

# 一致與分裂投票：方法論之探討

黃 紀

中正大學政治學系教授

跨黨投票的現象在國內外都頗受矚目，然而文獻中對此一重要議題，從概念界定、測量指數，到數據資料、分析方法等，卻眾說紛紜。本文是從方法論的角度，觀照「一致與分裂投票」的基本概念及其使用的個體和集體資料、分析方法等層面，一方面比較各種研究途徑的優缺點；另一方面則提出台灣本土研究應有的反思。本文認為：研究者應該把個體和集體數據視為是同一個現象在不同觀察層次所呈現出來的訊息 (information)，而學界的目標，就是要整合不同形態的訊息，使之交相為用、相輔相成，並建立一套能跨越並貫穿微觀、中觀、宏觀各層次的一致與分裂投票理論架構，使不同分析層次的切入點非但能關照同一現象的不同層面，而且各點均能推導出邏輯一致、上下連貫的論證，前後呼應。

**關鍵詞：**一致投票、分裂投票、混合勝算對數模型、跨層（區位）推論、訊息理論

國人對跨黨投票的現象，應該並不陌生。例如 1994 年台北市選舉，戰況激烈，結果民進黨的陳水扁當選市長，而市議會卻無一政黨贏得過半數席次；比較一下國、民、新三黨各自之市長與市議員候選人的得票率，發覺都有很大一段差距（分別為 -13.2%、13.6%、8.5%）（洪永泰，1995: 131）。到了 1998 年選舉，國民黨的馬英九入主市府，議會卻仍是三黨不過半，儘管三黨得票互有消長，但市長與市議員候選人的得票率差距仍然很大（分別為 10.5%、12.1%、-18.3%）（吳怡銘，2000: 43），其絕對值總和比起 1994 年不減反增。

顯然在這兩次選舉時，有相當比例的選民沒有把票投給同一個政黨的兩項公職候選人。這種「分裂投票」(split-ticket voting)的現象本身，就很值得探討，更何況分裂投票一旦佔相當比例，也使得分立政府(divided government)的可能性大為提高，對府會關係、政策績效等，都會有所影響(黃紀、吳重禮，2000)。然而文獻中，對此一看似單純的重要議題，卻眾說紛紜，從概念界定、測量指數，到數據資料、分析方法等，均有爭論。

本文的目的，是從方法論的角度，觀照「一致與分裂投票」的基本概念及其使用的數據資料、分析方法等層面，一方面檢討既有文獻，以釐清概念，並比較各種研究途徑的優缺點；另一方面則展望未來，提出台灣本土研究應有的反思，並建議若干努力的方向。

## 壹、概念、數據與分析方法

社會科學的研究主題，常起因於學者對某個實際政治社會現象的好奇。該實質議題一旦提出並受到矚目，便會引發學界的重視，形成一股風潮，相關之學說風起雲湧，百家爭鳴。為了印證學說理論，研究者必須更有系統的匯集相關之實證數據資料。不可諱言，數據資料採用之初，往往始於容易取得者(如政府或其他組織既有之出版品)，而此一數據資料的形態，也難免型塑了研究分析的方法。但是當這些數據與方法無法充分解答重要的問題(甚或提出相互矛盾的答案)時，研究者不得不回歸議題與理論的基本面重新檢討，從而釐清概念，並發展新的測量與資料蒐集方式，而這些新數據又刺激了新的思維、新的分析方法與應用，如此週而復始，形成一種議題、理論、數據與方法之間的循環辯證過程。(黃紀，2000b: 113-114)一致與分裂投票的研究，就是一個典型的例子。

### 概念界定與分類

如果有幾個固定任期的公職同日改選，而大多數選民都按照自己的政黨偏好，死心塌地的把這幾項公職的選票一致投給了該黨候選人，大概根本就不會有「分裂投票」這個名詞出現。不過早在1957年Campbell與Miller

(1957: 293) 就指出美國在聯邦制下，選民分裂投票的情形頗為普遍，只是他們覺得選舉結果的累計數字無法顯示究竟有多少這種選民，所以採用了全美選舉研究 (American National Election Studies, 簡稱 NES) 的民意調查個體資料比較 1952 與 1956 年的總統選舉，發覺確實有相當比例的選民在聯邦與州的選舉採分裂投票。1965 年 Burnham 在一篇著名的論文中，採用累計的集體數據，研究美國自十九世紀中葉以後將近一個世紀的種種選舉趨勢，發現分裂投票率是在 1890 年代的政黨重組之後，才顯著提高的。此後三十餘年間，一致與分裂投票之研究便與政黨體系之發展與演變（包含消長、分裂、解組、重組等）、一致與分立政府之形成等議題密不可分，相關文獻已累積甚多，可謂卷帙浩繁，然而其重要性非但未曾稍減，反倒隨著美國政黨對選民的影響力漸低 (Wattenberg, 1996)、分立政府之出現頻仍而日益凸顯 (Wattenberg, 2000: 313)，近年 Fiorina (1996: 129) 更大聲疾呼美國政治學界應該以分裂投票做為研究之核心焦點，足見其受重視的程度。

然而一致與分裂投票在概念定義、測量指標、數據形態及分析方法上，卻仍不時引發爭辯。Feigert (1979) 有鑑於文獻中常因對分裂投票的界定含糊、測量之操作化不夠精確，而有誇大分裂投票率之虞，因此針對幾個最有爭議的問題，整理出四項應共同遵循的規則，以釐清基本概念：<sup>1</sup>

1. 選民必須針對兩種公職都投了票，其投票抉擇才有所謂一致或分裂的對比；未投票或僅投其中一項者，因缺乏比對，故不應涵蓋在內。
2. 必須在同一天選舉數種公職，選民才有選擇一致投票或分裂投票的機會；不同時間舉行的單項公職選舉，選民並非針對兩種公職同時做投票抉擇，其考量因素與同日投票不盡相同，因此不宜與一致或分裂投票相提並論。
3. 一致與分裂投票是以實際投票抉擇界定的變數，純粹是比較同一選民是否把兩項公職的選票投給同一個政黨的候選人；不能因其投票支持的政黨與其認同的政黨不一致，就認定是分裂投票。

---

1 這四項規則中，前兩項是針對使用集體數據的研究而發，而後兩項則是針對使用民調個體資料的文獻而發。

4. 測量投票抉擇的問卷題目，應針對各項公職分題詢問受訪者有否投票及投給誰（或那個黨）；若以單題直接問選民是採一致還是分裂投票，常引起「應該選人不選黨」的聯想，無意間引導受訪者答稱自己採分裂投票，結果往往高估其比例。

這幾項忠告，確有釐清概念的作用，而且現今學界所採的定義，也多與之相符（例如 Rose, 2000: 311）。因此本文把「一致投票」（straight-ticket voting）界定為：選民在同一次選舉中把不同公職的選票都投給了同一個政黨的候選人；而「分裂投票」則是指選民在同一次選舉中把不同公職的選票分別投給了不同政黨的候選人。至於將數個時間點納入研究設計進行縱貫式（longitudinal）分析，觀察選民在連續兩次或多次選舉中把票投給同一個黨派的傾向與程度，文獻中多將之歸入「選舉穩定與變遷」（electoral stability and change）<sup>2</sup> 的範疇，留待另文討論。

在台灣，分裂投票之現象在院轄市長回復民選後也備受矚目，如 1994 年台北市及 1998 年北、高兩市之市長、市議員選舉等（洪永泰，1995；陳義彥，1995；吳怡銘，2000），未來相關之研究也必然會持續增加。國外文獻對闡述一致與分裂投票的性質，當然很有幫助，但是還不宜逕行移植到對台灣的研究。我國除了憲政體制並非美式總統制外，總統四年一任，立法委員三年一任，即使未發生解散立院之事例而在每逢 12 年的週期同年改選（例如 2004 年），總統與立委選舉也不是在同一天舉行，與美國聯邦政府極有規律之四年一大選、兩年一期中選，頗不相同，因此我國在中央層級，分立政府與分裂投票的關係脫鉤，而應以選舉變遷中動態之選票流動或轉移（vote switching or electoral transition）的概念來觀察。此外，我國在地方政府層級，也有類似狀況，截至目前為止，全台灣僅有台北、高雄兩院轄市之市長、市議員定期同日改選，其餘 21 縣市之行政首長選舉與縣市議員選舉，相隔數月，並非同日舉行。<sup>3</sup> 因此在台灣對一致與分裂投票的研究，不僅必須依本土特性

---

2 例如 Converse (1969)、Crewe and Denver (1985)、Bartolini and Mair (1990)、Leithner (1997) 等。國內對投票穩定度之應用研究，見張益超 (1999) 等。

3 但是北、高兩市以外的縣市，亦偶有因特殊狀況而同時選舉數項公職的情形發生（黃德福，1991；黃紀、張益超，2001）。據報載（聯合報，2001/1/2: 1），中央選舉委員會曾表示，「如

慎選案例，而且進行研究時更需考量國情、制度，重新思考相關概念與理論。

除了憲政體制與選舉週期之外，政黨體系與選舉制度的差異，在對一致與分裂投票下定義和做分類時，也應納入本土研究的考量。美國採單一選區制（single-member districts, 簡稱 SMD），並已發展成相當穩定之兩大黨，故一致與分裂投票的分類比較單純，在剔除未投票及單項棄權（roll-off）的選民後，其一致投票只有 DD、RR 兩種模式，而分裂投票只有 DR、RD 兩種模式（D 代表民主黨，R 代表共和黨）。

我國各級民代機構的選制屬於「複席單記不可讓渡投票制」（single nontransferable vote in multimember districts, 以下簡稱 SNTV/MMD），不僅與英美之小選舉區制不同，也與德日之兩票制不同；而政黨制度則仍在發展演變之中，但傾向多黨。由於市長選舉全市為單一選區，因此通常實力較大的政黨才會參選；但市議員選舉為 SNTV/MMD 制，當選門檻較低（王業立，1996: 124），故小黨也會選擇性的在部分議員選區提名候選人，更何況還有無黨籍的候選人，使得市議員參選者的黨派色彩，往往比市長候選人複雜，選民一致與分裂投票的可能模式也大增。假定甲、乙、丙三黨均有候選人角逐市長與市議員這兩項公職，如表 1 所示，在這種情況下，一致投票的定義還挺清楚（即表 1 主對角線上加網底的甲甲、乙乙、丙丙三個類型），唯一要注意的是：若市長與市議員都有無黨籍候選人（如 1994 年北高兩市選舉及 1998 年的高雄市選舉），則選民即使把兩張選票都投給了無黨籍的候選人，也不能視為一致投票，因為「無黨」並非同一個組織，除非兩者之間有結盟的關係，否則並無「一致」的共同基礎可言。

但分裂投票的定義，就值得再推敲。例如 1998 年台北市選舉，若第一選區有位選民，市議員投給綠黨候選人張振宇，而市長則投給民進黨的陳水扁，應否歸入分裂投票？從前述「在同一次選舉中把不同公職的選票投給了不同政黨候選人」的定義來看，似應歸入。但是另一方面，這位選民若是綠黨的

---

果鄉鎮市長停選適時修法通過，中選會將建議把〔年底之〕縣市長、立委和縣市議員選舉合併，『三合一』舉行。」此一擬議顯然未及在 2001 年實現，但若將來定案，則縣市長、縣市議員選舉亦將成為一致與分裂投票之研究案例。

表 1 我國直轄市選舉各種可能之投票模式

|             |     | 市 長         |             |             |     |    |
|-------------|-----|-------------|-------------|-------------|-----|----|
|             |     | 甲黨          | 乙黨          | 丙黨          | 無黨籍 | 棄權 |
| 市<br>議<br>員 | 甲黨  | 甲甲型<br>一致投票 | 甲乙型<br>分裂投票 | 甲丙型<br>分裂投票 |     |    |
|             | 乙黨  | 乙甲型<br>分裂投票 | 乙乙型<br>一致投票 | 乙丙型<br>分裂投票 |     |    |
|             | 丙黨  | 丙甲型<br>分裂投票 | 丙乙型<br>分裂投票 | 丙丙型<br>一致投票 |     |    |
|             | 其他黨 |             |             |             |     |    |
|             | 無黨籍 |             |             |             |     |    |
|             | 棄權  |             |             |             |     |    |

\* 本表假定：

- 1. 市長與市議員各選區均有候選人角逐的政黨有甲、乙、丙三黨。
- 2. 在 SNTV/MMD 制下，市議員的參選政黨數大於市長部分。

忠誠支持者，則她/他是「被迫」跨黨投票，因為綠黨並未提名市長候選人，使這位選民根本沒有選擇一致投票的餘地。依照 Feigert（1979）所強調，選民必須在有選擇的機會下做抉擇才有所謂一致與分裂的區別，那麼這位選民把票投給不同黨，可能是別無選擇，似不能說他/她「選擇」了分裂投票。當然，如果這位選民本是民進黨的認同者，則確實是選擇了分裂投票。換言之，在我國目前的政黨與選舉制度下，分裂投票的含意常會出現一灰色地帶，故本文將分裂投票分為廣義與狹義兩種：廣義的分裂投票，是指凡在兩項公職都有投票，且分別投給了不同黨的候選人、或至少有一票投給了無黨籍的候選人，一律都視為分裂投票（即表 1 中以虛線框起來但無網底的部分，共 17 個類型）；至於狹義的分裂投票，則其範圍只限於兩種公職都有提名候選人的政黨得票（即表 1 左上角以粗黑線框起來但無網底的部分，共有甲乙、甲丙、乙甲、乙丙、丙甲、丙乙等 6 個類型），換言之，選民若把其中一票投給只提名一項公職的政黨候選人、或至少有一票投給了無黨籍候選人者，均不計入分裂投票。

本文以下的討論採狹義之定義，<sup>4</sup> 其研究標的均滿足以下幾個條件：

1. 有兩項公職在同一天舉行公民投票。
2. 至少有兩個政黨均有候選人角逐這兩項公職。
3. 選民就滿足條件 2 的政黨候選人，針對兩項公職都投下了有效票，才屬於本文一致或分裂投票的研究範圍。

## 數據形態與分析方法

研究選舉與投票行為，幾乎無法規避個體與集體分析層次的挑戰，因為現代的普選，選民以無記名的方式個別投下自己的選票，但是選舉結果的揭曉，不論是政黨或候選人的勝負，皆取決於選票的集體累計。如果累計的總數還可以回溯到每張個別的選票和投下那張票的選民，那麼選舉研究就輕鬆得多了。但自十九世紀中葉以後，秘密投票漸成為民主政治的規範，政府公告的選舉結果便無法再追溯到個別選民。

一致與分裂投票的研究方法之所以常引發爭論，原因之一，就是理想的觀察單位和實際上可蒐集到的資料之間有一道鴻溝。投票者既是個別的選民，因此選舉行為的研究理應蒐集以選民為觀察單位的個體資料（individual-level or micro data）。但是秘密投票既是民主政治的常軌，研究者不可能取得個人層次的投票紀錄，所以目前主要的實證資料來源有二：一為民意調查，一為政府部門公布出版之選舉實錄。前者係針對選民母群抽樣後，以問卷詢問受訪者，由受訪者自己報告（self-report）對各項公職之屬意人選或政黨後，經訪員記錄所得到的個體資料；而後者則是由政府部門以各級地理區域（如投開票所、村里、鄉鎮市區、縣市等）為單位，公布各候選人之得票數與得票率，屬於集體數據。民調相較於其他蒐集資料方法的優缺點，文獻中已論之甚詳，在此不擬贅述。但用民調資料分析一致與分裂投票，首先抽樣之母群必須配合兩種公職中最小的選區，才會有足夠的樣本數達到代表性的要求。然而國內較嚴謹之民調多屬全國性，若以全國性民調結果分析台

---

4 瞭解了狹義定義的分析方法後，不難推廣至廣義定義的分析。重要的是，分析前應先釐清其定義。

北或高雄市之一致與分裂投票，樣本數顯然不足。此外，測量投票行為的可靠度與有效度是另一個顧慮，即選前民調拒答或回答「尚未決定」之比率常偏高，而選後民調則回答投給實際當選人之比率常偏高，若遇此情況，測量已有偏差，很難從分析上予以補救。

在民調興起之前，政治學者運用行政與地理區位之集體資料（group-level, aggregate, or macro data）為研究素材，其來有自。例如過去 Key (1949) 對美國南方各州的研究，就是採用政黨在選區中的得票紀錄，分析的單元主要是以市、鎮、州為主。惟自從社會學家 Robinson (1950) 顯示以集體數據計算之「區位相關」(ecological correlations) 常與個體資料分析的結果不符，人文區位研究便開始一蹶不振，幾乎被民意調查之個體資料所取代。但是在將近半個世紀後，許多研究者開始體會到個體與集體兩者間其實相輔相成，不可偏廢。尤其是分明有可靠之集體數據，卻只因缺乏民調個體資料，便需放棄重要議題之研究，顯然本末倒置。因此近年來，運用集體資料推論個人行為之方法，即跨層次推論（cross-level inference，或稱區位推論 ecological inference）方法，又重獲重視，並有長足之進步。(Achen and Shively, 1995; King, 1997; King, Rosen and Tanner, 1999; Rosen, Jiang, King and Tanner, 2001)

儘管如此，區位推論仍有其限制。故充分瞭解個體資料及集體數據的產生過程（data generating processes），掌握其特性並慎選適當的分析及推論方法，進而思考如何結合民調與區位資料，截長補短，實為當前方法論的重要課題與挑戰。

## 貳、個體數據之分析方法

民調之研究設計種類繁多，單次者可在選前或選後進行，多次者則可能每次皆獨立抽樣，但也可能採取定群追蹤研究（panel study）。但不論何者，就一致與分裂投票而言，最適當的問卷題型，是分題詢問受訪者對各項公職的投票傾向或實際投票抉擇。假定研究的目的是分析選民針對兩種同時選舉之公職的投票行為，且假定有甲、乙、丙三黨均有候選人角逐這兩項公職，

則根據受訪者對兩題均表態回答之投票抉擇，可建立一  $3 \times 3$  之二維正方形交叉分析表（以下簡稱正方表 square table）如表 2：<sup>5</sup>

表 2 一致與分裂投票模式：個體數據（樣本數  $n_{++}$ ）

|             |    | 公 職 二                |                      |                      |          |
|-------------|----|----------------------|----------------------|----------------------|----------|
|             |    | 甲黨                   | 乙黨                   | 丙黨                   |          |
| 公<br>職<br>一 | 甲黨 | 甲甲型一致投票數<br>$n_{11}$ | 甲乙型分裂投票數<br>$n_{12}$ | 甲丙型分裂投票數<br>$n_{13}$ | $n_{1+}$ |
|             | 乙黨 | 乙甲型分裂投票數<br>$n_{21}$ | 乙乙型一致投票數<br>$n_{22}$ | 乙丙型分裂投票數<br>$n_{23}$ | $n_{2+}$ |
|             | 丙黨 | 丙甲型分裂投票數<br>$n_{31}$ | 丙乙型分裂投票數<br>$n_{32}$ | 丙丙型一致投票數<br>$n_{33}$ | $n_{3+}$ |
|             |    | $n_{+1}$             | $n_{+2}$             | $n_{+3}$             | $n_{++}$ |

$I \times I$  的正方表構成了分析一致與分裂投票的主要數據型態。假定抽樣設計只固定成功樣本數  $n_{++}$ ，則正方表之邊緣總和（marginals） $n_{i+}$  及  $n_{+j}$  未固定，理論上  $I \times I$  細格內之次數呈多項分佈（multinomial distribution），其聯合機率參數為  $\{\pi_{ij}\}$ （Agresti, 1990: 39）。在文獻中，正方表的運用甚廣（Bishop, Fienberg and Holland, 1975: Chapter 8），例如醫學中對左、右眼視力之比較，社會學中以社會流動表（social mobility tables）（Hout, 1983）顯示代間階層之變化等。這些方法，頗有可借鏡之處。

## 初探分析<sup>6</sup>

把正方表應用於一致與分裂投票研究，我們不難建立幾個理念類型

5 本文所舉的交叉表例子，刻意假設有甲、乙、丙三個主要政黨，目的在更接近國內之實際狀況，以有別於美國文獻中大多預設僅有兩大黨的情況；不過本文所討論之方法，其實均適用於兩黨或多黨制。在符號方面，本文以  $i$  標示橫列（ $i=1, 2, \dots, I; I \geq 2$ ），以  $j$  標示縱行（ $j=1, 2, \dots, J; J \geq 2$ ）， $n$  代表樣本數， $N$  代表母群數。依慣例，表示聯合分佈時，第一個下標為橫列數，第二個下標為縱行數，而  $+$  代表加總，例如  $n_{1+} = \sum_{j=1}^J n_{1j}$ 。

6 本文所謂「初探分析」，是採 Tukey (1977) 之 exploratory data analysis (EDA) 的觀念，強調於後續之檢證分析 (confirmatory data analysis) 之前，應先對數據本身作充分之理解。

(ideal types)。首先，若所有受訪者把票都投給了同一政黨的候選人，則僅有主對角線（main diagonal）上的次數（frequencies） $n_{ii}$  會大於 0，其餘細格的次數  $n_{ij}(i \neq j)$  皆為 0，形成一對角方陣（diagonal matrix），可稱之為「完全一致投票」模式。反之，若所有受訪者都把兩項公職的選票投給了不同政黨的候選人，則主對角線上的次數皆為 0，把正方表切割成了左下方與右上方兩個三角方陣，可稱之為「完全分裂投票」模式。

準此，摘述一致投票率最簡單的方式，就是把正方表主對角線上估計的聯合機率（estimated joint probability） $p_{ii} = \frac{n_{ii}}{n_{++}}$  加總，即  $\sum_{i=1}^I p_{ii}$ ；而分裂投票率自然就是主對角線以外所有細格中之聯合機率的總和，即  $\sum_{i \neq j} p_{ij} = 1 - \sum_{i=1}^I p_{ii}$ 。不過這兩個估計式，並未考慮到樣本中純屬隨機之一致與純屬隨機之分裂。心理測量學發展的Cohen's Kappa（Cohen, 1960; Fleiss, Cohen and Everitt, 1969; Agresti, 1990: 366-367），又稱「同意係數」（coefficient of agreement），就是個試圖把隨機部分剔除的摘要統計量：

$$\hat{k} = \frac{\sum_i p_{ii} - \sum_i p_{i+} \cdot p_{+i}}{1 - \sum_i p_{i+} \cdot p_{+i}}$$

但是要格外小心的是，同意係數所比對的基點（null model）是獨立（independence）模型，因此在「完全一致投票」模式下，尚可得其極大值 + 1.0，但在「完全分裂投票」模式下，卻是負數，且可能小於 -1.0。此外，同意係數的值受正方表邊緣總和分佈的影響，也不能直接用來比較兩個正方表一致的程度，應用時需依照交叉表邊緣總和分佈先計算最大、最小可能值，再據以修正原來之 Kappa 值。

當然，實際的數據資料，多介於完全一致與完全分裂這兩個極端之間。例如政黨標籤若有相當影響力的話，那麼主對角線上一致投票的選民應居多，其餘選民則散佈於其他細格內。另一方面，若發生棄保效應之策略投票，則因當選無望而被棄者，其所屬政黨之一致投票率便降低，選票朝「退而求其次」之被保者流動。針對正方表進行之初探分析，就是希望能一窺選票一致與分裂模式之分佈結構。

由於表 2 中兩項公職同日選舉，並無自變數、依變數之分，<sup>7</sup> 所以最直接了當的初探分析，就是呈現正方表內之三種百分比：

1. 聯合百分比 (joint percentages)  $p_{ij} = \frac{n_{ij}}{n_{++}}$ ：估計聯合機率，也就是在所有樣本中，公職一的選票投給了第  $i$  個黨、而且公職二方面又把票投給第  $j$  個黨的比例。
2. 橫列百分比 (row percentages)  $p_{ji} = \frac{n_{ij}}{n_{i+}}$ ：估計橫列條件機率，也就是把公職一的選票投給第  $i$  個黨的樣本中，在公職二方面也把票一致地投給該黨 ( $i=j$ )、或是把票分裂給其他政黨 ( $i \neq j$ ) 的比例。
3. 縱行百分比 (column percentages)  $p_{ij} = \frac{n_{ij}}{n_{+j}}$ ：估計縱行條件機率，也就是在公職二方面把票投給第  $j$  個黨的樣本中，在公職一也把票一致地投給該黨 ( $i=j$ )、或是把票分裂給其他政黨 ( $i \neq j$ ) 的比例。

上述三種百分比的效用，依實際研究的焦點而異。例如橫列或縱行百分比在分析某項公職（如行政首長）的候選人是否會對另一項公職（通常為民意代表）的同黨候選人產生拉拔效應（coattail effect）時，較為有用。若是目的在瞭解選民一致與分裂投票模式的整體分佈，一般而言以聯合百分比最容易解讀；但如果要估計把公職一的選票投給某個黨的選民之中，在公職二方面卻把票分裂給了其他政黨的條件機率，則無疑橫列百分比最適當。

倘若兩種公職均有多黨（三黨及以上）候選人競選，可能不太容易以目測法直接從上述百分比看出端倪，此時可進一步將橫列及縱行百分比結合邊緣百分比 (marginal percentages)，運用對應分析 (correspondence analysis)（見 Greenacre, 1984; Clausen, 1998）將兩種公職之各黨得票相對位置放在座標軸上，檢視其聯合分佈有無特定之結構模式。

除了目測法外，對數線型模型 (log-linear model) 中準對稱 (quasi-symmetry)、準獨立 (quasi-independence) 等概念，亦有助於偵測正方表有無結構上之特徵（參見 Powers and Xie, 2000: 114-117）。例如表 2 中若六種分裂投票類型大致呈跨對角線對稱，便可檢驗其是否與準對稱結構無顯著差異。另一方面，若分裂投票果真如一些學者所說（如 Beck, Baum, Clausen

7 若是進行兩個時間點  $t$  與  $t+1$  之間選票轉移的分析，則較早的時間點  $t$  就視同為自變數。

and Smith Jr., 1992; Born, 1994; Forgette and Platt, 1999)，純粹是選民對兩種公職的個別考量所產生之副作用，而非刻意的一種策略選擇，那麼正方表的結構應該比較接近準獨立的結構。

## 解釋性分析

前述對交叉表的初探分析及摘要統計，只是充分描述、彰顯民調資料中一致與分裂投票模式的分佈狀況。不過大多數的研究，還想要進一步解釋選民為何作了這樣的抉擇。在文獻中，許多研究在此一階段便逕行將選民之投票模式一分為二：一致與不一致（分裂）投票，然後進行二分依變數之迴歸分析（binary regression）。這種做法，在早期或可說是因為二分依變數迴歸較易於處理，可用的統計程式也較多。但是此種積習卻延續至今，連近年出版的論文都充斥這種分析法（例如 Born, 2000; Mattei and Howes, 2000），就未免顯得過度簡化。

如前節所述，即使是狹義的一致與分裂投票，在概念上也屬於多類的（polytomous）變數，其民調數據資料基本上屬於多項分佈。質言之，即便在兩黨制的國家，甲甲型一致投票模式和乙乙型未必相同，而甲乙型分裂投票模式和乙甲型也未必一樣，更遑論多黨制的國家。研究者必須證明歸併的類別間同質，否則歸併後估計出來的參數可能有偏誤。

其實以表 2 之交叉表為基礎，無須強行歸併一致與分裂之次類，就至少可以朝三個方向進行解釋分析：多變數之對數線型模型（multivariate log-linear model）、多項之勝算對數或機率單元模型（multinomial logit or probit model）、混合式之勝算對數模型（mixed logit）。

多變數（或稱多維 multidimensional）之對數線型模型，或許是二維交叉表最自然的延伸。例如研究者的理論架構若認為決定一致與分裂投票模式的自變數主要有三個：省籍（分為  $K$  類）、政黨偏好（分為  $L$  類）、年齡（分為  $M$  類），則分析策略上，是分析  $K \times L \times M$  類的自變數對原  $I^2$  二維之交叉表之關係有無顯著之影響（參見 Christensen, 1997）。

倘若研究者對多維之對數線型模型不很熟悉，而較習慣迴歸式的分析，則可採用多項之勝算對數模型（如 Sigelman, Wahlbeck and Buell, Jr.,

1997)，從正方表的  $I^2$  個投票類型中擇一為參照類，以  $I^2 - 1$  個勝算之對數 (log of the odds) 為依變數，與自變數進行迴歸分析 (Lloyd, 1999: 323-324)。多項之勝算對數模型與多維之對數線型模型關係極為密切，只是概念上將正方表的  $I^2$  個投票類型明確設定為依變數。二者皆屬於推廣之線型模型 (generalized linear models, 簡稱 GLMs) (參見 黃紀, 2000a)。

不過多項之勝算對數模型建立在一個很強的假定上，即  $I^2$  個選項之間，彼此獨立，各不相干。<sup>8</sup> 這個假定對多數選舉而言，太過於牽強。因此近年來研究多黨或多位候選人競爭之選舉，紛紛改採多項之機率單元模型 (例如 Alvarez and Nagler, 1998a; 1998b; 2000; Lacy and Burden, 1999; Quinn, Martin and Whitford, 1999)，因為此一模型假定隨機功效 (random utility) 之誤差項間呈多變數之常態分佈，不僅未作「選項之間彼此獨立」之假定，還允許選項之間可有彼此替換之模式 (substitution patterns)，較符合實際多黨或多位候選人競選之實際狀況。

然而就分析一致與分裂投票而言，最有潛力既兼顧選項間之相互關係，又能配合檢證學理的，是由應用環境經濟學 (environmental economics) 所發展出來的混和勝算對數模型，又稱「隨機參數勝算對數模型」 (random-parameters logit)<sup>9</sup> (Revelt and Train, 1998; Train, 1998, 1999; Greene, 2000: 871-875; Glasgow, 2001)。此一模型將誤差成分 (error components) 與前述之勝算對數模型結合，不僅免除了後者「選項之間彼此獨立」之過強的假設，還可更有彈性的設定、估計與檢證套疊狀勝算對數模型 (nested logit)。倘若選民果真如 Fiorina (1996) 等學者所主張，刻意的選擇分裂投票，則相關的幾型分裂投票便應屬於同一個套疊 (nest) 之內，彼此有顯著相關，但與其他套疊內之投票模式則無顯著關係。同理，若選民有棄保之策略投票行為，則相關的幾型投票模式 (如「甲丙型分裂」與「丙丙型一致」)

8 在選擇理論中稱之為 independent from irrelevant alternatives (簡稱 IIA)，參見 Ben-Akiva and Lerman (1985: 51)。

9 當然，除了 mixed logit model 之外，還有其他將隨機參數結合 GLMs 的模型，如「推廣之線型混合模型」 (generalized linear mixed models, GLMMs)，參見 Breslow and Clayton (1993)、McCulloch and Searle (2001: Chapter 8) 等。

便應屬於同一個套疊之內，彼此呈顯著相關，反之則無。因此混和勝算對數模型提供了一個檢證一致與分裂投票學理的有效工具。此外，該模型還可朝多層次分析 (multilevel analysis)<sup>10</sup> 的方向發展，將選民居住所在之區域環境因素納入經驗分析，使民調之個體數據與集體資料交相為用、緊密結合 (Shields and Huang, 2001)。

參、集體數據之分析方法

投票行為之集體數據，主要來自政府各級選委會公告及出版之開票結果。如前所述，基於秘密投票之原則，這類數據只能以地理行政區域為單位，呈現各黨及候選人之得票數，無法顯示個別選民之投票抉擇，其數據型態與資料內容，自然與民調之個體數據，大不相同，分析方法也隨之而異。例如某市第  $k$  個里，其兩項公職同時選舉後所公布之三黨得票結果，也可據以建立一交叉表 (參見 King, 1997: 13)，如表 3 所示，則該市共有  $K$  個類似的正方表。

表 3 一致與分裂投票模式：集體數據 (第  $k$  里有效票數  $N_{++k}$ )

|             |    | 公 職 二                   |                         |                         |            |
|-------------|----|-------------------------|-------------------------|-------------------------|------------|
|             |    | 甲黨                      | 乙黨                      | 丙黨                      |            |
| 公<br>職<br>一 | 甲黨 | 甲甲型一致投票數<br>$N_{11,k}?$ | 甲乙型分裂投票數<br>$N_{12,k}?$ | 甲丙型分裂投票數<br>$N_{13,k}?$ | $N_{1+,k}$ |
|             | 乙黨 | 乙甲型分裂投票數<br>$N_{21,k}?$ | 乙乙型一致投票數<br>$N_{22,k}?$ | 乙丙型分裂投票數<br>$N_{23,k}?$ | $N_{2+,k}$ |
|             | 丙黨 | 丙甲型分裂投票數<br>$N_{31,k}?$ | 丙乙型分裂投票數<br>$N_{32,k}?$ | 丙丙型一致投票數<br>$N_{33,k}?$ | $N_{3+,k}$ |
|             |    | $N_{+,1,k}$             | $N_{+,2,k}$             | $N_{+,3,k}$             | $N_{++k}$  |

10 有關多層次分析之介紹，參見 Jones and Duncan (1998)。

表 3 也是  $I \times I$  的正方表，乍看之下與表 2 極為相似，只不過代表樣本的  $n$  改為代表母群的  $N$  而已。但是二者最大的差異，在於表 3 的  $I^2$  個細格中各型投票數  $N_{ij}$  均為未知，故以問號標示之，僅有該表邊緣上之各黨得票總數  $N_{i+}$  及  $N_{+j}$  為已知。換言之，若有個體層次之數據，則表 2 細格中之聯合分佈可輕易計算該表之邊緣分佈；但若僅有集體層次之數據，則表 3 只有邊緣分佈，卻無從得知表內細格之聯合分佈。也就是說，唯有當研究者能取得母群個體層次之數據時，交叉表中細格次數及百分比才能精確得知；試圖以已知的邊緣總和數，去精確推知該里各型未知的一致與分裂模式投票數，是不可能的。但若能蒐集所有  $K$  個里的集體數據，則推知全市的未知參數，似乎露出一線曙光。的確，試圖從集體層次之區位數據（如  $N_{i+}$  及  $N_{+j}$ ），重建、推估原來必須有個體數據才能得知的訊息（如  $N_{ij}$ ），就是所謂區位推論（ecological inference）<sup>11</sup> 問題。因此，區位推論是一種描述性推論，目標在重建未直接觀察到的個體層次事實，故不屬於因果解釋的推論。然而當個體數據不可得或不可靠時，以集體數據跨層次推估個體層次之事實，在邏輯上反而優先於因果解釋型的推論。

## 論「區位謬誤」

一般對跨層次推論有兩種各趨極端的誤解。一種認為，集體單位既是由個體所組成，集體數據和個體數據無異，可直接替代後者。這種誤解，在 Robinson (1950) 顯示「區位相關」（ecological correlations）往往與個體數據計算之相關不符後，已經較少發生；但是該文卻引起了另一種誤解，將 Robinson 的提醒與警告解讀為禁令，認定以集體數據跨層次推估個體層次之事實，若非緣木求魚，徒勞無功，就是煉金詐術，欺世盜名。

11 Ecology 多譯為「生態學」，因其研究生物與其周遭之自然環境及其他生物之間的互動關係，故測量之數據常以地理區域為單位。1930 年代芝加哥都市社會學派（the Chicago school of urban sociology）自創「人文區位學」（human ecology）一詞，特別強調人群在都會中之空間分佈關係（spatial relationships）。(Ellen, 1996: 207; Clark, 1998: 191; Johnston *et al.*, 2000: 352-353) 儘管該學派趨於式微，但在社會科學方法論中，ecological 一字幾乎已與 aggregate 同義，故本文將“ecological”譯為「區位」。

其實討論集體數據跨層次推估個體層次事實之風險，Robinson 並非第一人（參見 Achen and Shively, 1995: 5-10 之歷史回顧），而且其 1950 年的論文中，也未曾使用「區位謬誤」(ecological fallacy) 一詞。惟自 Selvin (1958: 615) 率先創用該詞後，便有人望文生義，將之與演繹邏輯上的「謬誤」混為一談，此一誤解還「口耳相傳」、變本加厲，甚至被寫進方法論的教科書，扣上了「區位謬誤『定理』」(ecological fallacy “theorem”) 的大帽子 (Miller, 1995: 155)，儼然視區位推論為學術禁地，其不求甚解，令人喟歎。

區位推論的困難，實源於個體層次的訊息在匯集成集體數據的過程中 (aggregation process) 有一部份流失。表 3 細格中的問號，就是第  $k$  里個別選票累計後失去的訊息。跨層次推論的挑戰，就在正視此訊息之缺漏，設法重建或補充之，不僅竭力將推論之假定與程序明確化，並將其限制及風險透明化。從訊息理論 (information theory) 的角度來看，推論縱使不能得知百分之百精確之真相，但只要能減低無知與不確定的程度，未嘗不是一種訊息之增進。所以區位推論的風險有程度之分，而非演繹邏輯「非真即偽」之二分。換言之，研究者仍應珍惜集體數據，充分運用，重點是：使用時必須拿捏得當，否則過度推論固然造成謬誤，因噎廢食，又豈是明智之舉？

## 初探分析

以集體數據研究分裂投票，最常用的摘述統計就是 Burnham (1965: 9) 指數，因其與「想當然耳」之直覺判斷最為接近，但該指數卻也是使用集體數據時必須格外謹慎的絕佳例證。由於 Burnham 討論的是美國兩黨制下的選舉，因此只需選定其中一黨（如民主黨  $D$ ），計算該黨在同一次選舉中兩種公職的得票率之差，然後取其絕對值即可，亦即  $|N_{D+} - N_{+D}|$ 。（亦參見 Rusk 1970: 1224）若將此一指數應用到類似表 3 之  $I \times I$  ( $I > 2$ ) 正方表，則須針對多黨的情況調整如下：

$$\text{多黨參選之下，Burnham 分裂投票總數之指數} = \frac{\sum_{i=1}^I |N_{i+} - N_{+i}|}{2} \quad \text{。}^{12}$$

12 類似之集體數據測量指標，在文獻中以各種不同的面貌出現，且應用甚廣，因此值得在此

若將上式之分裂投票總數之指數除以  $N_{++}$ ，則成為分裂投票率指數。

但是誠如 Cowart (1974: 111-113) 指出，Burnham 指數常低估（雖然絕對不會高估）兩黨制下之分裂投票率。上式中應用於多黨之分裂投票指數亦有相同之缺陷，這可以從表 4 的例子一窺端倪：假設表 4 的行與列邊緣比例分佈均為 0.5, 0.3, 0.2，則 Burnham 指數顯示分裂投票率為 0，意即符合前節所述之「完全一致投票」模式。但是實際上，卻有可能甲乙型及乙甲型分裂投票率各佔 0.3，而甲丙型及丙甲型分裂投票率各有 0.2——正好符合「完全分裂投票」模式，亦即真正之分裂投票率可能高達 100%，而非 0%！這確實是使用集體數據推論一致與分裂投票之陷阱，不可不慎。

表 4 一致與分裂投票率之上下限（一個假想的極端例子）\*

|             |    | 公 職 二                                  |                                        |                                        |
|-------------|----|----------------------------------------|----------------------------------------|----------------------------------------|
|             |    | 甲黨                                     | 乙黨                                     | 丙黨                                     |
| 公<br>職<br>一 | 甲黨 | 甲甲型一致投票率<br>$0 \leq \pi_{11} \leq .50$ | 甲乙型分裂投票率<br>$0 \leq \pi_{12} \leq .30$ | 甲丙型分裂投票率<br>$0 \leq \pi_{13} \leq .20$ |
|             | 乙黨 | 乙甲型分裂投票率<br>$0 \leq \pi_{21} \leq .30$ | 乙乙型一致投票率<br>$0 \leq \pi_{22} \leq .30$ | 乙丙型分裂投票率<br>$0 \leq \pi_{23} \leq .20$ |
|             | 丙黨 | 丙甲型分裂投票率<br>$0 \leq \pi_{31} \leq .20$ | 丙乙型分裂投票率<br>$0 \leq \pi_{32} \leq .20$ | 丙丙型一致投票率<br>$0 \leq \pi_{33} \leq .20$ |
|             |    | .50                                    | .30                                    | .20                                    |
|             |    | 1.00                                   |                                        |                                        |

\* 本表之 Burnham 分裂投票率指數等於 0；但依照 Duncan and Davis 上下限法計算，本表之分裂投票率可能是介於 0 與 1.0 之間的任何值。

詳加討論。例如熟悉「選舉穩定與變遷」相關文獻的讀者不難察覺：若將公式中的兩項公職改為兩個時間點  $t$  與  $t+1$ ，並將之應用於英國的保守黨與工黨兩大政黨歷次巴力門選舉之得票率，則此一公式就成了 Butler 的「選票鐘擺」（swing）指數（見 Berrington, 1965: 17; Butler and Van Beek, 1990: 178; Norris, 1997: 23）；若應用於歐陸國家多黨的情況，則又成為 Pedersen (1979: 4; 亦參見 Bartolini and Mair, 1990: 20; Katz, 1997: 83) 之「選舉變動」（electoral volatility）指數。故本文對 Burnham 指數之評論，亦皆適用於 Butler 的「選票鐘擺」及 Pedersen 之「選舉變動」指數。

因此，Burnham 指數應解讀為：I×I 表整體而言，分裂投票總數之下限。證明如下：

首先，I×I 表之分裂投票總數等於主對角線以外所有細格次數之總和，亦即  $\sum_{i \neq j} N_{ij}$ 。

其次，第  $i$  個政黨在兩項公職得票數之差的絕對值為

$$\begin{aligned} |N_{i+} - N_{+i}| &= \left| \sum_{j=1}^I N_{ij} - \sum_{j=1}^I N_{ji} \right| \\ &= \left| \sum_{j \neq i} N_{ij} - \sum_{j \neq i} N_{ji} \right| \leq \left( \sum_{j \neq i} N_{ij} + \sum_{j \neq i} N_{ji} \right) \\ &\because N_{ij} \geq 0 \text{ and the triangle inequality} \end{aligned}$$

最後，所有 I 個政黨在兩項公職得票數之差的絕對值總和為

$$\begin{aligned} \sum_{i=1}^I |N_{i+} - N_{+i}| &= \sum_{i=1}^I \left| \sum_{j \neq i} N_{ij} - \sum_{j \neq i} N_{ji} \right| \leq \sum_{i=1}^I \left( \sum_{j \neq i} N_{ij} + \sum_{j \neq i} N_{ji} \right) = 2 \sum_{i \neq j} N_{ij} \\ \therefore \frac{\sum_{i=1}^I |N_{i+} - N_{+i}|}{2} &\leq \sum_{i \neq j} N_{ij} \quad \text{and} \\ N_{++} - \frac{\sum_{i=1}^I |N_{i+} - N_{+i}|}{2} &\geq \left( N_{++} - \sum_{i \neq j} N_{ij} \right) = \sum_i N_{ii}. \quad \text{Q.E.D.} \end{aligned}$$

故得證 Burnham 指數會低估（但絕對不會高估）分裂投票總數，因此該指數除以  $N_{++}$  自然也會低估分裂投票率。同理，若以  $N_{++}$  減 Burnham 指數來估計一致投票總數，則會高估（但絕對不會低估）一致投票總數（及除以  $N_{++}$  後之一致投票率）。

Burnham 指數既然只能估計分裂投票之下限，僅以此估計值進行跨地區或跨時間點的比較就十分危險，因為在沒有其他訊息的情況下，理論上分裂投票率之上限可高達 100%，而分裂投票率的真值可能是介於上下限之間的任何數值。例如台北市市長與市議員選舉，1994 年 Burnham 指數估計之全市（狹義）分裂投票率為 14.16%，1998 年則為 19.03%，但我們不能據此即斷定台北市選舉 1998 年之分裂投票率必高於 1994 年，因為僅知道這兩年之下限，我們尚無法排除 1994 年實際的分裂投票率反而高於 1998 年的可

能。換言之，要在上下限之間找出最可能的真值，必須使用稍後將討論的「描述性推論」方法。不幸的是，Burnham 當初創用此一指數，目的就在直接比較美國十九世紀中葉至二十世紀中葉的投票及參與趨勢，而且其後亦有多位學者援引（如 Rusk, 1970 等），運用甚廣；然而跨年比較的測量指數既有問題，結論自然另人存疑。

當然，除了有低估分裂投票、高估一致投票總數之虞，且不宜直接做跨地區、跨時間之比較外，Burnham 指數最大的缺點是無法估計表 3 細格內各型之一致與分裂投票數（率），故應用上不僅陷阱重重，且頗多限制。因此使用集體數據分析一致與分裂投票，勢必要參酌運用其他的方法。

Burnham 指數的限制，其實只是反映了跨層次推論的基本挑戰：僅從一個交叉表之邊緣分佈，無從精確推知表內之聯合分佈。但這絕不表示研究者就應束手無策、棄械投降，因為我們仍可根據邊緣分佈推知各個細格內次數  $N_{ij}$  真值座落之區段範圍，這就是早在 1953 年 Duncan 和 Davis 所提議之「上、下限推算法」（method of bounds）。儘管該篇兩頁不到的短文並未明確列出上下限之計算公式，但不難將之重建如下：<sup>13</sup>

$$\begin{aligned} \because \quad & \max(N_{ij}) = \min(N_{+j}, N_{i+}); \\ \text{and} \quad & \min(N_{ij}) = N_{+j} - \max\left(\sum_{i' \neq i} N_{i'j}\right) \quad i' \in I \\ & = N_{+j} - \min\left(N_{+j}, \sum_{i' \neq i} N_{i'+}\right); \\ \therefore \quad & N_{+j} - \min\left(N_{+j}, \sum_{i' \neq i} N_{i'+}\right) \leq N_{ij} \leq \min(N_{+j}, N_{i+}) \\ \text{or} \quad & \max\left[0, \left(N_{+j} - \sum_{i' \neq i} N_{i'+}\right)\right] \leq N_{ij} \leq \min(N_{+j}, N_{i+}). \end{aligned}$$

將上述公式用於表 4 之邊緣分佈，則可計算出表內各細格聯合機率  $\pi_{ij}$  之上下限。因此即使在表 4 的極端情況下，研究者若先推算表內各細格  $\pi_{ij}$  之上下限，就不至於被表 4 之 Burnham 指數所誤導而驟下其必為「完全一致投票」之錯誤推論。

13 關於  $2 \times 2$  表細格內次數之上下限計算公式，見 King (1997: 302-303)。

「上、下限推算法」所求出之區段為確定 (deterministic)，亦即  $N_{ij}$  之真值必然落於該區段之內，與統計上或然之信心區間 (confidence interval)，截然不同。 $N_{ij}$  之區段求出後，只要除以適當之邊緣總和，自然也得出其三種（聯合、橫列、縱行）百分比的上下限區段。例如橫列條件機率  $\pi_{j|i}$  若下限大於 0% 而上限小於 100%，就得到了比邊緣分佈本身更有用的新訊息；區段越窄，訊息越精確。誠然，上下限法未必能把表內所有參數之區段縮得夠窄，不過任何有用的訊息，都不應輕言放棄。

如果集體單位的總數  $K$  不太大，則上述之測量指數及上下限可應用至每一個正方表。但若集體單位的總數甚大，則逐表推算，常有見樹不見林之憾，此時若能將所有的集體單位都呈現在一座標圖上，便較易窺其全貌。最直接了當的圖示法，就是把表 3 已知之邊緣總和換算成各黨得票率（比例），然後將各黨在  $K$  個里的兩項公職得票率劃在座標圖上，以目測法大致判斷該地區一致與分裂投票的傾向。如果只有兩黨參選兩項公職，則一般之二維座標軸已足。但若有三黨參選（如表 3），則需使用三維之立體座標，或用較為簡潔之三角圖 (ternary diagram, or triplot) (Aitchison, 1986: 5-6; Katz and King, 1999: 19-20)。必須注意的是，圖示法固然能把許多集體單位的數據同時呈現在座標圖上，一目了然，但畢竟呈現的只是表 3 的邊緣比例分佈，不能據以推論表內細格的參數值，因此若不輔以其他的方法，仍有犯區位謬誤之虞。

以圖示法呈現細格內的橫列條件機率  $\pi_{j|i}$  之上下限，並非沒有可能。若只有兩黨候選人角逐兩項公職，則表 3 簡化為  $2^2$  正方表，再假定最左邊兩個細格之橫列條件機率  $\pi_{1|1}$  及  $\pi_{1|2}$  為研究者最感興趣之參數，則每個集體單位針對該兩個參數所算出的上、下限區段，皆可以線段劃在一個「零至一正方座標圖」(unit square) 之內，這就是 King (1997: 81-82) 的「斷層掃描圖」(tomography plot)。就這一對參數而言，同時呈現  $K$  個里的一幅斷層掃描圖已包含了整體資料內的所有確定訊息，King (1997) 的區位推論模型就是以此圖做為進一步推估參數值的基礎。但是就本文關心的多黨參選兩種公職的情形，則研究者必須先將  $I^2$  表歸併成  $I(I-1)$  個  $2^2$  正方表，才能建立  $I(I-1)$  個斷層掃描圖，程序極為繁複。

## 描述性推論

就在 Duncan and Davis (1953) 發表上下限推算法的同一期 *American Sociological Review* 中, Goodman (1953: 663-664) 也針對 Robinson 區位相關的警語提出一個「適用於特殊狀況下」的解決之道, 此即日後著名的 Goodman 區位迴歸 (ecological regression) 模型 (詳見 Goodman 1959)。基本上, 此一模型若運用到一致與分裂投票, 則表 3 不再視為聯合分佈之交叉表, 而視為是列聯表 (contingency table) ——兩項公職之中有一項須設定為自變數, 另一項為依變數, 然後進行 (I-1) 個線型迴歸, 估計條件機率。設若表 3 之中, 公職一 (橫列) 設定為自變數, 並以丙黨為參照類, 則表 3 兩個區位迴歸式的依變數, 就是甲、乙兩黨在公職二 (縱行) 的得票率, 而其自變數就是甲、乙兩黨在公職一的得票率。在這兩個線型迴歸式中截距的估計值, 就是「全市平均而言公職一投給丙黨的選民之中, 在公職二卻投給了甲、乙黨的比例」(即橫列條件機率  $\pi_{1|3}$  與  $\pi_{2|3}$  之估計值)。截距與兩式中甲黨自變數之係數估計值的和, 就是「全市平均而言公職一投給甲黨的選民之中, 在公職二投給甲、乙黨的比例」(即橫列條件機率  $\pi_{1|1}$  與  $\pi_{2|1}$  之估計值), 依此類推。(亦參見 Achen and Shively, 1995: 34-38)<sup>14</sup> 誠如 Goodman 自己所說, 此一簡單之模型, 只有當各個集體單位間之行爲模式同質性高 (homogeneous) 時, 其推論才有效。實際運用時, 此一同質性假定卻常與實情不符, 而區位迴歸之估計值常呈大於 1.0 或小於 0 之不合理機率。儘管其後學者對 Goodman 模型有若干延伸與修正, 但成效有限。

King (1997) 近年提出的區位推論模型 (簡稱 EI model) 頗受矚目。他基本上針對 Goodman 的模型提出兩大修正：

1. 把 Duncan and Davis 上下限法求出的確定區段, 以斷層掃描圖的形式納入 EI 模型, 以克服參數估計值超出 0 與 1 之間的問題。
2. 把隨機參數模型納入 Goodman 的簡單區位迴歸, 以放寬 Goodman

14 當然, 研究者亦可剔除截距, 而直接以各黨在公職一之得票率為自變數, 其迴歸係數估計值便是表 3 之 Goodman 模型橫列百分比估計值。

模型過度嚴苛之同質性假定。

這兩大修正，還有一重要之副產品，亦即 EI 模型也可以估算第  $k$  個集體單位內的參數值  $\pi_{j|i, k}$ 。例如本人指導碩士生吳怡銘 (2000) 以 EI 模型分析台北市 1998 年選舉中，各里選民在市長、市議員兩項公職的一致與分裂投票，不僅估算全台北市各型投票模式之比例，還進一步估算 435 個里內各型一致與分裂投票模式之橫列條件機率，並據以進行以里為單位之背景脈絡分析 (contextual analysis)。

King 的模型及其後續延伸 (King, Rosen and Tanner, 1999; Rosen, Jiang, King and Tanner, 2001) 確實提供了一個強而有力的工具，可應用到一致與分裂投票的分析上 (如 Burden and Kimball, 1998)。但是正如所有的分析法，使用者也必須充分瞭解其限制。誠如 Cho (1998: 149) 所指出：King 的 EI 模型仍然假定沒有集結之偏誤 (aggregation bias) 存在。倘若有集結偏誤，King 的基本 EI 模型估計值也會產生偏誤，除非研究者能洞悉造成集結偏誤的原因，並正確的將之納入 EI 之擴展模型 (extended EI model) 之中，否則參數的估計值並不正確。換言之，該方法再有用，也絕對無法取代研究者對研究標的實質的瞭解與理論的思考。

## 肆、結合個體與集體數據之嘗試

在少數情況下，研究者或許同時擁有關於研究地區投票行為的民調資料與集體數據。文獻中，雖然也有些研究同時運用了集體數據及民調之個體資料，但大多是將二者視之為兩種平行的方法，個別進行分析，並未建立一貫的學理基礎，作有系統之整合。

英國地理學家 Johnston、Hay 及 Pattie (Johnston and Hay, 1982; Johnston and Pattie, 2000) 根據 Wilson (1970: Chapter 2) 在區域研究中引用來推算交通運輸量的 entropy-maximizing (以下簡稱 EM)<sup>15</sup> 法則，加

---

15 Entropy 一字源於熱力學，因此在社會科學中常引起莫名之神秘感與誤解，但此處所指，實為訊息理論創始人 Shannon (1948: 392-396) 之定義，旨在測量機率分佈之不確定性

以調整後提出了一套將民調資料與集體數據結合，推算表 3 內各型一致與分裂投票數的方法。這個方法的目的和 King 的 EI 模型一樣，是把 Duncan 與 Davis 的「上下限推算法」更細緻化，希望在表 3 每個細格的上下限區段內，進一步算出最可能為真的一個數值。

EM 法與 King 的 EI 模型不同之處，是 EM 法並不設定參數  $N_{ij,k}$  之機率分佈模型，而是先將 K 個表 3 堆疊起來，形成一個  $I \times I \times K$  的三維數據立方體 (data cube)，設定其各維的邊緣總和，然後運用非線型最適化 (nonlinear optimization) 的方式，求該立方體內  $I \times I \times K$  個細格的最可能 (即最大或然 maximum likelihood) 數值  $\tilde{N}_{ij,k}$ 。其基本步驟如下：

1. 先把政府公布的 K 個選區的各黨得票數 (集體數據) 做為 K 個表 3 的兩維邊緣總和  $N_{i+,k}$  及  $N_{+,j,k}$ 。
2. 再以民調樣本 (個體數據) 估計出表 2 之聯合百分比  $p_{ij}$ ，以之估算母群之票數分佈  $\hat{N}_{ij,+} = N_{++,+} \times p_{ij}$  後，得出第三維的邊緣總和，連同步驟 1 的兩維邊緣總和，即構成了  $I \times I \times K$  三維數據立方體的所有邊緣總和  $N_{i+,k}$ 、 $N_{+,j,k}$  及  $\hat{N}_{ij,+}$ 。
3. 將這三組邊緣總和設定為已知之約制條件 (constraints)，依照以下之 EM 原則，求兩項公職均有效之總票數  $N_{++,+}$  分佈在立方體  $I \times I \times K$  個細格內的票數  $N_{ij,k}$  之最大或然組合。假定每一位選民落入每一細格的機率都相同，則  $N_{++,+}$  在該立方體細格內的組合 (combination) 數為：

$$W(\{N_{ij,k}\}) = \frac{N_{++,+}!}{\prod_{i,j,k} N_{ij,k}!}.$$

EM 法就是取此一組合數之對數值的極大化：

$$\max_{\{N_{ij,k}\}} \log \left( \frac{N_{++,+}!}{\prod_{i,j,k} N_{ij,k}!} \right),$$

---

(uncertainty)。Wilson 與 Bennett (1985: 237) 認為改稱 EM 法為「機率極大化法」(probability maximizing) 更簡單明瞭。

且需滿足下列三維邊緣總和之約制條件：

$$N_{i+,k} = \sum_j N_{ij,k}$$

$$N_{+,j,k} = \sum_i N_{ij,k}$$

$$\hat{N}_{ij,+} = \sum_k N_{ij,k}$$

取有約制條件之極大化 (constrained maximization)，原則上並不困難，但實際應用時，常需以反覆比例修正 (iterative proportional fitting，簡稱 IPF) 的數值法，解出該三維立方體內  $I \times I \times K$  個細格的最大或然值  $\tilde{N}_{ij,k}$ 。

EM 法的優點，在以訊息理論中 entropy 的概念為基礎，將已知之個體資料與集體數據一併納入考量，求數學上最可能產生前述三組邊緣總和的各細格參數值的組合，因此就像 King 的 EI 模型一樣，EM 法也可推算出  $K$  個集體單位內所有投票類型之票數。不過 EM 法在應用上，仍有下列問題亟待克服：

1. 民調個體資料可能有測量誤差的問題 (如隱瞞、拒答、記錯等)，已如前述，即使以加減權 (weighting) 調整後，前述步驟 2 估計之聯合機率值  $p_{ij}$  仍有抽樣誤差，此種誤差對步驟 3 推算之  $\tilde{N}_{ij,k}$  影響如何，尚待研究，而這方面的探討才剛剛起步 (如 Johnston and Pattie, 1991)。
2. 以 EM 原則推算的交叉表，基本上已經假定各維的變數之間沒有任何的交互影響 (interaction) (Good, 1963: 920; Cleave, Brown and Payne, 1995: 63)。換言之，EM 法隱含了一個有重大實質意涵的假定：同一選區內選民在考慮兩項公職選舉的投票抉擇時，政黨偏好與標籤沒有影響，而且與各個選區之間也沒有交互的影響。衡諸實際，此一假定恐與多數的實際情況不符，若不明就裡，強加於政黨標籤有相當影響力的地區之上進行分析，結果可能會低估一致投票數而高估分裂投票數。

因此，EM 法就像所有其他方法一樣，應用時不應將之視為無往不利的工具，機械式的套用，而需仔細思考其背後的假定與實際研究對象之實況是否相符。例如除了三個維的邊緣總和之外，若還有其他已知之重要訊息，便應設法納入極大化公式的約制條件內，一併考量。

此外，EM 法是在英國發展出來的，目前也僅應用到實施小選舉區及兩票制的國家。若應用至我國的北高兩市分裂投票研究，則至少需將台灣 SNTV/MMD 制下投開票所→里→區→市議員選區→全市（市長選區）的多層級區劃結構納入考量，建立多維的資料表，然後以 EM 法進行多層次之系統分析（multilevel systems analysis）。

## 伍、結論

截至目前為止，以民調為主流的個體數據學派已經完全取代了 1930-1940 年代的人文區位研究，而稱霸於學界。但是從以上的討論，可以得知：個體資料與集體數據之應用於選舉研究，各有擅長。本文所持的觀點是：以集體數據重建與推估原需有母群之個體數據才能精確得知的參數，必然有風險存在；這就好比以樣本推論母群參數，一定要冒抽樣誤差的風險。然而不論數據的層次來自個體或集體，追根究底，一致與分裂投票背後的機制（underlying mechanism）都是選民的投票抉擇，研究的焦點與學理的基礎絕不致於因資料形態的不同而南轅北轍，故研究者理應捐棄成見，思考截長補短之道。本文認為：研究者應該把個體和集體數據視為是同一個現象在不同觀察層次所呈現出來的訊息（information），而學界共同的目標，正是要努力整合不同形態的訊息，使之交相為用、相輔相成，幫助我們透視無法直接觀察到的政治社會現象。

然而要打通個體與集體數據的任督二脈，學界勢必要跳脫目前兩種途徑各自為政、分道揚鑣的格局，回歸議題的基本面，重新思考分裂投票的根本機制與數據資料的產生過程（data generating processes），據以發展能貫穿個體與集體之整合研究設計（integrated research design）。就短程而言，或可採「兩路並進、分頭夾擊」之隧道鑽鑿策略，一方面在民調之抽樣設計上，由下而上，更綿密的扣緊投開票所→里→區→市議員選區（→立委選區）→全市（市長選區）的多層次結構，並將各層次內之樣本數與研究資源做最適調配，使個體資料在樣本結構上，就與選委會公布之選舉結果（集體數據）最契合、最利於兩者之比較與整合；如此，不僅在個體數據分析時可充分運

用第二節所述之多層次模型，而且也提供了最佳之個體、集體搭配架構，俾便彙整民調個體資料與各里或選區之人口特徵、社經條件等集體數據，進行脈絡與成分之分析（contextual and compositional analyses）。另一方面，從事集體數據蒐集與分析的學者，也不能夠再以政府目前已公佈之選舉結果為已足，而應由上而下，充分掌握從市一直到投開票所等多層級區劃的決定因素與程序，思索將個別選民劃歸各行政區域的集結過程（aggregation process），對累計起來的集體數據會產生什麼影響，對跨層次推論分裂投票又會有什麼後果。此外，還需思考在不違反民主政治秘密投票的原則下，督促政府釋出更詳細的投開票紀錄，並授權或委託學術機構收錄、儲存並釋出更詳盡之人口背景（含政府已不再統計之省籍、職業等變數）、社經指標等集體數據，供學界及民間單位使用參考。

但欲練就爐火純青的火候，學者終須面對無可逃避之艱鉅挑戰：也就是綜合短程「兩路夾擊」的成果，建立一套能跨越並貫穿微觀、中觀、宏觀各層次（micro-, meso-, and macro-levels）的一致與分裂投票理論架構，使不同分析層次的切入點非但能關照同一現象的不同層面，而且各點均能推導出邏輯一致、上下連貫的論證，前後呼應。（黃紀，2000b: 129）建立這樣的跨層次理論，誠非易事，但可行的途徑之一，或許是從個體著手，先以理性選擇理論建立個人分裂投票模型，再根據個別選民聚集成各層級集體單位（如投開票所、里、議員選區、全市等）的集結過程與劃分標準，建立關鍵變數之宏觀分佈，據以逐步建構 Stoker（1993: 1830-1832）所主張的「微觀-宏觀合一模型」（micro-macro models）。待此類模型成形後，個體資料與集體數據的藩籬才會漸趨消弭，而一致與分裂投票的研究趨於圓熟。

## 參考資料

王業立

1996 《我國選舉制度的政治影響》。台北：五南圖書出版公司。

吳怡銘

2000 〈台北市選民分裂投票之研究——民國八十七年市長與市議員選舉之分析〉，國立

中正大學政治學研究所碩士論文。

洪永泰

1995 〈分裂投票：八十三年臺北市選舉的實證分析〉，《選舉研究》2(1): 119-145。

陳義彥

1995 〈台北市選民的分裂投票分析〉，〈台北市民眾市政意向之研究〉，頁 29-40。台北：政大選舉研究中心。

許增如

1999 〈一九九六年美國大選中的分裂投票行為：兩個議題模式的探討〉，《歐美研究》29(1): 83-126。

黃紀

2000a 〈質變數之計量分析〉，見謝復生、盛杏媛（編），《政治學的範圍與方法》，頁 387-411。台北：五南圖書出版公司。

2000b 〈實用方法論芻議〉，見何思因、吳玉山（編），《邁入廿一世紀的政治學》（《政治學報》特輯，31 卷），頁 107-139。台北：中國政治學會。

黃紀、吳重禮

2000 〈台灣地區縣市層級「分立政府」影響之初探〉，《台灣政治學刊》，4: 105-147。

黃紀、張益超

2001 〈一致與分裂投票：嘉義市一九九七年市長與立委選舉之分析〉，見徐永明、黃紀（編），《政治分析的層次》，頁 183-218。台北：韋伯文化。

黃德福

1991 〈台灣地區七十八年底選舉分裂投票之初探研究：以台北縣、雲林縣與高雄縣為個案〉，《政治學報》19: 55-80。

張益超

1999 〈選民投票穩定程度之研究——嘉義市第四屆與第五屆市長選舉之分析〉，國立中正大學政治學研究所碩士論文。

Achen, Christopher H. and W. Phillips Shively

1995 *Cross-Level Inference*. Chicago: The University of Chicago Press.

Agresti, Alan

1990 *Categorical Data Analysis*. New York: Wiley.

Aitchison, J.

1986 *The Statistical Analysis of Compositional Data*. London: Chapman and Hall.

Alvarez, Michael R. and Matthew M. Schousen

1993 "Policy Moderation or Conflicting Expectations? Testing the Intentional Models for Split-Ticket Voting," *American Politics Quarterly* 21(4): 410-438.

Alvarez, Michael R. and Jonathan Nagler

1998a "When Politics and Models Collide: Estimating Models of Multiparty Elections," *American Journal of Political Science* 42(1): 55-96.

1998b "Economics, Entitlements, and Social Issues: Voter Choice in the 1996 Presidential Election," *American Journal of Political Science* 42(4): 1349-1363.

2000 "A New Approach for Modelling Strategic Voting in Multiparty Elections," *British Journal of Political Science* 30(1): 57-75.

Bartolini, Stefano and Peter Mair

- 1990 *Identity, Competition, and Electoral Availability: The Stabilisation of European Electorates 1885-1985*. Cambridge: Cambridge University Press.

Beck, Paul Allen, Lawrence Baum, Aage R. Clausen and Charles E. Smith, Jr.

- 1992 "Patterns and Sources of Ticket Splitting in Subpresidential Voting," *American Political Science Review* 86(4): 916-928.

Ben-Akiva, Moshe and Steven R. Lerman

- 1985 *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge: The MIT Press.

Berrington, H. B.

- 1965 "The General Election of 1964," *Journal of the Royal Statistical Society, Series A* 12(1): 17-51.

Bishop, Yvonne M. M., Stephen E. Fienberg and Paul W. Holland

- 1975 *Discrete Multivariate Analysis: Theory and Practice*. Cambridge: The MIT Press.

Born, Richard

- 1994 "Split-Ticket Voters, Divided Government, and Fiorina's Policy-Balancing Model," *Legislative Studies Quarterly* 19: 117-125.  
2000 "Policy-Balancing Models and the Split-Ticket Voter, 1972-1996," *American Politics Quarterly* 28(2): 131-162.

Breslow, N.E. and D.G. Clayton

- 1993 "Approximate Inference in Generalized Linear Mixed Models," *Journal of the American Statistical Association* 88(421): 9-25.

Burden, Barry C. and David C. Kimball

- 1998 "A New Approach to the Study of Ticket Splitting," *American Political Science Review* 92(3): 533-544.

Burnham, Walter Dean

- 1965 "The Changing Shape of the American Political Universe," *American Political Science Review* 59(1): 7-28.

Butler, David and Stephen D. Van Beek

- 1990 "Why Not Swing?: Measuring Electoral Change," *PS: Political Science & Politics* 23(2): 178-184.

Campbell, Angus and Warren E. Miller

- 1957 "The Motivational Basis of Straight and Split Ticket Voting," *American Political Science Review* 51(2): 293-312.

Cho, Wendy K. Tam

- 1998 "If the Assumption Fits . . . : A Comment on the King Ecological Solution," *Political Analysis* 7: 143-164.

Christensen, Ronald

- 1997 *Log-Linear Models and Logistic Regression*, 2<sup>nd</sup> edition. New York: Springer.

Clark, Audrey N.

- 1998 *The Penguin Dictionary of Geography*, 2<sup>nd</sup> edition. London: Penguin Books.

Clausen, Sten-Erik

- 1998 *Applied Correspondence Analysis: An Introduction*. Thousand Oaks, CA: Sage Publications.

Cleave, N., P.J. Brown and C.D. Payne

- 1995 "Evaluation of Methods for Ecological Inference," *Journal of Royal Statistical Society, Series A* 158(1): 55-72.

Cohen, Jacob

- 1960 "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement* 20(1): 37-46.

Converse, Philip E.

- 1969 "Of Time and Partisan Stability," *Comparative Political Studies* 2(2): 139-171.

Cowart, Andrew T.

- 1974 "A Cautionary Note on Aggregate Indicators of Split Ticket Voting," *Political Methodology* 1(1): 107-126.

Crewe, Ivor and David Denver

- 1985 "Introduction: Electoral Change in Western Democracies: A Framework for Analysis," in Ivor Crewe and David Denver (eds.), *Electoral Change in Western Democracies: Patterns and Sources of Electoral Volatility*. New York: St. Martin's Press.

Duncan, Otis Dudley and Beverly Davis

- 1953 "An Alternative to Ecological Correlation," *American Sociological Review* 18: 665-666.

Ellen, R.F.

- 1996 "Ecology," in Adam Kuper & Jessica Kuper (eds.), *The Scoail Science Encyclopedia*, 2<sup>nd</sup> edition. London: Routledge.

Feigert, Frank B.

- 1979 "Illusions of Ticket-Splitting," *American Politics Quarterly* 7(4): 470-488.

Fiorina, Morris P.

- 1994 "Response to Born," *Legislative Studies Quarterly* 19(1): 117-125.  
1996 *Divided Government*, 2<sup>nd</sup> edition. Boston: Allyn and Bacon .

Fleiss, Joseph L.; Jacob Cohen, and B. S. Everitt

- 1969 "Large Sample Standard Errors of Kappa and Weighted Kappa," *Psychological Bulletin* 72(5): 323-327.

Forgette, Richard and Glenn J. Platt

- 1999 "Voting for the Person, Not the Party: Party Defection, Issue Voting, and Process Sophistication," *Social Science Quarterly* 80(2): 409-421.

Glasgow, Garrett

- 2001 "Mixed Logit Models for Multiparty Elections." *Political Analysis* 9(2): 116-136.

Good, I.J.

- 1963 "Maximum Entropy for Hypothesis Formulation, Especially for Multi-national Contingency Tables," *Annals of Mathematical Statistics* 34(3): 911-

931.

Goodman, Leo A.

- 1953 "Ecological Regressions and Behavior of Individuals," *American Sociological Review* 18(6): 663-664.
- 1959 "Some Alternatives to Ecological Correlation," *American Journal of Sociology* 64 (March): 610-625.
- 1984 *The Analysis of Cross-Classified Data Having Ordered Categories*. Cambridge: Harvard University Press.

Greenacre, Michael J.

- 1984 *Theory and Applications of Correspondence Analysis*. London: Academic Press.

Greene, William H.

- 2000 *Econometric Analysis*, 4<sup>th</sup> edition. Upper Saddle River, NJ: Prentice-Hall.

Huang, Chi (黃紀) and Todd G. Shields

- 2000 "Interpretation of Interaction Effects in Logit and Probit Analyses: Reconsidering the Relationship between Registration Laws, Education and Voter Turnout," *American Politics Quarterly* 28(1): 80-95.

Hout, Michael

- 1983 *Mobility Tables*. Beverly Hills, CA: Sage Publications.

Jacobson, Gary C.

- 1990 *The Electoral Origins of Divided Government: Competition in U.S. House Elections, 1946-1988*. Boulder, Colorado: Westview Press.

Johnston, R.J. and A.M. Hay

- 1982 "On the Parameters of Uniform Swing in Single-Member Constituency Electoral Systems," *Environment and Planning A* 14: 61-74.

Johnston, R.J., Derek Gregory, Geraldine Pratt and Michael Watts, eds.

- 2000 *The Dictionary of Human Geography*, 4<sup>th</sup> edition. Oxford, UK: Blackwell.

Johnston, Ron and Charles Pattie

- 1991 "Evaluating the Use of Entropy-Maximising Procedures in the Study of Voting Patterns: 1. Sampling and Measurement Error in the Flow-of-the-Vote Matrix and the Robustness of Estimates," *Environment and Planning A* 23: 411-420.
- 2000 "Ecological Inference and Entropy-Maximizing: An Alternative Estimation Procedure for Split-Ticket Voting," *Political Analysis* 8(4): 333-345.

Jones, Kelyvn and Craig Duncan

- 1998 "Modelling Context and Heterogeneity: Applying Multilevel Models," in Elinor Scarbrough and Eric Tanenbaum (eds.), *Research Strategies in the Social Sciences: A Guide to New Approaches*. Oxford: Oxford University Press.

Katz, Jonathan N. and Gary King

- 1999 "A Statistical Model for Multiparty Electoral Data," *American Political Science Review* 93(1): 15-32.

Katz, Richard S.

- 1997 "Changing Patterns of Electoral Volatility," *European Journal of Political Research* 31 (February): 83-85.

Key, V. O. Jr.

- 1949 *Southern Politics in State and Nation*. New York: Alfred A. Knopf.

King, Gary

- 1997 *A Solution the Ecological Inference Problem: Reconstruction Individual Behavior from Aggregate Data*. Princeton, NJ: Princeton University Press.

King, Gary, Ori Rosen and Martin A. Tanner

- 1999 "Binomial-Beta Hierarchical Models for Ecological Inference," *Sociological Methods and Research* 28(1): 61-90.

Lacy, Dean and Barry C. Burden

- 1999 "The Vote-Stealing and Turnout Effects of Ross Perot in the 1992 U.S. President Election," *American Journal of Political Science* 43(1): 235-255.

Leithner, Christian

- 1997 "Of Time and Partisan Stability Revisited: Australia and New Zealand 1905-90," *American Journal of Political Science* 41(4): 1104-1127.

Lloyd, Chris

- 1999 *Statistical Analysis of Categorical Data*. New York: Wiley.

McAllister, Ian and Robert Darcy

- 1992 "Sources of Split-Ticket Voting in the 1988 American Elections," *Political Studies* 40(4): 695-712.

McCulloch, Charles E. and Shayle R. Searle

- 2001 *Generalized, Linear, and Mixed Models*. New York: Wiley.

Mattei, Franco and John S. Howes

- 2000 "Competing Explanations of Split-Ticket Voting in American National Elections," *American Politics Quarterly* 28(3): 379-407.

Miller, M. L.

- 1995 "Quantitative Methods," in David Marsh and Gerry Stoker (eds.), *Theory and Methods in Political Science*. London: Macmillan Press.

Norris, Pippa

- 1997 *Electoral Change in Britain since 1945*. Oxford, UK: Blackwell.

Pedersen, Mogens N.

1979. "The Dynamics of European Party Systems: Changing Patterns of Electoral Volatility," *European Journal of Political Research* 7 (March): 1-26.

Powers, Daniel A. and Yu Xie

- 2000 *Statistical Methods for Categorical Data Analysis*. San Diego, CA: Academic Press.

Quinn, Kevin M., Andrew D. Martin and Andrew B. Whitford

- 1999 "Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models," *American Journal of Political Science* 43(4): 1231-1247.

Revelt, David and Kenneth Train

- 1998 "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level," *The Review of Economics and Statistics* 80(4): 647-657.

Robinson, W.S.

- 1950 "Ecological Correlations and the Behavior of Individuals," *American Sociological Review* 15(3): 351-357.

Rose, Richard

- 2000 "Ticket Splitting," in Richard Rose(ed.), *International Encyclopedia of Elections*. Washington, D.C.: Congressional Quarterly.

Rosen, Ori, Wenxin Jiang, Gary King, Martin A. Tanner

- 2001 "Bayesian and Frequentist Inference for Ecological Inference: the  $R \times C$  Case." *Statistica Neerlandica* 55(2): 134-156.

Rusk, Jerrold G.

- 1970 "The Effect of the Australian Ballot Reform on Split Ticket Voting: 1876-1908," *American Political Science Review* 64(4): 1220-1238.

Selvin, Hanan C.

- 1958 "Durkheim's Suicide and Problems of Empirical Research," *American Journal of Sociology* 63: 607-619.

Shields, Todd G. and Chi Huang (黃紀)

- 2001 "Toward a Multilevel Analysis of Registration-Voting Decision," 見徐永明、黃紀(編),《政治分析的層次》,頁49-80。台北:韋伯文化。

Sigelman, Lee, Paul J. Wahlbeck and Emmett H. Buell, Jr.

- 1997 "Vote Choice and the Preference for Divided Government: Lessons of 1992," *American Journal of Political Science* 41(3): 879-894.

Shannon, C.E.

- 1948 "A Mathematical Theory of Communication." *The Bell System Technical Journal* 27(3): 379-423.

Stoker, Thomas M.

- 1993 "Empirical Approaches to the Problem of Aggregation Over Individuals," *Journal of Economic Literature* 31(4): 1827-1874.

Train, Kenneth E.

- 1998 "Recreation Demand Models with Taste Differences over People," *Land Economics* 74(2): 230-239.

- 1999 "Mixed Logit Models for Recreation Demand," in Joseph A. Herriges, and Catherine L. Kling(eds.), *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*. Cheltenham, UK: Edward Elgar.

Tukey, John W.

- 1977 *Exploratory Data Analysis*. Reading, MA: Addison Wesley.

Wattenberg, Martin P.

- 1996 *The Decline of American Political Parties: 1952-1994*. Cambridge, MA: Harvard University Press.

Wattenberg, Martin P.

2000 "Ticket Splitting in the United States." in Richard Rose(ed.), *International Encyclopedia of Elections*. Washington, D.C.: Congressional Quarterly.

Wilson, A.G.

1970 *Entropy in Urban and Regional Modelling*. London: Pion.

Wilson, A.G. and R.J. Bennett

1985 *Mathematical Methods in Human Geography and Planning*. Chichester: Wiley.

# **Straight- and Split-ticket Voting: Methodological Reflections**

**Chi Huang**

Department of Political Science, National Chung-Cheng University

## **ABSTRACT**

The practice of endorsing candidates of different parties while marking a ballot, called split-ticket voting, has long attracted the attention of scholars interested in electoral studies. However, despite its prominence, the literature of this subject suffers from a lack of agreement on the appropriate research methodology. This paper first provides a comprehensive review of how straight- and split-ticket voting is conceptualized and practised, and then examines various analytical approaches for individual and aggregated data. We argue that micro-level surveys and macro-level statistics should be treated as information observed at different levels of the same studied phenomenon, i.e., vote choice. The long-term aim is to construct a unified theory for ticket splitting on both the micro and macro level, that integrates the motives of individual voters and the voting patterns of the electorate as a whole. Such a micro-macro model can only be developed, however, if researchers think carefully about the underlying mechanism of voting behavior as well as the data generation process of the levels of information at hand.

**Key Words:** straight-ticket voting, split-ticket voting, mixed logit model, cross-level (ecological) inference, information theory, entropy maximization