

探索選民的投票行為變化： 應用機率分配模型的預測方法^{*}

張順全^{**}、莊文忠^{***}

《本文摘要》

過去許多有關選舉預測的研究發現，選舉民調中的贏家最後卻未能勝出，箇中隱藏不少值得深入探究的問題。本研究從方法論的觀點，嘗試利用過去選舉經驗資料，改進單純使用民意調查結果所做之選舉預測。就預測方法論而言，本文提出的選舉預測模型乃符合貝氏統計架構，此一架構所欲掌握的是選民投票行為中所潛藏的異質性，據以探索選民的投票行為昨日、今日、明日的變化。綜言之，本篇文章的研究所得價值有三：(1) 成功應用貝塔—二項機率分配 (beta-binomial distribution) 模式改進單純使用民意調查結果，發展輔助民調預測模型，及提供作為評估該次選舉是與過去開票結果不同的「變天型」抑或是「維持型」的預警指標；(2) 同時利用此一機率分配模型處理未表態資料，預測選舉民調中未表態受訪者的投票意向；(3) 將本研究所建立之模型應用在國內外不同選舉，檢證本模型的適用情境，無論是 2004 年美國總統大選、2006 年台灣高雄市長選舉、抑或是 2008 年台灣總統選舉的候選人得票率預估，都獲得相當準確的效果。此外，本研究結果也同時指出若干有待思考的問題，說明未來值得繼續發展的方向。

關鍵詞：選舉預測、機率分配模型、民意調查、投票行為、貝氏統計

^{*} 兩位作者感謝台灣大學政治系洪永泰老師、交通大學經營管理所唐瓊璋老師以及兩位匿名審查人對本篇論文所提出諸多的寶貴修正建議。尤其是本篇論文的構思靈感，源自於洪永泰教授前於本期刊創刊初期發表的政治版圖選舉預測模型，以及唐瓊璋教授所講授的「行銷工程」課程內容。當然，本文文責由作者自行承擔。

^{**} 交通大學經營管理研究所博士候選人。E-Mail: b110710@mail.nhitb.gov.tw。

^{***} 世新大學行政管理學系助理教授。E-Mail: jwj@cc.shu.edu.tw。

壹、前言

無論是參與選舉或是關心選舉的人都知道，選舉過程中如何精確地預測選民投票意向具有一定的難度。每次選舉預測之所以深具挑戰性，一方面源自於特定選舉處於不同的時空系絡，另一方面則來自選民個人迥異的主體特質。簡言之，選民的投票行為具有異質性 (heterogeneity)。一般來說，當我們利用民意調查來探究選民投票意向時，其中的問項常以影響選民投票行為的顯性因素（如選民人口統計變項、政黨屬性、經社指標等）為主，這些因素乃屬於可直接觀察的選民異質性 (observed heterogeneity)。但是，牽涉到選民內隱的價值觀等不可直接觀察的異質性 (unobserved heterogeneity)，雖深知它會左右投票行為，但卻是以民意調查方式為基礎的選舉預測所不易掌握的。職是之故，無論國內外的選舉民調，除了存在訪問失敗的問題或抽樣及估計方式的偏差，再加上民調內容對選民異質性的掌握不佳，導致許多選舉預測經常是與選舉結果相違，或者即使是成功預測由誰當選，但是預測得票率和實際得票率仍相去甚遠。¹

無論如何，選舉預測仍是選舉過程中極其重要的資訊依據（洪永泰 1994）。換言之，包括政黨、候選人或媒體在內，為了即時掌握民意，在選前不同時間階段，皆會大量運用民意調查來探索當下的民意取向，以評估選戰競爭的程度，又被稱為「賽馬式」(horse race) 民調。他們所關注的焦點常是為瞭解候選人受選民青睞的消長趨勢，並且據以預測未來選舉結果。值得注意的是，當各種民調充斥於媒體報導中，許多選民接觸到媒體所披露的最新選情時，往往認為那確實是反映出當下真正的民意，但是，選舉結束後，當我們檢視投票結果時，也不時發現選舉民調中的贏家，最後卻是未能勝出，尤其是戰況膠著的選舉，許多選舉預測方法的準確性更是倍受質疑。以美國為例，2000 年和 2004 年的總統選舉空前的激烈，2000 年甚至發生選民得票數多的一方和選舉人團得票多的一方不同，充滿爭議的佛州選票格式和驗票訴訟，使得最後必須以史上罕見的司法判決方式來決定白宮主人。令人氣餒的是，在這兩次總統選舉中，因為不少州的選戰陷入膠著狀態，選前預測錯誤的民調早已司空見慣，也引起各界對選舉預測方法的討論。

同樣的民調預測失準情節在台灣亦可說有過之而無不及，一個有趣的例子是 2002 年與 2006 年的高雄市長選舉，不少媒體民調都「猜錯」了當選人（如表 1），另外，2008 年的國內總統選舉剛結束，選前評估兩方相當接近的街頭耳語和不同機構的民調結果也是相當分歧。就結果論而言，一方面可以發現民意調查的確有其侷限性，似乎有許多選民內隱的訊息不是單純從民意調查可以獲得；另一方面，值得吾人省思的是，不同時點所做的民

¹ 有一主因乃是即使研究人員常設計題項來偵測未表態選民的投票意向，台灣的選舉民調受訪者不願意表態的比例仍經常高達三、四成，而美國選舉民調也常存在近一成未表態的樣本。

調用來作為瞭解候選人支持度的穩定程度，也似乎勝過於用於準確地預測選舉結果。

表 1 高雄市長選舉民調與實際選舉結果

	選前 1 個月	選前 5 天	選前 3 天	選舉 結果
2002 年	2002/11/4	2002/12/2	2002/12/4	
樣本數	N=903	N=931	N=1041	
謝長廷（民）	42.0%	39.0%	39.0%	50.04%
黃俊英（國）	27.0%	42.0%	43.0%	46.82%
張博雅（無）	9.0%	4.0%	3.0%	1.75%
施明德（無）	4.0%	2.0%	2.0%	1.13%
黃天生（無）	-	0.2%	0.7%	0.26%
未決定	18.0%	13.0%	12.0%	-
2006 年	2006/11/5	2006/12/4	2006/12/6	
樣本數	N=830	N=1086	N=957	
黃俊英（國）	52.0%	51.0%	50.0%	49.27%
陳菊（民）	29.0%	31.0%	36.0%	49.41%
羅志明（台聯）	2.0%	2.0%	2.0%	0.86%
林志昇（無）	0.2%	0.0%	0.2%	0.23%
林景元（無）	0.4%	1.0%	0.1%	0.23%
未決定	16.0%	16.0%	11.0%	-

資料來源：作者自行整理。

雖然民意調查在執行層面無可避免有許多問題存在，不過，純粹站在統計學估計的角度，無論是政黨、候選人或是媒體仍熱衷於在不同時點做民調，其背後的原因通常是基於傳統統計學即頻度學派 (Frequentist) 的決策觀點。因為傳統統計學古典機率的定義是建立在可以重複多次試驗的發生頻率的基礎上，即用以推論真正的選舉結果的資料是樣本民調結果，並且排除事前可能的任何主觀判斷。但是，如果只是用民調結果在緊要關頭去研判勝利的歸屬，當勝負可能只是在 1 個百分點之內（包括 2000 年美國總統選舉、2004 年台灣總統大選和 2006 年高雄市長選舉均是顯例），包含了未表態樣本的民調結果是否真的能命中選舉結果得票率實在是有待商榷。有鑑於此，近年來，有識者認為僅依賴民調數據進行選舉預測的風險實在相當高，有必要加入其他的資訊輔助預測，例如加入過去選舉開票

結果就是可能的途徑之一。在此一系絡下，重視過去所發生的經驗的統計分支貝氏學派 (Bayesian) 的多種統計推估方法，伴隨著電腦運算功能提升，也漸之受到重視，如國外已有一些學者將貝式統計分析的理念應用於比較政治及民意政治學域（蔡佳泓 2007, 6）。

事實上，貝氏學派和頻度學派的統計推論乃是建立在不同的統計決策程序觀點 (Berger 1985)。頻度學派的機率觀點及其統計推論是以重複多次試驗的理念為基礎；貝氏學派的觀點則是承認以過去結果當成先前的主觀經驗，據以修正現況的觀察結果，並且依據最新的證據不斷更新。貝氏觀點通常也可以表示為決策機率來自觀察資料且經過事前資訊的相互校正 ($\text{posterior} \propto \text{prior} \times \text{likelihood}$)，亦即事後決策機率的陳述，也反映出給定的狀況下對研究者先前主觀信念的修正。據此，在貝氏統計的架構下，母體選民不可觀察的異質性適可納入先前機率的經驗假設中。要言之，就統計學而言，具體代表異質性的作法就是考慮隨機變數的機率分配代表事前資訊，這也正是引發本文作者尋找合適的機率分配來描述母體選民不可觀察之異質性的研究初衷。

本篇文章所提出的論點，就資料層面即嘗試利用過去選舉經驗，提供我們改進一般單純使用民意調查結果所做之選舉預測；另外，就預估方法論層面，本文提出符合貝氏統計觀點的新架構進行選舉預測。同時，本文擬建構的選民投票行為機率分配模型，將據以分辨該次選舉是與過去開票結果不同的「變天型」抑或是「維持型」選舉，以及用來偵測出可能變盤的關鍵時間點以作為擬定選戰策略的參考。綜言之，本篇文章的研究目的有三：(1) 找尋合適的機率分配模式改進單純使用民意調查結果，發展輔助民調預測模型及提供評估選舉結果是否變盤的預警指標；(2) 採用貝氏統計觀點連結過去先驗知識和當前民調結果，建構投票行為的機率分配模型須同時處理未表態資料；(3) 將本研究所建立之模型應用在國內外不同選舉情境中，以檢證本模型的適用性，並提出未來的研究方向。

貳、文獻探討

基本上，有關選舉預測的研究主要可以分為兩大方向 (Holbrook and DeSart 1999)：一是預測模型的建構，二是預測模型的評估。就選舉預測模型的建立而言，建立預測模型的重要步驟包括應該挑選那些變數進入模型之中，以及利用何種統計估計方法進行預測。以美國為例，選舉預測在 1950 年代之後即成為歷次總統選舉的重要觀測指標，各個調查機構亦以選舉預測的準確性作為提升專業與建立形象的重要手段，因而積極開發各種選舉預測的模型。雖然有關選舉預測的批評不曾間斷，不過也已取得具有學術研究價值的正當性 (Holbrook and DeSart 1999)。進一步檢視美國各種選舉預測模型可以發現，政黨支持、政治意識形態等長期因素，與經濟狀況、候選人的特質等短期因素，結合這兩種因

素所建立的選舉預測模型通常具有較高的預測力 (Holbrook and DeSart 1999)。再者，美國有關總統預測模型的建構中 (Soumbatiants et al. 2006)，一般所謂建立「投票函數」(vote functions)，也是利用影響投票行為的重要因素作為解釋變數，據以解釋整體的選舉結果的因果關聯。

觀察國內自民國 83 年開放省市長民選之後，隨著政治環境的日漸開放，有關選舉預測的研究在過去 10 年間也蓬勃發展。有鑒於民調中候選人得票資料屬於類別變項，因為結果變項屬於是否支持某一候選人的二分法問題，其應用統計技術，一般會應用到二項分配 (binominal distribution)，若有多名候選人，則是採用進階的多項分配 (multi-nominal distribution)，如文獻中有「洛基模型」(logit model) (張紘炬與林顯毓 1995；張紘炬與黃男璋 2000)、「多元洛基模型」(multi-logit model) (盛杏媛 1998)、「洛基迴歸分析」(logistic regression) (陳義彥、鄭宇庭與蔡孟熹 1998) 等，也被運用於選舉預測，這些屬於迴歸模型的預測方法都牽涉到對解釋變數的選取，也大部分都可以利用民意調查的問項得到相關變數的訊息。

但是，就預測模型的評估而言，研究者最關切的卻是在形形色色的選舉預測模型之中，有什麼標準可以作為判斷預測模型好壞的依據？又模型中參數估計是否有助於解釋選戰實質意義？對此一問題，Holbrook 與 DeSart(1999) 指出，一個好的預測模型有幾個構成要素：(1) 必須對候選人的得票率提供精確的估計；(2) 必須在選舉結果發生之前即時地做出事前預測；(3) 預測模型最好是儘可能地簡單易懂和方便操作。換言之，一個好的選舉預測模型應該滿足三個條件：精確 (accuracy)、事前 (priority)、簡約 (parsimony)，此三個條件必須兼顧理論與實務需要。

預測模型的建構是科學也是藝術，迴歸模型如果用以預測而非解釋因果關聯，通常使用少量的變數以符合簡約 (parsimony) 原則即可，但是我們往往發現許多研究所採用的迴歸模型為了提高預測的說明效果，只要有貢獻的變數都可能被納入以提高有限的模型適配度 (R^2)，這不僅讓單次所得具有高度解釋力的自變項不一定適用於歷次選舉，也可能過度誇飾許多根本是多餘之自變項的解釋力。舉例言之，即使沒有考慮使用解釋投票行為的重要因素作為自變數的情況下，也就是不使用常見觀測選民投票行為的影響因素來解釋與預測選舉結果的傳統迴歸作法，Holbrook 與 DeSart(1999) 也曾嘗試利用 1992 和 1996 年美國總統選舉期間各州選前的民意調查結果和前次的投票結果來預測各州的選舉結果，其結果發現候選人如果在 9 月份的民調中領先可以贏得 11 月份的選舉。Holbrook 與 DeSart(1999) 除了利用交叉分析方法比較各州的民調結果與實際選舉結果的準確性外，亦利用這兩個變數建立選舉預測方程式，如下 (1) 式所示：

$$\text{Vote}_i = \alpha + \beta_1 (\text{poll}_i) + \beta_2 (\text{prior vote})_i$$

V_i : 某一黨候選人在第 i 州的得票率

poll_i : 某一黨候選人在第 i 州的民調支持率

prior vote : 上次選舉中某一黨候選人得票率

(1)

Holbrook 與 DeSart 兩位作者以 1992 和 1996 年總統選舉民主黨候選人 Clinton 的得票率作為依變數，驗證此一預測模型的預測效果。如果只利用上次選舉的得票率作為預測變數，1992 年和 1996 年的模型解釋力為 0.70 和 0.76；如果只利用民調結果作為預測變數，1992 年和 1996 年的模型解釋力為 0.81 和 0.88；如果同時利用上次選舉的得票率和民調結果作為預測變數，1992 年和 1996 年的模型解釋力提高至 0.86 和 0.92，而各州選舉結果的正確預測率為 90% 和 92%。

再者，Brown 與 Chappell(1999) 亦利用民調資料及其歷史資料來建立選舉預測模型，其特色是最新的民調資料可即時地更新預測模型對選舉結果的預測值。此一預測模型考慮到過去的得票實力和影響選舉的重要因素，以此作為預測的基礎，再利用民調資料反映最新的競爭態勢，最後以兩者的加權平均組合作為整體的預測值，且隨著新民調資訊的產生，可以不斷地修正預測值，也可說是同時兼顧傳統與當前的因素。根據其實證的結果，越接近投票日，其預測的精確度高於 60 天或 90 天以前的預測結果。不過，此一模型的缺點是所有的參數是根據過去選舉和民調的資料所估計出來，是固定不變的，無法隨著時間的變化而彈性修正。

Soumbatiants 等人 (2006) 曾利用美國選前各州的民調結果，並選取選前 15 天和選前 1 天的民調資料作為比較分析，以 Monte Carlo 模擬方法為基礎來預測 2004 大選中民主黨候選人的可能得票率，並且用以預測 2004 年的美國總統選舉中各州的選舉結果與選舉人團 (Electoral College) 的投票結果。其預測途徑具有三大特色：(1) 原始數據來自於各州的民調資料，因而可以單獨預測各州的選舉結果與選舉人團的投票情形。(2) 可以將各種可能結果以機率的形式呈現，因此，可以考慮到選前全國性或各州導致選舉結果產生變化的事件。(3) 可以用來評估各州在選舉人團投票中扮演關鍵角色的可能性。

以上屬國外文獻分析部分，通常在沒有考慮與選民結構直接有關的解釋變項的情況下，學者會以選前的民意調查結果和過去的投票結果來預測該次的選舉結果。不過，這些預測方法在台灣本土選舉中的應用可能會面臨兩個不同的情境：(1) 絕大多數有關台灣總統選舉的民調都是以台灣整體地區為調查範圍，除非是剛好有其他選舉合併舉行，如縣市長選舉或立法委員選舉，否則台灣總統選舉較不容易也不需要同時取得 23 個縣市的個別總統選舉的民調資料。(2) 台灣總統選舉屬於直接選舉制，由選民投票產生總統，各縣市

的輸贏並不會影響選舉結果；反觀美國為間接選舉制，各州是以「贏者全拿」的規則來決定全州的選舉人團票由哪一位候選人獲得，因而預測美國各州的選舉結果對於掌握選舉人團的選票歸屬相對重要。

相較於國外學者加入歷史紀錄的統計分析來改良民調結果，已廣泛應用於選舉預測與比較政治等研究領域，國內相關研究的文獻仍不多見。洪永泰（1994）所發展的政治版圖預測模型 (Aggregate Data Assisted Model, ADAM) 即是其中先驅，事實上 ADAM 也隱含了貝氏觀點，其作法混合使用選區內的民意調查和前幾次的投票結果整體資料，以民意調查為主，整體資料為輔，希望能發揮兩者之長，互補其短。因為貝氏觀點是建立在先前的主觀經驗判斷和對試驗結果的觀察為修正基礎，ADAM 模型則是表態的部分以民意調查結果做估計，未表態的部分把民意調查的資料（通常是：「如果現在就投票，您最有可能投給哪一位候選人？」這一題的答案分布）和政治版圖（亦即受訪者所在的集群所屬）交叉列表後，進一步把每一集群內回答「未決定」、「不知道」、「無意見」、「未回答」等選項的人數依照整體資料中各該集群的歷次選舉資料予以研判後分配給各候選人。

除此之外，國內近年來從事選舉研究學者隨著統計分析功力的提升，已有學者利用貝氏方法進行統計推論，如俞振華與蔡佳泓（2006）應用 Park、Gelman 與 Bafumi(2004) 所發展的多層次的貝式統計，以事後分層的概念，從人口統計資料得出各個人口特徵的交叉資料，並且以上一次的總統選舉結果作為各個縣市的參數，成功的預測各個縣市的選舉結果。此外，Yu 與 Tsai(2006) 亦利用「台灣選舉與民主化」2003 年的調查資料進行貝式統計分析，指出各縣市在政府角色及政策議題等民意上差異不大，只有在台灣與中國之間的互動問題上有顯著的南北差異。

蔡佳泓（2007）將貝氏定理應用在「小區域估計」² 則是另外一例。為了在各區域 (domain) 或團體取得夠多的樣本，傳統的分析方法是在進行調查時膨脹樣本數，以提高小區域估計的效度和信度。此一作法係利用機率抽樣的原則，即樣本估計的期望值，將等於各個界域的樣本估計期望值的總和，也就等於母體的參數，另外，也可以從迴歸係數往回推算各界域的參數估計，因此被稱之為「直接估計」。不過，如果有些層次或界域的樣本數非常小，便必須採取「間接估計」的方法，借助其他界域的參數進行估計。作者即是採取「間接估計」的方法，以 2006 年的「反貪倒扁運動」為例，利用貝式統計中的階層模型估計二元事件的機率，透過迴歸分析的方法，考慮界域之間的變異，估計出台灣七大地理區域對反貪倒扁運動的支持度。

² 所謂的「小區域估計」，係指分析母體中各區域或團體等層次的參數，其在公共政策的用途之一即在於運用各區域或團體的資料，協助瞭解特定地區或團體的政策需求，在政策設計時針對不同民眾的需求作出回應（蔡佳泓 2007, 2）。

由前述的文獻檢視中不難發現，現代選舉可說已無法離開民意調查。在現代民主社會講求政治行銷的氛圍下，如何探求民之所欲，並提出令人認同的政策或設定議題方案，部署整個選戰的戰略方向與戰術原則，甚至競選團隊要進行組織配票，均有賴民意調查。不過，吾人不應忽略此一基本假定的存在：選民在民意調查中均誠實且無保留地回答支持對象。試想民調中倘若全部受訪者皆表態，就不同時間點而言，吾人即可將民調作為評估特定期間選戰策略的參據，候選人也可以根據民調結果評估資源應該投入哪些邊際報酬率最高（即最有助於勝選）的區域，以達到有效的資源分配 (Soumbatiants et al. 2006)。但是實際上民調永遠存在有部分未表態的資料須處理，尤其是在台灣的特殊政治系絡中，未表態部分一直是準確預測的一大障礙。尚且，如果為了進一步探索選民投票行為的過去、現在與未來的變化，那麼只有看到代表現況的民調數據顯然是不夠的。此外，從前述的文獻中亦可以得知，看重過去選舉整體資料或先驗知識 (prior knowledge) 可鑑往知來，能同時兼顧傳統與當前之數據進行選舉預測工具中，其所運用的統計方法不一而足。

另一方面，從過去選舉整體資料出發，也是瞭解今昔選舉是否出現變盤最好的出發點。其立論基礎與經濟學者 P. Samuelson(1938) 提出來的顯示性偏好理論 (revealed preference theory) 說法相近，其所呈現之基本精神皆是民衆在一定條件下所作所為會暴露他內在的偏好傾向。換言之，顯示性偏好理論告訴我們可以根據選民過去的投票實際作為描述選民無可觀察的內在偏好，此一偏好具有異質性，至於在統計實作方面，代表投票行為異質性的具體作法則可以考慮應用機率分配作為選民過去投票實際作為的資訊。本研究將從直效行銷 (direct marketing) 管理領域中發展應用上已相對完備的貝塔—二項機率分配 (beta-binomial distribution) 模式 (Morwitz et al. 1998) 為基礎，並於下章節加以推演其適用情境以符合選戰現實。

參、貝塔—二項機率分配模式在選舉預測的適用性

歸納前述文獻可以得知，除了使用影響選民投票行為的可觀測異質性 (observed heterogeneity) 來解釋與預測選舉結果的傳統迴歸作法外，尤有進者，亦有進一步考量如何精緻量化選民投票行為中不可觀察的異質性 (unobserved heterogeneity)，例如利用貝氏統計來描述選民的投票作為等。以下，筆者將推演貝塔—二項機率分配模式，思考如何將其應用於描繪選舉預測，以及提供我們分析選民的過去、現在與未來投票行為變化的整體架構。

一、民意調查結果以二項分配表示

假設先將民意調查回答樣本設定為來自 k 個區隔 (segments)， k 可從 1 到大至每一個樣本自屬 1 個區隔，此區隔也可以是行政區域：如美國的州、台灣的縣市或鄉鎮市區；或來自以變數集群劃分 (cluster analysis) 的特定選民市場區隔，如以政治勢力分隔的區域；當然，區隔也可以設為代表不同民調的時間點。如果每一個區隔內的回答樣本數假設為 M_i ，在區隔 i 內的選民平均特定的候選人支持率為 P_i ， X_i 為該區隔內所得到特定候選人的支持人數，則具有二項分配的特性，其平均數及變異數如下：

$$P(X_i = x | P_i, M_i) = \binom{M_i}{x} P_i^x (1 - P_i)^{M_i - x}, \quad x = 0, 1, \dots, M_i, \quad (2)$$

$$E(x_i | P_i, M_i) = P_i M_i, \quad Var(x_i | P_i, M_i) = P_i (1 - P_i) M_i. \quad (3)$$

二、各區隔選民支持率的異質性以貝塔分配表示

由於區隔間差異大，區隔內差異小，因此假設每一個區隔 i 內都有一個不同的 P_i ，也就是說假設 P_i 是來自一個特定的分配，在常見的機率分配中，貝塔分布 (beta distribution) 是少數取值在 0 至 1 的區間的分布，正好合適用以描述 P_i 。其原理如下式 (4) 所表示，其中 $B(\alpha, \beta)$ 是貝塔函數 (beta function) (Abramowitz and Stegun 1972)， P_i 跨所有區隔 i 具有貝塔分配的特性，其平均數及變異數如下：

$$g(P_i | \alpha, \beta) = \frac{P_i^{\alpha-1} (1 - P_i)^{\beta-1}}{B(\alpha, \beta)}, \quad 0 < P_i < 1,$$

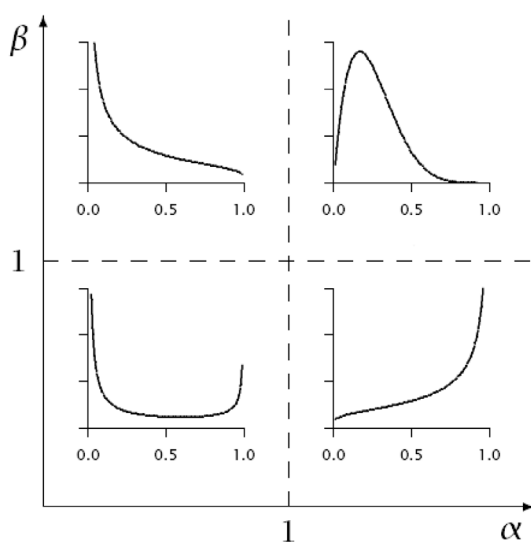
$$\text{其中 } B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt; \quad (4)$$

$$E(P_i | \alpha, \beta) = \frac{\alpha}{\alpha + \beta}, \quad (5)$$

$$Var(P_i | \alpha, \beta) = \frac{\alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}.$$

根據上述的公式，貝塔分配會隨著 α, β 的變化而改變其形狀，其機率分配如下圖 1 所示。當 $\alpha, \beta < 1$ 時呈現 U 形；如果 α, β 其中只有 1 個參數小於 1，則呈現 J 形或反 J 形； α, β 相等時則呈現對稱分布。值得注意的是，在 $\alpha, \beta > 1$ 時，符合我們對支持率 P_i 分布的直觀經驗。換言之，特定候選人過去的歷次得票率通常會類似一個介於 0 至 1 之間

的單峰機率分布，這時候圖形的最高點即眾數為 $\frac{\alpha-1}{\alpha+\beta-2}$ (Abramowitz and Stegun 1972)。總括而言，(5) 式中 $\frac{\alpha}{\alpha+\beta}$ 代表平均得票率，在 β 不變下， α 越大，圖形的最高點或平均得票率皆會越遠離均勢 (50%) 而向得票率等於 1(100%) 靠攏，此情況若發生於雙雄對決的情境，表示傳統上有一方勝出的機會恆大於另一方。



資料來源：作者自行整理。

圖 1 貝塔分配不同參數的圖形變化

舉例言之，以國內的民進黨為例，2008 年總統選舉前，民進黨在近幾次重要選舉（包括 2004 年總統選舉，2005 年縣市長選舉，2008 年立委選舉政黨得票等）的平均得票率為 44.99%，最高達 50.20%，最低為 41.88%（如表 2 所示）：

表 2 2004 至 2008 年選舉民進黨的總得票率

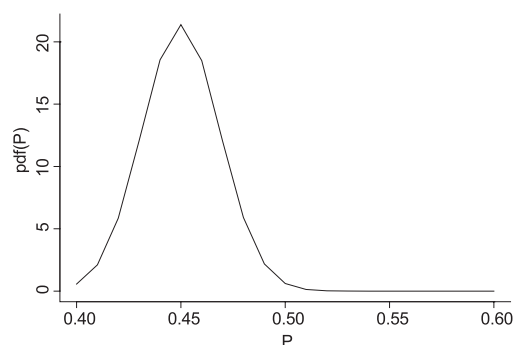
選舉類型	得票率 (%)
2004 年總統選舉	50.20
2005 年縣市長選舉	42.90
2008 年立委選舉政黨得票（第 2 階段）	41.88
2004-2008 平均值	44.99

資料來源：作者自行整理。

說明：1. 本表格之數據不包括金門縣和連江縣。

2. 2005 年縣市長選舉中，台北市和高雄市為 2006 年市長選舉。

利用上述資訊並且假設 95% 信賴區間介於 41.88% 到 50.20%，即可代入貝塔的理论分配取得估計相關參數，當作取值在 0 至 100% 之間的選民母體投票取向之機率表示（如圖 2），也就是說，可以用貝塔機率分配來描述選民過去的投票實際作為，此一機率分配即反映出民進黨在國內大型選舉整體可能得票率的機率分布。當然，隨著歷次選舉得票率的變動，由於平均值與上、下限的變動，此一機率分配亦會隨之改變，此正好可以反映出民進黨政治實力的消長。值得注意的是，由於納入考慮的選舉次數不同，其平均數與上下限可能會有所差異，因而影響到得票率的估計，是以，如果某次選舉中出現脫黨參選的候選人（如宋楚瑜在 2000 年總統選舉中脫離國民黨，以獨立候選人參選），可能造成低估該政黨的得票率。另一個變數則是，如果納入考慮的選舉所橫跨的時間太長，也可能因為時空環境已有大幅度的改變，導致政黨得票率的偏差估計。在這兩個因素的限制下，分析者應可考慮選擇近幾年內（如 5 年內）的選舉作為估計貝塔機率分配的依據，且選舉中如有政黨分裂的情形，應考慮該次選舉結果作為政黨得票率的代表性，考慮做其他的處理，如將脫黨獨立參選的候選人得票率加總，或以其他選舉取代。另外，貝塔的理论分配相關參數的估計，也可採用其他數理模擬的方式，例如 MCMC 法或 Bootstrap(Efron 1979) 來估計，因非本研究所欲處理的焦點，故不再贅述。



資料來源：作者自行整理。

圖 2 2004 至 2008 年選舉民進黨得票率以貝塔機率分配示意圖

三、連結傳統（貝塔分配）和當前（二項分配）訊息改進民意調查結果作預測

由於每一個區隔內都有一個未知的 P_i 必須估計，當我們假設 P_i 是來自一個特定的分布機制所產生，也就是說回答支持特定候選人的樣本數 X_i 可以由貝塔分配和二項分配組合推得下式 (6)，由於可從民調資料可以得到 M_i 和 x ，只有 α, β 是必須估計的部分，估計方式實務上可以是參考專家的經驗值或是過去同性質選舉的得票率（如同圖 2 的示例）。如果以前開票結果的基本資料不可考，或研究者不信其代表過去實際競爭實力（例如可能有人認為 2004 年台灣總統大選結果，是民進黨行政首長選舉得票率的異數），也可退而求

其次改採用民調資料，研判影響當前情勢的機制是同樣來自影響過去的政治勢力的「一個看不見的手」的話，亦可採用數理方法中最大概似法 (Kalwani, et al. 1980) 估計 α, β ，例如：利用一組民調數據可以綜合不同時間點得到 α, β 的估計，並假設此一機制不但影響過去、現在、也會繼續影響未來 P_i 的大小，只是這假設往往會因為民調品質的優劣而影響結果。換言之，許多民調或因未表態選民的意向混沌，或因選樣結構的偏差，也未能充分反應當前情況，更遑論影響過去和現在可表示為同一機制。尚且，欲清楚掌握傳統勢力的梗概，倘若只有一組民調資料，那麼最多也只能產生一組的 α, β 估計，另外必須靠數理模擬的方式，例如 Bootstrap，才能產生多組的 α, β ，進而求其平均趨勢，只是這樣的作法，都不如直接用過去選舉的得票結果來的自然，方便與讀者溝通。

$$P(X_i = x | \alpha, \beta, M_i) = \int_0^1 P(X_i = x | P_i, M_i) f(P_i | \alpha, \beta) d(P_i) \quad (6)$$

$$= \binom{M_i}{x} \frac{B(\alpha + x, \beta + M_i - x)}{B(\alpha, \beta)}.$$

四、連結傳統（貝塔分配）和當前（二項分配）描繪投票行為屬貝氏機率模型

由於我們假設是來自一個特定的貝塔分布機制所產生，也就是當決定後，每一個區隔內都有一個未知的 P_i 必須估計， P_i 事實上可以用貝氏定理推導如下式 (7)。

$$f(P_i | X_i = x, M_i) = \frac{P(X_i = x | M_i, P_i) g(P_i)}{\int P(X_i = x | M_i, P_i) g(P_i) dP_i} \quad (7)$$

$$= \frac{1}{B(\alpha + x, \beta + M_i - x)} P_i^{\alpha + x - 1} (1 - P_i)^{\beta + M_i - x - 1}.$$

得到是來自一個特定的符合參數為 $\alpha + x, \beta + M_i - x$ 的貝塔分布，期望值如下式 (8)。

$$E(P_i | X_i = x, M_i) = \frac{\alpha + X_i}{\alpha + \beta + M_i} \quad (8)$$

$$= \frac{M_i}{\alpha + \beta + M_i} \left(\frac{X_i}{M_i} \right) + \frac{\alpha + \beta}{\alpha + \beta + M_i} \left(\frac{\alpha}{\alpha + \beta} \right), \quad (9)$$

$$= W \left(\frac{X_i}{M_i} \right) + (1 - W) \left(\frac{\alpha}{\alpha + \beta} \right).$$

將 (9) 式拆解後， P_i 估計期望值可視為是來自傳統和當前加權平均數，權重分別是 $1-W$ 和 W ，如此我們就可以設定區隔內的變盤指標為 $W/1-W$ ，當 $W/1-W$ 越大於 1 就表示

P_i 估計期望值越受到當前數據的影響，也就是表示區隔 i 較易「變天」，當 (α, β) 越大時 W 越小， P_i 估計期望值會越接近原始基礎數據。排序每一個區隔 i 的 $E(P_i | X_i = x, M_i)$ 尚可決定特定候選人的最強到最弱區的政治版圖，據此，可進一步掌握選戰應投入的重點。如果區隔設為代表不同民調的時間點，承 (9) 式， $E(P_i | X_i = x, M_i)$ 尚可觀察某特定候選人在不同時點的 $1-W$ 和 W 不同消長趨勢。

肆、本研究模型的應用與實證

在許多推論性統計中，選擇適當的機率分配適用在實際資料上，是估計的一大課題，經由前一節有關貝塔二項機率分配模式的理論表述，有助於吾人理解如何利用貝氏機率的原理將歷史資訊與最新資訊相互結合後進行選情研判，同時也反映了選舉競爭的過去與現在，甚至得以預測未來，改進單純使用民意調查結果進行預測的風險，進而提升民調預測的表現。再者，正如同前文所述，貝式統計強調事前機率的重要性，因為事前機率類似一種權值，權值越小代表事後機率越接近樣本估計。因此，貝式統計可以幫助研究者找到更小的參數估計值的標準誤，而驗證事後機率估計的成果，更是貝式統計研究不可或缺的部分（蔡佳泓 2007, 2）。由於僅停留在概念上的說明並無法驗證此一做法的可行性與可靠性，是以，以下先利用美國 2004 年總統選舉為例，分為三個部分加以驗證，最後再應用台灣的選舉資料突顯本模型的實證價值。

一、美國 2004 年總統選舉民主黨整體得票率的預估

本研究所建立的投票行為機率模型係以美國 2004 年總統選舉選前的民調結果為基礎，再採用貝氏觀點加入先驗機率進行修正而成。在檢證的程序上，首先，本研究選擇 ARG、Gallup 和 Rasmussen 這三家民調機構在 2004 年 5 月及 9 月所進行的二次總統選舉民調作為模型檢驗的對象，三家民調機構調查所得的民主黨候選人³ 支持度如表 3 調查結果一欄所示。其次，由於事前機率的 (α, β) 必須以歷次選舉的經驗值代入，而前文已指出，如果納入考慮的選舉所橫跨的時間太長，與 2004 年的現實環境有很大的差異，會導致政黨得票率的偏差估計，因此，本研究選擇近 10 年內的美國總統選舉作為估計事前機率分配的基礎資料，亦即 2000 年、1996 年的民主黨得票率。由表 3 可知，民主黨候選人得票率的經驗值最高達 49.23%，最低為 48.38%，平均得票率為 48.81%，利用此一資訊並且假設 95% 信賴區間介於 49.23% 到 48.38%，即可代入貝塔的理論分布取得估計相關參

³ 本研究以民主黨候選人的得票率為例作為說明，讀者當然也可以共和黨候選人的得票率作為預測變數，其原理是一致的。

數 $(\hat{\alpha}, \hat{\beta}) = (2000, 2048)$ ，貝塔機率分配當作取值在 0 至 1 之間的母體選民投票取向之機率表示，也就是說，可以用貝塔機率分配來描述選民過去的投票實際作為，其分配如圖 3 所示，此一機率分配即反映出民主黨在過去總統選舉可能得票率的機率分布。當 α 和 β 決定後，吾人即可利用公式 (8) 進行民主黨候選人在 2004 年總統選舉的得票率估計，其預測結果如表 3 的模型推估一欄，舉例來說，承公式 (8)，ARG 民調機構在 2004 年 5 月的調查數據利用模型推估值計算如下：

$$\frac{770}{2000+2048+770} (.45) + \frac{2000+2048}{2000+2048+770} \left(\frac{2000}{2000+2048} \right) = 48.20\% \quad (10)$$

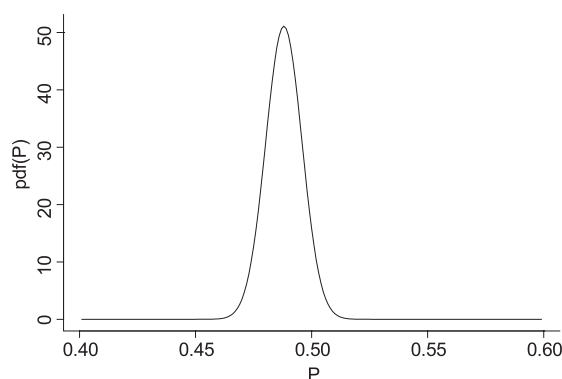
相較於 2004 年實際的得票率 (48.27%)，利用本研究所建立之模型修正三家民調機構所得結果預測民主黨得票率介於 47.20% 和 48.57% 之間，其中修正後 Gallup 民調機構和 ARG 民調機構的預測得票率都接近實際得票率，惟 Rasmussen 民調機構所得結果較差，究其原因係 (8) 式中樣本數越大，模型推估值越會受民調樣本大小影響所致，表示理論上其他條件不變下，民調樣本數越大，民意調查結果越會被研究者所採信，倘若實際上民調結果本身既已無法反映當前景況，仍可利用本模型進行修正，縮小誤差範圍的結果推論。要言之，除優質的民意調查本身是選舉預測準確性的必要條件外，本模型在估計程序也發揮了一定的修正效果。

表 3 2004 年美國總統選舉民主黨候選人得票率預測

民調資料						
來源	ARG		Gallup		Rasmussen	
	調查結果	模型推估	調查結果	模型推估	調查結果	模型推估
2004/5	45%(770)	48.20%	45%(579)	48.33%	44%(1500)	47.52%
2004/9	47%(600)	48.57%	44%(758)	48.05%	45%(3000)	47.20%
基礎資料						
1996 年選舉民主黨得票率			49.23%			
2000 年選舉民主黨得票率			48.38%			
2004 年選舉民主黨得票率			48.27%			

資料來源：作者自行整理。

說明：括號內代表該次調查總樣本數。



資料來源：作者自行整理。

圖 3 美國總統選舉民主黨得票率以貝塔機率分配示意圖

二、美國 2004 年總統選舉結果是「變天型」或「維持型」的判斷

由於本研究利用貝氏觀點所建立的預測模型，同時考量現在和傳統的因素，兩者的權重分別是 W 和 $1-W$ ，如公式 (9) 所示。因此，我們不但可以推估政黨的整體得票率，還可以判斷該次選舉在全國或特定地區是屬於「變天型」或「維持型」的選舉。換言之，如果將選舉區域依某一原則（如行政區、政治版圖）劃分為 i 個區隔（若以全國為單位， $i = 1$ ），可以設定區隔內的變盤判斷指標為 $W/1-W$ ，且由公式可知， (α, β) 估計值越大時， W 越小而 $1-W$ 越大，代表估計期望值受到當前因素的影響越小，受到傳統因素的影響較大。因此，當 $W/1-W$ 等於 1 時，即表示傳統和現在的因素對估計期望值 P_i 的影響力是勢均力敵；若大於 1，即表示估計期望值 P_i 受到當前因素的影響較大，也就是表示區隔 i 較易「變天」；當 $W/1-W$ 小於 1 時，估計期望值 P_i 受到傳統因素的影響較大，預測結果較接近原始基礎數據。為了便於說明此一判斷的方式，本研究以 ARG 民調機構的四次調查結果為例，計算 W 和 $1-W$ 的權值及其比值，其結果如表 4 所示，模型改良四次民調結果的民主黨得票機率分配如圖 4 所示，其高峰逐漸向基礎資料的結果靠攏。由表 4 的 $W/1-W$ 均小於 1 可以判斷，民主黨的得票率受到選民傳統投票行為的影響較大，即 2004 年總統大選民主黨的得票率傾向是「維持型」，與 2000 年差異不大，2004 年 10 月調查的 W 值增為 0.268 也顯示開票前民主黨得票率的調查結果越接近最後模型推估值。

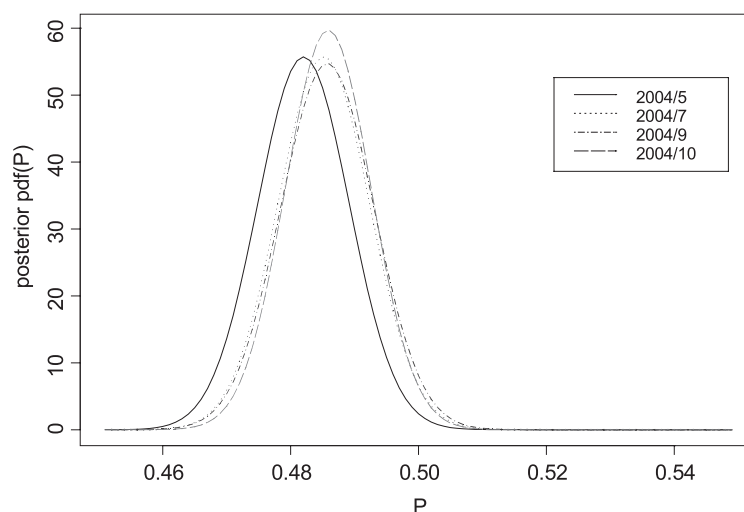
表 4 2004 年美國總統選舉民主黨得票率預測的權重變化 (ARG)

調查時間	調查結果	模型推估	W	1-W	W/(1-W)
2004/5	45%(770)	48.20%	0.158	0.842	0.187648
2004/7	47%(773)	48.52%	0.159	0.841	0.189061
2004/9	47%(600)	48.57%	0.128	0.872	0.146789
2004/10	48%(1500)	48.59%	0.268	0.732	0.366120

資料來源：作者自行整理。

說明：1. W 為來自民調的影響權重，(1-W) 為來自基礎資料的影響權重。

2. 括號內代表該次調查總樣本數。



資料來源：作者自行整理。

圖 4 2004 年美國總統選舉民主黨得票率按月推估變化圖

三、美國 2004 年總統選戰激烈地區的關鍵選票推估—未表態資料的處理

美國總統選舉為間接選舉制，且各州是以「贏者全拿」的規則來決定全州的選舉人團票由哪一位候選人獲得，因此，誰能在選民投票取得多數的支持，等於囊括了該州的全部選舉人團票，因而預測美國各州的選舉結果，對於掌握選舉人團的選票歸屬十分重要，這也是美國總統選舉民調除了全國性的調查與預測外，經常會以各州為調查或預測對象的原因，尤其是競爭激烈的州。對政黨及候選人而言，掌握各地區的競爭激烈程度往往是決定資源投入的判準之一，競爭越激烈的地區，為求勝選必須投入較多資源。因此，如何預測這些選戰膠著地區的得票率，是一大挑戰。

首先，本研究將「競爭激烈」定義為領先的前二位候選人的得票率差距在 1% 以內，以此一定義為標準，在 2004 年的美國總統選舉中，有 5 個州符合此一定義，即 Florida、

New Mexico、Wisconsin、Iowa 及 Oregon 等州（如表 5 所示），其中，又以 Florida 州最爲重要，因該州的選舉人團票多達 25 張，幾乎爲其他 4 州的總和，可說是超級戰區。因之，本研究以 Florida 州爲例作爲說明。在 2000 年的美國總統選舉中，民主黨在 Florida 州拿下 48.84% 的選票（即使在 1996 年的選舉也有 48.02% 的實力），在該次選舉中，充滿爭議 Florida 州 25 票選舉人團得票最後由最高法院判定歸屬共和黨的小布希，致使共和黨因而保住勝果。此一激烈競爭的選情在 2004 年的選舉中是否重演，受到各方高度的關注。

表 5 2000 年美國總統選舉競爭激烈的地區

州 名	選舉人團 票數	民主黨 得票率	共和黨 得票率	兩黨得票率 差距
Florida	25	48.84%	48.85%	-0.01%
New Mexico	5	47.91%	47.85%	0.06%
Wisconsin	11	47.83%	47.61%	0.22%
Iowa	7	48.54%	48.23%	0.31%
Oregon	7	46.96%	46.52%	0.44%

資料來源：作者自行整理。

其次，在競爭激烈的選區中，任何一位選民的投票意向都有可能是左右選舉結果的關鍵票，因此，在民調中除了已表明投票意向的選民外，對未表態的選民處理便相當重要。換言之，在戰況膠著的地區必須進一步思考，如何更細緻地處理未表態的受訪者，即使美國選舉民調中未表態的比例可能不到 4%，但此一比例在競爭激烈地區實已足以左右選舉結果了。以本研究的模型爲例，當未表態者存在而未加以處理時，造成 (9) 式 $W(\frac{X_i}{M_i}) + (1 - W)(\frac{\alpha}{\alpha + \beta})$ 中 $\frac{X_i}{M_i}$ 的 X_i 仍被低估，較易造成最後預測得票率偏低。爲了修正此一偏誤，本研究使用二階段預測的做法，第一階段以求得正確的 X_i （該區隔內特定候選人所得到的支持人數）爲目標，在加入未表態總數（假設爲 M_i^* ）後，由於 M_i^* 爲 M_i 不可忽略的一部分，當所有受訪者均加以處理後所求得的 X_i ，應可讓民意調查結果更接近實際的當下選民投票意向。

以 Florida 州計算爲例，假設未表態總數 (M_i^*) 中會投給民主黨的票數 (X_i^*) 的機率是來自傳統和當前態勢會投給民主黨候選人的兩股力量在角力，亦即如下式 (11) 所示，未表態總數 (M_i^*) 中會投給民主黨候選人的票數 (X_i^*) 的機率 = $\pi \text{Pr}(\text{傳統}) + (1 - \pi) \text{Pr}(\text{當前})$ ，即：

$$P(X_i^* = x) = \pi P(X_i^* = x | \alpha, \beta, M_i^*) + (1 - \pi) P(X_i^* = x | P_i^*, M_i^*) \quad (11)$$

其中 π 為代表傳統勢力的權重； $\Pr(\text{傳統}) = P(X_i^* = x | \alpha, \beta, M_i^*)$ ，乃利用 (6) 式展開計算， α, β 等屬於原模式用過去得票基礎資料推得的參數； $\Pr(\text{當前}) = P(X_i^* = x | P_i^*, M_i^*)$ 則利用 (2) 式展開計算，其中 P_i^* 為區隔 i 中已表態會投給民主黨的比例。

以上的 P_i^* 和 α, β 皆為已知參數，此後使用最大概似法推估使得未表態總數 (M_i^*) 中會投給民主黨候選人的票數 (X_i^*) 的機率最大，也就是以最大概似法推估可以得到 π 的估計值。由於 X_i^* 的可能值為 0, 1, ..., M_i^* ，如果還有其他可靠的參考資訊，吾人還可以再縮小範圍，否則要加總 0, 1, ..., M_i^* 所有觀測到的樣本可能性來估算 π 。舉例來說，承公式 (11)，ARG 民調機構 2004 年 7 月以本模型校正未表態資料方式，未表態總數 (M_i^*) 中會投給民主黨候選人的票數 (X_i^*) 的機率組合 log-likelihood function 如下：

$$\sum_{i=1}^{24} \text{Log} [\pi P(X_i^* = x | \alpha, \beta, M_i^*) + (1 - \pi) P(X_i^* = x | P_i^*, M_i^*)] \quad (12)$$

以此式求算滿足此式最大值時 π 的估計，接著代入 (11) 式就可計算未表態總數 (24) 中會投給民主黨候選人的票數 X_i^* 的可能值為 0, 1, ..., 24，其中以 12 的機率值最大，因為這是未知數 X_i^* 中最有可能出現的值，本方法在此過程中掌握了最大概似估計法的精神。是以，先處理好未表態的資料，也就是讓民意調查更接近實際當前的民意，其結果如表 6 所示，再據以進入研究模型，融合過去、推估未來。

表 6 利用本模型校正 2004 年美國總統選舉民主黨在 Florida 州民調中未表態比例

調查時間	機構	ARG			Gallup		
		民調 已表態	民調 未表態	未表態 分配數	民調 已表態	民調 未表態	未表態 分配數
2004/7		49%(600)	4%	2.00%	49%(1138)	5%	2.41%
2004/8		46%(600)	7%	3.33%	46%(674)	3%	1.50%
2004/9		50%(600)	5%	2.50%	46%(671)	4%	2.00%
2004/10		47%(600)	6%	2.83%	46%(699)	3%	1.50%

資料來源：作者自行整理。

說明：括號內代表該次調查總樣本數。

接著，進入第二階段利用原模型的 (8) 式進行全部選民的推估。如同前文所述，(8) 式乃是同時考慮到政黨候選人過去的得票實力和當前民調資料反映的最新競爭態勢，最後以兩者的加權平均組合作為最後的預測值。本研究同樣以 ARG 和 Gallup 這兩家民調機構的民調資料作為預測的依據，模型推估如表 7。表 7 顯示 ARG 這家民調機構的民調結果資料偏高，Gallup 則接近開票的調查結果較接近實際選舉結果。本模型在此發揮了修正的作用，雖然最後普遍都仍有 1% 左右的誤差，但這兩家民調機構經模型校正後，大都顯示民主黨無法超越前次 2000 年選舉結果的 48.84% 得票率，且依 (7) 式計算大於 48.84% 的機率均小於 1/2，足以表示雙雄對決時民主黨很難超越過去，即 2004 年民主黨不易在 Florida 州肉搏戰中脫穎而出。

表 7 2004 年美國總統選舉民主黨候選人在 Florida 州選舉的得票預測

調查時間	機構	ARG		Gallup	
		調查結果	模型推估	調查結果	模型推估
2004/7		51.00%	48.75%	51.41%	49.07%
2004/8		49.33%	48.54%	47.50%	48.29%
2004/9		52.50%	48.94%	48.00%	48.37%
2004/10		49.83%	48.60%	47.50%	48.29%
2004 年選舉得票結果		47.09%			

資料來源：作者自行整理。

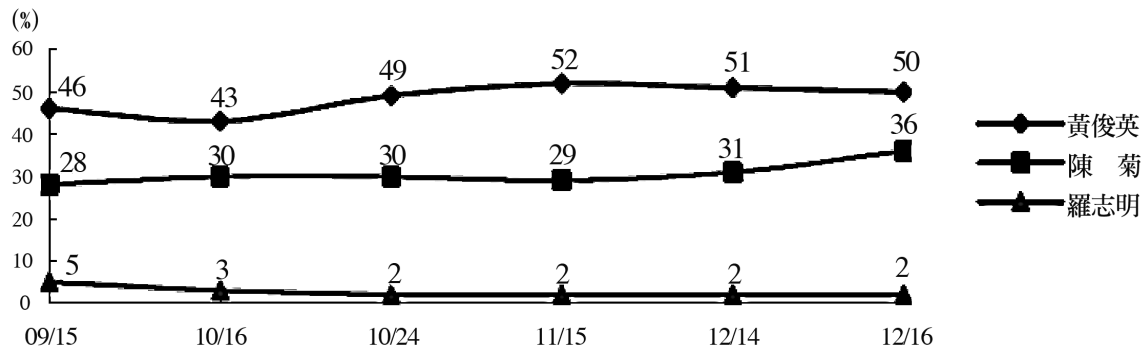
說明：調查結果包括表 6 中已表態和未表態預測的加總。

四、2006 年台灣高雄市長兩大政黨黨整體得票率的預估

雖然，台灣的總統選舉或縣市長選舉為直接選舉制，與美國有所不同，因而只需要估計整體的得票率，不需要特別估計特定選區的得票率，但對於未表態受訪者的處理仍有其必要性，尤其台灣的選舉民調中未表態的比例，少者一至二成，多者高達三至四成，即使是非激烈競爭的選舉，未表態受訪者的真實投票意向已經足以影響選舉結果，更遑論及在激烈競爭的選舉中，未表態受訪者的投票意向若不加以處理，實是選舉預測非常危險的事情，是以，吾人若能更精確地研判未表態選民的投票意向，可以大大提升選舉預測的準確度。在這個章節中，筆者將以 2006 年高雄市長選舉⁴ 作為檢證本模型的效度，藉以瞭解此一以機率分配為基礎的預測模型是否適用於台灣的地方性選舉，下一個章節再以 2008 年總統選舉為例，進一步檢視此一模型應用在全國性選舉的效度。

⁴ 本研究未以 2006 年台北市長選舉為例的考量是，國民黨與民進黨的候選人在民調中的支持度差距太大，非屬於激烈競爭的選舉，無法有效檢證本研究建立模型的預測能力。

一如所知，民調最令人詬病之處在於：民調結果的贏家，最後反而是實際選舉的輸家。舉例言之，圖 5 就是 2006 年高雄市長選舉中媒體民調失準的個案之一，在選前 3 個月至投票日前的民調中，國民黨候選人黃俊英一直是民調中的領先者，與民進黨候選人陳菊的差距最多達 20%，即使是投票前 3 天所完成的民調，黃俊英仍獲得半數的民調支持度，領先陳菊的 36%，兩人差距仍有 14%，即使把未表態的比例全部歸給陳菊，也是落後黃俊英，不過，實際的選舉結果卻是陳菊獲勝，顯然只根據已表態民調結果進行預測，「猜錯」的可能性是很高的，必須對民調結果做修正，包括對未表態受訪者的處理。因此，本研究以此一民調機構最後 3 次的調查結果為依據（如表 8 所列），利用本研究所建立之模型對國民黨與民進黨的得票率進行預測。



資料來源：作者自行整理。

圖 5 2006 年高雄市長候選人民調支持度趨勢圖

首先，如同前面的做法一樣，本研究採取二階段進行選舉預測，在第一階段中先處理未表態的部分。在國民黨方面，國民黨在 1998 年高雄市長選舉的得票率為 48.13%，2002 市長選舉結果為 46.82%，假設國民黨候選人得票率的經驗值最高達 48.13%，最低為 46.82%，平均得票率為 47.48%，利用此一資訊並且假設 95% 信賴區間介於 48.13% 到 46.82%，代入貝塔理論分布取得估計相關參數 $(\hat{\alpha}, \hat{\beta}) = (1800, 1989)$ ，據以代表國民黨的傳統勢力。在民進黨方面，民進黨在 1998 年高雄市長選舉的得票率為 48.71%，2002 年市長選舉結果為 50.04%，假設民進黨候選人得票率的經驗值最高達 50.04%，最低為 48.71%，平均得票率為 49.38%，利用此一資訊並且假設 95% 信賴區間介於 50.04% 到 48.71%，代入貝塔理論分布取得估計相關參數 $(\hat{\alpha}, \hat{\beta}) = (1800, 1845)$ ，據以代表民進黨的傳統勢力。其次，由於代表 $\Pr(\text{傳統}) = P(X_i^* = x | \alpha, \beta, M_i^*)$ 中的 $(\hat{\alpha}, \hat{\beta})$ 數值，及 $\Pr(\text{當前}) = P(X_i^* = x | P_i^*, M_i^*)$ 的 P_i^* 的值為已知，承公式 (11)，(12) 的方式以最大概似法推估決

定未表態受訪者中會投給兩黨候選人的人數，據以分配未表態的部分，其結果如表 8 所示。

表 8 利用本模型校正 2006 年台灣高雄市長兩大政黨候選人在民調中未表態部分

調查時間 \ 政黨	國民黨			民進黨		
	民調 已表態	民調 未表態	未表態 分配數	民調 已表態	民調 未表態	未表態 分配數
2006/11/15	52%(830)	16%	7.59%	29%(830)	16%	7.95%
2006/12/4	51%(1086)	16%	7.63%	31%(1086)	16%	7.92%
2006/12/6	50%(957)	11%	5.75%	36%(957)	11%	5.12%

資料來源：作者自行整理。

說明：括號內代表該次調查總樣本數。

第二階段針對整體得票率進行選舉預測。利用本模型的 (8) 式進行全部會去投票的選民的投票行為推估，預測結果如表 9 所示。由表 9 可以得到幾點研究發現：(1) 國民黨的得票率預測為 49.17%，非常接近實際得票率 (49.27%)，由此可知，本模型對國民黨得票率預測的校正效果相當好，也間接證實此一機構的民調結果傾向高估國民黨的得票；(2) 民進黨的得票率預測是 47.66%，與實際的得票率 (49.47%) 仍有將近 2% 的差距，此一差距最有可能的原因是最後來自同屬於泛綠陣營的台聯候選人羅志明的支持者，在國民黨候選人可能當選的危機意識下，在選舉投票階段大部分都轉投給民進黨陳菊。換言之，泛綠選民在選舉最後一刻若是出現「棄羅保陳」效應，將是造成民進黨的得票率陡增 2% 左右，因羅志明在選前 3 波民調的支持度均在 2% 以上，最後卻不及 1%。不過，此僅止於推論「棄保效應」發生的可能性，並無直接的證據支持，就學術的嚴謹性而言，如果在選舉前後進行固定樣本追蹤調查 (panel study)，才可檢證此一棄保效應是否存在。此外，表 9 的推估結果也間接證明，本研究模型中民進黨的支持者中自始至終不含有民調中支持羅志明的選票，換言之，就台灣選舉特性，未來也可考慮模型內融入棄保效應，只是這些都是必須視個案觀察，由於本模型民意調查結果以二項分配表示，本就較合適描述雙雄對決的選舉情境。(3) 由於本研究僅針對整體得票率進行預測，並未將高雄市分為數個選區單獨進行預測，未來若結合政治版圖將高雄市依政黨傳統實力進行區隔，再利用本模型個別做預測後加總，其效果應會更好，因為配合集群分政治版圖會縮小部分估計誤差，此為日後值得深入研究的方向。

表 9 2006 年台灣高雄市長選舉兩大政黨候選人的得票預測

調查時間 \ 政黨	國民黨		民進黨	
	調查結果	模型推估	調查結果	模型推估
2006/11/15	59.59%	49.68%	36.95%	47.08%
2006/12/4	58.63%	49.98%	38.92%	46.98%
2006/12/6	55.75%	49.17%	41.12%	47.66%
2006 年選舉得票率	49.27%		49.47%	

資料來源：作者自行整理。

說明：調查結果包括表 8 中已表態和未表態預測的加總。

五、2008 年台灣總統選舉民進黨整體得票率的預估

由於高雄市長選舉屬於地方性選舉，因此，在這個章節中，吾人將以 2008 年台灣總統選舉作為檢證本模型的效度，藉以瞭解此一預測模型是否適用於全國性選舉的預測。值得說明的是，由於這次總統大選中僅有民進黨和國民黨推出候選人參選，是典型的雙雄對決選舉，因此本研究僅對民進黨的長昌配（謝長廷與蘇貞昌）的得票率估計即可反推國民黨的馬蕭配（馬英九與蕭萬長）的得票率估計。

首先，本研究仍以同一家媒體民調機構的調查數據為例，如同前面的做法一樣，採取二階段進行選舉預測。在第一階段中先處理未表態的部分，在民進黨長昌配方面，利用表 2 民進黨近年選舉的總得票率的資訊並且假設 95% 信賴區間介於 50.2% 到 41.88%，代入貝塔理論分布取得估計相關參數 $(\hat{\alpha}, \hat{\beta}) = (320, 391)$ ，據以代表民進黨的傳統勢力。其次，由於代表 $\Pr(\text{傳統}) = P(X_i^* = x | \alpha, \beta, M_i^*)$ 中的 $(\hat{\alpha}, \hat{\beta})$ 數值，及 $\Pr(\text{當前}) = P(X_i^* = x | P_i^*, M_i^*)$ 的 P_i^* 的值為已知，承公式 (11)，(12) 的方式以最大概似法推估選前 3 天和選前 1 天 π 為 0.6 和 0.605，並據以決定未表態的受訪者中會投給長昌配的人數來分配未表態的部分，選前 3 天和選前 1 天分別為 91/205 和 109/242，其結果如表 10 所示。

表 10 利用本模型校正 2008 年台灣總統選舉民進黨長昌配在民調中未表態部分

調查時間 \ 調查內容	民調	民調	未表態	
	已表態 (%)	未表態 (%)	分配數 (%)	分配後 (%)
2008/3/19	30(1138)	18	7.99	37.99
2008/3/21	29(1100)	22	9.87	38.91

資料來源：作者自行整理。

說明：括號內代表該次調查總樣本數。

第二階段針對整體得票率進行選舉預測。利用本模型的 (8) 式進行全部會去投票的選民的投票行為推估。預測結果如表 11 所示。由表 11 可以得到兩點研究發現：(1) 民進黨的選前最後得票率預測為 41.30%，非常接近實際得票率 (41.55%)，由此可知，本模型對本次台灣全國性的總統選舉預測的校正效果相當好。(2) 表 11 推估中結果，由於 2 次時間點所計算之 W 皆為 0.61，代表估計期望值受到當前因素的影響大於受到傳統因素的影響。因此，本次總統大選是個「變天型」的選舉，「變天型」代表民進黨候選人在本次總統選舉的得票率會偏離過去選舉的平均得票率 (44.99%)，遑論雙雄對決的選舉中要當選需要獲得過半數的支持。

表 11 2008 年台灣總統選舉民進黨長昌配的得票預測

調查時間	調查內容	民調 結果 (%)	模型 推估 (%)
選前 3 天		37.99	40.69
選前 1 天		38.91	41.30

資料來源：作者自行整理。

說明：調查結果包括表 10 中已表態和未表態預測的加總。

伍、結論

檢視各種選舉研究的結果可以發現，選民的投票行為的確具有異質性，亦即民調中影響選民投票行為的顯性因素等，並無法完全掌握選民的投票意向，這些屬於不可直接觀察的異質性與選民過去的投票行為模式有所關聯。因之，在民意調查中，若僅是依賴可觀測的選民屬性進行選舉預測，將難以準確地預測選舉結果，此一現象在台灣的選舉中尤其是如此。

台灣當前多數非學術性民調機構進行選舉預測的常見作法主要有二：(1) 忽略未表態受訪者的投票意向，只以已表態受訪者的比例進行預測；(2) 根據歷次民調結果的平均值作為預測當選對象的依據。無論是何者，預測得票率和實際得票率仍有差距。持平而論，這兩種作法均不是真正的選舉預測，充其量不過是民調的變型罷了。至於學術界最常見的選舉預測亦有二種作法：(1) 利用已表態受訪者的特徵來預測未表態受訪者的可能投票意向；(2) 利用政治版圖來預測未表態受訪者的可能投票意向。前者除了假定未表態和已表態的受訪者沒有結構性的差異外，所反映的是「現在」的政黨實力，對於「變天型」的選舉可能有較佳預測效果；後者所反映的則是「傳統」的政黨實力，對於「維持型」的選舉

可能有較好的預測力。

由於機率分配的機制可以有效的描繪「不可看見卻永遠存在的選民投票行為異質性」，尤其貝塔分配是少數常見機率分配中適合用以描述得票率的分配。因之，本研究即以此一分配為基礎，連結了傳統政黨得票勢力的訊息，針對民調的結果進行選舉預測。換言之，本研究所建立之模型則是企圖同時將這兩個因素納入考量，並進一步依據其權值—— W 和 $(1-W)$ ——來判斷屬於哪一種類型的選舉。因此，本研究除了提供建構完整的選舉預測模型論述外，並推廣該機率分配模型應用於處理選舉預測中「未表態者的投票意向」的難題，在本研究即是將未表態視為傳統和當前的角力，並進行估計。

使用常用迴歸分析的經驗告訴我們，模型中永遠有誤差項的存在，如能掌握誤差項的分配模式，能決定模型的預測效果，遠勝於挑選了多少自變項，就如本文的選舉預測模型，用機率分配理論上若能有效掌握未表態選民的投票意向，即能提高選舉預測的準確度。事實上，具有貝氏統計觀點的機率分配模型（例如貝塔一二項分配）在不少學術論文中已見廣泛應用 (Morwitz et al. 1998)，但是在台灣，將此一機率分配模型應用在選舉研究中仍為少見，相信未來尚有很大的發展空間。

* * *

投稿日期：97.01.25；修改日期：97.04.07；接受日期：97.04.29

參考文獻

I. 中文部分

- 洪永泰，1994，〈選舉預測：一個以整體資料為輔助工具的模型〉，《選舉研究》，1(1): 93-110。
- 俞振華、蔡佳泓，2006，〈如何運用全國性民調推估地方民意？多層次定理估計模型與分層加權的應用〉，《台灣政治學刊》，10(1): 5-38。
- 張紘炬、林顯毓，1995，〈台北市長選舉投票傾向的 Logit 模式分析〉，《民意研究季刊》，192: 1-11。
- 張紘炬、黃男瑋，2000，〈臺北市市長選舉預測模型之比較〉，《民意研究季刊》，211: 1-14。
- 盛杏媛，1998，〈選民的投票決定與選舉預測〉，《選舉研究》，5(1): 37-75。
- 陳義彥、鄭宇庭、蔡孟熹，1998，〈選舉預測之研究：民國八十六年台南縣、市長選舉之分析〉，民調、策略、廣告與選舉研究論文研討會，3月14-15日，台北：世新大學民意調查研究中心、世新大學公共傳播系。
- 蔡佳泓，2007，〈全國共識或南北差異？反貪倒扁運動的地域支持度分析〉，公民意識與政治行動學術研討會，5月3-4日，台北：中央研究院人文社會科學研究中心政治思想研究專題中心。

II. 外文部分

- Abramowitz, M., and I. A. Stegun. 1972. *Handbook of Mathematical Functions*. New York: Dover Publications.
- Berger, J. 1985. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag.
- Brown, Lloyd B., and Henry W. Chappell, Jr. 1999. "Forecasting Presidential Elections using History and Polls." *International Journal of Forecasting* 15: 127-135.
- Efron, Bradley. 1979. "Bootstrap Methods: Another Look at the Jackknife." *The Annals of Statistics* 7(1): 1-26.
- Holbrook, Thomas M., and Jay A. DeSart. 1999. "Using State Polls to Forecast Presidential Election Outcomes in the American States." *International Journal of Forecasting* 15: 137-142.

- Kalwani, Manohar U. 1980. "Maximum Likelihood Estimation of Zero-Order Models Given Variable Numbers of Purchases Per Household." *Journal of Marketing Research* 17(4): 547-55.
- Soumbatiants, Souren, Henry W. Chappell, Jr., and Eric Johnson. 2006. "Using State Polls to Forecast U.S. Presidential Election Outcomes." *Public Choice* 127: 207-223.
- Morwitz, Vicki G., and David C. Schmittlein. 1998. "Testing New Direct Marketing Offerings: The Interplay of Management Judgment and Statistical Models." *Management Science* 44(May): 610-628.
- Park, David K., Andrew Gelman, and Joseph Bafumi. 2004. "Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls." *Political Analysis* 12: 375-385.
- Samuelson, P. 1938. "A Note on the Pure Theory of Consumer's Behaviour." *Economica* 5(17): 61-71.
- Yu, Chen-hua and Chia-hung Tsai. 2006. "Reaching into the Grassroots: A Comparative Study of Local Public Opinions in Taiwan." Presented at the Conference Group of Taiwan Studies at the Annual Meeting of the American Political Science Association, Philadelphia.

Change in Voting Behaviour: Applying an Election Forecasting Model of Probability Distributions to Modify the Accuracy of Poll Outcomes

Shun-Chuan Chang^{*} · Wen-Jong Juang^{**}

Abstract

Due to undervotes, misvotes, or switchvotes bias, many polling data users felt frustrated in using the past polling outcome to forecast the new election. It is commonplace for voters to note an early front-runner in polls will be doomed to fall in the real election outcome. A beta-binominal distribution is suggested to model the accuracy of early poll outcome which strategically influences the polling data users such as political parties, candidates, and mass media in implementing the election campaign. We demonstrate the advantages of probabilistic distribution and Bayesian reasoning, and how to estimate the parameters from past data, in modifying the accuracy of prior poll outcomes. In comparison with the traditional frequency approach, beta-binominal mixture distribution imposes a statistical-adjusting framework with ability to proportionate a coherent mechanism that synthesizes the performances of prior votes. The empirical data sets include the 2004 US presidential election in Atlas Web and TVBS polls in 2006 Kaohsiung mayor election and 2008 presidential election in Taiwan. This paper describes the general fitting of beta-binomial distribution on both datasets and discusses fruitful avenues for future research.

Keywords: Public Polling, Voting Behavior, Election Forecasting, Mixture Distribution, Bayesian Statistics

^{*} Ph. D Candidate of Graduate Institute of Business and Management, National Chiao Tung University.

^{**} Assistant Professor of Department of Public Policy and Management, Shih Hsin University.