

## 第五章 結論

本研究認為，在調查訪問中，選民是否表態並不是隨機產生的，一旦不表態的選民是因為某些因素而導致其不表態，只以表態者所建立的模型便存在選樣偏誤的問題。因此本研究便是希望經由處理「選樣偏誤」的問題之後，可以得到一個足以代表母體的正確投票模型，筆者將之稱為「選樣偏誤模型」，再利用「選樣偏誤模型」中校正後的參數估計值去推估未表態選民的投票意向，以得到更準確的選舉預測結果。

### 第一節 研究發現

在本研究中，筆者共挑選 2001 年台北縣長選舉、2002 年台北市長選舉、2002 年高雄市長選舉等三次地方首長選舉，以及 2000 年、2004 年兩次總統選舉，以檢驗選樣偏誤模型是否可以適用在不同層級的首長選舉，以及不同勝負差距的選舉上。模型預測結果與實際得票率的誤差詳見表 5 - 1。

表 5 - 1 傳統 probit 模型及選樣偏誤模型誤差一覽表

	LR test ( $\rho = 0$ )	傳統 probit 模型誤差	選樣偏誤模型誤差
2001 年台北縣長選舉	3.87*	4.66	0.71
2002 年台北市長選舉	1.18	8.46	8.46
2002 年高雄市長選舉	3.82 <sup>s</sup>	1.94	1.16
2000 年總統選舉	5.87*	7.68	0.59
2004 年總統選舉	3.73 <sup>s</sup>	6.19	0.65

註：\*表  $p < .05$ ，<sup>s</sup>表  $p < .1$ 。

根據表 5 - 1 的數據可以發現，對於  $\rho = 0$  的概似比檢定（likelihood ratio test, LR）結果顯示，除了 2002 年台北市長選舉之外，若只單純對結果方程式進行估計，的確會因為選樣偏誤而造成估計偏差的情況，我們利用選樣偏誤模型進行選舉預測，都得到比傳統 probit 模型更好的預測結果，2001 台北縣長選舉的誤差僅有 0.71%；2002 年高雄市長選舉的誤差也僅有 1.16%；2000 總統選舉的誤差更只有 0.59%；2004 年總統選舉，雖然當選人的預測錯誤，不過在得票率的預測誤差上，也僅有 0.65%。

在 2002 年台北市長選舉中，對於  $\rho = 0$  的概似比檢定（likelihood ratio test, LR）結果顯示，有無校正選樣偏誤，基本上沒有差異，也就是說選樣偏誤模型並沒有發揮校正的效果，不過這並不代表在該次選舉中就沒有選樣偏誤的問題，可能只是因為筆者所設定的選樣方程式並沒有掌握到真正影響選民表態的因素。

從五個模型中，變數參數估計值的變動來看，都顯示出當我們未校正選樣偏誤時，我們可能高估給予不同候選人較高評價者之間的差距、藍綠政黨認同者的之間的差距、統獨支持者之間的差距；同時也高估了政府首長施政滿意度與省籍的影響。因為願意回答自己投票對象的受訪者，往往是政治偏好相當確定且較強烈的選民，當我們在建構模型時，若只掌握了這些政治偏好較強或較確定的表態者，而忽略政治偏好較弱或較不明確的不表態者，我們極可能高估自變數對應變數的影響。這同時也說明了為什麼在勝負懸殊的選舉中，研究者往往過度高估勝者與敗者的差距，導致過度高估勝者的得票率。在台北市的模型中，就出現了上述的情況，因此必須要適當的校正選樣偏誤的情形，只可惜在本研究中，我們沒有掌握好影響選民在勝負懸殊的選舉中表態與不表態的因素。

整體來說，本研究所設定的選樣方程式能夠在勝負差距較小的選舉中，掌握到影響選民是否表態的因素，進而使得選樣偏誤模型發揮校正的功能，在勝負差距較小的選舉中，選樣偏誤模型所預測的候選人得票率，比傳統 probit

模型來得準確，四個模型的預測得票率與實際得票率的誤差，最小為 0.59%，最大也只有 1.16%而已。

筆者認為，一個好的選舉預測模型必須要能夠經得起多次選舉的考驗，以本研究所設定的選樣偏誤模型來說，在四次勝負差距較小的選舉中，都得到不錯的預測結果，由於本文在結果方程式的變數選擇上，三次地方首長選舉都選擇了同樣的變數，總統選舉也盡可能的選擇同樣的變數，在選樣偏誤問題存在的前提下，相同的模型在不同的選舉中，都可以發揮良好的預測力，也說明了「選樣偏誤模型」應該是一個經得起考驗的選舉預測模型。

## 第二節 檢討與建議

當然，在本文的研究過程中，還是有許多地方受到限制，以致於無法盡善盡美，接下來筆者就針對本研究受限及不足之處進行檢討，提供給未來相關研究做為參考。

### 一、結果方程式應變數的限制

本研究主要採取 Dubin 與 Rviers 所發展出來的方法，結果方程式的應變數僅能夠處理二分變數，在美國的兩黨政治之下，候選人通常只有兩位，因此這個方法在研究美國選舉上，並不會出現太多的問題，但是在台灣的選舉之中，即使是首長選舉，也常常不只有二組候選人，以本文所研究的範圍來說，僅有 2002 年台北市長與 2004 年總統選舉，是只有兩組參選人，其餘的參選人數都超過二組，也因此 2001 年台北縣長選舉及 2002 年的高雄市長選舉中，都必須將實力較弱的候選人剔除，而無法去估算所有參選者的得票率。在 2000 年的總統選舉中，必須妥協的地方更大，因為三組候選人的實力都不弱，無

法刪除任何一組，也使得筆者不得不將應變數改為投扁、不投扁二類，雖然得到的結果還算滿意，但是無法進一步計算三組候選人的得票率，也讓模型的價值失色不少。因此筆者認為，為了因應台灣選舉的特殊性，改善 Dubin 與 Rivers 的方法，讓它可以適用於多分類的應變數，是未來研究亟需努力的方向。

## 二、變數建構方式無法統一

在前面幾章筆者曾經提到過，本研究雖然盡可能以相同的變數來建構模型，但是受限於資料的限制，因此在變數的建構方式上，無法做到每次選舉都是一樣的，以候選人評價為例，幾乎每一次選舉的問卷題目都不相同，在 2001 年選舉中，是比較選民對二位候選人的個別評價，再加以區分；2002 年台北市長選舉中，則僅能以選民認為哪一位候選人可以解決他所認為的重要問題來建構。雖然筆者已盡可能依照變數的概念，來尋找最適合的問卷題目，不過還是可能有效度上的問題。當然，這個問題並不是筆者所能夠解決的，只能期待在未來的研究中，對於這方面的題目能夠有更一致的問法。

## 三、選樣方程式變數的補強

在本研究中，選樣方程式的建構是一個相當重要的部分，筆者所選擇的變數有：性別、年齡、教育程度、是否具有政黨認同、候選人評價問題回答情形、媒體使用的情形、政治功效意識；但由於本研究所蒐集到的資料，並不是專門為了處理選樣偏誤問題而設計的，因此有部分筆者認為可能會影響選民表態的變數，並沒有出現在問卷當中，例如政治知識、選民對電訪的信任度、政治信任感、對選舉的興趣等變數，如果這方面的變數能夠得到補強，或許能夠更增強「選樣偏誤模型」的預測力。