

第三章 研究設計

本章將針對本文依據研究架構所提出的研究假設，在方法上如何進行檢證進行說明。其中包括所要採取之統計分析的研究方法、總體及個體層次資料的來源及蒐集方式，以及各項變數的測量與建構。

第一節 研究方法

投票抉擇屬於個人行為，多數傳統投票行為研究由於擔心可能的「區位謬誤」，多採取個體層次的資料進行統計分析。但就統計方法而言，單採個體資料來分析選民的投票行為，忽略總體層次的系絡效果，即是假設模型中的迴歸係數在每一個系絡中皆相同(Luke 2004, 7)。這在社會環境相當一致的地區，或許這種假設尚可成立。但過去許多政治行為研究卻常發生相同的理論模式可適用在某些地區或國家，卻在其他地區發生矛盾或經驗上的爭議，即可從社會環境的總體差異提出解釋(Anderson 2007, 601)。如果學理上存在系絡間的差異，研究者卻遺漏了總體層次上的重要變數，忽略了群內個體之間的相關性，將會低估統計檢定上的標準誤，進而膨脹「第一型錯誤」(Type I Error)的發生機率，高估自變數在假設檢定上的顯著效果(Hox 1998, 147; Steenbergen and Jones 2002, 219-220; 溫福星 2006, 1-6)。²⁵也就是說，部分在統計模型中對選民投票行為有顯著影響之個體層次變數，若將學理上對投票行為具有影響力的總體層次變數納入考量，同時放入模型中進行估計，則部分個體層次的變數將可能因此喪失原有的顯著影響力。Tilly 與 Goodin(2006, 24)也表示，我們必須瞭解哪一個面向的系絡因素發生作用，並在我們的研究中將其控制在常數之下。唯有當這些系絡因素在我們的研究中沒有實際發生作用，我們才可以安全的忽視它們。

晚近的研究為了將系絡因素同時納入考量，嘗試運用各種不同的解決方法。最簡單的系絡分析可以利用「交叉表」來進行，比較兩個（或更多）團體之交叉表的内容差異，甚至是畫出二維圖形來解釋兩條線之「截距」與「斜率」的差異

²⁵ Hox(1998, 147-148)利用「設計效果」(design effect)來解釋，當「群內相關」(intraclass correlation)愈大，其設計效果愈大，進而導致有效樣本數愈低，此時研究者若忽略「群內相關」，而使用一般的統計方法，將會低估「抽樣變異」(sampling variance)，使得顯著性的檢證發生錯誤（即係數容易顯著）。Hox 引用 Tate 與 Wongbundhit(1983)實驗結果的發現指出，研究者若忽略「群內相關」的迴歸係數估計並不會發生偏誤，但是卻會產生虛假的顯著性檢定結果。

(Iversen 1991)。這種方式雖然易於分析與解釋，但由於交叉表所能控制的變數相當有限，仍然無法將諸多變數同時納入模型。有些研究則是將總體層次的資訊「分解」(disaggregate)至個體層次，並在統計模型中以個體層次為分析單位進行各項變數的解釋與預測。但此種方法仍因個體屬於相同的系絡而具有「誤差項相關」(correlated errors)，違反了迴歸模型對於分析單位必須相互獨立的基本假設(Luke 2004, 6-7)。此外，以「虛擬變數」或「交互作用」途徑，將總體層次的資訊納入個體層次的模型進行分析也是常見的處理方式，但 Steenbergen 與 Jones(2002, 220-221)則認為，虛擬變數的方法僅能檢視總體層次確實存在差異，卻無法解釋為何存在差異的「因果異質性」問題。至於交互作用的方法雖然可以在理論上解釋因果異質性的問題，但在統計上卻錯誤的假設總體層次間不存在殘差項(error components)，以致迴歸模型違反「變異一致性」(homoskedasticity)的基本假設。即便將這兩種傳統途徑合而為一，也常會因模型「自由度」的不足，而無法正確估計系絡效果與虛擬變數的效果，故這兩種途徑在理論與統計方法間存在無法兩全的「取捨」問題。²⁶

1980 年代開始統計學者逐漸發展出多層模型，試圖解決上述統計方法上的難題。這種分析方法起源於教育學中對學生學習成效的探討。研究者欲探討學生的學習成效，除了需考量學生個人資質及學習態度外，班級老師的教學方式，學校的教育方向，甚至是地區政府的教育政策，皆可能影響學生的學習效果。在這樣的分析層次下，學生屬於個人層次的單元，班級、學校、地區則屬於總體層次的單元。其中，學生隸屬於某一班級中，班級隸屬於某一學校中，學校則隸屬於某一地區中。各分析層次之間形成標準「階層性」(hierarchical)的「套疊結構」(nested structure)。而多層模型正是基於此種資料形式所提出的分析方法，這種資料層級的階層性，基於研究者的分析目的、資料蒐集或是學理依據，可能僅有二階層，亦可能在三層以上，完全依據理論的探討來決定。這種研究方法後續也逐漸運用在家庭、鄰居、自願性團體、宗教團體、公司行號、社會網絡等，與人類行為有關的社會學或政治學等系絡分析上(Iversen 1991)。

多層模型的基本學理，是在迴歸方法的基礎上，將個體與總體層次的階層屬性納入考量。首先建立起個體層次的迴歸模型（如公式 1）。其次，從學理上考量系絡因素的影響，若系絡因素會「直接」影響依變數的結果，則將迴歸模型中

²⁶ Steenbergen 與 Jones(2002, 220-221) 將這兩種途徑稱之為「虛擬變數模型」(dummy variable model)及「交互作用模型」(interactive model)。

的「截距」(β_0)獨立出來，依據可能影響的系絡變數來建立迴歸方程式（如公式 2）。若系絡因素對依變數的影響是「間接」的，也就是個體迴歸模型中的「斜率」(β_1)會受到特定系絡環境變數的影響，則將該斜率獨立出來，依據相關系絡變數建立迴歸模型（如公式 3）。最後，再將公式 2、3 代回公式 1 中（如公式 4），如此即可將個體與總體層次的相關變數同時納入模型，在考量變數之間的變異與共變異下，釐清影響依變數的真正變數。在公式 4 中， γ_{00} 是新迴歸模型中的截距， γ_{10} 是個體層次相關變數(X)的影響係數， γ_{01} 是總體層次相關變數(Z)的影響係數，而 γ_{11} 則是個體與總體層次變數間的交互作用。

第一層(Level-1)： i=個體(individual)；j=總體(group)

$$\text{公式 1..... } Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}$$

第二層(Level-2)：

$$\text{公式 2..... } \beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + \mu_{0j}$$

$$\text{公式 3..... } \beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + \mu_{1j}$$

合併後(combined)：

$$\text{公式 4..... } Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}Z_j + \gamma_{11}Z_jX_{ij} + \mu_{0j} + X_{ij}\mu_{1j} + \varepsilon_{ij}$$

由於多層模型是以線性迴歸方法為基礎，並在考量資料階層屬性下所發展而成，故學界多以「階層線性模型」(Hierarchical Linear Model，簡稱 HLM)稱之。²⁷以此次立法委員選舉而言，選民與選區之間正存在套疊關係，每一位選民必定隸屬於其所屬的選區，構成最標準的雙層模型分析架構。選民個人為第一層(level 1)即個體層次(individual level)的分析單位，選區環境則是第二層(level 2)屬總體層次(aggregate level)的分析單位。依據學理所提出的理論架構，藉由多層模型的分析方式，不但可以改善屬於個體層次相關影響因素的估計，建構並檢定總體層次

²⁷ 惟該多層模型在不同學科中亦有不同的名稱，在社會學領域中多稱之為「多層線性模型」(multilevel linear models)；在生物統計學領域中稱之為「混合效果模型與隨機效果模型」(mixed-effects models and random-effects models)；在經濟學領域中稱之為「隨機係數迴歸模型」(random-coefficient regression models)；在統計學領域中則稱之為「共變異成分模型」(covariance components models) (Raudenbush and Bryk 2002, 5-6)。

差異的影響外，更可以檢視不同層次之間的共變影響(Raudenbush and Bryk 2002, 7)。Steenbergen 與 Jones(2002, 219-220)更明確從實質上及統計上的動機，說明運用多層模型分析方法的重要性。在實質動機上，多層模型讓分屬不同層次的資料結合在單一完整的模型中(single comprehensive model)成為可能，並允許研究者探索系絡間的因果異質性，更提供對某項學理能否在比較研究的領域上建立「一般化能力」(generalizability)的檢驗。在統計動機上，則如前所述，若資料具有階層屬性，但研究者在分析時卻忽略它，將可能導致錯誤的估算模型變數中的標準誤，進而高估個體層次因素的影響效果。

由於「階層線性模型」是以依變數為連續變數之迴歸方法為基礎，並無法直接針對依變數為類別資料進行多層分析。之後統計學者將階層線性迴歸模型，進一步與「推廣線性模型」(Generalized Linear Models, 簡稱 GLM)結合，發展出「階層推廣線性模型」(Hierarchical Generalized Linear Models, 簡稱 HGLM)，利用不同的「連結函數」(link function)，得以針對各種型態的類別依變數進行多層模型分析 (Raudenbush and Bryk 2002; 黃紀 2008b)。²⁸本研究旨在探討台灣 2008 年立委選舉選民的投票行為，由於感興趣的依變數是選民在第一票及第二票的投票對象，以及兩張選票的抉擇是屬一致或分裂投票，皆屬於「二分類」(binary)或「無序多分」(nominal)的類別資料型態，此時「階層推廣線性模型」自然成為本研究最適當的統計方法。

第二節 資料蒐集

在資料的蒐集上，有關選民的人口背景特徵、心理態度認知或投票抉擇等個體層次的資料，本文將採用「台灣選舉與民主化調查」針對 2008 年立法委員選舉的選後面訪資料 (簡稱 TEDS2008L)。在 59 個單一選區中，總計完成 2,621 份成功樣本，訪問期間為 2008 年 1 月中旬至 3 月上旬。至於 TEDS2008L 的訪問執行可分為兩大部分，一部份為「定群追蹤調查」(panel study)，即針對 2004 年立法委員選舉選後面訪調查 (TEDS2004L) 之 2,507 份成功樣本全數進行追蹤訪問，總計完成 1,381 份。另一部份則為「橫斷面調查」(cross-section)之重新獨

²⁸ 亦有部分學者將「階層線性迴歸」與「推廣線性模型」結合後，稱之為「推廣階層線性模型」(Generalized Hierarchical Linear Models, 簡稱 GHLM)、「推廣線性混合模型」(Generalized Linear Mixed Models, 簡稱 GLMM)或「具隨機效果的推廣線性模型」(Generalized Linear Models with Random Effect)(Luke 2004, 53; Raudenbush and Bryk 2002, 292)。

立抽樣，在全國 73 個單一選區依選民人口比例抽取 38 個選區，並依「抽取率與單位大小成比例」(probability proportional to size, 簡稱 PPS)原則抽取村里後，再以等距抽樣抽出合格受訪對象來進行訪問，共計完成 1,240 份有效樣本。本研究同時將兩部分的資料合併，主要是考量提高選區與選民的樣本數，以利多層模型的建構與分析。但由於「定群資料」(panel data)有「非隨機性」之樣本流失與訪問效應的問題，在分析結果的推論上可能受到侷限(陳陸輝 1999；劉從葦、陳光輝 2006)。若資料的選樣標準與「依變數」有關係時，將使因果關係的係數估計發生偏誤，可能高估或低估自變數的影響效果。若是與「自變數」有關係時，則不會影響因果關係的係數估計，但在代表性的推論上將受到限制(King, Keohane and Verba 1994, 129-138)。為此，本文針對 TEDS2008L 之「定群追蹤」與「橫斷面調查」兩筆資料，在人口變項、重要政治態度及投票抉擇上進行卡方獨立性檢定。結果顯示(如附錄二)，兩筆資料僅在年齡及教育程度上有顯著差異，定群資料的年齡層偏高、教育程度偏低，但在重要政治態度及投票抉擇上則無統計上的顯著差異。這表示利用兩筆資料合併進行多層模型分析，並不會影響模型變數間因果關係的係數推估，但在推論上則有偏向高年齡與低教育程度選民的可能性。

至於各項總體層次變數方面，包括區域經濟發展、政黨勢力版圖、選區競爭程度及選區候選人資訊。主要是透過行政院主計處公開釋出的各縣市或各鄉鎮市統計資料、中央選舉委員會公布之歷次選舉投開票紀錄、學術單位或新聞媒體發佈之相關資訊來蒐集。而此次立法委員選舉單一選區的劃分是依據縣市人口比例進行。各縣市基本保障一個席次，每逾 31 萬餘人分配一席，縣市席次分配若超過二席者，原則上依據鄉鎮市區的行政區域來劃分選區。但部分鄉鎮市區因人口比例分佈不均，必須再以村里為單位進行跨鄉鎮市區的選區劃分。故有完整以單一縣市為一個選區(例如：基隆市、花蓮縣等)；有縣市內再完整以鄉鎮市區為單位劃分選區(例如：台北市第四選區包括「內湖區」、「南港區」，桃園縣第五選區包括「平鎮市」、「龍潭鄉」)；亦有部分選區需進一步以村里為單位劃分，導致同一個鄉鎮市區可能分屬在兩個不同的選區中(例如：台北市第一選區包括「北投區」及「士林區 13 個里」、台中縣第三選區則包括「太平市」及「大里市 25 個里」)。這種行政區劃與選區劃分不一致的結果，使本研究在部分總體層次的資料蒐集與變數建構上產生困擾。其中包括：(一) 政府提供的部分總體資料僅能以縣市為單位，導致以鄉鎮市區來劃分的選區缺乏實際數據；(二) 部分數據雖

能以鄉鎮市區為單位呈現，但對於以鄉鎮市區為劃分單位的選區而言，應如何整併出一個代表該選區的單一數據；(三)部分以村里為單位進行選區劃分者，在缺乏村里的相關總體數據下，如何提出該選區的總體數據。

針對上述選區劃分與總體層次數據單位不一致的問題，筆者採取下列方式解決之。首先，總體層次數據若僅提供以縣市為單位時（例如：失業率），則該縣市內的選區即以該數據作為參考依據。其次，若總體層次數據有提供到以鄉鎮市區為單位（例如：農牧人口比例），且選區是完整以鄉鎮市區來進行劃分時，即依據選區內各鄉鎮市區所佔選舉人口比例，針對該數據進行權重調整。最後，若選區有涉及以跨鄉鎮市區的村里來劃分時，若總體層次數據有提供到以村里為單位（例如：投票記錄），則直接以村里選舉人口比例進行權重調整，若僅提供到鄉鎮市區之數據，則依該鄉鎮市區所屬村里的選舉人口比例進行權重調整。

此外，由於金門縣與連江縣兩個選區的地理位置特殊，人文產業結構也與台灣地區存在著極大的差異。整體政治氛圍更是極度傾向泛藍的政治立場，長期以來的各項選舉也皆由泛藍政黨所掌控，民進黨在歷次選舉中未曾獲得一成以上的選票，可說是整體 73 個單一選區中的「極端值」。若將其納入統計分析，可能導致整體的分析結果受到這兩個「極端選區」的影響而發生偏誤。再加上，「台灣選舉與民化調查」歷年來所進行之全國性面訪調查，在實際執行成本與可行性的考量下，並未將金門縣及連江縣納入抽樣的母體範疇中。因此，從總體與個體層次而言，皆不適合將上述兩個選區納入分析，本研究的資料分析結果與推論也就不包括金門縣與連江縣。

第三節 變數測量

依據本文所提出的理論架構及研究假設，將同時從選民個體層次及選區總體層次分析選民在 2008 年立法委員選舉的投票抉擇，並採「階層推廣線性模型」進行統計分析。其中，模型內各項依變數及自變數的設定如表 3-1，後續並針對各項變數的測量方式提出說明。

表 3-1 各項統計模型之變數設定一覽表

投票抉擇	依變數	自變數
候選人票 (SMD)	二分類： 1. 國民黨候選人 2. 民進黨候選人	選民層次： 1. 省籍 2. 政治世代 3. 政黨認同 4. 政黨認同 * 認同強度 (交互作用) 5. 統獨立場 6. 族群認同 7. 陳水扁施政滿意度 8. 候選人評價 選區層次： 1. 農牧人口% 2. 失業率% 3. 2004 年立委選舉政黨得票率 4. 政黨得票變異 * 政黨認同 (交互作用) 5. 選區縣市長黨籍 6. 候選人現任與否
政黨票 (PR)	四分類： 1. 國民黨 2. 民進黨 3. 新黨 4. 台聯	選民層次： 1. 省籍 2. 政治世代 3. 政黨認同 4. 政黨認同 * 認同強度 (交互作用) 5. 統獨立場 6. 族群認同 7. 陳水扁施政滿意度 8. 政黨喜好程度 選區層次： 1. 農牧人口% 2. 失業率% 3. 2004 年立委選舉政黨得票率 4. 政黨得票變異 * 政黨認同 (交互作用) 5. 選區縣市長黨籍 6. 在單一選區提名候選人與否

表 3-1 各項統計模型之變數設定一覽表【續完】

投票抉擇	依變數	自變數
一致或分裂投票	<p>二分類：</p> <p>1. 一致投票</p> <p>2. 分裂投票</p>	<p>選民層次：</p> <p>1. 年齡層</p> <p>2. 教育程度（控制變數）</p> <p>3. 政治知識</p> <p>4. 政黨認同</p> <p>5. 政黨認同*認同強度（交互作用）</p> <p>6. 政黨認同與候選人評價一致與否</p> <p>選區層次：</p> <p>1. 有無脫黨參選或強勢獨立候選人</p> <p>2. 選區競爭程度</p> <p>3. 政黨在單一選區提名候選人與否</p>

一、依變數的測量

本文旨在探討台灣 2008 年立法委員選舉選民的投票抉擇，由於皆是屬於個人私密的投票行為，僅能藉由民意調查的訪問資料來蒐集。問卷中對於選民投票抉擇的測量題目為：「在選舉的時候，有很多人去投票，也有很多人因各種原因沒有去投票。請問在這次的立法委員選舉，您有沒有去投票？」若受訪者回答有去投票，即在選區候選人的第一票部分詢問：「請問您投給哪一位候選人？」，若受訪者回答不記得或拒答時，則追問「請問您（記不記得）是投給哪一政黨的候選人？」。至於在政黨比例代表制的第二票部分則詢問「下面我們列出這次參加立法委員選舉的所有政黨，請問您投給政黨的那一票，是投給哪一個政黨？」。

在變數的建構上，由於目前台灣的政治環境主要呈現藍綠陣營相互競爭的態勢，雖然此次選舉許多小黨為了具有政黨不分區名單的提名資格，或保有立法院的席次，紛紛在選區中提名候選人參選。但在單一選區制度下，除了極少數的選區因國民黨或民進黨未提名候選人外，其餘政黨或無黨籍候選人幾乎沒有勝算可言。選舉結果也顯示在單一選區中，國民黨得票率達 53.48%，民進黨則為 38.65%，兩黨合計獲得高達九成二以上的選票。在 73 個單一選區中，兩黨囊括了 70 席。²⁹在第二票的政黨比例代表選舉，也僅有國民黨及民進黨得票超過 5%

²⁹ 另外三席分別是台中縣第二選區的顏清標、澎湖縣的林炳坤以及金門縣的陳福海當選。其中，在台中縣第二選區及澎湖縣國民黨皆未提名候選人參選，而是在協調後「禮讓」顏清標及林炳坤參選。

的門檻，分別以 51.23%及 36.91%的得票率獲得不分區席次的分配資格。就連較具規模的新黨及台聯也僅獲得 3.95%及 3.53%的選票，未能獲得不分區的席次。再就統計方法而言，若要將投給無黨籍或其他小黨的候選人各自單獨列為一個類別，更可能因為分析個數過少導致係數估計發生偏誤。故在考量此次選舉的實際競爭型態及統計方法的可行性後，決定在探討選民單一選區的投票抉擇時，以投給「國民黨」或「民進黨」候選人兩類進行分析；在政黨比例代表票部分，則將新黨及台聯納入，形成「國民黨」、「民進黨」、「新黨」及「台聯」四類。

至於選民對於兩張選票是否投給同一個政黨及其提名候選人，即一致投票或分裂投票。由於並非所有政黨皆會在各單一選區中提名候選人參選，以致部分選民面臨「無機會」採取一致投票的狀況。這種情形與所屬政黨在單一選區有提名候選人參選，讓選民「有機會」採取一致投票有所不同。此次立委選舉，同屬泛藍陣營的國民黨、親民黨及新黨，在選前協調達成合作共識。其中，親民黨完全加入國民黨的提名作業機制，不論單一選區或政黨比例代表皆以國民黨名義參選或納入國民黨提名名單中。³⁰新黨則是在單一選區參加國民黨的初選，候選人以國民黨名義登記參選，但在比例代表制中則保有自己不分區的提名名單。至於在泛綠陣營方面，民進黨與台灣團結聯盟選前並未達成合作共識。不但在比例代表制中各自提出自己的不分區名單，就連在部分單一選區中，兩黨也都提名自己的候選人參選。為了完整檢視此次選舉選民採取一致或分裂投票的情形及影響因素，本研究除了針對政黨在單一選區及比例代表都有提名，屬於「滿意者」(satisfied)的選民進行分析外，同時也針對某些政黨在部分單一選區中未提名候選人參選，以致無機會一致投票的「挫折者」(frustrated)選民進行分析。檢視在被迫採取分裂投票的選民中，是否也會如 Benoit 等人(2006)所言，在單一選區中投給同一聯盟(coalition-stickers)的現象。因此，在依變數的設定上，將以選舉公報之登記為主，選民只要兩張選票都投給同一政黨，即屬「一致投票」；若兩張選票的抉擇非同一政黨，即屬「分裂投票」。

二、個體層次自變數的測量

有關選民個體層次的各項變數測量，皆是來自 TEDS2008L 的問卷內容。在人口特徵上，選民的省籍背景在台灣傳統以父系為主的家庭文化下，將以父親的

³⁰ 親民黨僅提名林惠官參選連江縣立委選舉，但連江縣的政治環境較為特殊，且 TEDS2008L 亦未在該選區中抽樣訪問，故本文並未將其納入分析。

省籍為主，區分為「本省客家人」、「本省閩南人」以及「大陸各省市人」三類。至於「原住民」或「華僑」等由於樣本個數過少，在統計分析的考量下未納入模型中。另外，有關選民政治世代的劃分方式，過去學者曾依據台灣政治環境的歷史發展演進，包括光復前的日據生活經驗、國民黨遷台前的大陸生活經驗、國民黨威權統治時期、台灣後威權或民主化階段等，而有不同的劃分方式（陳陸輝 2000；劉義周 1993）。但隨著世代交替，具有大陸與日據時期之生活經驗的選民逐漸凋零，對於威權統治以及民主化過程的經驗成為主要的影響因素。再加上本研究旨在探討選民在立法委員選舉的投票抉擇，故除了選民成長時期的外在政治環境與經濟發展可能影響選民政治態度的形塑外，選民過去的投票經驗也必須納入考量。其中，1972 年的「第一屆增額立法委員選舉」與 1992 年的「第二屆立法委員全面改選」，可說是台灣中央民意代表選舉史上最重要的兩個里程碑。³¹1972 年的增額選舉，除了讓本省菁英有機會踏入立法院參與議事運作外，更重要的是該增額選舉必須「定期改選」，較為符合一般民主選舉的規範，不但讓黨外人士有組織化參選的動機，也讓選民對於民主選舉有進一步的認知。至於 1992 年的立法院全面改選，不但終結了立法委員終身職的民主怪象，也是立法院「首度且全面」符合「參與、競爭」原則的選舉。相對於過去歷次的選舉活動，此次選民的投票權行使，除了將影響中央執政權的所屬外，更意味著台灣政治民主化的落實。

因此，本文對於選民政治世代的劃分方式，以 1972 年及 1992 年選民是否成年首度取得投票權作為切割的時間點。即以 1952 年及 1972 年出生與否，劃分為三個政治世代。第一世代（即 1951 年以前出生），其成長與成年階段幾乎都是處在國民黨的威權統治下，甚至具有大陸或日據時期的生活經驗，政治上也經歷過白色恐怖的階段、二二八事件的發生。且台灣當時的經濟狀況相當貧乏，投票經驗也較為欠缺，更何況當時的選舉幾乎僅止於地方層級的基層選舉活動。第二世代（即 1952~1971 年間出生），在此階段成長的選民，雖然仍處於國民黨的威權統治，但政治上卻開始出現不同的聲音。例如：雷震的《自由中國》及籌組「中國民主黨」、高雄美麗島事件。在經濟上，此時台灣正值經濟起飛階段，民眾逐漸脫離過去的貧窮生活，九年國教的實施，也提升民眾的知識教育水平。尤其是該世代選民，在立法委員定期增額選舉後，首度取得立法委員的投票權，對於國

³¹ 雖然台灣在 1991 年曾進行第二屆國民大會代表選舉，但國民大會的職權內容較不符合西方民主國家的國會，且對行政權的監督與制衡主要仍是以立法院為主。

會民主選舉的體會較為深刻。第三世代（即 1972 年以後出生），其成長過程屬於國民黨威權統治的後期，感受相對較為模糊甚至未曾經歷，後續民進黨的成立、威權統制的結束，在立法院全面改選後首次取得立法委員選舉權，以及經歷政黨輪替等，皆讓該世代選民對政治的認知可說是相當的民主化，且在經濟生活層面相對富裕，自然形塑出不同的政治態度與價值觀。

至於在選民的心理態度認知上，包括政黨認同、統獨立場、族群意識、施政滿意度、候選人評價以及政治知識等，則藉由問卷中的下列題組來測量。

（一）政黨認同

問卷中詢問受訪者「目前國內有幾個主要政黨，包括國民黨、民進黨、新黨、親民黨，以及台灣團結聯盟，請問您是否偏向哪一個政黨？」、「請問是哪一個政黨？」以及「請問，您偏向這個政党的程度是很強，普通，還是只有一點？」。除了無政黨認同的中立者外，由於親民黨在此次選舉完全加入國民黨的提名機制，故在政黨認同上將與國民黨認同者合併。另外再區分民進黨、新黨以及台聯認同者。

（二）統獨立場

問卷中詢問受訪者「關於台灣和大陸的關係，這張卡片上有幾種不同的看法：1. 儘快統一、2. 儘快獨立、3. 維持現狀，以後走向統一、4. 維持以後走向獨立、5. 維持現狀，看情形再決定獨立或統一、6. 永遠維持現狀。請問您比較偏向哪一種？」。依據統獨立場的方向，將「儘快統一」以及「維持現狀，以後走向統一」歸併為「傾向統一」；將「儘快獨立」以及「維持現狀，以後走向獨立」歸併為「傾向獨立」；將「維持現狀，看情形再決定獨立或統一」以及「永遠維持現狀」歸併為「維持現狀」三類。

（三）族群認同

問卷中詢問受訪者「在我們社會上，有人說自己是「台灣人」，也有人說自己是「中國人」，也有人說都是。請問您認為自己是「台灣人」、「中國人」，或者都是？」。直接將其區分為「台灣人」、「都是」、以及「中國人」三類，惟由於回答「中國人」的比例過低，故在模型中將「都是」以及「中國人」歸併為一類。

(四) 施政滿意度

問卷內容為「如果以 0 表示非常不滿意，10 表示非常滿意，整體來看，對於陳水扁總統所領導的政府這四年來的表現，請問您的滿意程度是多少？」。呈現出 0~10 的連續變數，數字愈高表示選民愈滿意陳水扁政府過去四年來的施政表現。反之，數字愈接近 0，表示對陳水扁的施政滿意度愈低。

(五) 候選人評價

問卷中分別詢問受訪者對選區中主要候選人在「服務選區選民」、「爭取地方福利」、「清廉」、「監督政府」、「推動法案」以及「堅持政治理念」等六個面向的評價。利用 0~10 的測量尺度給予評價，數字愈高表示評價愈好。在操作上，將受訪者給每一位候選人在六個面向的數字取其「平均數」，建構出 0~10 的連續變數，作為候選人綜合評價的單一指標。

(六) 政治知識

分別詢問受訪者五道題目「請問現在的副總統是誰？」、「請問現在的美國總統是誰？」、「請問現在的中國大陸國家主席是誰？」、「請問哪一個機關有權解釋憲法？」以及「請問這一屆選出的立法委員一任幾年？」。每一道題目中，當受訪者回答出正確答案，則編碼為 1；回答錯誤、知道卻說不出正確答案、不知道或拒答等，則編碼為 0。最後再將受訪者的五道題目編碼進行加總，建構出 0~5 的連續變數。數字愈高表示受訪者的政治知識愈高（5 表示五道題目的回答完全正確）。反之，數字愈低表示選民的政治知識愈低（0 表示五道題目的回答完全錯誤或不知道）。

三、總體層次自變數的測量

本研究依據學理探討可能影響選民投票抉擇的選區環境變數，包括區域經濟發展、政黨勢力版圖、選舉競爭程度與候選人參選資訊等。各項變數的資料來源與測量方式如下：

(一) 區域經濟發展

國民黨長期以來透過地方農會等基層組織，掌控地方資源分配與地方派系運作，在選舉時也透過人際網絡進行組織動員，以致農業地區往往成為國民黨組織動員的重點所在。本研究對於各選區農牧人口所佔比例是依據行政院主計處於

2005 年所進行的農林漁牧普查結果。³²該項資訊的公布除了以縣市為單位外，亦可以鄉鎮市區為單位進行查詢。若選區的劃分含跨數個不同鄉鎮市區，或甚至在同一鄉鎮市區又以村里為單位進行劃分，筆者即依據選區內的行政區劃所佔選舉人口比例進行權重調整，建構出代表各選區農牧人口比例的單一變數。

此外，由於國民所得、國民生產毛額、經濟成長率等各項數據，皆是以國家為單位進行發佈，並無法據以檢視選區之間的變異程度。在失業率的統計上，行政院主計處則有以縣市為單位發佈各縣市的失業率，雖然無法呈現出每一個選區的變異程度，但各地方的經濟發展政策主要由縣市政府擬定與主導，將會降低各縣市內部不同選區之間經濟發展的差異性。尤其失業率除了自己所可能面臨的切身感受外，周遭親友同事的際遇或相關資訊的傳遞，皆可能引起選民的關注進而影響投票抉擇。因此，本研究將以行政院主計處所公布之 2007 年各縣市失業率作為各選區經濟發展的另一項指標。

（二）政黨版圖勢力

主要是以各政黨前一次（2004 年）立法委員選舉的得票率，作為此次選舉各政黨的版圖勢力，同時在模型中也可作為控制變數，以釐清其餘變數對選民投票抉擇的影響力。惟 2001 年是新黨最近一次以新黨名義提名參選的立法委員選舉，故在探討政黨比例代表票的抉擇因素時，將以 2001 年立委選舉的新黨得票率作為新黨的版圖勢力。而在探討選區候選人票的影響因素時，由於國民黨、親民黨與新黨在單一選區協調以國民黨名義提名參選，故國民黨的版圖勢力即是 2004 年國民黨（已包括新黨）與親民黨合計的得票率。另外，由於 2008 年立委選舉當時，各縣市行政首長皆由國民黨或民進黨擔任，故在縣市首長黨籍之自變數上，將採「虛擬變數」至入模型，以「民進黨縣市長」作為參照組。

政黨在選區中的版圖勢力穩定性上，則以各政黨在選區中 2000 年以後的四次中央層級選舉之得票變異程度進行測量。包括：2000 年總統選舉、2001 年立委選舉、2004 年總統選舉以及 2004 年立委選舉。在操作上，先取政黨在這四次選舉的標準差，再依據政黨的得票率進行權重調整。之所以調整權重主要是考量相同的得票率標準差，在不同平均得票率的選區中有其不同意涵。例如：某一政黨在 A 選區三次選舉的得票率分別為 10%、20%、30%，在 B 選區三次選

³² 相關資訊可查詢行政院主計處網站 <http://www.stat.gov.tw/ct.asp?xItem=18472&ctNode=555>。

舉的得票率則分別為 70%、80%、90%。經計算該政黨在 A、B 選區三次選舉得票率的標準差皆為 10，但其所代表的意涵卻大不相同。在平均得票率為 20%的 A 選區中，政黨在該選區的版圖勢力相當薄弱，得票率之標準差為 10，表示政黨在該選區的得票變異極大。反觀，在平均得票率為 80%的 B 選區中，政黨在該選區的版圖勢力具有絕對優勢，雖然得票率的標準差一樣為 10，但相較之下政黨的得票變異程度卻較小。故本研究將採取「變異係數」(coefficient of variation，簡稱 CV)，作為測量各政黨在選區中版圖穩定程度的指標。除了計算 2000 年以後四次中央層級選舉之得票率的標準差外，亦同時將政黨的得票率納入考量。變異係數的公式為： $CV_A = \widehat{\sigma}_A \times \frac{1}{\bar{A}} \times 100\%$ 。

$0 < CV_A < \infty$ 。其中， CV_A 表示 A 政黨在該選區之得票率的變異係數。 $\widehat{\sigma}_A$ 表示 A 政黨在該選區四次選舉之得票率的標準差。 \bar{A} 則表示 A 政黨在該選區四次選舉的平均得票率。而變異係數的值介於 0 至無限大之間。其中，由於泛藍的國民黨、新黨與親民黨不但在此次選舉達成合作協商聯合提名，在過去幾次選舉亦有多次合作競選的運作，故在政黨得票穩定程度的測量上，將合併為單一指標，將三個政黨過去的得票紀錄合併計算。

(三) 選區選舉競爭程度

過去用以偵測選區競爭程度的指標甚多，其中最常見的即是以選舉結果候選人之間的得票率差距來測量，得票差距愈小者，表示選舉競爭程度愈高。但 Reed(1999)認為這種「事後預測」(ex post)的測量方法並不合理。因為選舉結果是發生在選民投票之後，在因果關係上不應該以時間點「後發生」的事實當作「先發生」之行為的因。選民應是先透過各種資訊瞭解選區競爭程度後，才決定如何投下手中的選票。故 Reed 提出以「事前預測」(ex ante)的方式，包括利用前一次選舉候選人的得票情形，及此次選舉候選人的競選支出等來建構選舉競爭程度。這不但在學理上更為合理，在統計分析上也具有更精準的預測結果。但由於此次立法委員選舉不但採取全新的選舉制度，而且席次減半，選區也重新劃分，前一次立法委員選舉的得票數據實難以作為選區競爭程度的指標。另外，候選人雖然必須在選後向監察院申報競政治獻金收入與競選支出的數據，但這種申報內容可說形式重於實質，可信度與真實性時常令人質疑。

在前一次選舉結果與此次選舉候選人的競選支出，皆不適宜作為選區競爭程度的指標下，如有「選前」的調查資料可以偵測各選區候選人的勝選機會，且其調查方法也具有代表性時，即是符合「事前預測」邏輯的適當指標。政治大學預測市場研究中心曾針對 2008 年立委選舉進行「市場預測」。對 73 個單一選區在「未來事件交易所」網站，於選前一個半月至投票日前一天，提供民眾以類似股票線上買賣的方式，預測各選區各候選人當選與否的可能性(0~100)，並於選後針對該項預測指標與選舉結果進行分析比對，不但在 73 個選區的當選命中率高達八成五外，其與候選人實際當選機率的吻合度也相當高(國立政治大學預測市場研究中心 2008)。故本文即以此項調查在各選區中，民眾預測當選率最高之前兩位候選人的差距作為測量選區競爭程度的變數，差距最大為 100 (領先之候選人預測當選率為 100，次之候選人為 0)，差距最小為 0 (選區前兩名領先之候選人的預測當選率皆為 50)。為讓此一指標更貼近一般的認知，再將兩候選人預測當選率的差距減去 100 並取絕對值，形成 0~100 連續指標，數字愈大，表示選區競爭程度愈高。

(四) 候選人參選資訊

有關各單一選區候選人的參選資訊，包括參選人的黨籍背景、是否脫黨參選、是否屬於強勢之獨立候選人、是否屬於現任者，以及各主要政黨是否在單一選區中提名候選人參選等。除了利用選舉公報、中央選舉委員會之選舉資料庫外，也透過選前選後各新聞媒體的報導、政黨或候選人所屬網站等方式，蒐集候選人的相關資訊來建立。

針對上述各項選區總體層次指標的資料來源、變數測量以及建構方式，經整理後，各項指標的整體描述性統計量等資訊，如表 3-2 所示。至於各選區各項總體層次的完整資訊，請參閱附錄一。

表 3-2 選區各項總體層次變數之描述性統計量

變數名稱	個數	平均數	標準差	最大值	最小值
失業率%	71	3.91	0.16	4.20	3.70
農牧人口%	71	15.03	14.56	45.09	0.17
2001 立委選舉新黨得票率	71	2.50	2.95	12.93	0.00
2004 立委選舉民進黨得票率	71	36.21	6.77	53.57	22.65
2004 立委選舉台聯得票率	71	7.99	4.56	23.46	0.00
2004 立委選舉國親新三黨得票率	71	46.35	10.99	69.15	0.00
國親新三黨得票變異係數	71	14.65	11.11	92.40	4.20
民進黨黨得票變異係數	71	20.17	6.92	39.40	9.80
選區競爭程度	71	28.05	29.89	99.62	0.02

資料來源：(不含金門縣及連江縣)

- (一) 失業率：行政院主計處網頁「主計處統計專區→就業、失業統計」，網址：
http://win.dgbas.gov.tw/dgbas04/bc4/manpower/103_2f.asp。
- (二) 農牧人口%：行政院主計處網頁「主計處統計專區→農林漁牧業普查」，網址：
http://win.dgbas.gov.tw/agr/ics_main.asp。
- (三) 各年度選舉政黨得票率：中央選舉委員會網頁「選舉資料庫」，網址：
<http://210.69.23.140/cec/cehead.asp#>。
- (四) 選區選舉競爭程度：國立政治大學市場預測研究中心公布之「2008 年立法委員選舉結果之初步分析」(版本日期：2008.01.18)。網址：
<http://nccupm.wordpress.com/2008/01/19/preliminary-analysis-of-legislator-election-market-results/>。