣本之間對於選舉行為與態度是否互相影響？本文認為，若在一個空間單元中要進行三回合抽樣調查，那麼，三組調查人馬，在同一個區域分頭抽樣訪談，不如一組調查人馬，分別在三個區域內進行抽樣訪談，也許投入人力有限，能採樣到的受訪戶較少，但是在不相鄰的區域內進行訪談，較能取得該空間單元內，客觀的選民選舉行為與態度的資訊。作者在表三中，將兩種三式的空間抽樣方法所推估的結果，與 TEDS 的推估值進行顯著性差異檢定，得到空間抽樣方法的推估值與 TEDS 的推估值無顯著差異，這表示，我們的操作方法，將舊有的調查資料作篩選，去掉有鄰近關係影響的樣本，並無不完善之處，反而透過空間抽樣方法的推估來修正抽樣偏差，能夠提升調查資料的準確性。

傳統抽樣方法在人文因子之外，也會考慮到以「選區」、「鄉鎮市區」來分層，這種分層方式的確考慮到「空間」上的分配，但是行政區的分級考量並不一定能真的做到樣本在空間上的分布均勻。

然而這些調查研究案除了了解選舉民意之外，同時也在問卷設計中羅列了許多可以作為選舉研究探測的議題與變數，若是這些變數因著樣本之間的自相關性而影響到分析結果，那麼對於後續使用者，需要以空間特性來探討這些議題與變數時，可能要透過空間自相關的計算，刪除掉某些樣本，像是本文所使用的空間再抽樣方法 2，或者透過其他方式來篩選避免樣本間的自相關性。

當然，臺灣聚落的分布型態，本身就屬於「聚集」，如何從空間中呈現聚集的母群體抽取具有代表性的樣本，且樣本本身還能兼顧傳統抽樣中，人文分層的訴求，本文認為可以從抽樣設計開始就加入樣本空間特性的考量，像是空間抽樣方法（C），目的是達到樣本符合隨機分布，將「空間」也是「距離」這個因子放入操作程序中：選擇母群體→選擇抽樣結構→決定樣本大小→進行空間抽樣。而母群體的選擇，會依照不同的選舉目的作改變，例如：總統大選的空間框架是全臺灣；縣市長選舉的空間框架是單一縣市；立委選舉的空間框架則是單一選區（以區域立委為例）。以總統選舉為例，若是以選區為抽樣單位，容易造成只有單一選區的縣市在隨機抽樣的程序中落選，像是花蓮市、臺東市、宜蘭市、新竹市與基隆市，因此本文認為，在選擇空間結構時，需要考慮到行政區與選區的分配特性。

樣本數的大小，與傳統抽樣相同，可以依照人口數設定合適的樣本比例。最後操作空間抽樣時，第一個樣本為隨機，往後的抽樣樣本的抽取原則，則根據所有選區與獲選選區之間的距離來考慮，

在距離上已經與獲選樣本有空間相關性時，則不列入抽取的對象中，當抽樣結束後，再對獲選樣本進行空間相關性的分析，確定為隨機分布。

當研究議題有空間分布的特性時，空間抽樣技術便能應用在抽樣設計之中，讓獲選樣本符合討論前提，則在後續使用上，讓每一筆樣本都能有效的使用。

最後，文中的空間抽樣方法，是探討將「空間特性」考慮進樣本的選擇中，希望透過空間特性的篩選，取得空間分布獨立的樣本；然而，這些樣本的代表性、與如何回推其所能代表的群體（母體中的哪些族群），是在未來延伸中，可以繼續探討的有趣議題。作者以為，一直以來，相關空間分析的研究，主要從二維的空間面向來討論，「受到距離的影響，越鄰近的家戶所表現出來的選舉行為會越相近」，在這個空間尺度中，我們所關心的距離，是每一個門牌號碼，彼此間所在位置的直線距離，但隨著人口密度增加、都市化程度提升，聚落的建築形式，逐漸從平房轉變為華廈與大樓，家戶位置的分布型態，也許從三維的空間面向來表示，會更為貼切；那麼，兩組相近的門牌號碼，可能會因為分屬不同大樓、社區（例如：幸福街 15 號 13 樓與幸福街 41 號 7 樓），增加 Z 軸的計算後，這兩個樣本在空間上的關係是獨立的。目前，調查研究實地進行面訪時，主要是以獲選門牌號碼的一樓為取樣對象，比較是將母體分布降維度抽樣；但是，電訪的資料就不一定侷限於二維的空間尺度了，因此電訪資料較面訪資料的推估效果佳，不只是樣本數較多，也較能從正確的三維家戶分布中抽取樣本。因此，作者認為，未來在進行調查研究時，除了可以將空間特性帶入抽樣設計中，若能從正確的三維家戶母體分布中抽取有效的樣本，必定有助於提升推估成果。

參考書目

- Earl, B. 1998. *The Practice of Social Research*. 9th ed. Belmont, CA :Wadsworth Pub. Co.
- Liu, F. C.S. and Y. S. Su. 2015. “Adjusting Survey Response Sistributions Using Multiple Imputation: A Simulation with External Validation.” *Research Article* 34: 8-32.
- Wang, J.F., A. Stein, B. B. Gao and G. Yong. 2012. “A Review of Spatial Sampling.” *Spatial Statistics* 2: 1-14.
- Wang, J.F., C. S. Jiang, M. G. Hu, Z. D. Gao, Y. S. Guo, L. F. Li, T. J. Liu, and B.Meng. 2013. “Design-based Spatial Sampling: Theory and Implementation.” *Environmental Modelling and Software* 40: 280-288.
- 中央選舉委員會。2010。〈選舉資料庫—2010、99年直轄市長選舉〉。
<http://db.cec.gov.tw/histMain.jsp?voteSel=20101101B1>。2013/08/10。（Central Election Commission. 2010. “Election Database of The Mayor of the Municipality.” in <http://db.cec.gov.tw/histMain.jsp?voteSel=20101101B1>. Latest update 10 August 2013.）
- 中央選舉委員會。2012。〈選舉資料庫—2012 第十三任總統與副總統選舉〉。
<http://db.cec.gov.tw/histMain.jsp?voteSel=20120101A1>。2013/08/10。（Central Election Commission. 2012. “Election Database of The Thirteenth President and Vice Presidential Eletion.” in <http://db.cec.gov.tw/histMain.jsp?voteSel=20120101A1>. Latest update 10 August 2013.）

- 朱雲漢。2012。〈2009 年至 2012 年「臺灣選舉與民主化調查」三年期研究規畫(3/3)：2012 年總統與立法委員選舉面訪案 (TEDS2012)〉。國科會專題研究計畫報告 (NSC 100-2420-H-002-030)。台北：國家科學委員會。(Chu, Y. H. 2012. "Taiwan's Election and Democratization Study, 2012 (TEDS2012) : The Survey of Presidential and Legislative Elections." National Science Council Research Project (NSC 100-2420-H-002-030). Taipei: National Science Council.)
- 林昌平、吳怡慧、徐永明。2015。〈空間推論與政治行為：最大熵方法於調查研究的應用，TEDS2012〉。《地理學報》76：69-95。(Lin, C. P., I. H. Wu and Y. M. Hsu. 2015. "Spatial Inference and Political Behavior : The Application of Maximum Entropy Method for TEDS2012 Survey Data." *Journal of Geographical Science* 76: 69-95.)
- 黃秀端。2010。〈2009 年至 2012 年「選舉與民主化調查」三年期研究規劃(2/3)：民國九十九年直轄市市長選舉面談訪案〉。國科會專題研究計畫報告 (NSC 99-2420-H-031-002)。台北：國家科學委員會。(Hawang, Shio-Duan. 2010. "Taiwan's Election and Democratization Study, 2010 (TEDS 2010C): Taipei, Taichung, and Kaohsiung Cities Mayoral Elections." National Science Council Research Project (NSC 99-2420-H-031-002). Taipei: National Science Council.)
- 黃紀。2012。〈2009 年至 2012 年「臺灣選舉與民主化調查」三年期研究規畫(3/3)：2012 年總統與立法委員選舉電訪案 (TEDS2012-T)〉。國科會專題研究計畫報告 (NSC

- 100-2420-H-002-030)。台北：國家科學委員會。(Huang, C. 2012. “Taiwan’s Election and Democratization Study, 2009-2012(III): Telephone Interview of the Presidential and Legislative Elections, 2012(TEDS2012-T).” National Science Council Research Project (NSC 100-2420-H-002-030). Taipei: National Science Council.)
- 姜成晟、王勁峰、曹志冬。2009。〈地理空間抽樣理論研究綜述〉。《地理學報》64，3：368-380。(Jiang, C. S., J. F. Wang and Z. D. Cao. 2009. “A Review of Geo-Spatial Sampling Theory.” *Acta Geographica Sinica* 64, 3: 368-380.)
- 童振源、周子全、林繼文、林馨怡。2011。〈2009年臺灣縣市長選舉預測分析〉。《選舉研究》18，1：69-94。(Tung, C. Y., T. C. Chou, J. W. Lin, and H. Y. Lin. 2011. “Analysis on the Prediction Results of the 2009 Magistrate and Mayoral Election in Taiwan.” *Journal of Electoral Studies* 18, 1: 69-94.)
- 楊佳芳。2002。〈利用EM演算法分析公元2000年總統大選民調資料〉。國立中央大學統計研究所碩士論文。(Yang, J. F. 2002. “Using EM Algorithm to Analyze the Survey Data of 2000 Taiwan Presidential Election.” Master Thesis National Central University.)
- 賴進貴、葉高華、張智昌。2007。〈投票行為之空間觀點與空間分析—以臺灣2004年總統選舉為例〉。《選舉研究》14，1：33-60。(Lay, J. G., K. H. Yap and C. C. Chang. 2007. “Spatial Perspectives and Analysis on Voting Behavior- A Case Study of the 2004 Taiwan Presidential Election.” *Journal of Electoral Studies* 14, 1: 33-60.)
- 葉高華。2006。〈近朱者赤？近墨者黑？臺灣總統選舉藍綠變遷的

鄰近效應，1996-2004〉。《2006 年臺灣政治學會年會暨「再訪民主：理論、制度與經驗」學術研討會》。2016 年 11 月 25-26 日。台北：台北大學。(Yap, K. H. 2006. “The Neighborhood Effect of Taiwan’s Presidential Election.” *Proceeding of 2006 Annual Meeting of Taiwanese Political Science Association and Democracy Revisited Conference*. 25-26 November 2006. Taipei: National Taipei University.)

盛杏媛。1994。〈Selection Bias in Vote Choice Models〉。《選舉研究》1，2：221-250。(Sheng, S. Y. 1994. “Selection Bias in Vote Choice Models.” *Journal of Electoral Studies* 1, 2: 221-250.)

盛杏媛。1998。〈選民的投票決定與選舉預測〉。《選舉研究》5，2：37-75。(Sheng, S. Y. 1998. “Voting Decision and Election Prediction.” *Journal of Electoral Studies* 5, 2: 37-75.)

盛治仁。2001。〈總統選舉預測探討—以情感溫度計預測未表態選民的應用〉。《選舉研究》7，2：75-108。(Sheng, C. J. 2001. “President Elections Forecast- Using Feeling Thermometer to Predict Undecided Voters.” *Journal of Electoral Studies* 7, 2: 75-108.)

Comparing Validities of Spatial Sampling Methods — A Case Study of 2010, 2012 TEDS Data

Yung-Ming Hsu * I-Hui Wu **

This study aims to develop several spatial sampling methods to check with TEDS data. This spatial technique is considered valuable for further research for several reasons. First, it is a way of optimal design for existing samples. Second, it avoids spatial autocorrelation and spatial heterogeneity to produce a better measurement. This study applies voting data from the 2012 presidential election and 2010 mayoral election of TEDS to test validity. As expected, the two-dimension spatial sampling method is a more accurate estimation for election studies.

Keywords: spatial sampling, electoral study, presidential election, mayoral election, voting behavior, TEDS

* Professor, Department of Political Science, Soochow University.

** Doctor Candidate, Institute of Ecology and Evolutionary Biology, National Taiwan University.