

# 別訪問我！我對政治沒興趣： 主題效應與 TEDS 高估投票率之研究

蔡奇霖\*

## 《本文摘要》

「主題效應」(topic effect) 指出，對調查主題越感興趣的人越可能接受訪問，所以和主題有關的估計值容易被誇大。台灣選舉與民主化調查 (TEDS) 是一個以政治及選舉為主題的訪問，而投票與否這個問項又和這些主題息息相關，因此主題效應很可能是 TEDS 高估投票率的一個原因。本文用兩個假設來檢驗這個想法，首先在總體層次的分析上，本文發現 TEDS 高估投票率的情況比台灣社會變遷基本調查 (TSCS) 的更嚴重。由於兩個調查在主題外有許多相似性，這個結果表示兩者在投票率估計值上的差異很可能是主題效應所致。其次在個體層次的分析上，本文也發現在 TEDS 的成功樣本裡高投票意願者比低投票意願者更可能接受 TEDS 的訪問，符合主題效應的論述。整體來說，總體和個體層次的分析都支持本文的看法。不過，本文的結論是立基在特定前提之上，未來 TEDS 應蒐集更多資料來進一步確認及處理主題效應和高估投票率的問題。

關鍵詞：主題效應、高估投票率、過度代表、過度回報

---

\* 國立中正大學政治學研究所碩士。E-mail: kirin\_guess@yahoo.com.tw。

## 壹、前言

投票與否是選舉研究裡相當重要的變數，但由於真實紀錄難以取得，研究者們紛紛透過民意調查來蒐集這項資料。只是民調的投票率估計值往往高於官方公佈的數據，使得研究偏誤 (bias) 的陰影始終揮之不去。舉例來說，如果高估投票率是有投票的人在樣本中被過度代表 (over-representative) 所致，研究者便會過分偏重他們的意見而錯估民意；如果是未投票者過度回報 (over-report) 所致，<sup>1</sup> 研究者則會誤把他們納入分析，導致選舉預測失準，或者誤解投票行為的特性而影響選舉動員的研究。

高估投票率的影響層面甚廣，因此吸引許多學者投入研究，有的探討成因 (Bernstein, Chadha, and Montjoy 2001; Burden 2000; Health and Taylor 1999; McDonald 2003; Silver, Anderson, and Abramson 1986; Stocké 2007; Stocké and Stark 2007)，有的謀求改善之道 (Belli, Moore, and VanHoewyk 2006; Belli et al. 1999; Fullerton, Dixon, and Borch 2007; McDonald 2007)，就連調查單位也致力於蒐集相關資料，譬如：美國選舉研究 (American National Election Studies, ANES) 就在 1964、1972 至 1980 年，以及 1984 至 1990 年進行「投票確認」(vote validations)，利用官方紀錄確認受訪者的回答是否屬實。

在台灣，台灣選舉與民主化調查 (Taiwan's Election and Democratization Studies, 以下簡稱 TEDS) 是學界非常重要的研究資料。TEDS2004L 也曾針對高估投票率的問題進行問項之間的比較，可惜未達預期效果 (Wu 2006)。在此之外，學界對於高估投票率的研究並不多，因此本文嘗試分析 TEDS (獨立抽樣的面對面訪問調查) 高估投票率的狀況，<sup>2</sup> 並且檢驗「主題效應」(topic effect) 和 TEDS 高估投票率的關聯性，亦即投票意願較高的人是否比較可能接受 TEDS 這個以政治和選舉為主題的訪問，因而導致投票率被高估。

或許有人認為，既然有官方公佈的數據，民調高估投票率有什麼關係？其實本文真正關心的是總體數據背後的個體資料，因為投票率被高估代表個體資料可能有系統性誤差；如果本文最後發現主題效應是 TEDS 高估投票率的原因之一，未來在調查規劃上就能作出因應措施。不過既有資料極為有限，難以直接驗證主題效應是否發生，因此本文將累積各種間接證據來推論主題效應和 TEDS 高估投票率的關係。在接下來的內容裡，本文將先討論民調高估投票率的各種原因，然後再說明主題效應和這些原因的關係，藉以形成研究假

<sup>1</sup> 文獻裡常看到的「過度回報」一詞有時是總體的概念，意指民調投票率高於官方投票率的現象，有時則是個體的概念，指涉沒有投票的受訪者謊報或誤報有去投票的情況。本文統一用「高估投票率」代表總體概念，用「過度回報」代表個體概念。

<sup>2</sup> 除了 2005NA 與 2008L 有執行電話訪問之外，TEDS 皆為面訪調查。本文分析的是 TEDS 面訪調查高估投票率的情況；除非有特別說明，否則本文所謂的 TEDS 皆指獨立抽樣的選後面訪調查。

設：接著分別利用總體資料與個體資料檢驗假設，最後則提出結論和建議。

## 貳、主題效應與高估投票率

### 一、高估投票率的原因

民意調查高估投票率的主要原因有四種：第一、過度回報：沒有投票的受訪者故意謊報或不小心誤報為有去投票 (Belli, Moore, and VanHoewyk 2006; Belli et al. 1999; Bernstein, Chadha, and Montjoy 2001; Silver, Anderson, and Abramson 1986; Stocké 2007)。第二、霍桑效應 (Hawthorne effect)：訪問本身改變了受訪者的行為，例如：受訪者的投票意願受到選前訪問的刺激而提高，選後再次訪問這些受訪者的追蹤調查便容易高估投票率 (Burden 2000)。第三、過度代表：有投票者在樣本裡的比例高於他們在母群裡的比例 (Burden 2000; Health and Taylor 1999; Voogt and Saris 2003)。第四、分母不一致：以美國為例，普查局 (Bureau of the Census) 採用「投票年齡的人口」(voting-age population) 估計選民人數，ANES 則採用「有投票權的人口」：兩者在計算投票率時使用的分母不同，計算結果也就會有差距 (McDonald 2003)。<sup>3</sup>

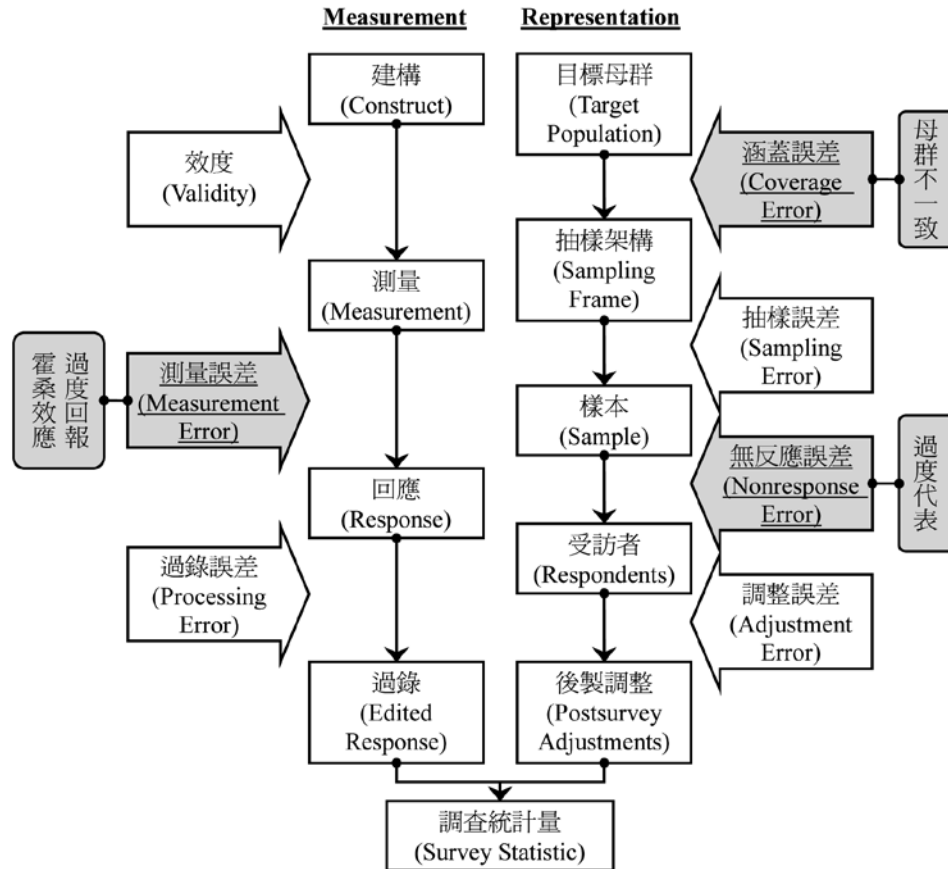
對於這四種原因，使用「總誤差方法論」(Total error methodology) (Groves et al. 2004, 48) 可以「定位」它們各是在哪個環節上發生問題，又該從哪裡著手改善。由圖 1 來看，分母不一致屬於「涵蓋誤差」的問題，過度代表屬於「無反應誤差」的問題，而過度回報與霍桑效應則屬於「測量誤差」的問題。<sup>4</sup>

針對 TEDS 的情況來看，本文分析的是選舉之後才進行訪問的獨立抽樣樣本，因此霍桑效應與高估投票率沒有關係。台灣的戶政資料較為完備，TEDS 所定義的選民人口和中央選舉委員會所定義的非常接近，所以相較於美國的調查資料，TEDS 也比較沒有分母不

<sup>3</sup> McDonald(2003) 指出 ANES 和普查局的分母差異主要在於如何處理「不合格選民」(ineligible voter)，亦即已滿投票年齡但無投票權的人口。普查局的分母包含了「非公民」(noncitizen) 和「非住家人口」(the institutional population)，ANES 的則一律排除。另外，兩者的分母都沒有包含「海外公民」(overseas citizens)，可是普查局的分子卻包含海外公民的投票數，ANES 的則沒有。

<sup>4</sup> 本文是先假定 (assume) 調查單位已盡可能處理了自己所能掌握的部分，再定位四種高估投票率的原因。舉例來說，過度回報可能來自於問項建構的問題（如問項的用字遣詞讓受訪者誤解而誤答），也可能是測量過程的問題（如未投票者誤答或謊稱有去投票）。然而，問項建構的問題是調查單位在訪問前較能盡力克服的；相反的，測量過程的問題則是調查單位比較難以掌控的，因此本文把過度回報定位為測量誤差。以 TEDS 的嚴謹程度來說，這個假定應該頗為合理。

一致的疑慮。<sup>5</sup> 由此可見，過度回報與過度代表才是 TEDS 高估投票率的主因；相關研究應著眼於測量誤差與無反應誤差。



資料來源：作者修改自 Groves 等 (2004, 48)。

說明：圖中專有名詞之定義請見 Groves 等 (2004, 39-65)。

圖 1 高估投票率的原因及其來源

<sup>5</sup> TEDS 和中選會在計算投票率時所界定的分母都是「有投票權的民眾」，TEDS 的母群清冊和中選會的選舉人名冊也都是使用戶籍資料，因此本文認為 TEDS 較無分母不一致的疑慮。不過，本文也感謝審查人提醒有關戶籍登記不實的問題；如果「籍在人不在」的比例夠高且該族群的沒投票者夠多，那的確也會使高估投票率的問題惡化。這是值得研究的課題，可是基於兩個原因，它並不會影響本文的結論。首先，許多文獻 (Olson 2006; Peytchev, Baxter, and Carley-Baxter 2009) 指出調查偏誤主要是來自於不願意接受訪問的人 (refusing to participate) 而非沒辦法被接觸到的人 (cannot be contacted)。由於籍在人不在屬於後者，它對投票率估計值的影響力相對較小。其次，本文在總體層次的分析裡比較「台灣社會變遷基本調查」(Taiwan Social Change Survey, TSCS) 和 TEDS 高估投票率的情況。兩者的抽樣清冊都來自於政府的戶籍資料，因此籍在人不在的情況類似，不至於影響兩者的比較。在個體層次的分析裡，本文直接使用訪問成功的樣本討論拒訪傾向和投票意願的關係，因此也不會受到籍在人不在的影響。

## 二、主題效應

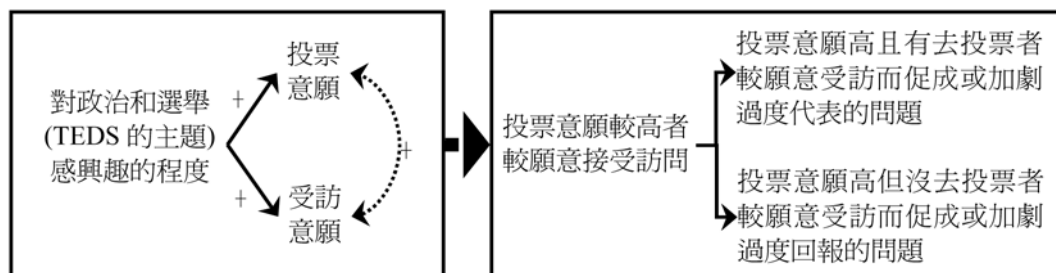
雖然過度回報（測量誤差）和過度代表（無反應誤差）的發生原因並不相同，本文卻注意到兩者有一個交集之處，亦即主題效應。Groves 等 (2006, 728-731) 曾進行過的一項調查實驗 (survey experiment) 可以清楚的說明何謂主題效應。他們設計了兩份問卷，第一份的調查名稱是「Survey on Birds, Bird-Watching, and Birding」，第二份則是「Survey on the Design of Indoor Shopping Malls」。這兩份問卷都有一個問項用於測量受訪者在近半年內的賞鳥經驗。結果發現，對賞鳥活動越有興趣的受訪者，在第一份問卷的受訪率越高，該問卷裡回答有賞鳥經驗的比例也越高。這個實驗想傳達的概念正是訪問主題所引發的調查偏誤，也就是所謂的主題效應。詳言之，由於調查主題是受訪者接不接受訪問的重要判準，對主題越感興趣的受訪者，就越可能接受訪問，反之則越可能拒訪；在這種情況下，與主題有關的問項估計值也就容易被誇大。

本文認為主題效應相當適合用來解釋 TEDS 為何會在測量投票參與時發生過度代表和過度回報的問題。具體而言，TEDS 的中文正式名稱是「台灣選舉與民主化調查」，舉凡事前通知受訪者將有訪員前來拜訪的明信片、信件、來訪未遇卡，或是訪員的聘書、識別證、問卷……等都有註明這十個字。它們明白的告訴受訪者 TEDS 是以政治和選舉為主題的訪問，因此對於政治和選舉較有興趣的人會比其他更可能接受訪問。再者，由於政治興趣對於投票意願的影響甚大 (Brady, Verba, and Schlozman 1995; Campbell et al. 1960)，這就表示投票意願較高的人也會比較可能接受 TEDS 的訪問，從而造成兩種後果：第一、投票意願較高，真的有去投票的機率也比較大，所以當調查主題促使受訪意願和投票意願呈正相關時，有投票者就容易被過度代表。第二、即使投票意願較高的人因為某些原因而沒去投票，他們在接受訪問時也可能會過度回報，因為研究指出投票可能性較高卻沒去投票的受訪者最容易過度回報 (Bernstein, Chadha, and Montjoy 2001; Silver, Anderson, and Abramson 1986)。由此可見，主題效應和 TEDS 高估投票率的兩大主因有著密切的關係。

上述論點可摘述為圖 2，而其中的關鍵之處就在於調查主題使得受訪意願和投票意願呈正相關。事實上，除了主題效應的論述之外，許多國內外的研究也都顯示這兩種意願有正相關的可能，例如：Cheng(2006, 164) 發現拒訪者與受訪者的年齡、教育程度、省籍、政黨偏好及統獨立場都明顯不同，而這些因素又與投票意願有關，尤其是年齡與教育程度 (Tsai 2001, 149; 吳俊德與陳永福 2005, 87; 陳敦源 2000, 48; 楊孟麗 2003, 171-172; 謝邦昌與江志民 1998, 44)。另外，洪永泰 (2003, 43) 在比較 TEDS2001 的成功樣本與「追蹤樣本」後也發現訪問失敗者的投票率較低。<sup>6</sup> Voogt 與 Saris(2003) 更是直

<sup>6</sup> 這裡的「追蹤樣本」是指追蹤訪問失敗的受訪者，改用比較簡短的問卷再試著說服他們接受訪

接利用社會參與 (social involvement) 解釋政治興趣、訪問參與，以及投票參與之間為何會呈正相關。這些研究都說明了本文的假設並非空穴來風，主題效應很可能是 TEDS 高估投票率的一個重要原因。



資料來源：作者自行繪製。

圖 2 主題效應、過度代表與過度回報

不過，Cheng(2006) 分析的是政治大學選舉研究中心的面訪資料，而且沒有直接分析投票意願和受訪意願的關係。Voogt 與 Saris(2003) 雖有分析兩種意願的關係，卻也沒有使用 TEDS 的資料。針對 TEDS 進行研究的洪永泰（2003）則只分析 TEDS2001 的資料，而且追蹤樣本的數目只有 100 個。不同於這些文獻，本文援引主題效應，從理論層面來說明受訪意願和投票意願為何會呈正相關，並為 TEDS 高估投票率的現象提供一個解釋。

### 三、待驗假設

簡單來說，在調查主題的作用下，受訪者的投票意願和受訪意願呈正相關，因此投票意願較高的人比較可能進入 TEDS 的訪問成功樣本；他們之中的兩種成員（有投票者和沒投票者）又分別促發或加劇過度代表和過度回報的問題，使得 TEDS 的投票率估計值被誇大。基於這個想法，本文假設主題效應是 TEDS 高估投票率的原因之一。

想要檢驗這個假設，最直接的方法是仿效 Groves 等 (2006) 的調查實驗。譬如：設計兩份名稱不同，內容完全相同的問卷。控制組的問卷仍叫「台灣選舉與民主化調查」，實驗組的問卷名稱則和政治或選舉無關。最後比較兩組資料在投票率估計值上是否有顯著差異。不過，TEDS 未曾作過這種調查設計，在缺乏相關資料的情況下，本文改用兩個較為間接，但能夠檢驗的假設來分析 TEDS 的投票率估計值是否發生主題效應：

假設一：與主題不是政治或選舉的訪問相比，TEDS 高估投票率的程度比較嚴重。

問。換言之，它是「follow-up sample」而非「panel sample」。

假設二：就 TEDS 的成功樣本來看，投票意願較高者比投票意願較低者更可能接受 TEDS 的訪問。

## 參、研究設計

爲了克服資料不足的困難，本文分別透過總體和個體層次的兩個假設來分析 TEDS 的投票率估計值是否發生主題效應。兩個假設分述如下：

### 一、總體層次——TEDS 和 TSCS 之比較

在總體分析上，本文比較「台灣社會變遷基本調查」（以下簡稱 TSCS）與 TEDS 高估投票率的程度（附錄一列出本文使用的資料）。由於 TEDS 的主題是政治與選舉，TSCS 在名稱上則和政治或選舉沒有明顯關聯，如果 TEDS 高估投票率的程度比 TSCS 的更嚴重，那就可能表示主題效應發生在 TEDS 的投票率估計值上。之所以會挑選 TSCS 和 TEDS 進行比較是因為兩者在調查名稱之外有著許多相似性，它們都是國內大型的學術調查，也都由學術單位負責執行；在樣本抽取上都用政府戶籍資料作爲抽樣清冊，也都先用人口資料對鄉鎮市區作分層（但分層變數和數目並不一定相同），再用「抽取機率與抽出單位大小成比例」（Probability proportional to size, PPS）進行抽樣；在訪員組成上兩者皆無固定的訪員團隊，每次調查都要重新招募訪員（但 TEDS 的訪員以學生居多，TSCS 則否）；在訪問執行上都規定受訪者必需經過至少三次接觸失敗才能放棄，也都有檢查訪員舞弊的機制（但 TEDS 有再測信度分析，TSCS 則沒有）。

儘管兩個調查頗爲相似，仍有三個問題必須釐清。第一、母群差異：歷次 TSCS 的推論母群並不一定是選民，因此本文在計算投票率時都將利用受訪者的出生時間來篩選出有投票權的受訪者，以盡可能減輕母群不同所帶來的干擾。<sup>7</sup> 第二、機構效應：TSCS 由中央研究院社會學研究所主持，TEDS 則由台大、政大、中正、中山、東吳……等政治相關系所輪流執行。雖然執行機構不同，它們其實都是學術單位，而且訪員組成、抽樣設計和訪問執行的方式也都頗爲相似，所以受訪者對於機構差異的感受不至於太明顯。再者，TEDS 是由各校輪流執行，所以 TSCS 的執行單位（中研院社會所）和這些大學系所全都發生機構效應的可能性是相對較低的。若真要說有機構效應，本文認爲機構的名稱才是比較重要的因素。由於受訪者可能會利用執行單位的名稱來猜測訪問內容，對於政治或選舉較無興趣的選民就可能傾向於拒絕「某某大學『政治』學研究所」所執行的訪問，亦即

<sup>7</sup> 這裡只作原則性說明。在後面的分析裡將以註釋的方式更詳細的說明資料處理方法，以便陳述和閱讀。

拒絕 TEDS。即便如此，那也是政治興趣較高者較會接受 TEDS 訪問的一種證明。基於這些因素，本文認為如果兩者高估投票率的程度真有差異，機構效應可能不是主因。

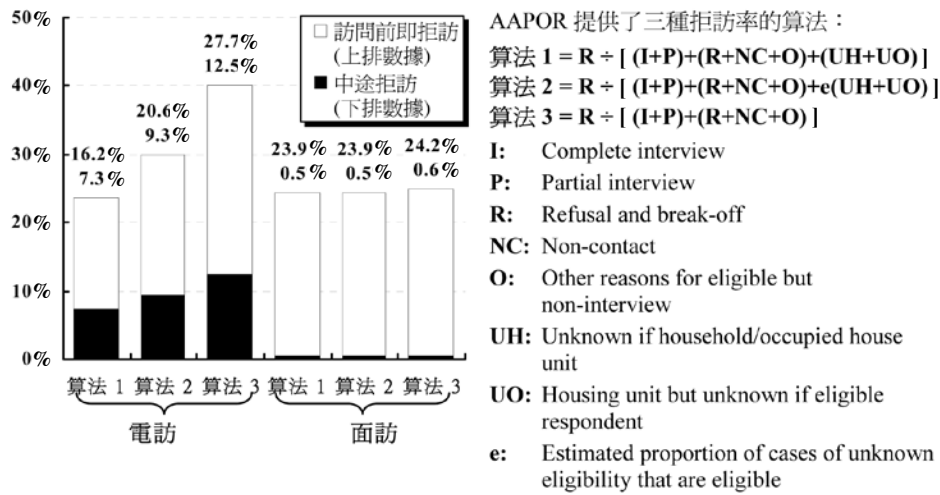
第三個必須釐清的問題是「多元主題」。雖然 TSCS 統稱為「台灣社會變遷基本調查」，它其實有許多不同「內容」的問卷，而且每個都是獨立的調查，譬如：第五期第三次調查就有兩份問卷，一份的內容是社會階層，另一份則是休閒生活。TSCS 曾經調查過的還有公民權組、國家認同組、公民與國家組、工作與生活組、綜合議題組……等，可見有些問卷內容仍是政治性的。對於政治或選舉不感興趣的受訪者也許不會馬上拒訪，但是在一連串政治性的問題後，他們還是可能在回答投票與否的問項之前就中途拒訪 (break-off)。如此一來，主題效應對於某些組別的 TSCS 和 TEDS 就會有類似的影響，即使兩者高估投票率的程度不同，那也無法證明是主題效應的作用。

的確，如果 TSCS 是電話訪問，這會是個大問題。不過，TSCS 是面訪調查，因此本文認為這種情況的影響有限。事實上，「見面三分情」一直都是面訪的重要優勢。圖 3 以美國民意調查學會 (American Association for Public Opinion Research, AAPOR) 在 2008 年所修訂的方法計算 TEDS2008L 面訪與電訪的拒訪率 (AAPOR 2008)。<sup>8</sup> 中途拒訪在電訪裡佔全部拒訪的 31.10% (中途拒訪率 / 拒訪率，以算法 3 為例： $12.4892\% / (12.4892\% + 27.6725\%) = 31.0973\%$ )，在面訪裡只佔 2.21% (以算法 3 為例： $0.5483\% / (0.5483\% + 24.2157\%) = 2.2140\%$ )。由此可知，即使面訪的受訪者在訪問過程中展現出較低的受訪意願，真的會中途拒訪的人也不多。也就是說，雖然有些 TSCS 的問卷內容是政治性的，它們仍統稱為「台灣社會變遷基本調查」；一旦開始進行訪問，對於政治或選舉不感興趣的受訪者也不太會在回答投票與否的問項前就因為一連串的政治性問項而中途拒訪，所以本文認為多元主題對於總體分析的影響有限。<sup>9</sup>

<sup>8</sup> 圖 3 純粹是為了解說「中途拒訪比較容易發生在電訪，但很少發生在面訪，因此 TSCS 的多元主題應該不會影響本文的分析」。只是 TSCS 並無電訪的資料，無法呈現出兩種訪問方法的差異，因此才改用兩種訪問資料都有的 TEDS2008L 作圖。特此說明，以免讀者混淆。

<sup>9</sup> 基於同樣的理由，本文也不會特別期待公民權組、國家認同組、公民與國家組的 TSCS 會比工作與生活組、綜合議題組的 TSCS 更容易高估投票率。





資料來源：作者根據 AAPOR 的規則和 TEDS2008L 的資料自行繪製，詳細的訪問結果請見附錄二。

圖 3 面訪與電訪的中途拒訪比例

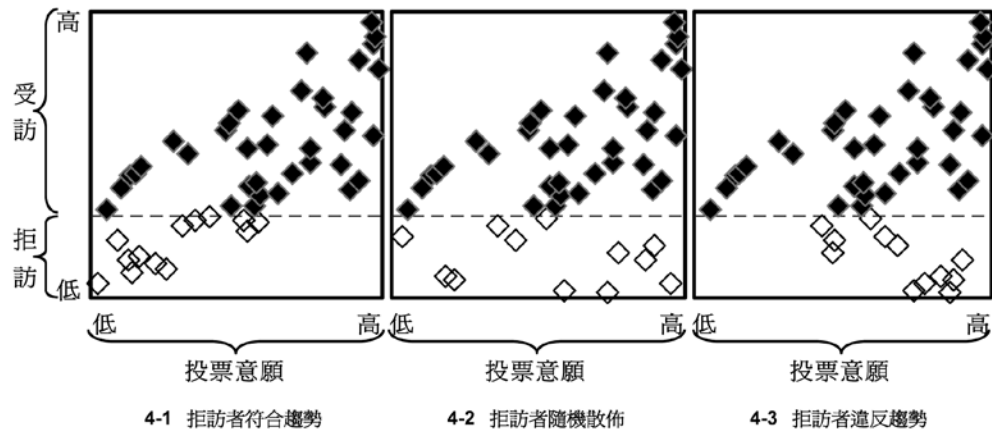
雖然總體假設只是比較 TEDS 和 TSCS 的投票率估計值，基於上述說明，分析結果仍有助於瞭解主題效應和 TEDS 高估投票率的關係。然而，即使已盡可能排除各種干擾因素，TEDS 和 TSCS 之間難免還是會有調查名稱以外的差異。單憑兩者在投票率估計值上的比較，並不足以作為主題效應的證明，因此本文將另外進行個體層次的分析。

## 二、個體層次——高投票意願者和低投票意願者之比較

TEDS 的投票率估計值有沒有發生主題效應，關鍵就在於受訪者的投票意願和受訪意願是否呈正相關，而這也是本文想要檢驗的第二個假設。對此本文選用受訪者在投票與否這個問項上的回答作為投票意願的指標，這是因為絕大部分回答「有去投票」的人都是：第一、真的有去投票的人，第二、想要投票卻沒去投票的人 (Bernstein, Chadha, and Montjoy 2001; Silver, Anderson, and Abramson 1986)。也就是說，雖然回答有去投票的人並不一定真的有去投票，大部分卻都是投票意願較高的人；反之，回答沒去投票的人則大部分是投票意願較低的人，因此本文認為這個問項非常適合用來測量投票意願。至於受訪意願（或說拒訪傾向），<sup>10</sup> 本文則是使用訪員評估的資料作為指標。TEDS 從 2004L 起開始釋出訪員評估的資料（參見附錄三）。<sup>11</sup> 每當完成訪問，訪員就必須記錄訪問開始前及訪問過程中的情況，並對整體訪問進行評估，因此即使只有訪問成功的樣本，本文仍有資料可以用來測量他們的受訪意願。

<sup>10</sup> 本文把受訪意願和拒訪傾向視為一體兩面的概念，因此後續分析將交互使用這兩個辭彙。

<sup>11</sup> 其實 TEDS2002 至 TEDS2004P 的問卷也都有訪員評估，可是釋出資料檔並未包含這些資訊，因此本文只使用 TEDS2004L 至 TEDS2008P 進行個體資料的分析。



資料來源：作者自行繪製，圖中皆為虛擬資料。

說明：縱軸是受訪意願，橫軸是投票意願。虛線是受訪門檻，空心方塊是拒訪者，實心的則是受訪者。

圖 4 以成功樣本推論全部樣本的前提

話雖如此，個體層次的假設仍有兩個潛在的疑問必須釐清：第一、既是成功樣本，何來「受訪意願」或「拒訪傾向」？第二、為何訪員評估能夠作為受訪意願的指標？針對第一個疑問，本文認為雖然受訪與拒訪表面上是截然二分的結果，背後卻有程度之別。這個想法可以用圖 4 來說明。圖中的縱軸代表受訪意願，橫軸代表投票意願；實心方塊代表接受訪問的受訪者，而空心方塊則代表拒訪者；橫虛線是接不接受訪問的門檻 (threshold)，高於門檻的受訪者就會接受訪問。如圖所示，即便是接受訪問的人（也就是成功樣本），他們的受訪意願也不會完全相同，而是有高有低（如同實心方塊雖然都在橫虛線的上方，仍有高低之別），於是本文能夠只用成功樣本分析投票意願和受訪意願的關係。

針對第二個疑問，本文主張每個訪員評估的項目都是受訪意願的部分反映。首先，在訪問時的「合作程度」及「有無中途拒訪的傾向」作為拒訪傾向之指標的合理性不言而喻。其次，許多文獻都指出個人隱私、人身安全、時間成本、對訪問的興趣和瞭解程度是受訪者接不接受訪問的重要考量 (Brehm 1993, 58-62; Groves, Cialdini, and Couper 1992, 477-479)。在訪員評估裡，「表示不信任調查訪問」、「表示對訪問用途有疑慮」、「表示對資料能否保密有疑慮」、「表示對訪問單位 / 主持人有疑慮」、「質疑訪員的身份」、「有家人反對」、「表示題目太敏感」，以及「對問題的敏感程度」就反映了個人隱私和人身安全的考量；「表示自己太忙、無法受訪」、「表示家裡在忙、無法受訪」及「表示問卷題目太多」反映的是時間成本的考量；「表示對訪問不感興趣」、「覺得問卷題目太難」、「對問題的瞭解程度」、「對訪問內容感興趣的程度」和「對一般政治事務訊息的知悉度」則反映了受訪者對於訪問的興趣和瞭解程度。最後，「要求自行填寫問卷」、「要求看問卷」、「表示問卷題目太多」、「表示題意不清或定義不明」及「回答問題的可信度」這些項目和拒訪傾

向的關係也不難想像，譬如：要求自行填寫問卷的人會比其他人更容易拒訪，因為他們可能對一問一答的方式感到不耐煩；要求看問卷的人可能是聽不懂訪員在說什麼，這種溝通上的問題會妨礙訪問的流暢度，增加拒訪的可能性；題意不清或定義不明可能會讓受訪者產生答題的挫折感而降低受訪意願；不願意但勉強受訪的人可能會胡亂回答，因此受訪意願也能從答題的可信度中窺知一二。基於這些理由，本文認為訪員評估可以作為拒訪傾向的指標。

透過上述的研究設計，本文可以判斷 TEDS 的成功樣本是否出現主題效應的跡象。不過，缺少拒訪樣本的資料畢竟是個事實，想用成功樣本的分析結果推論全部樣本（包含受訪者和拒訪者）的情況還得仰賴一個前提 (assumption)，亦即投票意願和受訪意願的關係必須像圖 4-1 一樣，拒訪者們是順著成功樣本的趨勢集中在圖的左下方。如此一來，若成功樣本的分析結果顯示投票意願和受訪意願呈正相關，那麼主題效應就真的是高估投票率的原因。<sup>12</sup> 當然，拒訪樣本也可能會違反成功樣本的趨勢，如圖 4-2 或圖 4-3。如果拒訪者的投票意願是隨機分佈的（圖 4-2），整體樣本的受訪意願和投票意願仍是正相關，只是程度較輕。換言之，主題效應仍是高估投票率的原因，只是影響程度小於圖 4-1 的情況。但如果拒訪者的趨勢與受訪者的相反（如圖 4-3，成功樣本是受訪意願越高，投票意願也越高，拒訪樣本卻是受訪意願越高，投票意願越低），那麼主題效應就無法解釋 TEDS 為何會高估投票率。

然而基於以下兩個理由，本文認為把圖 4-1 當作個體分析的前提並不至於太荒謬。第一、就邏輯的合理性而言，圖 4-1 是較為合理的情況，因為民衆的態度與行為比較可能具有一種整體性的、連貫性的趨勢，就像圖 4-1 一樣。投票意願與受訪意願的相關趨勢在越過受訪門檻後就馬上改變（圖 4-2），甚至於相反（圖 4-3），都是比較違反常態的情形。第二、Groves 等 (2006) 已證實了主題效應的現象，許多研究也都曾討論高估投票率和主題效應的關係 (Voogt and Saris 2003)，這多少都說明了圖 4-3 是比較不切實際的情況，因為圖 4-3 的情況很可能會低估而非高估投票率，不符合這些文獻的研究結果。誠然，光憑這兩點還是無法徹底排除圖 4-2 與圖 4-3 的可能性，但至少能讓大家對圖 4-1 的前提產生一定程度的信心。

<sup>12</sup> 這個前提的重點是「拒訪者們會『順著成功樣本的趨勢』集中在圖的左下方」，而非「投票意願和受訪意願呈正相關」，因為後者是本文的待驗假設，而且先前也已援引許多文獻來說明這個假設的合理性。

## 肆、分析結果

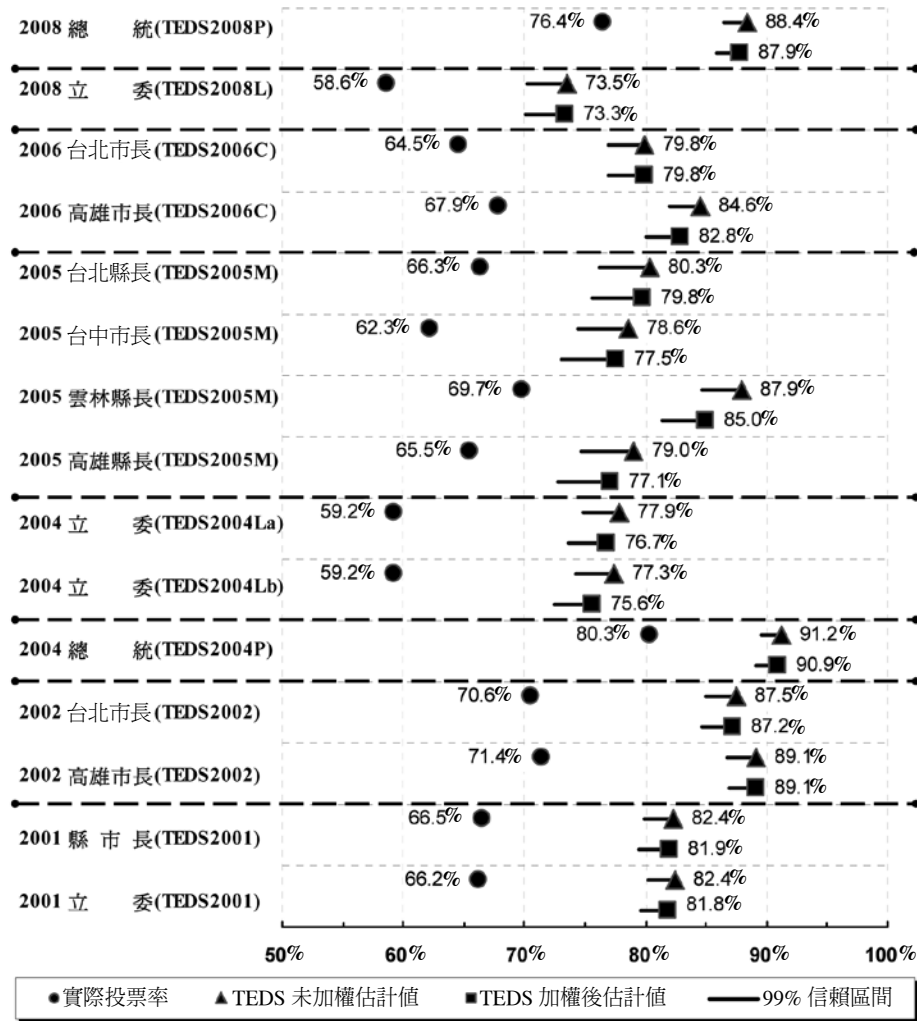
### 一、TEDS 高估投票率的情況

圖 5 呈現 TEDS2001 至 2008P 高估投票率的情況。首先，比較同類型的選舉可以發現實際投票率有逐漸下降的趨勢。其中，立委選舉的降幅最大，直轄市長選舉及總統選舉次之，縣市長選舉比較沒改變（2005 年縣市長選舉的整體投票率為 66.23%）。由於實際投票率越低，被高估的空間就越大，如果未來投票率持續下滑，這對選舉民調並不是個好消息。

其次，就未加權的投票率估計值來看，這 15 筆 TEDS 高估投票率的平均程度是 15.69 個百分點。其中最高的是 TEDS2004La 的 18.71 個百分點，最低的則是 TEDS2004P 的 10.91 個百分點。另外，我們也可以看到，實際投票率較低的立委選舉，高估的情況最嚴重，平均高估 17.02 個百分點；實際投票率較高的總統選舉，高估的情況則最輕微，平均只有 11.43 個百分點。如果顧及實際投票率的差異，這 15 筆 TEDS 的估計值平均是其實際投票率的 1.24 倍。最高的仍是 TEDS2004La 的 1.32 倍，最低的也是 TEDS2004P 的 1.14 倍。最重要的是 15 個未加權估計值的 99% 信賴區間都沒能觸及實際投票率，可見抽樣誤差並不足以解釋兩者的差距。

加權是處理樣本代表性問題的常用方法。在訪問、過錄、檢誤……完成後，TEDS 會檢定樣本在性別、年齡、地區、教育程度上的代表性，並加權有顯著差異的部分。圖 5 顯示，雖然加權後的投票率估計值都比加權前的更接近實際投票率，但是變動幅度卻都很小。加權後估計值的 99% 信賴區間依舊沒能觸及實際投票率，而加權前後的差異也不大。由此可見，TEDS 的加權程序無法有效的修正高估投票率的問題。至於加權效果不彰的原因可能是選用的加權變數和投不投票的關係有限。<sup>13</sup> 無論如何，圖 5 告訴我們的是 TEDS 的確有高估投票率的現象，而且目前的加權程序沒辦法解決這個問題。

<sup>13</sup> 加權變數和投不投票的相關程度越低，修正的效果就越小 (Brehm 1993, 119)。



資料來源：作者根據中央選舉委員會（2009）和 TEDS 的資料自行整理繪製，調查資料之說明請見附錄一與附錄四。

說明：實際投票率的計算都配合 TEDS 的推論母群。此外，估計值全都大於實際投票率，故只呈現 99% 信賴區間的左側，右側用來呈現估計值的數據。

圖 5 TEDS 高估投票率的情況

## 二、總體分析——TEDS 和 TSCS 之比較 1

在確定 TEDS 有高估投票率的問題後，本文將接著檢驗這個問題和主題效應的關係。圖 6 與圖 7 是 TEDS 與 TSCS 對於四次選舉的投票率估計值，它們分別是 2000 年與 2004 年的總統選舉，以及 2001 年與 2004 年的立委選舉。這些是 TEDS 與 TSCS 最有交集的選舉（請參閱附錄四）。以 2004 年總統選舉為例（圖 6-2），除了 TEDS2004P 之

外，TEDS2004La、2004Lb、2008L、2008P 也都有問項可以估計這次選舉的投票率；<sup>14</sup> TSCS 也有兩筆：第 4 期第 5 次的問卷 1（TSCS451 公民權組），及第 5 期第 1 次的問卷 1（TSCS511 綜合問卷組）。<sup>15</sup> 其次，圖中也標示出資料的訪問年度，以方便比較同一年執行訪問的調查。訪問的時間點越相近，時空背景越相似，比較同一年進行訪問的調查有助於排除許多主題效應以外的差異。再者，圖中用三角形代表估計值是從「投票與否」的問項計算而來的，如：圖 6-2 的 TEDS2004P；用正方形代表估計值是從「投票抉擇」的問項計算而來的，如：圖 6-2 的 TEDS2004La。<sup>16</sup> 最後，圖中皆為未加權的估計值。

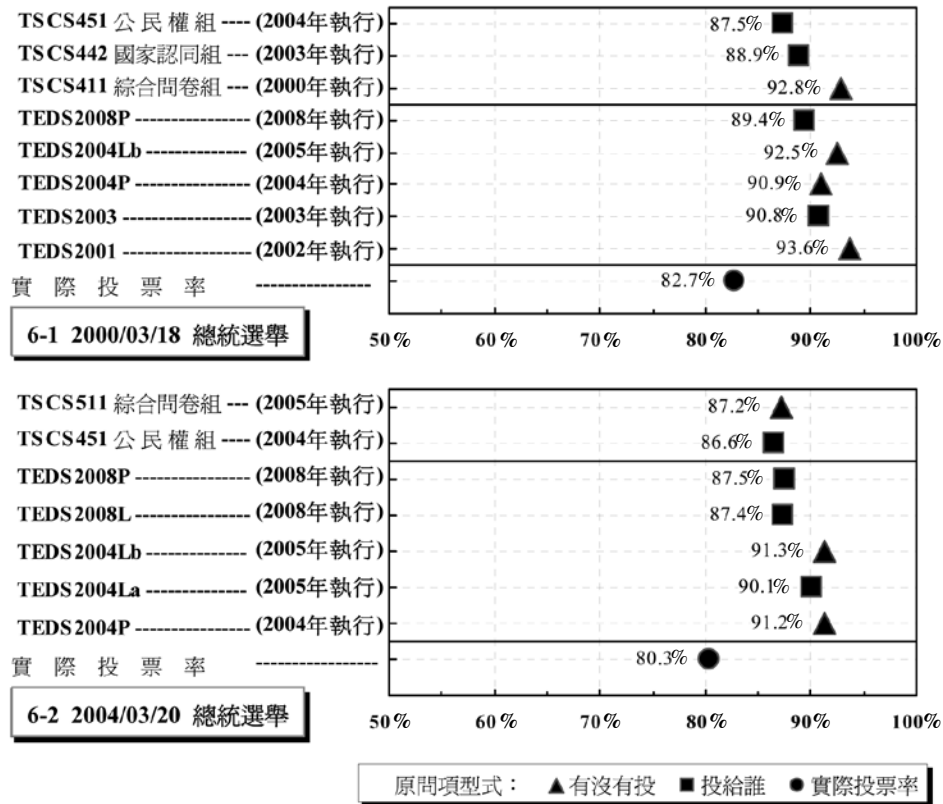
先就 2000 年總統選舉來看（圖 6-1），雖然 TSCS411 的估計值高於 2003 年起的四次 TEDS 的估計值（TEDS2003、2004P、2004Lb、2008P），<sup>17</sup> 在控制調查的時空背景後，TSCS 的估計值就全都小於 TEDS 的。具體來說，比較同在 2003 年執行訪問的 TSCS442 與 TEDS2003，以及比較同在 2004 年執行訪問的 TSCS451 與 TEDS2004P，TSCS 的高估程度都比較輕微。如果 2000 年執行訪問的 TSCS411 與執行時間最接近的 TEDS2001 相比，結果也是 TSCS 的估計值比較接近實際投票率。再就 2004 年總統選舉來看（圖 6-2），TSCS 的估計值不僅全部小於 TEDS 的，控制時空背景後兩者的差異更為明顯（TSCS451 vs. TEDS2004P; TSCS511 vs. TEDS2004La; TSCS511 vs. TEDS2004Lb）。

<sup>14</sup> TEDS2006Ctp（台北市長選舉調查）與 TEDS2006Cks（高雄市長選舉調查）有估計 2004 年總統選舉投票率的問項，可是這兩筆資料的推論母體只限於台北市或高雄市，因此不列入圖 6-2 的分析。同理，TEDS2002Ctp 與 TEDS2002Cks 也不列入圖 6-1 與圖 7-1 的分析。

<sup>15</sup> TEDS2004P 是針對 2004 總統選舉而執行的調查，推論母體即為在該次選舉具有投票權的選民，但其他調查的母群則有些出入。TEDS 只有提供受訪者的出生年，因此統一篩選出 1984 年出生的受訪者來計算 2004 年總統選舉的投票率。再以 2001 年立委選舉為例，除了 TEDS2001 之外，其他 TEDS 或 TSCS 的資料都篩選出 1981 年出生的受訪者來計算投票率。本文使用這個原則來處理母群不一致的問題。

<sup>16</sup> 投票與否的問項是指「在這一次舉行的總統大選中，有很多人去投票，也有很多人因各種原因沒有去投票，請問您有沒有去投票？1.『有』2.『沒有』91.『忘了』95.『拒答』，在估計投票率時把「忘了」和「拒答」設為遺漏值。投票抉擇的問項則是指「在 2004 年的總統選舉中，請問您投票給哪一組候選人？1.『陳呂』2.『連宋』3.『未投票』4.『投廢票』5.『當時無投票權』91.『忘記了』95.『拒答』98.『不知道』，先把「無投票權、忘記了、拒答、不知道」都設為遺漏值，再用 100% 減去「未投票」的比例即為投票率的估計值。

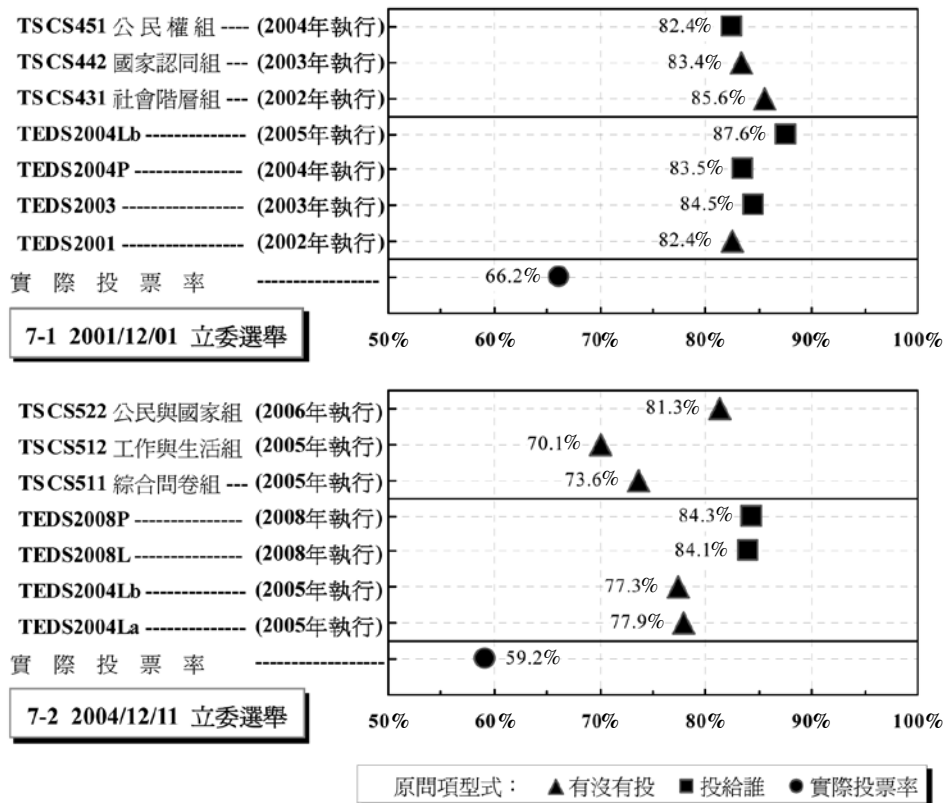
<sup>17</sup> 本文篩選 1980 年出生的受訪者來計算 2000 年總統選舉的投票率，然而 TSCS411 卻只有 1979 年 12 月之前出生的受訪者。比起其他調查，TSCS 少了 1980 年出生的年輕選民，而年輕選民正好是比較不會去投票的人（Evans 2004; Tsai 2001; Wolfinger and Rosenstone 1980），這可能是 TSCS411 高估投票率的情況較為嚴重的原因之一。



資料來源：作者根據中央選舉委員會（2009）、TEDS 和 TSCS 的資料自行整理繪製，調查資料之說明請見附錄一與附錄四。

圖 6 TEDS 與 TSCS 之比較（總統選舉）

接著，在 2001 年立委選舉裡（圖 7-1），2003 年執行訪問的兩個調查 (TEDS2003 vs. TSCS442) 和 2004 年執行訪問的兩個調查 (TEDS2004P vs. TSCS451) 都是 TEDS 高估投票率的程度比較嚴重，但 2002 年執行訪問的兩次調查 (TSCS431 vs. TEDS2001) 卻是 TSCS 高估較多。最後是 2004 年立委選舉（圖 7-2），雖然和不同年執行訪問的 TSCS522 相比，TEDS2004La 和 TEDS2004Lb 的估計值比較接近實際投票率，可是和同年執行訪問的 TSCS511 及 TSCS512 相比，TEDS2004La 和 TEDS2004Lb 高估投票率的程度卻又明顯地比較嚴重。值得注意的是，TSCS512 的估計值與圖 7-2 裡的四次 TEDS 估計值都有著相當大的差距，而 TSCS512 的問卷內容正好是和政治沒有明顯關係的「工作與生活組」。



資料來源：作者根據中央選舉委員會（2009）、TEDS 和 TSCS 的資料自行整理繪製，調查資料之說明請見附錄一與附錄四。

圖 7 TEDS 與 TSCS 之比較（立委選舉）

整體而言，雖然 TSCS 的估計值並非全都低於 TEDS 的估計值，兩兩之間的差異也沒有很大，圖 6 與圖 7 仍有些支持假設的證據：第一、在四張圖裡，TSCS 的簡單平均數都低於 TEDS。<sup>18</sup> 第二、在四次選舉裡，高估投票率最嚴重的都是 TEDS 的調查，最輕微的都是 TSCS。第三、無論哪一次選舉，在控制訪問的時空背景後，TSCS 的估計值都比 TEDS 小，唯一例外是 TSCS431 高估 2001 年立委選舉投票率的程度大於同一年執行訪問的 TEDS2001。除此之外，四次選舉都毫無例外的具有上述三個特點，這表示 TEDS 真的比 TSCS 更容易高估投票率。

<sup>18</sup> 以 2000 年總統選舉為例，3 個 TSCS 估計值的簡單平均數是  $(92.8+88.9+87.5)/3=89.7\%$ ，5 個 TEDS 估計值的則是  $(93.6+90.8+90.9+92.5+89.4)/5=91.4\%$ 。同樣的，在 2004 年總統選舉裡，TSCS 和 TEDS 的簡單平均數分別是 86.9% 比 89.5%；2001 年立委選舉是 83.8% 比 84.5%；2004 年立委選舉則是 75.0% 比 80.9%。



### 三、總體分析——TEDS 和 TSCS 之比較 2

爲了進一步確認圖 6 與圖 7 的分析結果，本文匯集 TEDS 與 TSCS 的投票率估計值（請參閱附錄四）進行迴歸分析。模型的依變數是「高估投票率的百分點」，亦即估計值減去實際投票率，自變數則是一個虛擬變數，TEDS 的估計值編碼爲 1，TSCS 的編碼爲 0。此外，模型也加入一些控制變數：第一、實際投票率越高，被高估的空間就越小，因此控制「實際投票率」的差異。第二、分析資料是由不同選舉聚集而成的，樣本數太少所以無法控制每次選舉的差異，只好改爲控制「選舉的類別」。第三、估計值有的是從投票與否的問項計算而來，有些則是從投票抉擇的問項計算而來，因此控制「問項別」。第四、訪問日與投票日的距離不同，估計值受干擾的程度也不同。對此，本文取每筆資料的訪問日中位數，計算它距離投票日的天數並取自然對數值，藉以控制「訪問日與投票日的相距時間」。<sup>19</sup> 第五、TEDS 有估計 1994 年直轄市長選舉、1996 年總統選舉……等 2000 年之前的選舉投票率，TSCS 並沒有。本文納入這些估值計以增加樣本數，但也控制它們的可能影響。第六、除了 TEDS2003 之外，每個 TEDS 都有自己的「目標選舉」，譬如：TEDS2004P 的是 2004 年總統選舉、TEDS2008L 的是 2008 年立委選舉。畢竟 TEDS 是專爲目標選舉而設計的，目標選舉的估計值和非目標選舉的可能會有系統性差異，再加上 TSCS 的調查全都沒有所謂的目標選舉，因此模型也控制這些差異的可能影響。

表 1 的模型 1-1 是針對圖 6 與圖 7 裡 29 個投票率估計值的分析，模型 1-2 則是增加樣本之後的分析。在模型 1-2 迴歸裡，本文採用 TEDS2001 至 2008 的資料，可是基於兩個原因只採用 TSCS 第 4 期及其之後的調查資料：第一、TSCS 第 4 期之前的推論母群有年齡的上限，譬如：20 至 64 歲、20 至 75 歲、25 至 65 歲、19 至 75 歲。年齡下限低於 20 歲的情況尚可解決（亦即刪去未滿 20 歲的樣本），年齡上限則難以處理。第二、TEDS2001 至 2008 都是 2000 年之後才開始執行的計畫，TSCS 第 4 期之前的調查則都是在 2000 年之前就執行的計畫，因此使用第 4 期及其之後的資料有助於排除訪問背景的干擾。

模型 1-1 顯示，TEDS 與 TSCS 高估投票率的程度有顯著的差別。當其他條件不變，TEDS 的高估程度會比 TSCS 多 3.45 個百分點。這確認了圖 6 與圖 7 的分析結果，亦即 TEDS 比 TSCS 更容易高估投票率。同樣的，樣本數相對較多的模型 1-2 也確認了這個結果：在其他條件不變的情況下，TEDS 與 TSCS 高估投票率的程度有 3.54 個百分點的顯著差異。此外，也許是樣本數增加，有些控制變數和依變數的關係也達到顯著。若其他條件

<sup>19</sup> 在資料處理時筆者發現 TEDS2008L 有八個成功樣本的訪問日是在訪員訓練日 (2008/1/19) 之前，其中有三個是在投票日 (2008/1/12) 之前就訪問完成，有一個是在投票日當天訪問完成，在資料分析時本文把這八個樣本的訪問日期修改爲其他成功樣本的訪問中位日期 (2008/2/10)。

不變，投票率每增加 10 個百分點，高估的程度就會減少 3.15 個百分點；立委選舉被高估的程度比總統選舉多 5.09 個百分點；地方首長選舉（直轄市長與縣市長）則比總統選舉多 3.15 個百分點。這些都與圖 5 的分析結果相同。

表 1 TEDS 與 TSCS 高估投票率之迴歸分析

自變數	模型 1-1 $\beta$ (S.E.)	模型 1-2 $\beta$ (S.E.)	模型 1-3 $\beta$ (Robust S.E.)
<b>TEDS 的估計值</b> (TEDS 為 1；TSCS 為 0)	<b>0.0345 (0.0140)*</b>	<b>0.0354 (0.0085)**</b>	<b>0.0354 (0.0078)***</b>
<b>實際投票率</b>	-0.1792 (0.2174)	-0.3146 (0.1138)*	-0.3146 (0.1038)**
<b>立委選舉</b> (參照類：總統選舉)	0.0758 (0.0420)	0.0509 (0.0206)*	0.0509 (0.0175)**
<b>地方首長選舉</b> (參照類：總統選舉)	沒有這個變數	0.0315 (0.0124)*	0.0315 (0.0126)*
<b>問項型式</b> (投票與否為 1、抉擇為 0)	0.0081 (0.0131)	0.0064 (0.0047)	0.0064 (0.0070)
<b>訪問距離投票日的天數</b>	-0.0006 (0.0191)	-0.0033 (0.0123)	-0.0033 (0.0125)
<b>西元 2000 年之前的選舉</b> (之前的為 1、之後的為 0)	沒有這個變數	0.0067 (0.0148)	0.0067 (0.0112)
<b>目標選舉</b> (目標選舉為 1、其餘為 0)	-0.0247 (0.0269)	-0.0244 (0.0159)	-0.0244 (0.0162)
<b>常數項</b>	0.2050 (0.1768)	0.3237 (0.0928)**	0.3237 (0.0816)***
n (TEDS; TSCS)	29 (18; 11)	59 (46; 13)	59 (46; 13)
F-Test (df1, df2)	16.93 (06, 23)***	22.59 (08, 50)***	25.96 (08, 50)***
Adjusted R <sup>2</sup>	0.77	0.75	0.78
VIF (mean; max; min)	6.70; 15.77; 1.49	4.31; 7.53; 1.63	
Normality <sup>a</sup>	Z = -0.90	Z = 1.90*	
Heteroskedasticity <sup>b</sup>	$\chi^2(1) = 4.28^*$	$\chi^2(1) = 7.07^{**}$	
Omitted Variables <sup>c</sup>	F(3, 19) = 5.21***	F(3, 47) = 0.73	

資料來源：作者根據中央選舉委員會（2009）、TEDS 和 TSCS 的資料自行分析與製表，調查資料之說明請參見附錄一與附錄四。

說明：1. 模型的依變數皆為高估投票率的百分點。

2. <sup>a</sup> 用 Shapiro-Wilk Test 檢定常態性，虛無假設是殘差服從常態分配。

<sup>b</sup> 用 Breusch-Pagan/Cook-Weisberg Test 檢定變異數同質性，虛無假設是變異數具有同質性。

<sup>c</sup> 用 Ramsey RESET Test 檢定遺漏變數的問題，虛無假設是模型無遺漏變數的問題。

3. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ 。

根據模型診斷 (Model diagnostics) 的結果，模型 1-1 和 1-2 都有違反迴歸假定的問

題。模型 1-1 有變異數異質性 (heteroskedasticity) 和遺漏變數 (omitted variables) 的情況，而且有兩個自變數的變異數膨脹因子 (Variance inflation factor, VIF) 超過 10，多元共線性 (multicollinearity) 非常嚴重 (王保進 2006, 455)。模型 1-2 則違反了常態性 (normality) 和變異數同質性 (homoskedasticity) 的假定。本文嘗試修正模型 1-2 違反迴歸假定的問題，因為它的樣本數相對較多，是分析的主要依據。針對變異數異質性的問題，本文採用 Huber 和 White 等人的穩健估計式 (robust estimator) 修正標準誤 (Standard error, S.E.) (Hamilton 2004, 256-258)。違反常態性假定會提高型一錯誤 (type I error) 的可能性，因此本文降低  $\alpha$  值，以較嚴格的標準解讀係數是否顯著 (王保進 2006, 451)。模型 1-3 是重新估計的結果，它與模型 1-2 極為類似，而且主要自變數的  $P$  值變得更小 ( $P$ -value $<0.001$ )，就算  $\alpha$  值降低也不影響原先的結論。<sup>20</sup>

透過上述分析可知 TEDS 高估投票率的程度比 TSCS 的嚴重。TEDS 以選舉為主題，而 TSCS 的名稱則與政治無關，再加上本文在研究設計和模型設定上排除了許多主題之外的干擾，因此主題效應可能得為兩者在投票率估計值上的差異負起部分責任。只是這個結果不應該被解讀為 TEDS 的品質比 TSCS 的差。根據主題效應的論述可知，TEDS 在估計投票率時本來就居於「先天的劣勢」。也許 TSCS 在估計與本身主題有關的變數時也會有這種劣勢，所以高估投票率的程度並不適用來評斷兩者的調查品質。

整體來說，總體分析驗證了本文的第一個假設：TEDS 比 TSCS 更容易高估投票率。基於研究設計，這個結果在某種程度上也說明了主題效應與 TEDS 高估投票率的關係。不過，兩個調查在主題之外仍有許多無法控制的差異，而且在圖 6 和圖 7 裡的差異不算太大，表 1 模型的樣本數也不夠多，因此接下來將繼續檢驗個體層次的假設，累積更多證據來支持本文的想法。

#### 四、個體分析——高投票意願者和低投票意願者之比較 1

在個體分析裡，本文使用投票與否的問項作為投票意願的指標，使用訪員評估的資料測量拒訪傾向（或說受訪意願），以此檢驗兩者在成功樣本裡的關係，並且在圖 4-1 的前提下（亦即在拒訪傾向和投票意願的關係上，成功樣本與拒訪樣本具有相同的趨勢），把分析結果推論至全部樣本。

<sup>20</sup> 雖然不是時間序列的資料，模型 1-1 和 1-2 也可能違反觀察值互相獨立的假定 (independent-errors assumption)，因為有些樣本是來自於相同的資料，譬如：TEDS2008P 就有 2004 年立委、2008 年立委、2000 年總統、2004 年總統和 2008 年總統等五次選舉的投票率估計值。為此，本文先依照訪問別把樣本分成許多集群 (clusters)，例如：上述五個樣本就都屬於 TEDS2008P 這個集群，再使用 Stata 提供的穩健估計式修正標準誤（詳情請參閱 Hamilton(2004, 258-261)）。不過重新估計後仍維持原先的結論，因此不再列出詳細數據，如有需要請與作者連絡。

表 2 拒訪傾向和投票意願的相關係數

各種拒訪傾向的指標 <sup>a</sup>	Kendall's Tau(Asymptotic S.E.) <sup>b</sup>	n
<b>訪問前的情況 (Tau-b)</b>		
表示自己太忙，無法受訪	-0.0063 (0.0099)	10,570
表示家裡在忙，無法受訪	0.0140 (0.0093)	10,570
<b>表示對訪問不感興趣</b>	<b>-0.0279 (0.0101)**</b>	<b>10,570</b>
表示不信任調查訪問	0.0104 (0.0094)	10,570
表示對訪問用途有疑慮	0.0101 (0.0095)	10,570
表示對資料能否保密有疑慮	0.0207 (0.0092)*	10,570
表示對訪問單位及主持人有疑慮	-0.0100 (0.0101)	10,570
質疑訪員身份	0.0054 (0.0096)	10,570
有家人反對受訪者接受訪問	-0.0155 (0.0102)	10,570
<b>訪問時的情況 (Tau-b)</b>		
要求自行填寫問卷	-0.0335 (0.0107)**	10,567
要求看問卷	0.0026 (0.0097)	10,569
中途拒訪的傾向	0.0178 (0.0093)	10,565
表示問卷題目太多	0.0329 (0.0094)***	10,565
表示題意不清或定義不明	0.0406 (0.0090)***	10,557
覺得問卷題目太難	0.0314 (0.0089)***	10,558
表示題目過於敏感	0.0729 (0.0079)***	10,552
<b>整體訪問的情況 (Tau-c)</b>		
合作態度	-0.0121 (0.0083)	10,567
對問題的瞭解程度	-0.0232 (0.0085)**	10,563
對問題的敏感程度	0.0548 (0.0084)***	10,555
<b>對訪問內容感興趣的程度</b>	<b>-0.1008 (0.0083)***</b>	<b>10,564</b>
回答問題的可信度	-0.0275 (0.0078)***	10,559
對一般政治事務訊息的知悉度	-0.0878 (0.0082)***	10,554

資料來源：作者利用 TEDS2004La、04Lb、05M、06C、08L、08P 的面訪獨立樣本自行分析與製表。

說明：<sup>a</sup>拒訪傾向的指標編碼方式請參照附錄三；投票意願則以歷次 TEDS 的目標選舉為準，0 代表低投票意願者，1 代表高投票意願者。

<sup>b</sup>相關係數為負表示高投票意願者比較願意接受訪問，符合假設；相關係數為正則表示高投票意願者比較容易拒訪。\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ 。

正如前述，筆者認為每一項訪員評估都是受訪意願的部分反映，不過，本文最關心的指標是「訪問前是否表示對訪問不感興趣」和「對訪問內容感興趣的程度」。根據主題效應，投票意願較高的人之所以比較可能接受訪問，而投票意願較低的人之所以比較可能拒訪，關鍵就是兩者對於訪問主題的興趣不同，也因此這兩個受訪意願的指標和投票意願之間的相關性最能夠檢驗主題效應是不是 TEDS 高估投票率的原因。表 2 的 Kendall's Tau

顯示，和低投票意願者相比，高投票意願者在訪問前較少表示對訪問沒興趣，對於訪問內容感興趣的程度也比較高。這兩個相關係數不僅在方向上符合預期，而且也都具有統計上的顯著性。

除了這兩個項目之外，表 2 也有其他指標和投票意願達到顯著相關。其中，「要求自行填寫問卷」、「對問題的瞭解程度」、「回答問題的可信度」、「對一般政治事務訊息的知悉度」，這四個項目也顯示高投票意願者比低投票意願者更可能接受訪問。相反的，「對資料能否保密有疑慮」、「表示問卷題目太多」、「表示題意不清或定義不明」、「覺得問卷題目太難」、「表示題目過於敏感」、「對問題的敏感程度」，這六個項目卻顯示高投票意願者比低投票意願者更可能拒訪。當兩種趨勢互相抵消之後，或許主題效應就不會對投票率的估計值產生影響。不過，情況可能沒有這麼樂觀，仔細觀察表 2 可以發現，符合預期趨勢的項目大多是整體訪問的評估，違反預期的則多為個別訪問狀況的紀錄。雖然每個項目都能反映拒訪傾向，整體訪問的評估應該比個別狀況的紀錄更具代表性。再者，正如先前圖 3 所示，面訪的中途拒訪原本就比較少，就算投票意願較高的人在訪問中展現出較低的受訪意願，真的會中途拒訪的也不多。基於這兩個理由，本文認為主題效應的影響並不會被完全抵消。

值得注意的是，表 2 顯示投票意願較高的人似乎對於政治比較敏感。在訪問開始前，他們比較擔心資料能否保密；在訪問當中或在整體訪問裡，他們對問卷的敏感度也比較高。政治敏感度越高就越會擔心什麼能說、什麼不能說。為了避免訪問帶來的心理負擔，拒訪傾向勢必會提高。不同於其他違反假設的項目，高投票意願者的政治敏感性不只反映在訪問前和訪問時的狀況，就連整體訪問的評估也出現這種現象，因此若說主題效應有可能被抵消，政治敏感度就是個值得注意的因素。

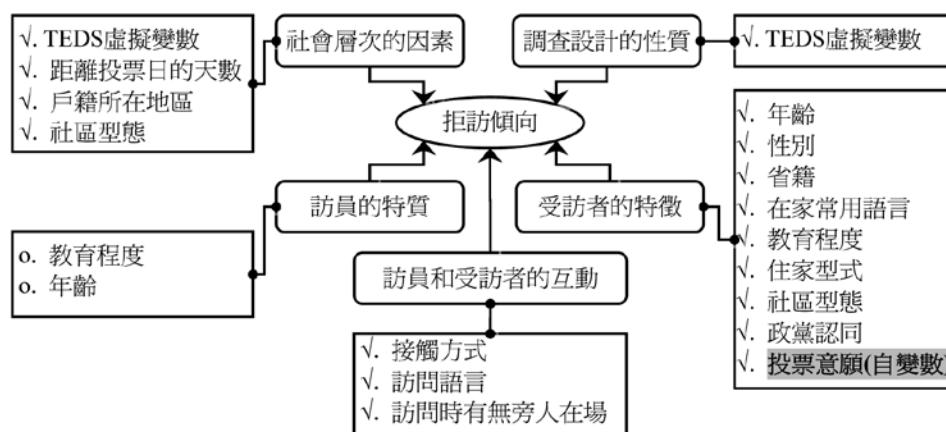
整體而言，表 2 初步驗證了本文的第二個假設，其中這兩個最重要的指標（訪問前是否表示對訪問不感興趣，以及對整體內容感興趣的程度）和投票意願的關係都符合預期。不過，投票意願較高的人對政治性訪問太敏感卻也可能削弱本文的論點，接下來本文將繼續進行多變數分析，以進一步確認這些分析結果。

## 五、個體分析——高投票意願者和低投票意願者之比較 2

### （一）模型設定

在多變數分析裡，自變數仍然是投票意願，依變數也還是各項訪員評估，它們反映的是受訪者的拒訪傾向，兩者的編碼方式與表 2 相同。除此之外，我們也將控制其他會影響拒訪傾向的變數。根據 Groves、Cialdini 與 Couper(1992, 477-479) 的研究，這些變數可分成五大類：社會層次的因素 (societal-level factors)、調查設計的性質 (attributes of the survey

design)、受訪者的特徵 (characteristics of the sample person)、訪員的特質 (attributes of the interviewer)，以及訪員和受訪者的互動 (respondent-interviewer interaction)。詳言之，社會層次的因素是指執行訪問時的社會凝聚力，大眾對民調的觀感、調查方法的使用頻率（過度調查的效應）……等時空背景。其次，接觸受訪者的模式、說服受訪者的策略、更換受訪者的規則、問卷的長度、調查主題……等都屬於調查設計的性質。再次，受訪者的特徵包括年齡、性別、收入、健康情形、家庭狀況、住家環境等，不同特徵會形成不同的心理特質 (psychological predispositions)，進而影響接不接受訪問的決定。接著，訪員的特質可分為外顯和內在兩種，外顯的特質如：性別、年齡、種族……等都是受訪者判斷訪員來意的依據；內在的特質如：當下情緒、訪問經驗，甚至是對訪問主題與內容的看法……等也都影響訪員的表現。最後，也是 Croves、Cialdini 與 Couper 最強調的因素就是訪員與受訪者的互動，他們認為受訪者的拒訪傾向會隨著互動而改變，而訪員也會依照受訪者的反應來選擇及變換說服的策略。



資料來源：作者自行繪製。

圖 8 影響拒訪傾向的因素及其指標

本文根據這些因素來挑選控制變數（詳細編碼方式請見附錄五）。如圖 8，在社會因素上，從 TEDS2004L 到 2008P 每筆資料的執行時間都有一段差距，因此使用 TEDS 的虛擬變數來控制這個時空差異。另外，每次選舉的氛圍會隨著時間逐漸減弱，訪問的敏感性也會逐漸降低，因此控制受訪日和投票日的相隔天數。除了時間層面的變數之外，本文也採用受訪者的戶籍所在地，以及訪問地點的社區型態來控制兩種不同範圍的訪問背景。其次在調查設計上，TEDS 虛擬變數也能控制這個因素，所以本文依據訪問時間與調查設計的不同劃分出十個 TEDS 的虛擬變數，它們同時具有控制社會因素和調查設計的功能。

在受訪者特徵上，控制變數有年齡、性別、省籍、在家常用語言、住家型式、社區型態。另外，受訪者對民調的認識與看法可能會隨教育程度或政治態度而不同，<sup>21</sup> 所以模型也控制這兩項特徵。不過，政治態度的變數甚多，沒辦法全都控制，所以用政黨認同作代表，因為它較能反映民衆的主要政治態度。至於受訪者的收入、健康情形與家庭狀況則因為資料限制而無法納入。<sup>22</sup> 接著，由於 TEDS 並未公開訪員資料，本文只能「間接控制」訪員的年齡與教育程度（大多數訪員都是年紀與學歷相仿的大學生或研究生）。<sup>23</sup> 最後，訪員和受訪者的互動情況本來就不容易記錄和量化，藉由「訪員接觸受訪者的方式」、「訪問語言」及「訪問時有無他人在場」這三個變數，本文希望能控制住一些雙方互動的影響。<sup>24</sup>

## （二）模型分析

表 3 是拒訪傾向和投票意願的勝算對數模型（每一列都是一個模型），先看最能反映主題效應的兩個指標：在「是否表示對訪問不感興趣」及「對訪問內容感興趣之程度」的模型裡，高投票意願者對低投票意願者的勝算比 (odds ratio) 都小於 1，而且係數都達到統計上的顯著水準。這表示當其他條件不變，高投票意願者接受 TEDS 訪問的勝算大於低投票意願者。除此之外，在雙變數分析（表 2）裡符合假設的項目，如：「對問題的瞭解程度」、「回答問題的可信度」、「對一般政治事務訊息的知悉度」，它們在表 3 裡也呈現出高投票意願者比較可能接受訪問的趨勢。唯一不同的項目只有「要求自行填寫問卷」，它和投票意願的關係變成不顯著。不過，「合作態度」和投票意願的關係卻從表 2 裡的不顯著

<sup>21</sup> 許多經驗研究都曾指出政治態度，尤其是政黨認同，和受訪意願的關聯性，譬如：Cheng(2006, 164) 發現拒訪者和受訪者的政黨認同對象明顯不同；Merkle 等 (1998, 轉引自 Groves(2006, 666)) 也發現相同的訪問誘因對於不同政黨的支持者有不同的影響力；Tang、Wang 與 Bautista(2009) 甚至於指出有些政黨曾經公開呼籲民眾不要參與特定訪問。為此本文利用政黨認同作為指標，控制政治態度對於受訪意願的影響。

<sup>22</sup> TEDS2004L 至 2008P 並沒有記錄受訪者的健康情形與家庭狀況。在收入方面，TEDS2004La 沒有這個變數；2004Lb 至 2006C 的分類方式則與 2008L、2008P 的不同。

<sup>23</sup> TEDS2006C 至 2008P 的資料有公開訪員的性別，TEDS2004L 與 2005M 的則沒有。換言之，在本文的分析樣本裡，幾乎有一半沒有訪員性別的資料，所以只好放棄這個變數。

<sup>24</sup> 本文依循王濟川與郭志剛 (2003, 238) 的建議，利用 VIF 檢查自變數（含控制變數）的多元共線性，也就是說先用一般線性迴歸取代勝算對數模型再計算 VIF。具體而言，本文用勝算對數模型來分析自變數和依變數的關係（即表 3），但用一般線性迴歸的模型診斷方法來檢查多元共線性。結果顯示 VIF 的平均值是 2.04，最高值是 6.87，次高的是 5.15，其他的都介於 1.03 和 3.72 之間。VIF 都遠低於 10，可見多元共線性並沒有非常嚴重（王保進 2006, 455）。值得說明的是，「在家常用語言」和「訪問語言」都是用複選題的方式建立各類別的虛擬變數（參見附錄五），這或許是兩者的共線性沒有想像中嚴重的原因。詳言之，在家常用的語言可能有好幾種，但訪問時不一定都會用到；相反的，訪問語言也可能有好幾種，但也不一定都是在家常用的語言。

變成表 3 裡的顯著，亦即高投票意願者的合作情況明顯優於低投票意願者。這一來一往之間並無損假設的正確性，因為「合作態度」應該比「要求自行填寫問卷」更能反映受訪意願的情況。

表 3 拒訪傾向和投票意願的勝算對數模型

依變數—各種拒訪傾向的指標 <sup>a</sup> (每一列是一個模型)	投票意願之估計值 <sup>b</sup>		LR $\chi^2$ (df)	Pseudo R <sup>2</sup>	n
	$\beta$ (S.E.)	Odds Ratio			
<b>訪問前的情況 (Binary Logit)<sup>c</sup></b>					
表示自己太忙，無法受訪	-0.0232 (0.0888)	0.9770	131.48 (45)***	0.02	9,904
表示家裡在忙，無法受訪	0.0298 (0.1132)	1.0302	178.43 (45)***	0.04	9,904
<b>表示對訪問不感興趣</b>	<b>-0.2901 (0.0762)***</b>	<b>0.7482</b>	<b>593.54 (45)***</b>	<b>0.07</b>	<b>9,904</b>
表示不信任調查訪問	0.0745 (0.1276)	1.0773	151.67 (45)***	0.04	9,904
表示對訪問用途有疑慮	0.0147 (0.0916)	1.0148	175.81 (45)***	0.03	9,904
表示對資料能否保密有疑慮	0.1151 (0.1011)	1.1220	174.56 (45)***	0.03	9,904
表示對訪問單位 / 主持人有疑慮	-0.1610 (0.1292)	0.8513	112.84 (45)***	0.03	9,904
質疑訪員身份	-0.0244 (0.1189)	0.9759	118.67 (45)***	0.03	9,904
有家人反對受訪者接受訪問	-0.1115 (0.1257)	0.8944	139.15 (44)***	0.04	9,802
<b>訪問時的情況 (Binary Logit)</b>					
要求自行填寫問卷	-0.1480 (0.1195)	0.8624	233.73 (45)***	0.06	9,903
要求看問卷	0.2094 (0.1008)*	1.2329	190.84 (45)***	0.03	9,904
中途拒訪的傾向	-0.0023 (0.1168)	0.9977	580.35 (45)***	0.12	9,903
表示問卷題目太多	0.0886 (0.0690)	1.0927	822.66 (45)***	0.08	9,901
表示題意不清或定義不明	0.0877 (0.0867)	1.1100	475.32 (45)***	0.06	9,894
覺得問卷題目太難	0.0343 (0.1226)	1.0349	734.47 (45)***	0.15	9,895
表示題目過於敏感	0.7086 (0.1121)***	2.0312	382.55 (45)***	0.06	9,890
<b>整體訪問的情況 (Ordered Logit)</b>					
合作態度	-0.1378 (0.0510)**	0.8712	1055.06 (45)***	0.04	9,904
對問題的瞭解程度	-0.3460 (0.0505)***	0.7075	3329.28 (45)***	0.12	9,901
對問題的敏感程度	0.2266 (0.0479)***	1.2544	577.75 (45)***	0.02	9,893
<b>對訪問內容感興趣的程度</b>	<b>-0.5015 (0.0502)***</b>	<b>0.6056</b>	<b>1823.26 (45)***</b>	<b>0.06</b>	<b>9,902</b>
回答問題的可信度	-0.2725 (0.0538)***	0.7615	1533.07 (45)***	0.08	9,898
對一般政治事務訊息的知悉度	-0.6081 (0.0531)***	0.5444	3169.75 (45)***	0.14	9,893

資料來源：作者利用 TEDS2004La、04Lb、05M、06C、08L、08P 的面訪獨立樣本自行分析與製表。

說明：1. <sup>a</sup> 受訪意願和投票意願的編碼方式與表 2 相同，其他控制變數的編碼方式請見附錄五。

<sup>b</sup> 本表只呈現「投票意願」這個自變數的估計值，省略控制變數的，如需完整數據請與作者連絡。

<sup>c</sup> 勝算比 (odds ratio) 小於 1 表示高投票意願者接受訪問的勝算 (odds) 高於低投票意願者，符合本文假設；反之，勝算比大於 1 則不符合假設。

2. 不仰賴「平行迴歸假定」(parallel regression assumption) 的模型請參閱附錄六。

3. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ 。



再看違反假設的項目。在「對資料能否保密有疑慮」、「表示問卷題目太多」、「表示題意不清或定義不明」、「覺得問卷題目太難」等模型裡，投票意願的係數都未達顯著水準。這表示原本在表 2 裡違反假設的趨勢已變得不明顯；控制各種影響拒訪的因素之後，投票意願較高的人並不會比投票意願較低的人更傾向於拒訪。不過，表 3 仍舊顯示投票意願較高的人對政治比較敏感。在訪問過程中或是整體訪問的情況，他們覺得題目太敏感的勝算仍明顯高於低投票意願者。另外，他們要求看問卷的勝算也明顯高於低投票意願者。這也違反了本文的假設，但由於中途拒訪較少在面訪中發生，這個項目違反假設對結論的影響有限。

最後必須說明的是表 3 的 pseudo  $R^2$  偏低、模型配適度 (goodness-of-fit) 不佳的問題。筆者認為主要原因有兩個：第一、模型只用訪問成功的觀察值，第二、模型使用的是拒訪傾向的間接指標。這兩個因素都來自於資料的限制，暫時難以克服。但就算模型配適度偏低，表 3 仍比表 2 多控制了一些干擾因素，讓分析結果更具說服力。

總括來說，表 3 更明確的驗證了本文的第二個假設。在控制其他影響拒訪傾向的因素後，絕大多數違反假設的趨勢已變得不顯著，支持假設的趨勢卻依舊存在，尤其是受訪者對 TEDS 是否感興趣的兩個項目，這個結果在某種程度上證明了主題效應是 TEDS 高估投票率的重要原因。不過，表 3 也再次顯示高投票意願者的政治敏感度的確比較大，這可能會抵消主題效應的影響。只是這麼多項指標都顯示投票意願較高的人比較願意接受訪問，光憑政治敏感度一項可能無法阻止主題效應發生。

## 伍、結論與建議

高估投票率是民意調查經常發生的問題，TEDS 也不例外。在台灣，這雖然不會妨礙真正投票率的估算，卻是個體資料發生系統性誤差的表徵。TEDS 高估投票率的兩大主因是過度回報與過度代表，雖然兩者的成因不盡相同，卻也不無交集之處，亦即主題效應。詳言之，主題效應使得投票意願較高的人比較可能接受 TEDS 的訪問；他們之中的兩種成員（想投票且有投票者和想投票卻沒投票者）分別促發或加劇過度代表和過度回報的問題，導致 TEDS 的投票率估計值被誇大，因此本文認為主題效應很可能是 TEDS 高估投票率的一項重要原因。不過既有資料有所侷限，本文改為分析兩個可以檢驗的假設，一個是總體層次的，一個是個體層次的，藉以窺探主題效應和 TEDS 高估投票率的關係。

研究結果顯示 TEDS 的確有高估投票率的問題。從 TEDS2001 至 2008P，15 筆面訪資料平均高估投票率的程度都遠超過抽樣誤差的範圍，而且既有加權程序的修正效果也極為有限。在假設的檢驗上，總體層次的分析指出，TEDS 高估投票率的程度的確比 TSCS

的嚴重，而 TSCS 的名稱（台灣社會變遷基本調查）正好與政治無關。個體層次的分析則顯示，在成功樣本裡投票意願較高的人確實比投票意願較低的人更可能接受 TEDS 的訪問，因為受訪意願的指標，包括：對訪問感興趣的程度，在訪問時的合作程度、對問題的瞭解程度、答題的可信度、對一般政治事務訊息的知悉度，都和投票意願呈正相關。綜合來說，總體與個體層次的分析都符合主題效應的預期。雖然本文無法直接檢驗主題效應是否發生，這些分析結果在某種程度上仍說明了主題效應是 TEDS 高估投票率的成因之一。

必須注意的是，這個結論是建立在特定前提之上。為了避免誤用這個結論，讀者必須留意各項分析的前提和侷限。在總體分析上，雖然本文已盡力排除主題以外的干擾因素，TSCS 與 TEDS 畢竟是不同的調查計畫，它們之間仍有許多差異，而且兩者的迴歸分析也有樣本不足的問題。在個體分析上，缺乏拒訪者的資料是主要的限制，在這種情況下，想用成功樣本的分析結果推論全部樣本的情況必須仰賴一個前提：在受訪意願和投票意願的關係上，拒訪樣本會順應成功樣本的趨勢（圖 4-1）。此外，在多變數模型裡，許多控制變數的指標也付之闕如。最後，高投票意願者比低投票意願者更具政治敏感性；本文無法確定這會抵消多少主題效應的影響。這些因素都增加了推論的不確定性。

雖有許多研究限制，本文仍希望拋磚引玉，吸引大家對民調高估投票率的注意，共同加入這個主題的研究，為此本文接著提出一些建議。首先，如果未來的研究想更進一步確認主題效應與 TEDS 高估投票率的關係，資料限制仍是最大的挑戰。或許 TEDS 可從下列幾點來解決資料的問題：第一、善用既有資料。TEDS 可以考慮釋出 2002 至 2004P 的訪員評估結果，以及歷次 TEDS 的訪員個人資料，譬如：性別、訪問經驗、語言能力……等。在缺乏拒訪者資料的情況下，這些仍是相關研究的重要資訊。第二、蒐集拒訪者的資料。TEDS2001 曾用較簡短的問卷蒐集拒訪者的資料。雖然成功率不高，卻是相當難得的資料。TEDS 可以考慮再次執行這種短卷調查。除此之外，TEDS 也許可以使用郵寄問卷、網路訪問或電話訪問來蒐集面訪拒訪者的資料。<sup>25</sup> 這些調查方式的成本較低，受訪者面臨的訪問壓力也較小，能夠提高他們的受訪意願。第三、進行主題效應的調查實驗。舉例來說，TEDS 可以考慮把受訪者隨機分為兩組，控制組接受的訪問仍名為「台灣選舉與民主化調查」，實驗組接受的訪問則用其他和政治沒有直接關係的詞彙命名。在此之外，兩組調查的方式與內容都完全相同。透過兩組資料的比較就能瞭解主題效應的影響，譬如：比較兩組資料在投票率的估計值或其他政治參與的估計值上是否有顯著差異，也能比較這種差異會不會影響各種政治參與的多變數分析（政治參與可以是自變數，也可以是依變數）。另外，還可以利用更複雜的實驗設計同時操控多個實驗變數，屆時就能夠解答政

<sup>25</sup> 或許訪員在確定被拒訪後可以馬上詢問拒訪者是否願意改用網路或電話訪問並留下連絡方式，藉以爭取蒐集拒訪者資料的機會，不願意者仍可利用郵寄問卷嘗試接觸。

治敏感度的問題是否會抵消主題效應對於投票率估計值的影響。這些都是未來 TEDS 在規劃時可以納入考慮的。

如果確定了主題效應有不可忽略的影響力，研究焦點就應轉向如何提高對政治沒興趣者的受訪意願。全面更改調查名稱是不切實際的，這不僅有違研究倫理，被蒙蔽而接受訪問的人在接觸到訪問內容後，他們的答題品質也可能會降低。由於調查主題只是受訪者決定是否接受訪問的其中一項依據，TEDS 可以在中選通知信、來訪未遇單、標準開頭語、訪員的教戰手則……等地方增加其他資訊或調整各項資訊的比例，藉以降低主題的影響力。舉例來說，強調 TEDS 是「行政院國家科學委員會的全國大型研究計畫」也許能夠壓低調查主題的相對比重，也可能有助於提高 TEDS 在受訪者心目中的正當性。除了心理誘因之外，實質誘因也很重要。與其再設訂更繁重的接觸規則，並以更高的酬勞彌補訪員的負擔，不如試著增加受訪者接受訪問的實質報酬，譬如：把禮物換為等值但應用彈性更大的禮券或現金。這在降低拒訪傾向的同時，其實也達到減輕訪員工作負擔的附加效果。

其實這些方法大多都是老生常談，不同的是，本文並不一味主張追求高受訪率，因為這不是解決主題效應的保證；盲目的提升受訪率不僅成本高、效果差，甚至可能適得其反。試想，如果提升受訪意願的努力大部分是作用在那些不關心政治的人身上那將會如何？如果對政治沒興趣的人正好也是厭惡政府的人，強調 TEDS 與行政院的關係勢必會有反效果；如果對政治沒興趣的人又正好是收入較高的人，現金的吸引力也不見得會比較大。為此，本文認為要針對那些不關心政治的人投其所好，優先提高他們的受訪意願。正因如此，未來的重要任務將是測量受訪者對於政治及 TEDS 的興趣，藉此找出目標群體的特徵，如此才能知道要對誰投其所好、又要如何投其所好。或許 TEDS 也可以考慮運用調查實驗來蒐集這些資料。

高估投票率幾乎可說是民意調查的通病，再加上主題效應的影響，以政治或選舉為主題的民調更難逃脫這個問題的糾纏，所以 TEDS 有此問題並不令人意外。然而，這並不表示 TEDS 的資料品質不佳；它只是指出一個值得深入探討的課題。TEDS 是台灣學界相當重要的調查計畫，絕對值得大家投入心力使其更加完備。

\* \* \*

投稿日期：98.09.14；修改日期：99.01.20；接受日期：99.09.02

## 附錄一：資料來源

簡稱	計畫名稱	國科會計畫代碼	主持人
TEDS 2001	「台灣選舉與民主化調查：民國九十年立法委員選舉全國大型民意調查研究」	NSC90-2420-H 194-001	黃紀 教授
TEDS 2002	「2002年至2004年『選舉與民主化調查』三年期研究規劃(I)：民國九十一年北高兩市選舉大型面訪案」	NSC91-2420-H 194-001-SSS	黃紀 教授
TEDS 2003	「2002年至2004年『選舉與民主化調查』三年期研究規劃(II)：民國九十二年民主化與政治變遷民調案」	NSC92-2420-H 001-004	朱雲漢 教授
TEDS 2004P	「2002年至2004年『選舉與民主化調查』三年期研究規劃(III)：民國九十三年總統大選民調案」	NSC92-2420-H 031-004	黃秀端 教授
TEDS 2004L	「2002年至2004年『選舉與民主化調查』三年期研究規劃(IV)：民國九十三年立法委員選舉大型面訪案」	NSC93-2420-H 004-005-SSS	劉義周 教授
TEDS 2005NA	「2005年台灣選舉與民主化調查：任務型國民大會代表選舉全國電話訪問案」	NSC92-2419-H 001-001-B9403	黃紀 教授
TEDS 2005M	「2005年至2008年『選舉與民主化調查』四年期研究規劃(I)：民國九十四年縣市長選舉大型面訪案」	NSC94-2420-H 004-008-SSS	黃紀 教授
TEDS 2006C	「2005年至2008年『選舉與民主化調查』四年期研究規劃(II)：民國九十五年北高兩市選舉面訪案」	NSC95-2420-H 031-011	黃秀端 教授
TEDS 2008L	「2005年至2008年『選舉與民主化調查』四年期研究規劃(III)：民國九十七年立法委員選舉面訪計畫案」	NSC96-2420-H 002-025	朱雲漢 教授
TEDS 2008P	「2005年至2008年『選舉與民主化調查』四年期研究規劃(IV)：民國九十七年總統選舉面訪計畫案」	NSC96-2420-H 004-017	游清鑫 教授
TSCS 411	台灣社會變遷基本調查計畫 第四期 第一次調查 問卷一：綜合問卷組	NSC89-2412-H 001-014-B1	章英華 教授
TSCS 431	台灣社會變遷基本調查計畫 第四期 第三次調查 問卷一：社會階層組	NSC91-2420-H 001-004	章英華 教授
TSCS 442	台灣社會變遷基本調查計畫 第四期 第四次調查 問卷二：國家認同組	NSC92-2420-H 001-002-B1	章英華 教授
TSCS 451	台灣社會變遷基本調查計畫 第四期 第五次調查 問卷一：公民權組	NSC93-2420-H 001-002-B1	章英華 教授
TSCS 511	台灣社會變遷基本調查計畫 第五期 第一次調查 問卷一：綜合問卷組	NSC94-2420-H 001-008-B1	章英華 教授
TSCS 512	台灣社會變遷基本調查計畫 第五期 第一次調查 問卷二：工作與生活組	NSC94-2420-H 001-008-B1	章英華 教授
TSCS 522	台灣社會變遷基本調查計畫 第五期 第二次調查 問卷二：公民與國家組	NSC95-2420-H 001-006-B1	傅仰止 教授
TSCS 531	台灣社會變遷基本調查計畫 第五期 第三次調查 問卷一：社會階層組	NSC96-2420-H 001-002	張苙雲 教授

資料來源：作者自行整理製表。

說明：以上為本文所使用的 TEDS 與 TSCS 資料，更詳細的資訊請參閱 TEDS 網頁：<http://www.tedsnet.org> (台灣選舉與民主化調查 2009)，以及 TSCS 網頁：<http://www.ios.sinica.edu.tw/sc> (台灣社會變遷基本調查 2009)。非常感謝上述機構及相關人員提供資料協助，惟本文之內容概由作者自行負責。

## 附錄二：TEDS2008 電訪 / 面訪結果

AAPOR 電訪 / 面訪結果分類	編號	TEDS2008L 電訪結果	n	TEDS2008L 面訪結果	n
1. Interview	1.000				
Complete (I)	1.100	訪問成功	3,843	訪問成功	1,240
2. Eligible, Non-Interview	2.000				
Household-level refusal (R)	2.111	接電話者即拒訪	2,723	家人代為拒訪	247
Known respondent refusal (R)	2.112	受訪者拒訪（無法再訪者）	153	受訪者本人拒絕一切訪問	548
Break-off (R)	2.120	受訪者中拒（非當日約訪者） 受訪者中拒（無法再訪者）	96 1,202	受訪者本人中途拒訪	18
Non-contact (NC)	2.200	受訪者訪問期間不在 受訪者不在（非當日約訪者）	130 1,988		
Unable to enter building/ reach housing unit (NC)	2.230			不得其門而入	49
No one at residence (NC)	2.240			無人在家	293
Respondent away/ unavailable (NC)	2.250			短暫外出、工作、求學、遊玩 暫時不方便接受訪問 外地工作、求學、遊玩 出國工作、求學、遊玩 籍在人不在（寄住戶） 原住戶搬走（遷移） 訪問期間不方便接受訪問 外出不知去向、失蹤 服役中（當兵） 受訪者出嫁 受訪者出家	81 3 258 90 193 133 40 20 16 13 1
Other (O)	2.300			訪員未依規定訪問（舞弊）	2
Dead (O)	2.310			受訪者死亡	6
Physically or mentally unable/ incompetent (O)	2.320	因生理因素無法受訪	195	身體、心理因素無法受訪	31
Language (O)	2.330	因語言因素無法受訪	63	語言不通	1

AAPOR 電訪 / 面訪結果分類	編號	TEDS2008L 電訪結果	n	TEDS2008L 面訪結果	n
<b>3. Unknown Eligibility, Non-Interview</b>					
Unknown if housing unit (UH)	3.100			房子改建中	5
Always busy (UH)	3.120	電話中	541		
No answer (UH)	3.130	無人接聽	6,589		
Telephone answering device (don't know if housing unit) (UH)	3.140	答錄機	220		
Unable to reach/unsafe area (UH)	3.170			放棄，因工作環境惡劣	9
Unable to locate address (UH)	3.180			查無此地址	28
Housing unit, unknown if eligible respondent (UO)	3.200	無法確定是否有合格受訪者	6		
<b>4. Not Eligible</b>					
Out of sample	4.100	已訪問過或非受訪地區	109		
Fax/data line	4.200	傳真機	1,426		
Non-working/disconnected number	4.300	電話停話改號故障空號	8,108		
Business, government office, other organization	4.510	宿舍機關公司營業用電話	1,296		
Regular, vacant residences	4.610			空屋、廢墟	35
No eligible respondent	4.700	戶中無合格受訪對象	86	移民	2
				服刑中（坐牢）	3
				該地址查無此人	91
Quota filled	4.800	配額已滿	1		

資料來源：作者根據 AAPOR 之電訪 / 面訪結果分類準則以及 TEDS2008L 的訪問結果自行整理製表。

說明：因空間限制，本表未列出 AAPOR 全部的訪問結果類別，只列 TEDS2008L 有發生的類別。完整的 AAPOR 訪問結果類別請見該單位出版的 *Standard Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys* (5th ed.)。

## 附錄三：訪問評估

	問項內容	本文的編碼方式
訪 問 前	1. 在訪問開始前，受訪者是否有下列情形？（可複選）	
	01. 表示自己太忙，無法受訪	這是一個複選題。除了「10. 沒有前述情形」之外，本文把每個選項都處理成二分變數 (binary variables)；每個變數都有兩個類別： 00. 代表沒發生這種情況 01. 代表有發生這種情況
	02. 表示因為家裡的事在忙，無法受訪	
	03. 表示對訪問不感興趣	
	04. 表示不信任調查訪問	
	05. 對調查訪問的用途有疑慮	
	06. 對受訪意見及個人資料能否保密有疑慮	
	07. 對訪問單位及主持人有疑慮	
	08. 質疑訪員身份	
	09. 曾有家人反對受訪者接受訪問	
10. 沒有前述情形		
訪 問 時	2. 在訪問進行中，受訪者是否有下列情形：	
	2a. 受訪者有無要求自行填寫問卷？	記錄訪問時狀況的問項也都是二分變數： 00. 代表沒發生這種情況 01. 代表有發生這種情況
	2b. 受訪者有無要求看問卷？	
	2c. 受訪者有無中途拒訪的傾向？	
	2d. 受訪者有無表示問卷題目太多？	
	2e. 受訪者有無表示題意不清或定義不明？	
	2f. 受訪者有無覺得問卷題目太難？	
2g. 受訪者有無表示題目過於敏感？		
整 體 訪 問	3. 您覺得受訪者的合作態度是：	00. 非常合作    03. 有點不合作 01. 相當合作    04. 相當不合作 02. 有點合作    05. 非常不合作
	4. 您覺得受訪者對問題的瞭解是：	00. 非常瞭解    03. 有點不瞭解 01. 相當瞭解    04. 相當不瞭解 02. 有點瞭解    05. 非常不瞭解
	5. 您覺得受訪者對問題的敏感（有政治性的敏感、抗拒或恐懼）是：	00. 非常不敏感    03. 有點敏感 01. 相當不敏感    04. 相當敏感 02. 有點不敏感    05. 非常敏感
	6. 您覺得受訪者對訪問內容的興趣是：	00. 非常有興趣    03. 不太有興趣 01. 相當有興趣    04. 相當沒興趣 02. 有點興趣    05. 非常沒興趣
	7. 您覺得受訪者回答問題的可信度是：	00. 很可信    02. 有些不可信 01. 大致可信    03. 大部分不可信
	8. 您覺得受訪者對於一般政治事務訊息的知悉程度是：	00. 很高    02. 低 01. 高    03. 很低

資料來源：作者根據 TEDS 的問卷內容自行整理製表。

## 附錄四：歷年選舉之投票率估計值

選舉 資料	1994 直轄 市長	1996 總統	1997 縣市 長	1998 直轄 市長	1998 立委	2000 總統	2001 立委	2001 縣市 長	2002 直轄 市長	2004 立委	2004 總統	2005 縣市 長	2006 直轄 市長	2008 立委	2008 總統
TEDS 2001		90.8%	87.4%	87.5%		93.6%	82.4%	82.4%							
TEDS 2002Cks	92.2%				93.3%	94.0%	89.4%		89.1%						
TEDS 2002Ctp	89.9%				92.6%	92.0%	88.1%		87.5%						
TEDS 2003						90.8%	84.5%								
TEDS 2004P		88.0%				90.9%	83.5%				91.2%				
TEDS 2004La										77.9%	90.1%				
TEDS 2004Lb						92.5%	87.6%			77.3%	91.3%				
TEDS 2005Mks														79.0%	
TEDS 2005Mtc														78.6%	
TEDS 2005Mtp														80.3%	
TEDS 2005Myl														87.9%	
TEDS 2006Cks									87.1%		89.9%		84.6%		
TEDS 2006Ctp									85.1%		87.5%		79.8%		
TEDS 2008L										84.1%	87.4%			73.5%	
TEDS 2008P						89.4%				84.3%	87.5%			79.9%	88.4%
TSCS 411						92.8%									
TSCS 431							85.6%								
TSCS 442						88.9%	83.4%								
TSCS 451						87.5%	82.4%				86.6%				
TSCS 511										73.6%	87.2%				
TSCS 512										70.1%					
TSCS 522										81.3%					
TSCS 531													81.9%	79.5%	

資料來源：作者使用 TEDS 和 TSCS 的資料自行統計和製表。



## 附錄五：勝算對數模型的變數操作化

1. 社會層次的因素	
1-1. TEDS 的虛擬變數	00【04La】 01【04Lb】 02【05Mtp】 03【05Mtc】 04【05Myl】 05【05Mks】 06【06Ctp】 07【06Cks】 08【08L】 09【08P】
1-2. 距離投票日的天數	受訪日期減投票日期，並取自然對數。
1-3. 受訪者戶籍所在地區	00【北】：台北縣市、宜蘭縣、桃園縣、新竹縣市、基隆市 01【中】：苗栗縣、台中縣市、彰化縣、南投縣、雲林縣 02【南】：嘉義縣市、台南縣市、高雄縣市、屏東縣、澎湖縣 03【東】：台東縣、花蓮縣
1-4. 訪問處的社區型態	00【住宅區】 01【住商混合】 02【農村】 03【眷村】 04【國宅】 05【其他（含工業區及公教宿舍）】
2. 調查設計的性質	
2-1. TEDS 的虛擬變數	與變數 1-1 是同一組變數。
3. 受訪者的特徵	
3-1. 年齡	訪問年減出生年，並取自然對數。
3-2. 性別	00【女】 01【男】
3-3. 省籍	父親的省籍為主、母親的為輔，歸為以下 4 類： 00【閩南】 01【客家】 02【大陸各省】 90【其他（含原住民）】
3-4. 在家常用語言	以複選題的處理方式建立四個二分變數： 【國語】、【閩南語】、【客家語】、 【其他（含原住民語、外省方言及外國語言）】。 在每個變數裡，類別 0 都是「在家不常用 __ 語」，類別 1 則是「在家常用 __ 語」。例如，若受訪者的答案是「國台語並用」，他在國語的和閩南語的虛擬變數裡都會被編碼成 1。
3-5. 教育程度	00【國小及以下】 01【國初中】 02【高中職】 03【大專】 04【研究所及以上】
3-6. 訪問處的社區型態	同 1-4。
3-7. 住家型式	00【公寓】 01【電梯大樓】 02【透天厝】 03【其他（含公教宿舍）】
3-8. 政黨認同	綜合「有無偏向哪個政黨」與「有無稍微偏向哪個政黨」兩個問項歸為以下 3 類（TEDS2004La 沒有這兩個問項，改用「通常支持哪個政黨」與「情感偏向哪個政黨」）： 00【泛藍】：國民黨、新黨、親民黨、泛藍 01【泛綠】：民進黨、建國黨、台聯、泛綠 02【都不認同】：沒有偏向任何政黨，連稍微偏向也沒有 【遺漏值】：其他政黨（總共只有 24 筆觀察值）。
3-9. 投票意願	00【沒投】：代表投票意願較低的人 01【有投（含投廢票）】：代表投票意願較高的人 【遺漏值】：忘了。

4. 訪員的特質			
沒有可以使用的變數			
5. 訪員和受訪者的互動			
5-1. 接觸方式	00【警衛】	01【對講機】	02【門鈴或敲門】
	03【直接詢問/進入】	90【其他】	
5-2. 訪問時有無別人在場	00【沒有】	01【有】	
5-3. 訪問語言	以複選題的處理方式建立三個二分變數： 【國語】、【閩南語】、【客家語】 在每個變數裡，類別 0 都是「訪問沒用到__語」，類別 1 則是「訪問有用到__語」。		

資料來源：作者自行製作。

說明：1. 把各變數裡的類別 0 當作參照類。

2. 把觀察值次數過少的類別併入「其他」，避免影響模型的估計，譬如：社區型態裡的「工業區」與「公教宿舍」、住家型式裡的「公教宿舍」、省籍裡的「原住民籍」……。不過在政黨認同裡，合併認同泛藍、泛綠以外的政黨後，觀察值次數仍然太少，只有 24 筆，所以刪去不使用。
3. 無反應選項，如：「拒答」、「很難說」、「不知道」……等皆視為遺漏值。
4. 在「有家人反對受訪者接受訪問」的模型裡自變數稍有不同。「其他省籍」會與其他變數發生完全共線性的問題導致模型無法估計，為此我們把「其他省籍」的觀察值設為遺漏值（損失 102 筆觀察值，約為 1.03%），省籍變數因而只剩下 3 類：本省客家、本省閩南、外省。

## 附錄六：推廣順序勝算對數模型

拒訪傾向	投票意願的估計值		Wald $\chi^2$ (df)	Pseudo $R^2$	NPP	n
	$\beta$ (Robust S.E.)	Odds Ratio				
<b>合作程度</b>			1267.47(164)***	0.05	86	9,719
(4+3+2+1) / (0)	-0.1356 (0.0567)*	0.8732				
(4+3+2) / (1+0)	-0.1203 (0.0690)	0.8866				
(4+3) / (2+1+0)	-0.2810 (0.1152)*	0.7550				
(4) / (3+2+1+0)	-0.3357 (0.2288)	0.7149				
<b>對問題的瞭解程度</b>			3562.90(210)***	0.13	68	9,716
(5+4+3+2+1) / (0)	-0.1358 (0.0737)	0.8730				
(5+4+3+2) / (1+0)	-0.4146 (0.0605)***	0.6606				
(5+4+3) / (2+1+0)	-0.5163 (0.0842)***	0.5967				
(5+4) / (3+2+1+0)	-0.7150 (0.1258)***	0.4892				
(5) / (4+3+2+1+0)	-0.6612 (0.2412)**	0.5162				
<b>對問題的敏感程度</b>			980.28(210)***	0.03	13	9,708
(5+4+3+2+1) / (0)	0.1565 (0.0894)	1.1694				
(5+4+3+2) / (1+0)	0.1893 (0.0585)**	1.2084				
(5+4+3) / (2+1+0)	0.2526 (0.0571)***	1.2873				
(5+4) / (3+2+1+0)	0.3492 (0.0933)***	1.4180				
(5) / (4+3+2+1+0)	0.1601 (0.1753)	1.1737				
<b>對訪問內容感興趣的程度</b>			2247.09(210)***	0.08	5	9,717
(5+4+3+2+1) / (0)	-0.4640 (0.1776)**	0.6288				
(5+4+3+2) / (1+0)	-0.4314 (0.0792)***	0.6496				
(5+4+3) / (2+1+0)	-0.5330 (0.0584)***	0.5868				
(5+4) / (3+2+1+0)	-0.6892 (0.0865)***	0.5020				
(5) / (4+3+2+1+0)	-0.7826 (0.1354)***	0.4572				
<b>回答問題的可信度</b>			1614.69(126)***	0.09	0	9,713
(3+2+1) / (0)	-0.2526 (0.0631)***	0.7768				
(3+2) / (1+0)	-0.3419 (0.0814)***	0.7104				
(3) / (2+1+0)	-0.3662 (0.1580)*	0.6934				
<b>對一般政治事務訊息的知悉度</b>			3125.79(126)***	0.15	0	9,709
(3+2+1) / (0)	-0.3206 (0.1010)**	0.7257				
(3+2) / (1+0)	-0.6836 (0.0610)***	0.5048				
(3) / (2+1+0)	-0.8205 (0.1065)***	0.4402				

資料來源：作者利用 TEDS2004La、04Lb、05M、06C、08L、08P 的面訪獨立樣本自行分析與製表。

說明：內文表 3 有關整體訪問情況的分析都採用「順序勝算對數模型」(ordered logit model)，這種模型的解釋雖然簡便，卻得符合「平行迴歸假定」。「推廣順序勝算對數模型」(generalized ordered logit model) 可放寬假定，但在解釋上較為麻煩且可能估計出負的機率值。兩種模型各有優缺點，故本文在內文呈現

順序勝算對數模型的估計，推廣順序勝算對數模型的則放在附錄六，其中的「NPP(Negative predicted probability)」就是指有多少個樣本遇到負機率值的問題。基本上，附錄六的變數操作化方式與內文表 3 相同，但推廣順序勝算對數模型要估計的係數更多。本文發現有些變數類別的樣本數太少導致估計值異常的大或小，故對樣本數少於 100 的類別作處理。第一、合作程度（依變數）有 66 筆的「非常不合作」併入「相當不合作」。第二、社區型態（自變數）有 98 筆的「其他（含工業區及公教宿舍）」，以及接觸方式（自變數）有 98 筆的「其他」設為遺漏值。第三、不使用「在家常用原住民語 / 其他外省語 / 外國語」這個自變數（79 筆）。由於表格太多，附錄六也只呈現高投票意願者對低投票意願者在拒訪指標上的勝算比（完整數據請洽作者）。此外，在內文表 3 裡無論依變數有幾個類別都只會估計出 1 個勝算比，放寬假定的附錄六則是依變數有  $K$  個類別就會有  $K-1$  個勝算比。以回答問題之可信度為例， $(3+2+1)/(0)$  是指「很不可信、不可信、可信」的勝算，亦即「很不可信 + 不可信 + 可信的機率 / 很可信的機率」； $(3+2)/(1+0)$  是指「很不可信、不可信」的勝算； $(3)/(2+1+0)$  則指「很不可信」的勝算。儘管附錄六的估計值與內文表 3 有些差異，但投票意願與各個拒訪指標的相關方向以及係數的顯著情況都類似，故本文仍維持原來的結論。

## 參考文獻

### I. 中文部分

- 中央選舉委員會，2009，〈選舉資料庫〉，中選會選舉資料庫網站：<http://210.69.23.140/cec/cehead.asp>，檢索日期：2009年9月14日。
- (Central Election Commission. 2009. "Election Database." Central Election Commission Database. <http://210.69.23.140/cec/cehead.asp> [accessed September 14, 2009].)
- 王保進，2006，〈中文視窗版 SPSS 與行為科學研究〉，台北：心理。
- (Wang, Bao-jin. 2006. *SPSS for Windows and Behavioral Science Research*. Taipei: Psychological.)
- 王濟川、郭志剛，2003，〈Logistic 迴歸模型：方法及應用〉，台北：五南。
- (Wang, Ji-chuan, and Zhi-gang Guo. 2003. *Logistic Regression Models: Methods and Application*. Taipei: Wunan.)
- 台灣社會變遷基本調查，2009，〈問卷與報告書〉，台灣社會變遷基本調查網站：<http://www.ios.sinica.edu.tw/sc/cht/scDownload2>，檢索日期：2009年09月14日。
- (Taiwan Social Change Survey. 2009. "Questionnaire and Report." The Website of Taiwan Social Change Survey. <http://www.ios.sinica.edu.tw/sc/en/scDownload2.php#fifth> [accessed September 14, 2009].)
- 台灣選舉與民主化調查，2009，〈TEDS 歷年計畫〉，台灣選舉與民主化調查網站：<http://www.tedsnet.org/cubekm1/front/bin/ptlist.phtml?Category=133>，檢索日期：2009年09月14日。
- (Taiwan's Election and Democratization Study. 2009. "Chronology of Major Events of the TEDS Project." The Website of Taiwan's Election and Democratization Study. <http://www.tedsnet.org/cubekm2/front/bin/ptlist.phtml?Category=11> [accessed September 14, 2009].)
- 吳俊德、陳永福，2005，〈投票與不投票的抉擇——2004年總統大選與公民投票的探索性研究〉，《台灣民主季刊》，2(4): 67-98。
- (Wu, Jun-deh, and Alexander C. Tan. 2005. "To Vote or Not to Vote, That Is the Question: An Exploratory Analysis of the Presidential and Referenda Vote in Taiwan." *Taiwan Democracy Quarterly* 2(4): 67-98.)
- 洪永泰，2003，〈原始樣本、替代樣本、與追蹤樣本的比較：「2001年台灣選舉與民主化調查研究」訪問失敗問題的探討〉，《選舉研究》，10(2): 37-58。

- (Hung, Yung-tai. 2003. "The Nonresponse Problems of the 2001 TEDS Survey." *Journal of Electoral Studies* 10(2): 37-58.)
- 陳敦源，2000，〈人爲何投票？——理性選擇觀點的緣起與發展〉，《民意研究季刊》，212: 31-64。
- (Chen, Don-yun. 2000. "Why People Vote? The Origin and Development of the Rational Choice Perspective." *Public Opinion Research Quarterly* 212: 31-64.)
- 楊孟麗，2003，〈投票意願與經濟不景氣：台灣的情形〉，《選舉研究》，10(2): 159-191。
- (Yang, Meng-li. 2003. "Voter Turnout and Economic Adversity in Taiwan." *Journal of Electoral Studies* 10(2): 159-191.)
- 謝邦昌、江志民，1998，〈民意測驗中的社經發展指標、人口特性與投票行爲之研究〉，《民意研究季刊》，204: 26-45。
- (Shia, Ben-chang, and Chih-ming Chiang. 1998. "Study on the Index of Socioeconomic Development Population Characteristics and Voting Behavior in Public Opinion Research." *Public Opinion Research Quarterly* 204: 26-45.)

## II. 外文部分

- American Association for Public Opinion Research. 2008. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (5th ed.). Lenexa, KS: AAPOR.
- Belli, Robert F., Sean E. Moore, and John VanHoewyk. 2006. "An Experimental Comparison of Question Forms Used to Reduce Vote Overreporting." *Electoral Studies* 25(4): 751-759.
- Belli, Robert F., Michael W. Traugott, Margaret Young, and Katherine A. McGonagle. 1999. "Reducing Vote Overreporting in Surveys: Social Desirability, Memory Failure, and Source Monitoring." *Public Opinion Quarterly* 63(1): 90-108.
- Bernstein, Robert, Anita Chadha, and Robert Montjoy. 2001. "Overreporting Voting: Why It Happens and Why It Matters." *Public Opinion Quarterly* 65(1): 22-44.
- Brady, Henry E., Sidney Verba, and Kay Lehman Schlozman. 1995. "Beyond SES: A Resource Model of Political Participation." *American Political Science Review* 89(2): 271-294.
- Brehm, John. 1993. *The Phantom Respondents: Opinion Surveys and Political Representation*. Ann Arbor, MI: The University of Michigan Press.
- Burden, Barry C. 2000. "Voter Turnout and the National Election Studies." *Political Analysis* 8(4): 389-398.
- Cheng, Su-feng. 2006. "Survey Participation in Taiwan: Evidence from the ESC Survey."

- Journal of Electoral Studies* 13(1): 163-195.
- Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. Chicago, IL: The University of Chicago Press.
- Evans, Jocelyn A.J. 2004. *Voters & Voting: An Introduction*. Thousand Oaks, CA: Sage.
- Fullerton, Andrew S., Jeffrey Dixon, and Casey Borch. 2007. "Bringing Registration into Models of Vote Overreporting." *Public Opinion Quarterly* 71(4): 649-660.
- Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70(5): 646-675.
- Groves, Robert M., Robert B. Cialdini, and Mick P. Couper. 1992. "Understanding the Decision to Participate in a Survey." *Public Opinion Quarterly* 56(4): 475-495.
- Groves, Robert M., Floyd J. Fowler, Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2004. *Survey Methodology*. Hoboken, NJ: John Wiley and Sons.
- Groves, Robert M., Mick P. Couper, Stanley Presser, Eleanor Singer, Roger Tourangeau, Giorgina P. Acosta, and Lindsay Nelson. 2006. "Experiments in Producing Nonresponse Bias." *Public Opinion Quarterly* 70(5): 720-736.
- Hamilton, Lawrence C. 2004. *Statistics with Stata: Updated for Version 8*. Belmont, CA: Brooks/Cole-Thomson.
- Health, Anthony, and Bridget Taylor. 1999. "New Sources of Abstention." In *Critical Elections*, eds. Geoffrey Evans and Pippa Norris. London: Sage.
- McDonald, Michael P. 2003. "On the Overreport Bias of the National Election Study Turnout Rate." *Political Analysis* 11(2): 180-186.
- 2007. "The True Electorate: A Cross-Validation of Voter Registration Files and Election Survey Demographics." *Public Opinion Quarterly* 71(4): 588-602.
- Merkle, Daniel, Murray Edelman, Kathy Dykeman, and Chris Brogan. 1998. "An Experimental Study of Ways to Increase Exit Poll Response Rates and Reduce Survey Error." Presented at the Annual Conference of the American Association for Public Opinion Research, St. Louis, Missouri.
- Olson, Kristen. 2006. "Survey Participation, Nonresponse Bias, Measurement Error Bias, and Total Bias." *Public Opinion Quarterly* 70(5): 737-758.
- Peytchev, Andy, Rodney K. Baxter, and Lisa R. Carley-Baxter. 2009. "Not All Survey Effort is Equal: Reduction of Nonresponse Bias and Nonresponse Error." *Public Opinion Quarterly* 73(4): 785-806.

- Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. "Who Overreports Voting?" *American Political Science Review* 80(2): 613-624.
- Stocké, Volker. 2007. "Response Privacy and Elapsed Time since Election Day as Determinants for Vote Overreporting." *International Journal of Public Opinion Research* 19(2): 237-246.
- Stocké, Volker, and Tobias Stark. 2007. "Political Involvement and Memory Failure as Interdependent Determinants of Vote Overreporting." *Applied Cognitive Psychology* 21(2): 239-257.
- Tang, Ge, Yeh-diing Wang, and Rene Bautista. 2009. "A Comparison Study of Exit Poll Methodologies between Taiwan and the U.S." Presented at the Annual Conference of the American Association for Public Opinion Research, Hollywood, Florida.
- Tsai, Chia-hung. 2001. "Why Do Taiwanese Vote?" *Journal of Electoral Studies* 8(2): 125-154.
- Voogt, Robert J.J., and Willem E. Saris. 2003. "To Participate or Not to Participate: The Link between Survey Participation, Electoral Participation, and Political Interest." *Political Analysis* 11(2): 164-179.
- Wolfinger, Raymond E., and Steven J. Rosenstone. 1980. *Who Votes?* New Haven, CT: Yale University Press.
- Wu, Chung-li. 2006. "Vote Misreporting and Survey Context: The Taiwan Case." *Issues & Studies* 42(4): 223-239.



# **Don't Ask Me! I'm Not Interested in Politics: An Analysis of Topic Effect and the Turnout Overestimate in TEDS**

Chi-lin Tsai\*

## **Abstract**

Literature of “topic effect” suggests that people who are interested in the survey topic are more willing to participate in the interview, so the estimates of the survey items that are associated with the topic tend to be exaggerated. Given that politics and election are the focal points of Taiwan’s Election and Democratization Study (TEDS), this paper explores whether the topic effect is one of the major causes to overestimate turnout rates in TEDS. The finding of the aggregate-level analysis shows that the turnout overestimate in TEDS is more serious than that in Taiwan Social Change Survey (TSCS). Since two survey projects have a lot in common except topics, their difference in turnout estimates may result from the topic effect. The individual-level analysis also indicates that the correlation between respondents’ propensity to vote and their propensity to participate in TEDS is positive, which conforms to the theoretical expectation of topic effect. Generally speaking, the hypotheses are confirmed by both analyses. However, the conclusions are based on some untested assumptions; therefore future TEDS could use special survey design to further check and cope with the topic effect and turnout overestimate.

Keywords: topic effect, turnout overestimate, over-representative, over-report

---

\* Master, Department and Graduate Institute of Political Science, National Chung Cheng University.