

# 廢票爲何發生？

## 兼論臺灣選舉無效票之空間效應，1992-2008

鄧志松\*、吳親恩\*\*、柯一榮\*\*\*

### 《本文摘要》

有關選舉無效票（即廢票）的研究很少，主要原因是無效票比率通常不高，對選舉結果的影響不大；而且無效票一般被視爲是「疏失」造成，並非選民「意志」的表現，因此很少人探討其發生的原因。作者考察我國 1992 至 2008 年，37 個不同類型的選舉，結果發現無效票率的空間分布並非隨機，且有一定的模式。不同類型的選舉，無效票率的高低有跡可尋，代表無效票率其實是可以被解釋的，其背後有社會與制度因素存在。經由無效票率的追蹤分析 (panel data analysis) 後，發現選舉法規、選舉複雜程度、選舉受重視程度、選舉年度及經社背景變數（教育年數、老年人口、人口密度比及原住民人口比例），約可解釋 34% 之無效票率變異。作者進一步以空間迴歸模型探討，發現無效票率存在正向之空間自相關情形，部分地區則有明顯空間異質現象，極有可能是該地區民眾長期對政治冷漠使然，如果我們要提昇民主的內涵，則這些無效票率特別高的地方值得關注。

關鍵詞：無效票率、空間自相關、空間異質、追蹤分析、空間迴歸

\* 國立臺灣大學國家發展研究所副教授。E-mail: terry@ntu.edu.tw。

\*\* 中央研究院政治學研究所助研究員。E-mail: chinen@gate.sinica.edu.tw。

\*\*\* 國立臺灣大學國家發展研究所碩士。E-mail: r96341011@ntu.edu.tw。

## 壹、前言

國內外有關選舉無效票（無效票）的研究相當少，主要原因是它通常被視為是一種「疏失」，而且無效票率通常不高，對選舉結果影響不大，因此不被重視。我國各類型選舉全國無效票率，大多約在 2% 以下，但 2004 年第 11 屆總統選舉卻是個例外。該次選舉，由於國民黨（連戰、宋楚瑜）與民進黨候選人（陳水扁、呂秀蓮）得票差距甚小（僅 29,518 票），而無效票卻異常地高（337,297 票），無效票似乎有影響最後選舉結果的可能性。在政治人物、輿論的推波助瀾下，選舉舞弊的傳聞一度喧囂塵上。<sup>1</sup> 其後經過全面驗票，發現選舉作業基本上無大礙，<sup>2</sup> 無效票率異常的原因，推測可能與「無效票認定標準」由寬轉嚴有關，選舉舞弊的爭議始逐漸平息。

經過此次選舉爭議，無效票引起作者的關注，無效票的空間分布是否值得研究？影響無效票的因素是什麼？它可以給我們什麼啓示？無效票基本上是投票失誤造成，也許是蓋錯章、或蓋錯地方、或選票污損、難以辨識，大部分情形是不小心造成的。依常理判斷，應該很少人願意自己的票「意外地」變成廢票，否則又何必辛苦來投票？但是，我們也不能排除另一種可能性，即故意投無效票，可能是基於惡作劇，但也可能是某種意見表達，例如對政治的不滿與厭惡。問題是，我們很難判斷無效票究竟是「有意的」，還是「無意」造成的。由於是匿名投票、開票，除非是真的有意投無效票，否則投票人亦無從得知自己所投之票是否被當成無效票處理。因此，傳統的投票行為研究途徑，不管是抽樣問卷或深入訪談，都難以針對無效票現象進行探討。分析總體的投票資料，往往成為唯一的途徑。

廢票可能純粹是「無意或失誤」造成，但為什麼不同的人失誤的機率不一樣？是什麼因素影響失誤的出現機會，既有文獻除了無效票認定標準這一因素之外，其他可能的因素並未被有效討論，本文將廣泛討論可能的選舉制度與社會經濟變項。就民主政治，擴大政治參與本是天經地義，無效票代表部分人（雖然比例不高）的意見被漠視了，如果存在某種類型的人，其意見特別容易被意外地「漠視」，那就代表投票方式可能有改善的空間，否則形同對這些人的歧視。除了制度與社會經濟因素之外，空間分析則提供另一種研究

<sup>1</sup> 選後有學者指出發現，無效票率越高之地區，連宋之得票率就越低，反之，而陳呂得票率卻偏高，並認為這是陳呂陣營作票之證據。

<sup>2</sup> 依據當時之民進黨副秘書長鍾佳濱、政策會副執行長梁文傑所述，民進黨黨團曾要求其律師團在參與驗票之同時，觀察若無效票認定標準未改，可辨識為投給何組候選人之無效票數有多少？結果發現，在未包括臺南縣之其他縣市的統計資料中，可辨識要投給一號（民進黨）卻被判為無效票之比例約為 46%；可辨識要投給二號（國民黨）卻被判為無效票之比例約為 26%；純粹無效票比例約為 28%。

的角度，其關注變數的空間分布差異，是否有特定的分布模式（空間隨機，或者空間聚集），如何解釋變數的地區性差異與空間分布？這也是我們要討論的部分。

另一方面，如果無效票是有意造成的，我們則要問，哪些地區的人特別容易有意地去投廢票？這可能是廢票運動的影響，也可能是對政治長期不滿造成的結果，這部分後文也會討論。

本論文運用中選會公布的選舉資料，從 1992 至 2008 年，總統、立委、國代、省長、省議員、縣市長及縣市議員選舉等，共計 37 個選舉。以 GIS（地理資訊系統）與空間分析方法，將選舉資料匯整至鄉鎮，接著比較不同地區間的差異，探討變數的空間分布模式，最後進行空間迴歸分析。因為探討多個選舉資料，以追蹤分析 (panel data analysis) 的方法進行。文章首先討論無效票的起源及相關研究，接著說明研究設計，界定依變數與自變數，以及分析方法；最後進行數據分析，並結論。

## 貳、無效票的產生與相關研究

### 一、既有研究

國內外有關選舉無效票之研究極少，主要原因是它通常被視為是一種「疏失」，而且它通常不高，影響選舉結果的可能性低，因此不被重視。2004 年第 11 屆總統選舉是個例外，該次選舉，由於國民黨與民進黨候選人得票差距甚小，而無效票卻異常地高，引起社會大眾的關注。駱明慶（2006）以第 10 屆（2000 年）與第 11 屆總統選舉各鄉鎮之選舉資料，以及 2000 年戶口普查之人口資料，檢驗選後盛行之「廢票運動」、「作票」與「無效票範圍擴大」三個假說。

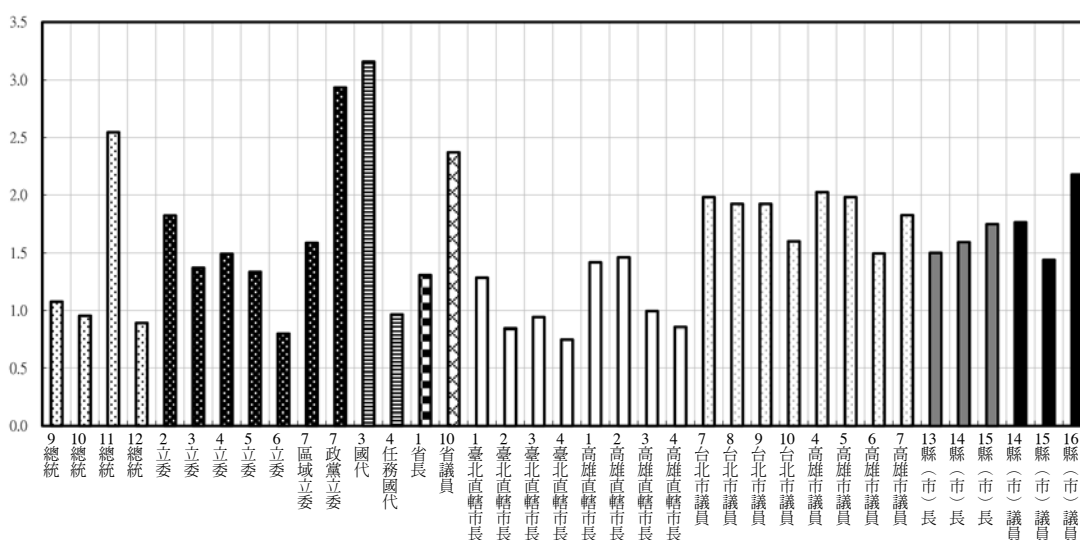
首先，「廢票運動」係指由工人立法行動委員會等民間團體所組成的「百萬廢票行動聯盟」發起，號召選民在 2004 年第 11 屆總統選舉投下無效票，以表達對於藍、綠陣營皆是「爛蘋果」之不滿。關於「作票說」，則是該次選舉後，國民黨中常會邀請學者馬自恆就其所提出之「饅頭曲線」作專題報告。「饅頭曲線」為總統選舉，藍、綠得票數差距與無效票率之相關圖，發現無效票率越高之地區，連宋之得票率就越低，而陳呂得票率卻偏高，認為這是作票之證據（林修卉 2004）。第三個假說為「無效票認定範圍擴大」。由於在該次選舉前，《總統副總統選舉罷免法》對於無效票之認定標準作了修正，相較之下，比起之前之認定嚴格許多，因此始導致無效票大增。

駱文研究發現，總統選舉無效票率增加較少之縣市，皆為直轄市或省轄市等都會型縣市；而無效票率增加較多之地區，則為都市化程度較低之縣市，此與「無效票運動」在都

市地區較顯著之預期不符合。關於「作票說」（即陳呂得票率越高，無效票比例越高），作者發現在控制第三變數，「教育程度」之後，陳呂得票率與無效票率間不再相關，亦即陳呂得票率與無效票率呈現正向的假相關，主要影響因素其實是「教育程度」。最後，作者指出各鄉鎮「不識字比例」和「無效票增加」二者呈高度正相關，也就是說，不識字比例較高之鄉鎮，在無效票認定標準從嚴之後，無效票之比例亦顯著增加，此與「無效票範圍擴大」之假說較為一致。因此其支持第三個假說，即無效票認定標準從嚴使得無效票「異常」增加。

鄧志松（2008）的研討會論文則以「空間分析」之觀點看無效票率分布之狀況。鄧文考察歷屆的總統（第9至11屆）與立委選舉（第3至7屆），看出無效票率的空間分布隱約存在某種模式，並非空間隨機。其中，第11屆總統選舉的無效票率空間分布尤其特殊，與其他選舉有明顯差異。如果扣除第11屆總統選舉，計算其他選舉的平均無效票率，以此為依變數，教育年、原住民比及宗教場所密度為自變數，則可解釋23%的變異。若只考慮第11屆總統選舉無效票率，同樣地，以教育年、老年人口比及宗教場所密度為自變數，則可解釋的比例高達54%。顯然，第11屆總統選舉無效票率的性質與其他選舉顯有不同，造成此次選舉無效票率特殊的原因應是有效票認定標準從嚴所致，此看法與駱明慶（2006）相同。亦即教育程度較低、年老者較多之地區，基於過去之投票習慣，較不適應新的改變。

除了廢票認定標準之外，我們特別想要知道的是，是否有其他的制度與社會經濟變項會影響一個地區廢票率的高低。為了回答這個問題，有別於過去僅針對幾個特定選舉的研究，作者完整蒐錄了1992至2008年，總統、立委、國代、省長、省議員、縣市長及縣市議員等，共計37個選舉的無效票率。圖1顯示各類型選舉無效票率的高低，發現一些有趣的現象。



資料來源：中央選舉委員會（2009）選舉資料庫，作者自行計算畫圖。

圖 1 各選舉全國無效票率長條圖

圖 1 顯示，無效票率大多介於 1% 到 2% 間，最高可超過 3%，但很少低於 1%，平均是 1.57%。比較不同型態之選舉，地方層級選舉之無效票率大於中央層級選舉；議員性質選舉又較首長性質的選舉高。然而，在中央層級的選舉中，有幾次選舉的無效票率較高，最特殊的是第 11 屆總統選舉（2004 年），其無效票率高於其他三次總統選舉甚多，如前所述，這可能是與有效票認定標準變嚴格有關。第 6 屆立委選舉無效票率變低，正好反映當時無效票認定標準變寬鬆。

第 7 屆立委選舉無效票率急速升高，尤以政黨票之部分最為明顯，這可能與立委選制調整有關。首次實施單一選區兩票制，選民熟悉程度較低，間接影響廢票的高低。第三屆國民大會代表選舉無效票率特別高，可能與憲政改造，職權縮水，不受人民重視所致。另外 2005 年三合一選舉（第 15 屆縣市長、第 16 屆縣市議員與鄉鎮市長）無效票率升高，尤以縣市議員為最，這可能與合併選舉易造成混淆，或是人民較不重視議員選舉有關。

## 二、制度與社經因素

綜上所述，無效票率的高低，除了廢票認定標準之外，也可能會受到其他制度性因素的影響，既有的文獻對此討論不多，我們區分幾個不同面向的影響因素，分別加以討論。

第一、法規與選制改變因素。無效票認定標準的差異最為重要，2004 年第 11 屆總統選舉，無效票率異常增加，原因是該次選舉無效票認定標準嚴格，為歷年之最。一般而言，有效票的認定標準，載明於該次選舉所適用之選罷法，大同小異。只要選民依據規定

至投票所領取選票，並使用選票專用之圈選章，圈選在可辨識要投給哪位候選人的地方，即認定為有效票。不過，認定標準歷經多次修正或適用法律改變後，仍有寬嚴之別，主要的差異在「圈選章應蓋於何處」方屬有效。2004年第11屆總統選舉規定，所圈位置於圈選欄內方屬有效，<sup>3</sup> 導致無效票大幅增加，因為有不少人習慣將章蓋在「圈選欄」外。第12屆總統選舉，再次改回原來較寬鬆的規定。

立法委員方面，第5屆以前的立法委員選舉，明定：「不圈在姓之上端者」<sup>4</sup> 視為無效，即須圈選在選票上印製候選人姓名之「姓的上端」處，始為有效票，較一般規定僅要圈選在可辨識投票對象處為嚴格。不過第6屆立委選舉（2004年）以後，認定標準統一適用《公職人員選舉罷免法》較寬鬆之規定，與省（市）長、省（市）議員、縣市長、縣市議員、選舉之認定標準一樣。至於國民大會代表選舉，圈選在可辨識投票對象處即視為有效票。<sup>5</sup>

另一個可能的影響因素是，該次選舉是否採取新的選舉制度，例如第七屆立委選舉，從SNTV改為單一選區兩票併立制。選區劃分從原本之中選區（一選區選出數名代表）變為小選區（一選區選出一名代表），投一票變成投兩票（分為區域票與政黨票）。選制變革有可能使得選民在不瞭解的情況下，增加出錯的機會。根據台灣選舉與民主化調查（TEDS）2008年第7屆立法委員選舉電訪人民對於新選制的認知，發現44.9%的民衆不知道一個選區可選舉出多少立委；30.2%的民衆不知道除了公投票以外，還可以領多少張票，<sup>6</sup> 對選制的不熟悉有可能造成無效票率的增加。

第二、無效票率的高低也與選舉的受重視程度有關，這也可以有幾個部分。選舉的層級越高（例如中央層級的總統、立委選舉），在選民心目中的重要性也較高，選民會重視自己的一票，謹慎地行使投票權，因此無效票率推測會比較低。其次，選舉的受重視程度也可能與選舉所產生之公職人員的職權變化有關係，當職權降低時，受到選民重視的程度也可能會相應降低，例如2005年任務型國代選舉，職權大幅縮減，受重視的程度自然也就跟著下降，因此無效票率可能會因此而增加。而選舉受重視的程度除了觀察選舉的層級與職權變化之外，也可以觀察投票率以及藍綠得票的勝差，投票率越高或藍綠得票的差距很小的選舉，代表競爭激烈，選民較可能會審慎行使投票權，因此廢票率預期可能會因此

<sup>3</sup> 《總統副總統選舉罷免法》2003年修訂，第60條第1項第3款規定：「所圈位置全部於圈選欄外或不能辨別為何組者，視為無效票」（法務部 2003）。

<sup>4</sup> 《立法院立法委員選舉罷免法施行條例》第44條第2項第4款之規定（法務部 1988a）。

<sup>5</sup> 在第3屆國代選舉之前，適用《國民大會代表選舉罷免法施行條例》之規定；2005年第4屆任務型國代，適用《國民大會代表選舉法》之規定，大致內容相同（法務部 1988b）。

<sup>6</sup> 詳細資料請參考「台灣選舉與民主化調查」網頁之「2008年立法委員選舉電訪變數清單」（台灣選舉與民主化調查 2008）。

降低。

第三、另一個影響無效票率的高低的因素，與選票的複雜度有關。這包括選票張數與候選人人數，某次選舉領票數越多（合併舉行之投票越多），或者某個選區候選人數越多（選票越長）越容易造成選民之混淆。領選票之多寡，牽涉到「投錯票匱」的問題，在選舉實務上，投錯票匱的選票仍屬有效，不過，2004年第11屆總統選舉例外，因為該次選舉與全國性公民投票同時舉行，中選會規定公投與總統票投錯票匱者，皆屬無效。2008年第12屆總統選舉規定放寬，投錯票匱仍屬有效，事實上，不論有效與無效，影響不大，因為採兩階段投票，投錯票匱的情況並不多見。倒是合併投票數目越多，領取選票越多，有可能犯錯的機率亦會增加。其次，選票的長短（亦即候選人數之多寡）亦可能有影響。如選票太長，要將選票摺疊始能投入票匱，則章漬複印變成廢票的可能性增加；或候選人（或政黨）越多，選民越容易混淆，於是在不知道該如何抉擇的情形下，直接投無效票。

第四、投有效票也可以視為一個「學習」的行為，從圖1的長條圖中，發現投越多次，廢票率就降低；換選制或是廢票認定標準，廢票率就上升，因為要重新學習，此種趨勢，立委與總統選舉就很明顯。另外與此有關的，我國幾乎年年有選舉，每次選舉如同一次公民教育，選民逐年累積投票經驗，如果「學習」有成效，廢票理應逐年減少。

第五、除了上述的制度性因素外，經社變數也可能有會影響人們「犯錯」的機會，主要的因素包括教育程度、老年人口比例、人口密度與原住民人口比。教育程度高者，對新資訊接受較多，投票犯錯的可能性較低，這部分的討論已見於駱明慶（2006）；老年人行動較為不便，容易在投票中出錯；另外人口密度高，代表都市化程度高，訊息傳播較多，選民投票犯錯的可能性較低；最後原住民可能對於政治較為冷漠，或因為資訊相對封閉，對投票程序較不熟悉，所以投廢票的機率可能較高。所以一個地區，教育程度普遍較低、老年人口比例較高、人口密度較低或原住民人口比較高，則廢票率都可能相應增高。

另外，無效票聯盟運動屬有意的無效票，其影響之強弱甚難估計，一般認為，廢票聯盟主要活動範圍在都會區，其影響範圍應以都會區為限。此外，在制度與社經因素之外，無效票率是否有「鄰近效應」(neighborhood effect)？即控制以上的變數之後，「鄰近地區」的無效票率會不會相互影響？亦即「鄰近地區」的無效票率高，「我」也會比較高？反之亦然？如果有這樣現象存在的話，要如何解釋？「空間異質」(spatial heterogeneity)則是另一個空間分析關注的議題，是不是有些地區其無效票率產生的因果機制與其他地區不同？這也是本文要關注的面向。

## 參、研究設計

### 一、分析架構與假設

本文旨在探討影響無效票率的因素，分析架構如圖 2 所示。影響無效票率的來源有三。

第一為「制度與結構性因素」，其下又可以分為：A、「法規與選制改變」（有效票認定的寬嚴、是否新選制）；B、「選舉受重視程度」（投票率、選舉層級、是否職權降低、及藍綠勝差）；C、「選舉複雜度」（合併投票數與候選人數）；D、「民主經驗」（選舉年度）。

第二個來源是「經社背景」變數。不同的人是否犯錯（投無效票）的機會不一樣，因此納入教育年數、老年人口比與人口密度（代表都市化程度的高低）、原住民人口比等社經背景變項。

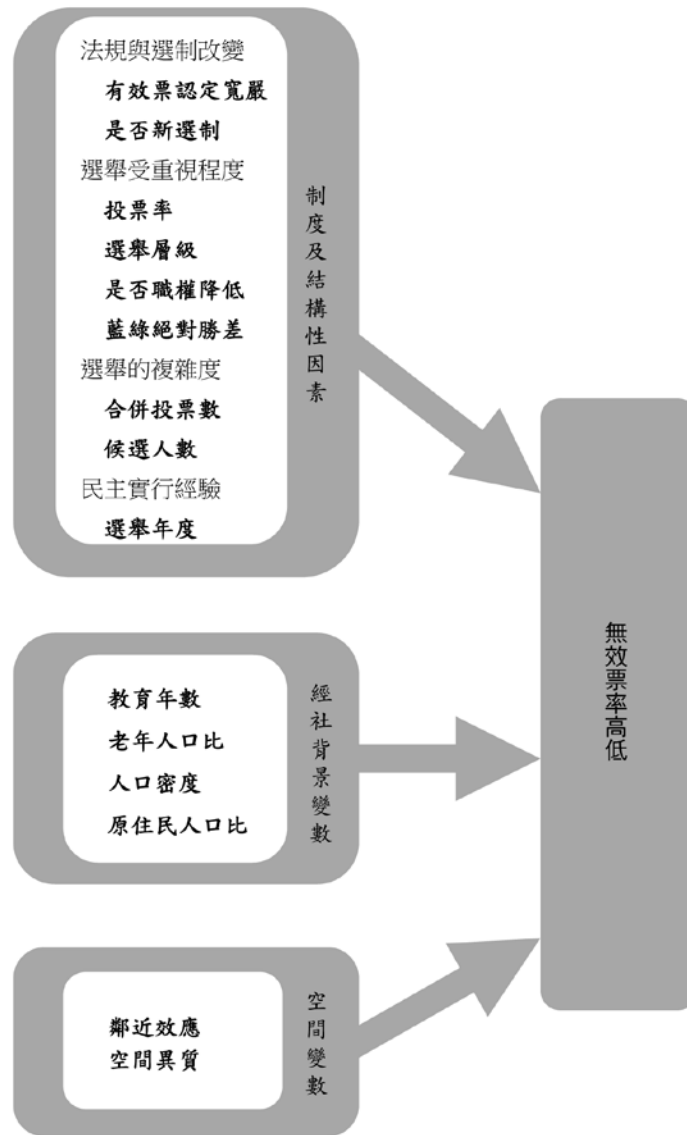
最後一個類型的變數是「空間變數」，用來探討鄰近效應或者空間異質是否存在，以及迴歸模型誤差的空間校正之用。

本文的研究假設，整理如下：

「選舉制度與結構性因素」方面：有效票認定越嚴、新選制實行、選舉越複雜（合併投票數越多、候選人越多）、投票率低、選舉層級越低（即地方選舉）、職權降低、選舉較不激烈（藍綠絕對勝差越大）、選舉年度越遠，則無效票率越高。

「經社背景變數」方面：教育水準越低（教育年數低）、都化市程度越低（人口密度低）、老年人口比、原住民人口比越高的地區，則無效票率越高。





資料來源：作者自行繪製。

圖 2 研究架構圖

## 二、研究方法

本文使用追蹤資料，共有 37 個選舉，7,518 個觀察值。主要的分析方法是將資料合併 (pooled) 進行迴歸分析，以探討自變數與依變數間的關係。首先必須說明的是，本研究以「鄉鎮」作為分析單位，基本上討論鄉鎮間的關係，符合某種區位特性的「鄉鎮」，其廢票率的表現會有某種傾向，而非說具有該屬性的「個人」投廢票的可能性較高或低。集體層次的關係，未必會等同於個人層次的關係，如果未詳細考察，輕易論斷，極可能會有區位（生態）誤謬的問題。

一般說來，個體層次的關係是集體層次關係的必要條件，也就是說，要先有個體層次的傾向，經過加總才會顯現集體層次的傾向；但並非充分條件，因為個體的關係極可能因加總之後，關係不見或減弱了，除非地區內同質，地區間異質，地區的變異才會保存。另外，也不能有「跨域流動」的現象，例如返鄉投票的人越多，個體與集體的關係可能會不一致，因為地區屬性量度的對象，可能是不同的一批人。

然而在理論建構時，卻無法那麼理想化，因為「個人」才是行為的主體，「地區」不會「行為」，所謂地區的「行為」其實是「個人」行為的匯總 (aggregate)，因此在思考「地區」間的關係時，免不了會參考「個人」層級的關係。但兩者仍不應混同，這是必須要注意的。

在計算工具上，本文採用 Elhorst(2010) 在 Matlab 上開發的空間追蹤分析模組，<sup>7</sup> 進行空間追蹤落遲模型 (SLM) 與空間追蹤誤差模型 (SEM) 的討論。空間迴歸與傳統迴歸 (最小平方法，OLS) 最大的差別在，前者假設誤差不是空間隨機 (spatial random)，亦即鄰近地區的誤差會彼此影響，而非獨立，違反傳統迴歸的基本假定。由於誤差並非彼此獨立，最小平方法估計會喪失有效性，導致自變數的檢定過度樂觀；同時，因為可能有遺漏變數存在，不偏性也難以維繫。

誤差是否空間隨機，一般用 Moran's I 檢定 (Anselin 1988)。由於空間分析方法在社會科學上的應用在台灣並不多見，作者在此略加說明，以幫助讀者瞭解。「空間隨機」是空間分析最基本的概念，它指變數值不受鄰近地區變數值的影響，亦即彼此獨立，與此相反的是「空間自相關」(spatial autocorrelation)，或稱「空間相依」(spatial dependence)，表示鄰近地區的變數值，彼此會相互影響。Moran's I 檢定的原理是，進行「各地區變數值」與「各地區的鄰近地區變數值」兩個變數的相關性討論，如果是「空間隨機」則相互獨立，相關係數應為零，「空間相依」則反之。

有兩個關鍵的問題必須先解決。第一，如何定義「鄰近地區」？第二，如何計算「鄰近地區變數值」的大小？最簡單的做法是給定每個地區一個「鄰近地區」的集合。而「鄰近地區變數值大小」則以「鄰近地區」集合內各元素的變數值，加權平均為之。如果要簡化問題，簡單平均即可。

判斷是否是「鄰近地區」，通常以距離遠近，或邊界是否相接壤為準則。本文以邊界有相接壤者定義為鄰近，並以簡單平均計算鄰居變數值大小。為計算方便，捨去金門與連江，而僅以台灣本島及澎湖等 358 個鄉鎮市區作為分析之範圍。<sup>8</sup> 如此，如圖 3 所示，各

<sup>7</sup> 請參考 Paul Elhorst 的網頁：<http://www.regroingen.nl/elhorst/software.shtml>，分析模組與範例可自由下載。

<sup>8</sup> 有些地區因沒有和其他地區有行政邊界的接壤，導致其在 GeoDa 軟體計算時出現無鄰居之狀

地區變數值標準化（平均值為零，單位為標準差）之後置於橫軸，縱軸為鄰近地區變數值（標準化之後）的大小。迴歸線的斜率即為 Moran's I 值。亦即，Moran's I 用來量度「地區」與「鄰近地區」變數值的相關程度。其值介於 +1 與 -1 之間，正表正的空間自相關，負表負的空間自相關。如果 Moran's I 為 0 則表示「空間隨機」，亦即無「空間自相關」。正的空間自相關習慣上有另一個說法「空間聚集」。Moran's I 為正，表示自己高的地方，鄰居亦高，此即「空間聚集」現象。

由此可知，Moran's I 與相關係數的概念極為類似，只不過它是用來探討「各地區的表现」與「其鄰近地區的表现」的相關性。其計算公式，如下：

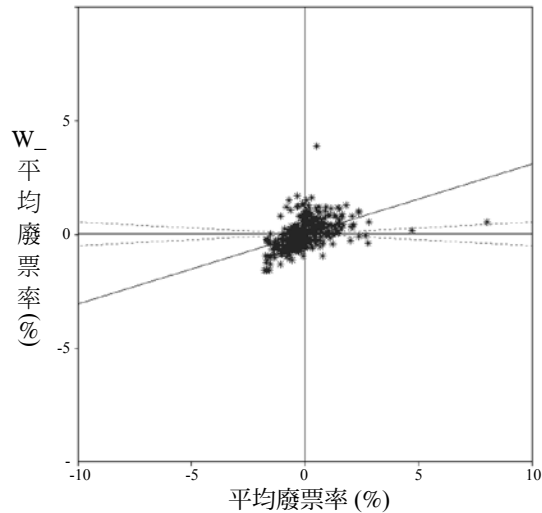
$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n W_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

其中， $n$  為觀察的地區數， $y_i$  和  $y_j$  為地區  $i$  和地區  $j$  的變數值，為該變數在所有地區中的平均值， $W_{ij}$  代表地區  $i$  和地區  $j$  的空間鄰近 (Spatial contiguity) 關係，如果地區  $i$  和地區  $j$  相鄰其值為 1，否則為 0 (Anselin 1988)。

圖 3 為歷次選舉的平均無效票率的 Moran's I 散布圖，圖中每個點代表一個地區，共有 358 個點，橫軸是標準化之後的變數，縱軸為相應的鄰近變數。通常橫軸變數用  $Y$  表示，縱軸變數用  $WY$  表示，<sup>9</sup>  $W$  稱為鄰近矩陣 (contiguity matrix)，散布圖的斜率即是 Moran's I 指數，其值為 0.3078。Moran's I 亦可進行檢定，Moran's I 的抽樣分布近似  $Z$  分布，因此由機率可判斷是否拒絕虛無假設（空間隨機）(Anselin 2003; 2005)。以本圖為例，Moran's I 為 0.3078，經過  $Z$  檢定，機率遠小於 0.05，達顯著水準，表示變數有空間自相關現象。

況，無鄰居時，鄰展的表现設為平均值。為避免此種狀況，這些「孤島」由作者自行給定鄰居。高雄縣旗津區，指定鄰居小港區、前鎮區；臺東縣蘭嶼鄉，指定鄰居綠島；而綠島之鄰居，則除蘭嶼外，還有臺東市；澎湖縣七美鄉，指定鄰居望安鄉；而望安鄉之鄰居則包括七美鄉與馬公市。

<sup>9</sup>  $Y$  為  $n \times 1$  矩陣， $W$  為  $n \times n$  矩陣，對角線為零，非對角線代表地區與地區的鄰近關係，如果是鄰近其值為 1，否則為零，通常  $W$  會經過列標準化，亦即每個列總和為 1。 $WY$  即為  $W$ （經過列標準化）與  $Y$  兩矩陣相乘，亦為  $n \times 1$  矩陣。



資料來源：中央選舉委員會（2009）選舉資料庫，作者自行計算。

說明：Moran's I=0.3078。

圖 3 平均無效票率 Moran's I 散布圖

上述之 Moran's I 關心的是整體 (Global) 的空間自相關趨勢。若我們關心的是某個地區與其鄰近地區之相關關係，則可選擇 Local Moran's I 指數，其定義為如下：

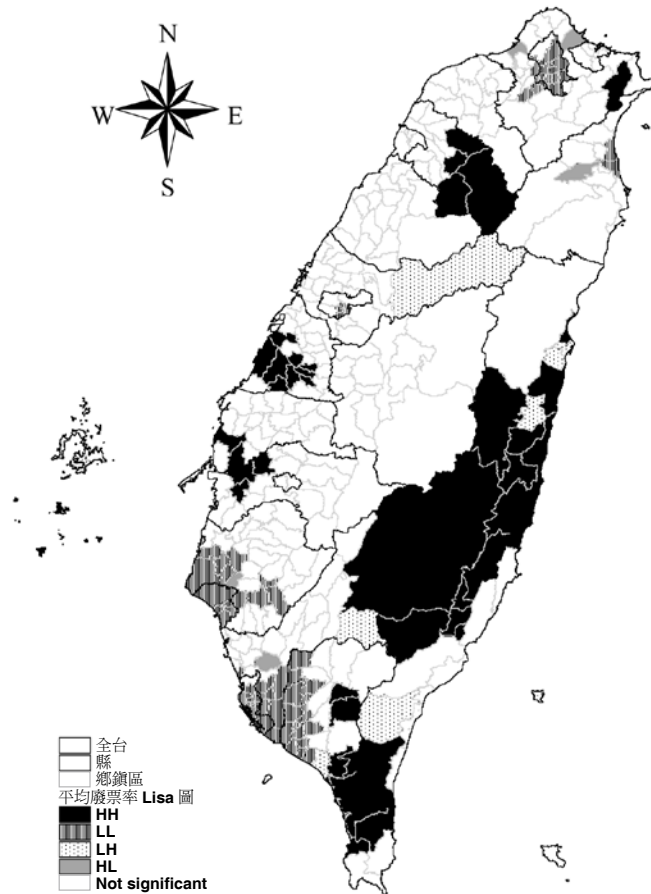
$$I_i = \frac{y_i - \bar{y}}{\sum_i (y_i - \bar{y})^2} \sum_j W_{ij} (y_j - \bar{y})$$

而 Local Moran's I 之值如果為正，代表該地區與其鄰近地區呈現正相關現象，負則為負相關現象。此外，各地區 Local Moran's I 值之平均，即等於 Global Moran's I 之值。

GeoDa 軟體之 LISA (Local Moran's spatial association) 分析，將 Local Moran's I 值達到顯著水準（機率值小於 0.05），依照其空間聚集之狀況，分為以下四種 (Anselin 2003; 2005)：

- 一、High-High (正正)：即自己高，鄰居也高的地區，亦稱為「熱區」，如圖 3 第一象限地區。
- 二、High-Low (正負)：即自己高，鄰居卻低的地區，亦稱為「冷區」，如圖 3 第四象限地區。
- 三、Low-Low (負負)：即自己低，鄰居也低的地區，如圖 3 第三象限地區。
- 四、Low-High (負正)：即自己低，鄰居卻高的地區，如圖 3 第二象限地區。

圖 4 為鄉鎮平均無效票率之 LISA 圖，可精確地看出北、高都會區是 Low-Low 之聚集狀況，而彰化與東南部地區，則為 High-High 之聚集情況。



資料來源：中央選舉委員會（2009）選舉資料庫，作者自行計算繪圖。

圖 4 平均無效票率的 LISA 圖

如果 OLS 迴歸的誤差無通過 Moran's I 檢定（表示誤差有空間自相關），就有進行空間迴歸的必要。本研究採取三種迴歸模型，除了傳統的 OLS 迴歸外，還有 SLM 與 SEM，略述如下：

#### （一）Pooled OLS（最小平方法迴歸）

Pooled OLS 迴歸模型，如下所示：

$$Y_{it} = X_{it}\beta + \varepsilon_{it}$$

Y 是依變數，X 為傳統自變數， $\varepsilon$  是誤差，i 代表地區，t 代表時間。 $\beta$  為待估計的參數，不因地區及時間而異，而誤差  $\varepsilon$  是 i.i.d.，即無自相關且變異齊一。

通常 Pooled OLS 會控制固定效應，以排除異質造成的影響。然而本研究有許多僅和時間有關，與地區無關之變數（例如選舉層級、新選制、職權降低、選舉年度等），因此無法控制時間固定效應；同樣地，亦無法控制地區固定效應，因經社背景變數之時間變異

太小 (slow changing variables)。

## (二) Pooled SLM (空間落遲模型)

如果 Pooled OLS 的誤差有空間自相關，而推測其產生原因是鄰近效應造成的，則可考慮空間落遲模型 (spatial lag model)，如下所示：

$$Y_{it} = \rho WY_{it} + X_{it}\beta + \varepsilon_{it}$$

$Y$  是依變數， $X$  為傳統自變數， $\varepsilon$  是誤差，為 i.i.d.， $\beta$  為待估計的參數。比較特殊的是多了一個新的自變數  $WY$ ， $\rho$  為其係數。 $W$  稱為「鄰近矩陣」，為一右上左下對稱的方形矩陣，欄與列數都等於地區個數（例如本研究共有 385 個鄉鎮），如果兩地區具有鄰近關係，其相應的元素值為 1，否則為 0，對角線也為 0（自己和自己不具鄰近關係）。不過此處之鄰近矩陣經過列標準化，亦即列總合為 1。 $WY$  其實就是鄰近地區的平均值，為一欄矩陣。空間落遲模型因為自變數多了一個空間變數 ( $WY$ )，有內生變數的問題，無法用傳統的最小平方法估計，否則會喪失不偏性與有效性，一般用最大概似法估計。

## (三) Pooled SEM (空間誤差模型)

如果空間變數不是出現在等號右邊的自變數，而是出現在誤差裡，則成了「空間誤差模型」(spatial error model)，如下式所示。基本上本模型是針對誤差進行空間校正，校正後的殘差 ( $u$ ) 會趨近空間隨機。

$$Y_{it} = X_{it}\beta + \varepsilon_{it}$$

$$\varepsilon_{it} = \lambda W\varepsilon_{it} + u_{it}$$

誤差  $\varepsilon$  不是 i.i.d.，加入空間變數 ( $W\varepsilon$ ) 之後，誤差  $u$  為 i.i.d.。同樣地，空間誤差模型亦無法用最小平方法估計，否則會喪失有效性。

另外，在進行空間追蹤資料分析時，比較特別的是每個橫切面的觀察個數必須一致 (balanced)，因此，本研究將直轄市長、市議員選舉，併入與其同時舉行之省長、省議員選舉或與其舉行時間點最接近之縣市長、縣市議員選舉，如此橫切面之觀察值數目方能相同。此外，鄰近矩陣  $W$  視為恆定，亦即鄰近關係不會隨著不同時間點而改變。空間誤差模型主要用來做誤差的空間校正，讓估計值具有有效性，但它也有另一個意義：代表有遺漏變數存在，而遺漏變數剛好具備空間自相關的特性，所以可以用空間誤差模型來校正。

另一個本文觸及的空間分析概念是「空間異質」，它代表變數關係的空間不穩定狀態，亦即變數大小或變數關係會隨區域而異。造成空間異質的原因，大多是遺漏自變數使然，正因為有許多可能的變數難以量化（例如歷史傳統、文化特殊性等），造成空間異質，顯現在誤差的異常上。

### 三、資料與變數

本研究考察 1992 到 2008 年間的 37 個選舉，以鄉鎮為分析單位（台灣地區，金馬不計，共有 358 個鄉鎮），<sup>10</sup> 進行空間追蹤分析 (spatial panel data analysis)。

#### （一）資料

資料有選舉資料及經社背景資料，其來源分別說明如下：

##### 1. 選舉資料

來自中央選舉委員會之選舉資料庫，共蒐錄 37 個選舉之資料。包括：

- （1）總統選舉 4 個：第 9 至第 12 屆總統選舉，舉行年度分別為 1996、2000、2004 及 2008 年。
- （2）立法委員選舉 7 個：第 2 至 7 屆立委選舉，舉行年度分別為 1992、1995、1998、2001、2004 及 2008 年。其中，在第 7 屆立法委員選舉部分，因為選制改為單一選舉兩票制之關係，而區分為區域與政黨兩者。
- （3）國民大會代表選舉 2 個：第 3 屆國民大會代表選舉，及第 4 屆任務型國代選舉，舉行年度分別為 1996 及 2005 年。
- （4）省長及省議員選舉 2 個：第 1 屆省長及省議員選舉，舉行年度為 1994 年。
- （5）直轄市長及市議員選舉 16 個：第 1 至第 4 屆北、高市長選舉、第 7 至第 10 屆台北市議員選舉，及第 4 至第 7 屆高雄市議員選舉，舉行年度分別為 1994、1998、2002 及 2006 年。
- （6）縣市長選舉 3 個：第 13 至 15 屆縣市長選舉，舉行年度分別為 1997、2001 及 2005 年。
- （7）縣市議員選舉 3 個：包括第 14 至 16 屆縣市議員選舉，舉行年度分別為 1998、2002、2005 年。

從中央選舉委員會公布的選舉資料，可計算無效票率、投票率、競爭激烈程度、選舉層級、投票複雜度等制度性變數。

<sup>10</sup> 本研究考察了 1992 到 2008 年間 37 個選舉，有些選舉只在部分區域舉行，並未擴及全國，例如直轄市長及議員、省長及議員、縣市長及議員，所以不同選舉必須配對歸併，才能湊齊每個鄉鎮都有變數值，以方便進行空間分析。因此，1994 年省長與 1994 年直轄市長；1994 年省議員與 1994 年直轄市議員；1998、2002、2006 年直轄市議員與 1998、2002、2005 年縣市議員；1998、2002、2006 直轄市市長與 1997、2001、2005 年縣市長，配對歸併之後，共計有 7,518 個觀察值 (21\*358)。

## 2. 經社背景資料

本文研究的經社背景變數，包括教育年數、老年人口比來自 1990 與 2000 年戶口普查，中間各年度變數值以內插法求得，2000 年以後之資料則與 2000 年同。人口密度資料取自於 1993 至 2006 年中華民國臺閩地區人口統計資料（2005 年後改稱「中華民國人口統計年刊」）。受限於作者可以取得的資料，1992 年之數值與 1993 年相同；而 2006 年以後之值則與 2006 年相同。事實上，以上三個經社背景變數，台灣各鄉鎮的時間差異並不大。

### （二）變數

依變數為無效票率，其定義為無效票數除以投票數，再乘以 100%。自變數分為制度與結構變數以及經社背景變數，分別說明之。

#### 1. 制度與結構變項

##### （1）法規與選制改變

- a. 有效票認定：指有效票認定標準之寬嚴程度，本文將無效票認定標準之寬嚴分為三個等級。2004 年第 11 屆總統選舉時之無效票認定標準最為嚴格（僅能圈選在圈選欄內），第 2 至 5 屆立法委員選舉次之（僅能圈選在姓之上端），而其餘選舉之無效票認定標準則大致相同（圈選在可辨識所投對象之處皆可），較為寬鬆。2004 年總統選舉最嚴，給定為 3；第 2 至 5 屆立委選舉次之，給定 2，其餘選舉則給定 1。無效票認定嚴格，對於那些已有特殊投票習慣的選民來說，投無效票之機率就會增加。
- b. 新選制：在本文研究之時間範圍內，有兩次選制之變革，一為 2005 年任務型國代選舉，自原來之 SNTV 改為政黨比例代表制；另一為第 7 屆立委選舉，自 SNTV 制改為單一選區兩票制。本文將以上二者給定其值 1，而其餘則給定 0。新選制之實施，使得選民在不熟悉新制度之情形下，投廢票之機率升高。

##### （2）選舉的複雜度

- a. 合併投票數：指選民在該次投票時所領到之票數，它可能隨著地區之不同而有差異。例如：以縣市長、縣市議員及鄉鎮市長三合一選舉為例，對於未選舉鄉鎮市長之省轄市而言，僅領到二張選票，而非三張。除此之外，2004 及 2008 年總統選舉，2008 年立委選舉，同時舉行之全國性公民投票（第一至六案）亦計入領票數。合併投票數可以衡量投票之複雜程度，選民所領票數越多，越容易混淆而投成無效票。
- b. 候選人數：亦即選票之長度。候選人數越多，選票當然也會越長。以鄉鎮市區為單位作分析時，有些議員型態的選舉（例如：立委選舉或縣市議員選舉等），會把同



一鄉鎮切割，分屬兩個不同的選區。碰到此種情形，本文以兩個選區候選人數的平均作為該鄉鎮的候選人數。候選人數越多，選民混淆的可能性增加，亦有可能因為選票太長，容易摺疊時章漬反印，而成為無效票。

### (3) 選舉受重視程度

- a. 投票率：即投票數除以選舉人數，再乘上 100%。而投票率越高，代表選舉越受到重視，理論上選民應會越小心，避免出錯成為無效票。
- b. 選舉層級：本文將所欲討論之選舉類型分為 6 個層級，並給定其值 1 至 6，數值越大層級越高，反之，數值越小，層級越低。總統為國家元首，其選舉層級最高，最受到重視。其餘公職人員則依據 1994 年 7 月 15 日修正後《公職人員選舉罷免法》第 2 條之規定，可分為中央公職人員（包括：立委及國代）及地方公職人員（包括：省（市）議員、縣（市）議員、鄉（鎮、市）民代表、省（市）長、縣（市）長、鄉（鎮、市）長、村、里長），依序遞減。立委及國代選舉屬同一層級，僅次於總統選舉。地方公職人員選舉之層級則依照《地方制度法》第 3 條之規定，其中，省及直轄市同級，縣市次之，鄉鎮市最低。層級越高的選舉，較受重視，無效票率可能會較低。各選舉所屬之層級，如下所示：
  - 層級一（最低）：縣市議員選舉。
  - 層級二：縣市長選舉。
  - 層級三：省議員及直轄市議員選舉。
  - 層級四：省長及直轄市長選舉。
  - 層級五：立法委員及國民大會代表選舉。
  - 層級六（最高）：總統選舉。
- c. 職權降低：指被選對象之職權較上屆降低之選舉，主要是針對第 3 屆國大代表選舉及 2005 年任務型國代選舉，職權皆明顯降低，因此處理方式為將上述二個選舉設為虛擬變數，給定值為 1，其餘選舉則給定為 0。職權降低代表被選職位已不如以前重要，因此選民可能會不重視，而增加無效票之機率。<sup>11</sup>
- d. 藍綠絕對勝差：藍營指國民黨、新黨及親民黨；綠營指民進黨及台灣團結聯盟，藍、綠營得票率差距之絕對值越大，代表選情越不激烈；反之，勝差越小，代表選

<sup>11</sup> 自第 3 屆國民大會起，由於總統已改為全民直選產生，因此國大代表不再有選舉正副總統之重要權力，此時國大之職權為補選副總統、提正副總統罷免案、議決監察院提出之總統副總統彈劾案、修改憲法、複決立法院所提之憲法修正案及對總統所提名之任命人員行使同意權。2005 年所選舉出來之任務型國代，其職權又比第 3 屆國代又降低許多。此時之國大已虛級化，僅有被動複決立法院所提之憲法修正案、立法院所提之領土變更案及議決立法院提出之總統副總統彈劾案等權力。

情越激烈。藍綠差距小之選舉推測應較受選民重視，投無效票之機率應越低。部分選舉僅有單一陣營推派代表，究其原因，係禮讓友好的第三黨派始然。碰到這些情況，本文將第三黨派候選人視為未提名陣營之代表，避免數據嚴重偏差。<sup>12</sup>

#### (4) 民主實行經驗

- a. 選舉年度：即該選舉舉行之年度，以民國紀元。第 11 屆總統選舉為例，民國 93 年舉行，給定其值 93，其餘類推。將選舉年度納入考慮，主要是看無效票率是否隨著時間之增加而有降低之趨勢，因為隨著民主實踐時間的拉長，民衆對於正確投票方式的理解勢必會逐漸改善，投錯票的情形勢必減少。

#### 2. 社經解釋變數

- (1) 教育年數：即該地區人口的平均受教育年數，受教育年數的計算為：國小假定為 6，國中為 9、高中 12，大學 16，研究所以上為 18。而一個地區之教育年數越高，其選民素質應越佳，投票失誤的可能性降低。
- (2) 老年人口比：即該地區 65 歲以上之人口，佔該地區總人口之比例，再乘上 100%。而老年人口可能因辨識力較差，使得其投無效票之機率較高。
- (3) 人口密度：即該地區總人口數（以「十萬人」為單位，本文之總人口數乃指戶口數）除以該地區之土地面積（平方公里）。而人口密度是最容易代表「都市化」程度之變數，都市地區因為發展較好、資源多，所以人口聚集。一個地區之都市化程度越高，代表接受資訊之能力越強，應越能瞭解選舉制度及投票規則，而使得其無效票率較低。
- (4) 原住民人口比例：原住民人口數除以該地區總人口，再乘以一百，成為百分比。資料來自 2003 及 2007 年客委會的客家人口基本調查（單一認定），2003 年之前的數據以 2003 年替代，2007 年之後者以 2007 年替代，其餘年度以內插法計算各年數據。

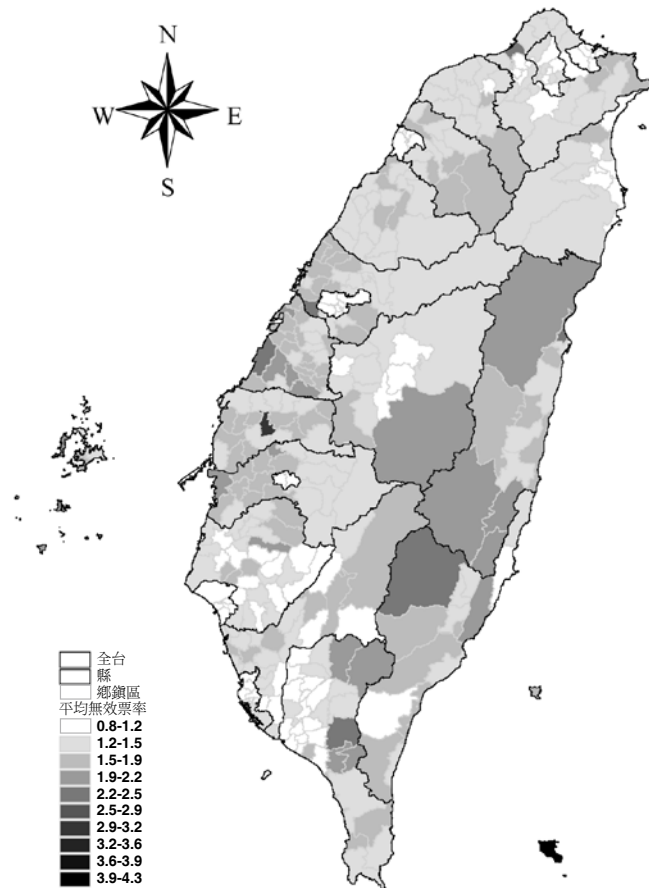
<sup>12</sup> 2004 年立委選舉，澎湖縣之藍營以無黨團結聯盟（以下稱無盟）之林炳坤取代；在 2008 年立委選舉，臺中縣第二選區，藍營以無盟之顏清標取代；臺南縣第二選區由無盟之李和順取代；屏東縣第二選區以無黨籍之蔡豪取代。在 1993 年縣市長選舉，臺東縣之綠營以無黨之陳益南取代；嘉義市之綠營以無黨籍之張文英取代。在 2005 年縣市長選舉，臺東縣之綠營以無黨籍之劉耀豪取代。

## 肆、實證分析

### 一、無效票率並非空間隨機

首先處理的第一個問題是，無效票率是否空間隨機？以鄉鎮市區為單位，將所有選舉的無效票率平均，如下圖 5 所示，無效票率高的地區，多集中在中部和東南部地區，而北、高、宜蘭等地則偏低，都會區也有偏低的趨勢。另外也可發現平均無效票率並非空間隨機，Moran's I 高達 0.3078，達顯著水準。Moran's I 值為正，代表平均無效票率有正向空間自相關，亦即鄰近地區之無效票率高，自己的無效票率亦會高，反之亦然。

接下來要問的問題是：為什麼它不是「空間隨機」？難道不同的人犯錯的機率不一樣？答案很可能是肯定的。那麼誰的犯錯機率比較大？或說影響無效票率高低的因素是什麼？下節，作者將先討論制度性因素的影響，接著是經社背景變數對無效票率的影響，最後則是討論空間因素的影響。



資料來源：中央選舉委員會（2009）選舉資料庫，作者自行計算。

圖 5 所有選舉平均無效票率主題圖

## 二、無效票率之迴歸分析

依變數為無效票率，自變數則包括制度與結構性因素，以及經社背景變數。分別跑了三個模型：Pooled OLS、Pooled SLM 及 Pooled SEM，結果如表 1 所示。

表 1 第二欄為 Pooled OLS 模型的結果，總共約可解釋 36% 的變異。變數的表現基本上與研究的假設符合，即有效票認定越嚴格、新選制、選舉層級越低、職權降低、藍綠勝差越大、投票的張數越多、候選人數越多（選票越長）、選舉年度越早、教育年數越低、老年人口比越高、人口密度越低、原住民人口比越高的地區，無效票率也越高。唯一例外的是投票率，其他條件一樣的情況下，投票率越高之地區，無效票率也越高。或許我們可以先從個體層次思考這個問題：投票率越高代表一些平常不積極參與投票的選民也出來投票，因為其對投票過程不太熟悉，出現投無效票的機會反而增加。以此解釋 2005 年的任務型國代選舉為代表就很有趣，一方面任務型國代職權降低，人們比較不重視，廢票率有增加的可能；但是，另一方面，投票率低（只有 23.4%），政治意識強烈的才出來投票，最後的結果，廢票率只有 0.96%，顯然後者的影響蓋過前者。這些會去投票的選民，都是行使投票權最積極的選民，對選舉有一定的重視與瞭解，選票無效的可能性自然比較低。我們的資料是集體層次的資料，個體層次與關係未必能完全等同於集體層次，然而個體的關係還是值得參考。

值得注意的是原住民人口比越高，廢票率亦越高，這是在控制所有制度、法規、經社背景變數之後的結果。亦即其他條件都一樣的情況下，原住民多的地方，廢票率就高，何以如此？原因可能是因為，這些地區普遍對政治冷漠，甚至反感，所以廢票率特別高，不過也可能是因為，原住民部落資訊較封閉，對於投票程序資訊較陌生，造成廢票率較高。不過不管是哪一種原因，都值得政府關注，是不是長期政策失衡造成這個結果？

第三欄為 Pooled SLM 模型的結果，大致與 Pooled OLS 模型相同，只是控制空間變數 WY 後，「人口密度」與「藍綠絕對勝差」變成不顯著，表示這兩個變數與空間變數相重疊，且 WY 係數高達 0.63，代表鄰近地區間有正向的關係存在，且影響力相當強。周圍地區的無效票率每增加一個百分點，我就增加 0.63 個百分點！何以如此？最直接的解釋是「鄰近效應」，亦即鄰近地區的無效票率彼此正向影響。不過這有點難以理解，因為我們很難說無效票會相互「傳染」、「模仿」、「學習」或「外溢」，這在理論上難以解釋，過去從來沒有人這樣子探討；但也不排除此種可能性，迴歸結果就是「鄰近效應」的明證；另一種可能則是「空間異質」，我們也不能排除此種可能。空間異質的成因是某些不清楚的遺漏變數。也就是說，我們控制的變數並不完全，因為有遺漏變數存在，造成誤差有正向的空間自相關，使得 SLM 模型的空間變數顯著。

第四欄為 Pooled SEM 模型的結果，與 Pooled OLS 模型差異不大。空間變數  $W\varepsilon$  係數亦相當高，達 0.66。這呼應我們的前一個猜測，是有遺漏變數存在，該變數存在某種空間特性，因此 SEM 模型顯著。不過，這個遺漏變數是什麼？從迴歸式中看不出來。只知道，它與「選票長度」重疊，所以控制了它之後，選票長度就不顯著了。其他自變數的顯著性與方向都不變，表示關係相當穩定。

表 1 無效票率之追蹤分析結果

	Pooled OLS	Pooled SLM	Pooled SEM
常數項	1.882 <sup>***</sup>	0.511 <sup>*</sup>	3.105 <sup>***</sup>
<b>制度變數</b>			
法規與選制改變			
無效票認定標準	0.571 <sup>***</sup>	0.212 <sup>***</sup>	0.579 <sup>***</sup>
新選制	0.706 <sup>***</sup>	0.318 <sup>***</sup>	0.628 <sup>***</sup>
受重視程度			
投票率 (%)	0.016 <sup>***</sup>	0.008 <sup>***</sup>	0.007 <sup>***</sup>
選舉層級	-0.163 <sup>***</sup>	-0.064 <sup>***</sup>	-0.157 <sup>***</sup>
職權降低	0.986 <sup>***</sup>	0.391 <sup>***</sup>	0.825 <sup>***</sup>
藍綠絕對勝差 (%)	0.002 <sup>***</sup>	0.001	0.002 <sup>***</sup>
選舉複雜度			
投票張數	0.678 <sup>***</sup>	0.247 <sup>***</sup>	0.621 <sup>***</sup>
選票長度	0.019 <sup>***</sup>	0.008 <sup>***</sup>	0.004
民主經驗			
選舉年度	-0.043 <sup>***</sup>	-0.015 <sup>**</sup>	-0.048 <sup>***</sup>
<b>社經變數</b>			
教育年數	-0.071 <sup>***</sup>	-0.051 <sup>***</sup>	-0.065 <sup>***</sup>
老年人口比	0.029 <sup>***</sup>	0.020 <sup>***</sup>	0.035 <sup>***</sup>
人口密度	-1.156 <sup>***</sup>	0.310	-0.987 <sup>***</sup>
原住民人口比	0.008 <sup>***</sup>	0.005 <sup>***</sup>	0.009 <sup>***</sup>
WY/W $\varepsilon$		0.626 <sup>***</sup>	0.663 <sup>***</sup>
N	7,518	7,518	7,518
R <sup>2</sup>	0.356	0.570	0.342

資料來源：中央選舉委員會（2009）選舉資料庫，作者自行計算。

說明：顯著水準：<sup>\*</sup> $p < 0.1$ ；<sup>\*\*</sup> $p < 0.05$ ；<sup>\*\*\*</sup> $p < 0.01$ 。

從前文圖 1 中可以看出，每一屆選舉中的廢票率，雖然看起來不高，但以選票張數這個變項為例，以表 1 的 Pooled OLS 模型的估計結果為例，每增加一張選票，廢票率會提高 0.68 個百分點，所以舉例來說，在一次選舉中若增加公投票兩張，廢票率會提高 1.36

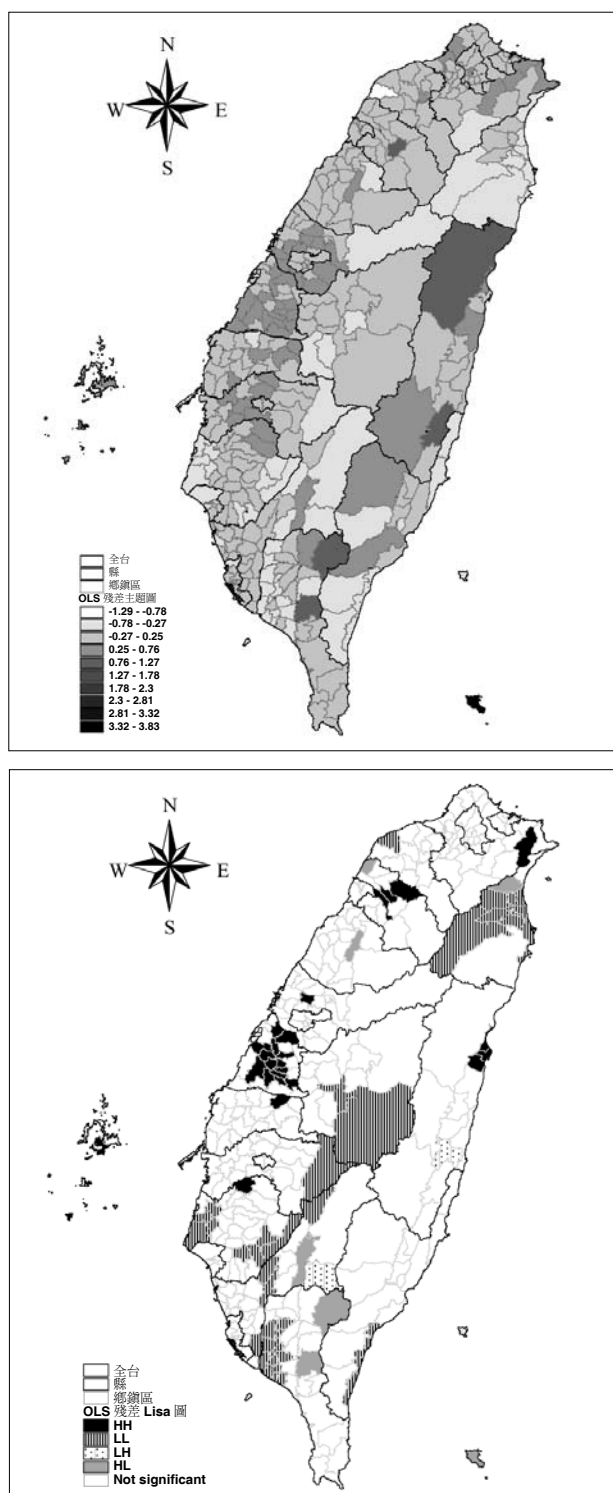
個百分點，這對於每個支持率百分點都錙銖必較的選舉來說，當然很重要，代表有一部分的民意無法正確表達出來，如能減少廢票比率，當更能體現選舉中的民意。

### 三、空間自相關與空間異質問題

從迴歸報表看來，絕大多數的變數的表現均合於預期。我們發現制度與法規的因素很顯著的影響廢票的發生，此外，我們也發現經社條件越劣勢的地方，廢票率越高。也就是說「犯錯」的機會並不均等，特定制度下與某些地方比較容易犯錯，不管此種犯錯是故意的，或者無意的，這都值得政府注意。

另一個待解的問題是，為什麼廢票率有空間自相關？其他條件都一樣的情況下，廢票率高的地方仍會聚在一起。聚在哪裡呢？如圖 6 上圖，平均 37 次選舉的 OLS 殘差值，發現誤差有空間聚集現象，Moran's I 值是 0.3259 達顯著水準。圖 6 下圖為 LISA Cluster 圖，■部分為「熱區」(High-High)，為無效票率低估，且鄰近關係達正向顯著者；■為「冷區」(Low-Low)，為無效票高估，且鄰近關係正向顯著者。熱區在北台灣、中台灣的幾個分散的聚點，冷區則聯成一氣，主要在南台灣，台南以及高雄、屏東，部分平原、山區交界處。與圖 5 相較，山區無效票率較高的地區，圖 6 的殘差圖中已被解釋掉，主要是我們放了一個自變數「原住民人口比」。不管是熱區或冷區，代表的都是與鄰近地區的正向關係，如前所述，這有可能是鄰近效應的影響，也可能是某個（些）遺漏變數未控制住的結果。由上可知，「鄰近效應」的認定不是那麼容易，就統計檢定言，只能判斷是否「空間自相關」。而造成空間自相關的原因，可能是「鄰近效應」也可能是「遺漏變數」或「空間異質」。除非我們有把握，所有該控制的變數都控制住了，此時 SLM 模型顯著，才可以大膽的說「的確！有鄰近效應存在」。

SEM 模型是顯著的，暗示有「遺漏變數」存在，圖 6 給我們一些線索，顯示這個（些）遺漏變數的地域分布。為什麼「熱區」或「冷區」會這樣分布？為什麼會集中在這裡？則是待解的謎團，未來可以進一步探索。

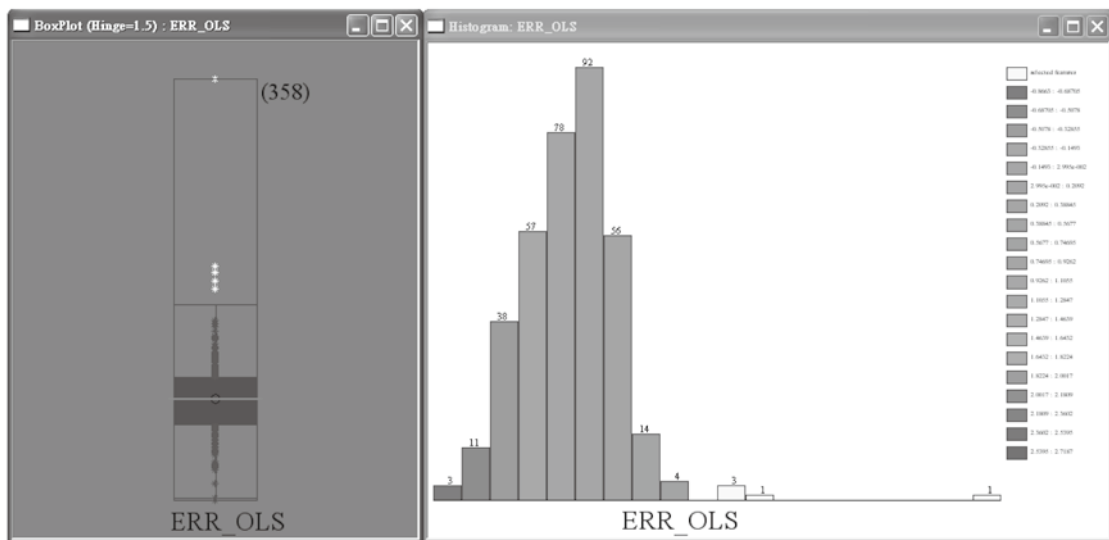


資料來源：中央選舉委員會（2009）選舉資料庫，作者自行計算。

說明：上圖為追蹤分析迴歸殘差平均值主題圖，下圖為 LISA Cluster 圖。

圖 6 殘差平均值的主題圖與 LISA Cluster 圖

圖 7 是 OLS 迴歸各地區殘差平均值的盒狀圖與直方圖，目的是想看殘差的分布。我們發現有五個地區在兩倍標準差以外，也就是說這五個地區是被我們（迴歸方程式）嚴重低估的，亦即他們的無效票率高於我們預期。這五個地區是：台東縣蘭嶼、台北縣淡水鎮、台東縣海端、台北市信義區、花蓮縣玉里、台北縣瑞芳。明顯地，它可以區分為兩個類型，台北市信義區、台北縣淡水鎮屬都會類型，蘭嶼、海端、玉里、瑞芳屬偏遠地區，我們沒有足夠資訊說明，這五個地區何以如此特殊。蘭嶼是其中比較可以解釋的，有可能是爭議多年的核廢料問題，導致對政府（不管國民黨或民進黨執政）政治冷漠或不滿，反應在廢票率上，即使控制原住民人口比，它的獨特性還是很明顯。就空間分析的觀點，任何地區，只要表現獨特，背後都是有原因的，除非是資料有誤，不然就表示有隱藏變數，尚未被我們察覺。真正的原因，恐怕須深入當地，實地調查，才能一窺究竟。



資料來源：中央選舉委員會（2009）選舉資料庫，作者自行計算。

說明：圖左為追蹤分析迴歸殘差平均值盒狀圖，圖右為直方圖。淺色的點粒與圖塊表示上歧異值，共有五個地區。

圖 7 殘差平均值的盒狀圖與直方圖



## 伍、結論與發現

本文考察台灣 37 個選舉，358 個鄉鎮的無效票率，共有 7,518 個觀察值。追蹤分析的結果發現，無效票率並非空間隨機，它是可以被解釋的。就制度與法規因素而言，有效票認定標準越嚴格、選舉重要性越低、選舉層級越低、選舉複雜度越高、選舉年度越前，則廢票率越高，基本上合乎我們的假設。經社背景變數也有影響，顯示不同的人投無效票的機會不相等，教育水準越低、老年人口比例越高，人口密度越低、原住民人口比越高的地區，廢票率也越高。如果民主選舉的信念是每人一票、每票等值，每個人都有同樣的機會透過選舉表達政治意志，則「犯錯」機率不等的現象就值得注意了。因為部分人投廢票的機率相對較高，形同是一種歧視，他們的意見被忽略了，卻不自知。特別是原住民，在其他經社背景都一樣的前提下，原住民仍然有較高的機會投廢票，此現象值得我們重視。

另外，本文雖然探討制度、法規、經社背景等變數，但是未處理組織動員可能的影響，例如地方派系介入選舉，是否有可能動員選民投廢票？選監票人員，廢票認定的地方性差異，也不在討論之列。此外，有意的廢票與無意的廢票，我們也無法區隔，只能放在一起討論，例如廢票運動的確切影響。這些都是研究的限制；不過，從迴歸結果看來，制度與社經變項的影響是很明確的，與我們的預期也大致吻合。

如果要減少廢票的發生，在可能的範圍內，首先要增加制度變革的宣導與減少選舉的複雜度，不論是採行新的選舉制度，或是有效票認定標準趨嚴，都可能會增加廢票率；而多種選舉合併舉行，選舉配合公投，選票太長（候選人太多）也是一樣（當然選舉人數這部分無法限制）。政府應加強宣導，避免選民在不知情下，喪失投票權。此外，地方層級的選舉比中央層級的選舉廢票率高，反應民衆對於地方層級選舉重視程度較低，如何能讓民衆能更加重視地方政治，也值得討論。

另外是「鄰近效應」，亦即廢票是否會「傳染」的問題。我們的研究發現，控制了所有變數之後，迴歸殘差仍然具有空間自相關，SLM 與 SEM 模型均顯著。嚴格地說，我們無法判斷殘差具「空間自相關」，是否就是「鄰近效應」的結果。因為造成殘差空間自相關的原因，除了「鄰近效應」之外，也可能是「遺漏變數」或「空間異質」。除非我們有把握，已經控制了所有該控制的變數，此時 SLM 模型顯著，才能說「的確！有鄰近效應存在」；否則，只能說不排除此種可能，或懷疑。

不過有一點倒是可以肯定，廢票率高低顯然存在「空間異質」現象。有些地區的廢票率特別高，例如：蘭嶼、海端、玉里、瑞芳、台北市信義區、淡水鎮等地區，即使控制了所有自變數（含原住民人口比）之後，仍然嚴重低估，顯示有未被發現的隱藏因素存在，而這個隱藏因素可能不是單一的，因為這幾個歧異的地區，彼此之間也很異質，有的是高

度都市化地區，有的是偏遠山區，其真正原因，可能要從當地特殊的歷史經驗去理解。總之，本文發現，廢票率的高低還是可以被解釋的，選舉研究，除了政黨競爭、投票傾向外，廢票率也是個有趣且值得研究的議題。

\* \* \*

投稿日期：2011.12.06；修改日期：2012.02.22；接受日期：2012.06.15

## 參考文獻

### I. 中文部分

- 中央選舉委員會，2009，〈歷屆公職選舉資料〉，中選會選舉資料庫網站：  
<http://210.69.23.140/cehead.asp>，檢索日期：2009年5月1日。  
(Central Election Commission. 2009. "Voting Results of Each Election." Cental Election Commission Database. <http://210.69.23.140/cehead.asp> [accessed May 1, 2009].)
- 台灣選舉與民主化調查，2008，〈2008年立法委員選舉電訪變數清單〉，台灣選舉與民主化調查網站：[http://www.tedsnet.org/cubekm1/front/bin/ptdetail.phtml?Part=Query2008LT\\_R&Category=124](http://www.tedsnet.org/cubekm1/front/bin/ptdetail.phtml?Part=Query2008LT_R&Category=124)，檢索日期：2009年12月25日。  
(Taiwan's Election and Democratization Study. 2008. "The List of Variables in Telephone Polling." Taiwan's Election and Democratization Study Website. [http://www.tedsnet.org/cubekm1/front/bin/ptdetail.phtml?Part=Query2008LT\\_R&Category=124](http://www.tedsnet.org/cubekm1/front/bin/ptdetail.phtml?Part=Query2008LT_R&Category=124) [accessed December 25, 2009].)
- 林修卉，2004，〈無效票饅頭 中看不重吃〉，新台灣新聞週刊網站：<http://www.newtaiwan.com.tw/bulletinview.jsp?bulletinid=17471>，檢索日期：2009年6月30日。  
(Lin, Shiou-huei. 2004. "Invalid-Votes Burns Look Nice but Not Taste Nice." New Taiwan Weekly. <http://www.newtaiwan.com.tw/bulletinview.jsp?bulletinid=17471> [accessed June 30, 2009].)
- 法務部，1988a，〈立法院立法委員選舉罷免法施行條例〉，全國法規資料庫：<http://lis.ly.gov.tw/ttscgi/lgimg?@923102;0197;0374>，檢索日期：2009年12月25日。  
(Ministry of Justice. 1988a. "Enforcement Rules of Legislator's Election and Recall Act." Laws & Regulations Database. <http://lis.ly.gov.tw/ttscgi/lgimg?@923102;0197;0374> [accessed December 25, 2009].)
- ，1988b，〈國民大會代表選舉罷免法施行條例〉，全國法規資料庫：<http://lis.ly.gov.tw/ttscgi/lgimg?@923102;0197;0374>，檢索日期：2009年12月25日。  
(-----, 1988b. "Enforcement Rules of Election and Recall Act of Representative of National Assembly." Laws & Regulations Database. <http://lis.ly.gov.tw/ttscgi/lgimg?@923102;0197;0374> [accessed December 25, 2009].)
- ，2003，〈總統副總統選舉罷免法〉，全國法規資料庫：<http://law.moj.gov.tw/LawClass/LawContent.aspx?PCODE=D0020053>，檢索日期：2009年12月25日。

(-----, 2003. "Presidential and Vice Presidential Election and Recall Act." Laws & Regulations Database. <http://law.moj.gov.tw/LawClass/LawContent.aspx?PCODE=D0020053> [accessed December 25, 2009].)

鄧志松，2008，〈無效票率與民主深化：台灣 358 鄉鎮無效票率的空間探索分析〉，民主深化與國家發展學術研討會，5 月 16-17 日，台北：政治大學。

(Teng, Chih-sung. 2008. "Invalid Votes and Democratic Strengthening: A Exploratory Spatial Data Analysis of Invalid Votes in Taiwan." Presented at the Conference on Democratic Strengthening and National Development, Taipei.)

駱明慶，2006，〈無效票率哪裡來？無效票定義範圍擴大對 2004 年總統選舉的影響〉，《人文及社會科學集刊》，18(4): 639-669。

(Luoh, Ming-ching. 2006. "The Effect of Broader Definition of Invalid Votes on 2004 Presidential Election in Taiwan." *Journal of Social Sciences and Philosophy* 18(4): 639-669.)

## II. 外文部分

Elhorst, J. Paul. 2010. "Spatial Panel Data Models." <http://www.springerlink.com/content/u8086626076458v0/> (accessed December 25, 2009).

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Netherlands: Kluwer Academic Publishers.

-----, 2003. *GeoDa 0.9 User's Guide*. USA: Center for Spatially Intergrated Social Science.

-----, 2005. "Exploring Spatial Data with GeoDa<sup>TM</sup>: A Workbook." <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.146.4634&rep=rep1&type=pdf> (accessed December 25, 2009).

# What Causes the Invalid Votes? With a Concurrently Discussion of Spatial Analysis of Invalid Votes in Taiwan's Elections, 1992-2008

Chih-sung Teng<sup>\*</sup> · Chin-en Wu<sup>\*\*</sup> · I-jung Ko<sup>\*\*\*</sup>

## Abstract

The issue of invalid votes receives few scholarly attentions. Many consider invalid votes as the result of unintended behavior and do not explore the factors that may influence the incidence of invalid votes. We examine 37 elections of different levels between 1992 and 2008 in Taiwan by using panel data analysis. We find that the distribution of invalid votes are not random but are influenced by institutional and socioeconomic factors. Regarding institutional factors, the elections codes, the complexity of elections, the importance of elections, and years after the democratic transition are the main influencing factors. Years of education, the percentage of elders, population density, and percentage of indigenous citizens are the main socioeconomic variables that affect the incidence of invalid votes. The two strings of factors can explain about 34% of the variance in invalid votes. In addition to the pooled OLS model, we also employ spatial lag model and spatial error model. The two models show that the distribution of invalid votes exhibit positive spatial autocorrelation. In addition, some areas also exhibit spatial heterogeneity, which is likely to be attributed to the alienation

---

\* Associate Professor, Graduate Institute of National Development, National Taiwan University.

\*\* Assistant Research Fellow, Institute of Political Science, Academia Sinica.

\*\*\* MA, Graduate Institute of National Development, National Taiwan University.

of voters in the districts. We might need to pay special attention to these areas to enhance the quality of democracy.

Keywords: invalid vote, spatial autocorrelation, spatial heterogeneity, panel data analysis, spatial regression