

「固定樣本連續訪談法」 (panel studies) 樣本流失問題之探討*

陳陸輝**

《 本文摘要 》

在固定樣本連續訪談法 (panel studies) 中，受訪者非隨機性的流失 (mortality)，對研究的效度會產生負面的影響。本文作者利用美國全國選舉研究 (American National Election Studies ; ANES) 在西元 1990 到 1992 年進行的固定樣本連續訪問的資料進行分析，發現了接受續訪的樣本 (panel sample) 與流失的樣本 (mortality sample)，在政治興趣、政治知識以及政治參與上有著顯著的差異。因此，當我們的研究目的是在推論母體在政治興趣、政治知識以及政治參與的分佈時，「樣本流失」的問題會對研究的「外在效度」 (external validity) 產生顯著的影響。

當影響受訪者是否參與後續訪問的因素，與我們研究模型中的變數相互關聯時，「樣本流失」也會對研究的「內在效度」 (internal validity) 造成傷害。本文作者檢視「樣本流失」對投票參與模型中，解釋變數的參數估計的影響時發現：「樣本流失」使我們低估了受訪者年齡以及對報上選舉消息注意程度這兩個變數，對其投票與否的影響力。當我們要解釋國會選舉中現任者的優勢時，「樣本流失」使我們高估了民主黨議員現任者優勢的重要性，也低估了共和黨現任議員的重要性。

由於受訪者對調查研究的興趣以及態度，會影響他們參與訪談的意願，因此，作者認為，調查研究工作者應該努力讓受訪者覺得接受訪問是一個愉快經驗，減少調查研究的拒訪率，使研究結果更為精確。

關鍵詞：內在效度、外在效度、樣本流失、固定樣本連續訪談法、政治參與、調查研究。

* 作者感謝 Darren Davis 對文章初稿提出的建議。本文曾宣讀於政治大學選舉研究中心主辦「世紀末的選舉」學術研討會 (1998年10月17及18日，政大公企中心國際會議廳)，承蒙論文評論人鄭宇庭教授以及與會學者提出諸多指正，作者深表感激。此外，兩名審查委員亦提供凱切的建議，在此一併致謝。如果文中仍有疏漏之處，自當歸咎於作者。本文所使用的訪問資料可經由 Inter-university Consortium for Political and Social Research (ICPSR) 取得，統計分析是使用 SPSS 軟體，分析的程式可向作者索取。

** 陳陸輝現為美國密西根州立大學政治系助教暨博士候選人。

樣本流失 (mortality) 是「固定樣本連續訪法」(panel studies) 研究效度的一大威脅。本文作者使用美國全國選舉研究 (American National Election Studies ; ANES) 的資料來探討流失樣本與續訪樣本在一些人口學變項、政治興趣、政治知識與政治參與上，是否存在顯著的差異，而影響研究的「外在效度」(external validity)。本文作者也提出一個模型來解釋為何在固定樣本連續訪談法中，受訪者不參加續訪的原因。研究發現，受訪者對參加訪問的興趣以及對訪問是否存在負面的態度，對他／她參與續訪的機率有很顯著的影響。而樣本流失的問題也影響我們研究中解釋變數的參數估計。所以，樣本流失的問題也會對研究的「內在效度」(internal validity) 有負面的作用。

壹、前言

正如 King, Keohane 以及 Verba 所言，科學研究的目的是在做推論 (make an inference) 「科學研究是以真實世界的經驗性資訊 (empirical information) 來作描述性與解釋性的推論 (descriptive or explanatory inferences)」(King, Keohane, and Verba, 1994 : 7)。而以經驗性資料來做推論的調查研究法，由於具有符號上地「科學性」與「代表性」(symbolically scientific and representative) 的特質，使得它成為民意研究的不同研究方法中，最熱門的研究法 (Herbst, 1993 : 38)。我們可以研究的觀察時間長短，將研究分為切面的 (cross-sectional) 與長期的 (longitudinal) 的研究。其中，「固定樣本連續訪談法」是長期研究中，瞭解個人態度的變化以及變數之間因果關係的最為適當的研究方法。在有關民意與選舉的研究中，我們可使用「固定樣本連續訪談法」來解釋民意的動力來源、政治態度的穩定與否以及政治社會化的過程 (Achen, 1975 ; Converse, 1964 ; Converse and Markus, 1979 ; Green and Palmquist, 1990 ; Jennings and Markus, 1984 ; Jennings and Niemi, 1981 ; Smith 1989)。

然而，「固定樣本連續訪談法」也有一些缺點。首先，由於該研究法需使用大筆人力與財力在不同時間 (也許一年或是很多年後) 重複訪問同一個受訪者，使得研究成本相當昂貴 (Kish, 1965 : 470 ; Mannheim and Rich, 1995 : 1330)。其次則是「污染」(contamination) 的問題。這是指「測試 (test) 或是訪問的問題破壞了受訪對象原先無知 (innocent) 或是無偏差 (unbiased) 的本質」(Kish, 1965 : 470)。例如，我們在1990年的訪問中，請受訪者指出一些政治人物目前所擔任的職務。這次訪問，可能對於受訪者在1992年回答類似的問題時，造成一些影響。因為受訪者有了1990

年受訪經驗的刺激後，也許會開始去注意一些政治人物目前擔任何種職務的資訊，而提高了他在1992年作答時的正確性。第三個問題則是「樣本流失」(mortality) 的問題。此指「固定樣本連續訪談」的研究中，具有不同特質的受訪者有著不同的流失程度(differential loss) (Campbell and Stanley, 1963 : 5)。在研究中，特定的實驗組(experimental group) 裡，不同背景的受訪者存在著不同的流失情形，則「樣本的流失」就會威脅研究的效度。「由於在前測(pretest) 中，實驗組是由不同的人所組成的，(「樣本流失」) 就造成了選樣的問題(selection artifact)」(Cook and Campbell, 1979 : 53)。「樣本流失」是一種「選樣的偏差」(selection bias)，它對我們研究的「內在效度」與「外在效度」皆具有殺傷力。

以1992年的「美國全國選舉研究」為例，該研究中有一半的成功受訪對象曾接受1990年的「美國全國選舉研究」以及1991年的「戰爭的政治後果的研究」的訪問。在該次(1990年到1992年) 的「固定樣本連續訪談」中，有1,980位受訪者參與1990年的訪談，在1992年的訪談中，1,359位繼續參加，其他的621位則流失。本文試圖探討，這些「流失樣本」與「續訪樣本」是否相同。如果「流失樣本」是隨機地(randomly) 流失，對研究的效度的影響不太。反之，如果兩種樣本具有不同的屬性，就會傷害我們的研究效度。首先，它會扭曲研究的「內在效度」。所謂「內在效度」，它是指「我們可以對變數之間的因果關係做出結論的程度」(Kidder et al., 1986 : 27)。此外，當「選樣偏差」與研究的變數之間存在的交互影響(interaction effects)，它更會危及我們的「外在效度」(Campbell and Stanley, 1963 : 5-6)。在這種情況下，我們無法將我們的研究，推論至真實的世界(generalize to the real world)。

貳、理論與研究假設：接受調查訪問是政治參與的一種方式

政治參與需要資源，而資源卻是不均等地分配給市民。因此，參與政治活動者，是不均等地來自於社會中教育程度較高、擁有更多資訊以及有錢的市民(Verba and Nie, 1972)。此外，擁有更多閒暇時間(free time) 以及公民技巧(civic skills) 的市民，也更傾向參與政治活動(Brady et al., 1995)。而運用隨機抽樣的原則以及適當的問卷設計所進行的調查訪問，使人民的意向得以透過調查訪問的方式而傳達。因此，人民參與調查訪問也是政治參與的一種方式(Verba, 1996)。不過，「接受調查訪問是不是也是需要一些資源？」以及「在『固定樣本連續訪談』的研究中，參與續訪的受訪者與流失的樣本之間，是否存在著差異？」是兩個很有趣的問題。本文的研究焦點是放在第二個問題，對於前一個問題，我們可以從 Brehm (1993) 的研究中發現，

「美國全國選舉研究」與「一般社會調查」(General Social Survey; GSS)的面訪樣本與「當前人口研究」(Current Population Studies; CPS)的普查結果相較,前兩者的受訪者中,年輕的、白人、男性以及富有的家庭的受訪者低於他們在實際人口中的比例(underrepresentative)。此外,「美國全國選舉研究」與「一般社會調查」的受訪者中,擁有高中畢業文憑的受訪者也比實際人口中的比例高(overrepresentative)。Brehm (1993: 38)認為,「調查訪問(survey)似乎是一個對政治上以及經濟上弱勢者的代表媒介(a vehicle for representation)」。我們想瞭解,是否在「固定樣本連續訪問」中,也有著同樣的現象。此外,Brehm也提出了一個模型來解釋為什麼受訪者要參與訪問。首先,他發現受訪者的「接觸難易程度」(accessibility)影響訪問的成功率。「老年人、家中有小孩以及較富裕的受訪者,比年輕人、較窮困以及家中無小孩者,來的容易接觸」(Brehm, 1993: 66)。在參與受訪上,受訪者本身的能力,也影響他參與訪談的意願。因此,訪員雖然容易接觸年長的受訪者,但後者較不願意參加訪問。此外,受訪者對陌生人、訪員、訪問及對自己的「自我印象」(self-image)等態度,都影響他接受訪問的機率(Brehm, 1993: ch. 4)。所以,我認為受訪者本身的政治興趣、政治參與的情況以及政治知識,都會影響他接受訪問的意願。特別對「固定樣本連續訪談」而言,受訪者已經有被訪問過的經驗,他們知道訪問是怎麼一回事。所以,他們對「訪問」這件事的態度,必定會影響他們參與訪談的意願。假設上次的訪談經驗是相當痛苦與無聊,他們必定不願意再接受另一次「折磨」。

社會科學家希望用隨機的抽樣調查來推論母體。不過,當「固定樣本連續訪談」的過程中,具有一定特徵的受訪者有系統地(systematically)流失時,我們要用這些樣本做推論就會出現一些問題。從Brehm的研究中可以發現,「當成功率不是百分之百,且受訪者與拒訪者(在一些背景上明顯地)不同時,訪問失敗率(nonresponse)會造成單變量統計(univariate statistics)的偏差」(Brehm, 1993: 94)。

社會科學家也希望解釋各變數之間的因果關係。而樣本流失的問題也會扭曲研究中的因果關係。在Brehm的書中,他描述「選樣的偏差」降低了研究的「內在效度」。「如果我們感興趣的方程式中的一個變數與人們參與訪談之間相互關連,我們對該變數的估計就會產生偏差」(Brehm, 1993: 117; 也請參閱Johnston and DiNardo, 1997: 436-441; King et al., 1994: ch. 4)。

政治參與需要資源以及一些「公民技巧」。參與「固定樣本連續訪談」是一種形式的政治參與,所以,也需要一些資源與技巧。我想探討在「固定樣本連續訪談」中影響人們參與訪談的因素。首先,是人口學背景是否會影響他們參與「固定樣本連續訪談」。我的假設是:男性、白種人、年長者、教育程度較高者以及家庭收入較高者,比

「固定樣本連續訪談法」(panel studies) 樣本流失問題之探討

較會參加續訪。其次，受訪者的政治興趣、政治知識、對其他政治活動的參與情況以及對訪問這件事的態度，都會影響他是否會接受訪談。我的假設是：受訪者有較高的政治興趣、較多的訊息、較常參與政治活動以及對訪問有正面的態度者，有較高的機率參與「固定樣本連續訪談」的接續訪問。

如果「續訪樣本」與「流失樣本」之間確實存在著顯著的差異，我將進一步去瞭解是否樣本流失的問題會影響研究的「內在效度」。我的假設是：參與續訪的受訪者有比較多的「資源」從事一些與政治參與相關的活動。因此，我們必須探討當我們利用這一群人來探討變數之間的關係時，會不會影響我們研究推論的效度。我以選舉研究常探討的兩個問題：選民會不會去投票以及在國會選舉研究中的「現任者優勢」(incumbent advantage) 這兩個模型，來檢視「樣本流失」對研究「內在效度」的影響

叁、資料與分析

本文利用「美國全國選舉研究」在1990年到1992年之間所做的「固定樣本連續訪談」的調查訪問資料。在1990年有1,980位受訪者參與該次訪談，而到了1992年，有其中的1,359位受訪者繼續參與了兩年後的訪問。我利用1990年的樣本編號(在1990年是V4；1992年是V3006)，合併兩筆資料，並將受訪者區分為「續訪」與「流失」兩種樣本。由於我的研究焦點在於探討影響受訪者參與連續訪談的因素，所以，我從1990年的資料即可看出這兩群日後分別接受訪問或是流失的樣本，是否原先就存在著差異。若是真的存在著明顯的差異，我將進一步探索，這種樣本流失的情況，對研究是否產生影響。

在本文中，大多數的變數都是類別資料，所以我將使用卡方的獨立性檢定(chi-square test for independence) 來看看「流失樣本」(樣本數為621) 與「續訪樣本」(樣本數為1359) 之間，是否存在顯著差異。我也將提出一個受訪者為什麼參與連續訪談的解釋模型。當然，我更會進一步探討樣本流失的問題對一般的投票行為的研究中常用的兩個解釋模式是否產生影響。由於這幾個模型的依變項都是二分類(dichotomous)，我將利用「對數成敗比模型」(logit model) 來加以分析。

一、參加續訪也需要一些「資源」

在「固定樣本連續訪談」的研究中，續訪的受訪者，並不是經由隨機抽樣的程序抽出，我們只是對第一次訪問成功的樣本加以追蹤，因此，也提供了我們一個比較接受續訪的與流失的兩種樣本之間差異的機會。從表1中我們可以發現：在續訪樣本中，男

性、白人、年齡低於65歲、高教育程度以及較高收入的受訪者較多。表1的發現，間接驗證了接受訪問需要一些能力與技巧。雖然，這兩群樣本之間，並不存在著統計上的顯著差異，不過，這個發現還是相當值得注意的。因為台灣地區進行的許多訪問工作，都是對照訪問成功的樣本與普查樣本之間，在人口學一些變數分佈的差異情況，來決定是否進行加權（例如，洪永泰與陳陸輝，1995）（註一）。我們在看到表1的結果後，也許會鬆一口氣地認為兩種樣本之間並不存在著差異。但是，很可能這兩群看似相似的受訪者，有著不同的政治態度。所以，我必須進一步釐清。

表1 「流失樣本」與「續訪樣本」在人口學變數上的差異，1990 ANES

	流失樣本%	續訪樣本%	所有樣本%
性別			
女性	57.5	53.6	54.8
男性	42.5	46.4	45.2
種族			
白人	86.2	86.8	86.6
黑人	13.8	13.2	13.4
年齡層			
18-30歲	25.6	26.1	26.0
31-45歲	31.1	32.2	31.8
46-65歲	23.3	24.8	24.3
66歲及以上	20.0	16.9	17.9
受教育年數			
少於12年	26.8	24.4	25.2
12年	36.7	32.5	33.8
13-15年	20.8	21.7	21.4
16年及以上	15.7	21.4	19.6
家庭收入			
低於 \$ 10,000	19.6	17.1	17.8
\$ 10,000 - 19,999	22.7	20.6	21.3
\$ 20,000 - 29,999	16.6	18.8	18.1
\$ 30,000 - 39,999	15.6	14.5	14.8
\$ 40,000 - 49,999	9.3	10.7	10.3
\$ 50,000 - 59,999	4.5	6.2	5.7
\$ 60,000及以上	11.7	12.1	12.0

說明：1.「流失樣本」（樣本數為621）係指受訪者在1990年至1992年的「固定樣本連續訪談」中，僅參加1990年「美國全國選舉研究」之訪談。「續訪樣本」（樣本數為1,359）則參加兩次（1990及1992）的訪談。「所有樣本」則包括前述兩種樣本（樣本總數為1,980）。以下個表同此說明。

2.表格內輸入的皆為直欄百分比（column percentage）。

3.卡方獨立性檢定：*： $p < 0.05$ ；**： $p < 0.01$ ；***： $p < 0.001$ 。

資料來源：美國全國選舉研究（ANES），1990年。

爲了瞭解兩個樣本之間的政治態度是否相同，我接著分析這兩群樣本在政治興趣上的差異。如果接受訪問是一種形式的政治參與，則政治興趣愈高的受訪者，應該愈有可能參加續訪。從表 2 中發現，續訪樣本擁有比較高的政治興趣，他們比較會對選舉產生興趣、注意報紙以及電視上的選舉消息。當然，在知道了這些相關的消息後，他們也樂於跟家人與朋友討論選舉。續訪樣本中，有接近七成對選舉 (campaign) 「有點興趣」(somewhat) 或是「很有興趣」(very much)，而流失樣本中則只有六成對選舉「有點興趣」或是「很有興趣」。由於 1990 年只有國會的期中選舉，當詢問受訪者對國會選舉 (congressional election) 的興趣時，受訪者中感興趣的比例較前一變數爲高。而其中流失樣本中，有接近八成受訪者對國會選舉感到興趣 (包括「有點興趣」以及「很有興趣」者)，續訪樣本中則有八成四感到興趣。從表 2 中也發現前述兩個關於選舉興趣的變數，兩種樣本之間，是存在統計上顯著的差異。至於閱讀與收看選舉消息上，續訪樣本比流失樣本較會注意這些競選新聞，不過兩者間的差異並未達統計上的顯著水準。在與家人及朋友討論政治上，續訪樣本比流失樣本高了約七個百分點，且兩者的差異達到統計上的顯著水準。因此，我們從表 2 可以發現：參與續訪的受訪者，是具有較高的政治興趣。

表 2 「流失樣本」與「續訪樣本」在政治興趣上的差異，1990 ANES

	流失樣本%	續訪樣本%	所有樣本%
對選舉的興趣**			
沒有興趣	38.4	30.9	33.2
有點興趣	42.7	46.8	45.5
很有興趣	18.9	22.4	21.3
對國會選舉的興趣**			
沒有興趣	21.6	16.0	17.8
有點興趣	33.9	34.5	34.3
很有興趣	44.4	49.5	47.9
閱讀報紙選舉消息			
否	35.7	34.6	34.9
是	64.3	65.4	65.1
收看電視選舉消息			
否	37.4	36.1	36.5
是	62.6	63.9	63.5
與家人朋友討論政治**			
否	35.6	28.8	30.9
是	64.4	71.2	69.1

說明：1. 見表 1。

2. 表格內輸入的皆爲直欄百分比 (column percentage)。

3. 卡方獨立性檢定：* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$ 。

資料來源：美國全國選舉研究 (ANES)，1990 年。

從上述兩種樣本在政治興趣上的差異，我們不難預測，流失樣本應該具備較低的政治知識。在1990年有關政治知識的測量上，我利用問卷中詢問受訪者七位國內外政治人物目前所擔任的職位的題目，來比較兩群受訪者的差異。表3則按照受訪者答對的高低順序，列出七位政治人物。從最後一欄的直欄百分比中可以發現，受訪者對 Quayle 以及 Gorbachev 最為熟悉，分別有八成多以及接近七成的受訪者，可以說出他們當時擔任的職位，而對當時參院民主黨的多數黨領袖 George Mitchell 的熟悉程度最差，只有不到三個百分點（2.3%）的所有受訪者，可以說出他當時擔任的職務。而從表3可以看出續訪樣本對政治人物所擔任的職位的知悉程度，比流失樣本高。兩群受訪者對前五個政治人物的熟悉程度，達到了統計上的顯著差異。續訪樣本比流失樣本高了三點六（Nelson Mandela 所擔任的職務）到七點二個百分點（Mikhail Gorbachev）的熟悉度。此外，我將受訪者在政治知識七個題目中答對的題數加總，建立一個政治知識總和

表3 「流失樣本」與「續訪樣本」在政治知識上的差異，1990 ANES

	流失樣本%	續訪樣本%	所有樣本%
Dan Quayle 的職務*			
答對	77.5	82.3	80.8
Mikhail Gorbachev 的職務**			
答對	64.1	71.3	69.0
Margaret Thatcher 的職務*			
答對	28.0	33.7	31.9
Nelson Mandela 的職務*			
答對	12.7	16.3	15.2
Tom Foley 的職務**			
答對	5.5	10.4	8.8
William Rehnquist 的職務			
答對	4.0	4.3	4.2
George Mitchell 的職務			
答對	1.8	2.6	2.3
政治知識總和***			
低	68.3	59.2	62.0
中	28.8	35.6	33.5
高	2.9	5.2	4.5

說明：1.見表1。

2.表格內輸入的皆為直欄百分比（column percentage）。

3.卡方獨立性檢定：*： $p < 0.05$ ；**： $p < 0.01$ ；***： $p < 0.001$ 。

4.「政治知識的總和」的編碼方式為：答對上述七個題目中兩題以及以下者為「低」；答對三至四題者為「中」；答對五題至七題者為「高」。

資料來源：美國全國選舉研究（ANES），1990年。

的指標。此指標將受訪者的政治知識歸為「低」、「中」以及「高」等三個類別，其中，答對兩題及以下者為「低」，三題至四題為「中」，而五題或是更多則為「高」。從表 3 最下面一區的幾個欄位中可以發現，在續訪樣本中，有四成以上具有中等程度或是高等程度的政治知識，而流失樣本中只有三成具有類似的政治知識程度。當然，很可能許多政治知識較低的受訪者，在1990年經歷這些政治知識問題的訪問過後，因為不願意再接受重複的「煎熬」與「出糗」，而沒參加1992年的續訪。不過，我卻必須排除另外一個可能，那就是一些受訪者在記取1990年的經驗後，開始注意時事，增加其政治知識。正因為本文僅採用1990年的資料，所以我們可以排除因為1990年訪問的這個刺激因素，對其1992年政治知識的影響 (Campbell and Cook, 1963)。

本文的一個重要觀點，是認為參與調查訪問，特別是接受續訪，也是一種政治參與的方式。我在上兩段中已經指出，流失樣本與續訪樣本在政治興趣以及政治知識上存在著顯著的差異，我要進一步探討兩群受訪者是否在一些政治參與上也有不同。對於政治參與的測量，我使用包括詢問受訪者是否曾經應他人要求而登記投票或是去投票、影響他人投票、配戴競選標誌或是貼紙 (wear campaign button or sticker)、參加政治性的集會與遊行、捐錢給候選人以及為政黨或是為候選人工作等六道問題。從表 4 中可以看出，除了「配戴競選標誌或是貼紙」這個舉動外，流失樣本參與這些政治活動的比例較續訪樣本少。不過，在這六項政治參與的問題上，只有「影響他人投票」以及「捐錢給候選人」上，兩群受訪者存在著統計上的顯著差異。我將六個政治參與的活動加總，作成一個「政治參與的總和」指標，該指標將是受訪者的政治參與分為四個類別，分別是「都沒有」、「低」(有任何一項政治參與)、「中」(有二至三項) 以及「高」(有四項及以上)。表 4 中可以發現，續訪樣本與流失樣本在「中」與「高」等程度參與政治活動上的差異並不大，主要的差異來自於「低」程度參與以及「都沒有」參與兩項。在「都沒有」參加任何上述的六項政治活動的比例上，流失樣本中比續訪樣本多了八個百分點。因此，參與續訪活動的受訪者也較常參與其他的政治活動。

任何的面訪活動，在訪員還沒有讓受訪者相信他「真的」是研究單位派出的訪員之前，他是以一位「陌生人」的身份，出現在受訪者的家門。因此，受訪者對「陌生人」的態度，自然影響他們回答訪問的意願。不過，訪員從一位「陌生者」的身份變為「訪員」的身份後，他不必然地可以成功地訪問受訪者。訪問成功與否必須視受訪者接受訪問的意願高低而定。從表 5 中可以看出，對「陌生人」疑慮以及對訪問興趣不高的受訪者，他們即使接受了第一次的訪問，他們比較容易拒絕我們的續訪。這兩個變數都達到統計上顯著的差異。

表4 「流失樣本」與「續訪樣本」在政治參與上的差異，1990 ANES

	流失樣本%	續訪樣本%	所有樣本%
他人會要求登記選舉或投票			
是	27.1	30.6	29.5
影響他人投票***			
是	13.0	19.3	17.3
配戴競選標誌或是貼紙			
是	7.0	6.5	6.7
參加政治性的集會			
是	6.0	6.1	6.1
捐錢給候選人*			
是	3.9	6.1	5.4
為政黨或是候選人工作			
是	2.3	2.7	2.6
政治參與的總和**			
都沒有	62.8	54.3	56.9
低	24.1	30.3	28.3
中	10.6	12.6	12.0
高	2.6	2.8	2.7

說明：1. 見表1。
 2. 表格內輸入的皆為直欄百分比 (column percentage)。
 3. 卡方獨立性檢定：*： $p < 0.05$ ；**： $p < 0.01$ ；***： $p < 0.001$ 。
 4. 「政治參與的總和」的編碼方式為：沒有參與上述六項政治活動中任喝一項為「都沒有」；參與一項為「低」；參與二至三項為「中」；參與四至六項為「高」。

資料來源：美國全國選舉研究 (ANES)，1990年。

表5 「流失樣本」與「續訪樣本」對「訪問」的態度之差異，1990 ANES

	流失樣本%	續訪樣本%	所有樣本%
對「陌生者」是否疑慮***			
否	77.2	83.6	81.6
是	22.8	16.4	18.4
受訪者對訪問的興趣***			
低	19.3	10.3	13.1
一般	44.8	38.8	40.6
高	35.9	50.9	46.2

說明：1. 見表1。
 2. 表格內輸入的皆為直欄百分比 (column percentage)。
 3. 卡方獨立性檢定：*： $p < 0.05$ ；**： $p < 0.01$ ；***： $p < 0.001$ 。

資料來源：美國全國選舉研究 (ANES)，1990年。

而在固定樣本連續訪談中，受訪者已經有接受訪問的經驗，所以，他對「訪問」的態度，攸關他是否接受下一次訪問的意願。在表 6 中我們可以發現，當受訪者覺得訪問太長、對訪問內容困惑、曾經提及要終止訪問、覺得訪問無趣、覺得訪問太複雜以及訪問過程中曾有負面的反應者，比較不會參加續訪。兩種不同的受訪樣本在上述的六種態度的差異皆達到統計上的顯著水準。值得一提的是，受訪者提及「覺得訪問太長」以及「覺得困惑」的比例，較其他四項為高。而我們更可以大膽推論，許多中途拒絕訪問或是根本不願接受訪問的受訪者，可能正是擔心自己作答問卷的能力，或是訪問時間太長會耽擱他們生活的正常作息所致。因此，身為調查研究的工作者，應該在如何使受訪者覺得訪問很愉快，而不會耽擱太多時間上多努力。以使調查研究的訪問成功率更高，樣本更具代表性。在表 6 的最下方，我也將上述六項對訪面的負面態度加總成一個指標，我們發現，續訪樣本中有八成二的受訪者並沒提及上述任何一種負面的反應，而流失樣本則有七成二，兩種樣本的差異接近十個百分點。

表 6 「流失樣本」與「續訪樣本」對「訪問」的負面態度，1990 ANES

	流失樣本%	續訪樣本%	所有樣本%
覺得訪問太長**			
提及	15.2	10.5	12.0
受訪者覺得困惑**			
提及	10.8	5.8	7.3
想要終止訪問***			
提及	8.3	3.3	4.8
覺得訪問太複雜***			
提及	6.5	2.4	3.7
覺得訪問無趣或乏味**			
提及	5.4	2.8	3.6
受訪者有負面的反應***			
提及	5.1	2.2	3.1
負面態度的總和**			
都沒有	72.4	82.0	79.0
一個	14.2	11.8	12.5
兩個	7.7	4.7	5.7
三個或是更多	5.7	1.6	2.9

說明：1. 見表 1。

2. 表格內輸入的皆為直欄百分比 (column percentage)。

3. 卡方獨立性檢定：* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$ 。

資料來源：美國全國選舉研究 (ANES)，1990年。

從上述的討論中可以發現，流失樣本與續訪樣本在人口學變數上，並不存在顯著的差異。不過，在政治態度與政治參與上，續訪樣本擁有較高的政治興趣、政治知識並且更傾向參加一些政治活動。而在對訪問的態度上，流失樣本對「陌生者」較容易產生疑慮、對訪問的興趣較低以及較容易對訪問有負面的評價。因此，樣本流失的問題，對研究的「外在效度」會是一個威脅。當我們的研究目的是在描述人們的政治興趣、政治知識以及政治參與時，使用續訪樣本的資料，會使我們高估了實際母體中的政治興趣、政治知識以及政治參與的分佈情況。而流失樣本與續訪樣本在人口學一些背景上不存在顯著差異，更值得我們注意。特別是我們以樣本中的這些變數與人口普查資料相對照，而要決定是否加權時，本文的一些發現，更是實際從事調查研究工作必須留意的。問題主要的關鍵在於這些看似相似的兩群受訪者，實際上是擁有不同的政治興趣與政治態度。所以，我們在做任何推論時，必須格外謹慎。不過，實際在進行調查訪問工作時的另一個問題是，我們沒有關於「母體」政治態度分佈的資訊。因此，當我們只進行切面的調查時，即使成功樣本與母體在一些人口學的變數分佈無顯著的差異，我們仍然無從檢證這些成功的受訪者是不是一群對政治興趣較高，對訪問不排斥的人。

二、一個解釋受訪者參與連續訪談的模型

流失樣本與續訪樣本在政治興趣以及對訪問的態度上，存在著一些差異。我想提出一個模型來解釋為什麼公民願意參加連續訪問。我的依變數為受訪者是否參加1992年的連續訪談，參加者為編碼「1」，不參加者為「0」。由於依變數是二分類，我使用的統計方法是對數成敗比模型。至於我的解釋變數，包括了受訪者的「政治興趣」、「政治參與」的情況、「對陌生者」的態度、「對訪問的態度」以及其「對訪問的負面態度」。在表7中我們可以預期的是，續訪樣本有著較高的政治興趣、較常參與其他政治活動者、對陌生人較不疑慮、對訪問較有興趣且較少對訪問有負面的態度。統計的分析結果請參考表7。其中，第一個模型與第二個模型之間的差異，在於模型一是一一估計六個對訪問負面態度的變數之係數，而第二個模型則使用「對訪問負面態度的總和」指標。

從表7的第一個模型中可以發現，大多數係數的方向與我們的假設預期相同，只有與家人或是朋友「談論政治」和「訪問太長」這兩個變數的方向除外。模型一的十一個變數中，有兩個達到統計的顯著水準，他們分別是「對訪問的興趣」以及「想要終止訪問」。「對訪問的興趣」是一個三分類的變數（分為「低」、「一般」以及「高」，請見表5），我們要估計該變數對受訪者參與連續訪問的機率之影響，可以將其他變數代入平均值，而操縱我們感興趣的這個變數（參見Hamilton 1992：ch.7）（註二）。

表 7 受訪者參與1992年續訪的預測模型，1990 ANES

	模型一 係數 (標準誤)	模型二 係數 (標準誤)
政治興趣		
對選舉的興趣	0.037 (0.078)	0.047 (0.078)
談論政治	-0.008 (0.120)	-0.076 (0.119)
政治知識	0.136 (0.096)	0.137 (0.096)
政治參與的總和	0.065 (0.069)	0.060 (0.069)
對陌生者的態度		
疑慮	-0.104 (0.133)	-0.114 (0.133)
對訪問的態度		
對訪問的興趣	0.298 (0.087) ***	0.301 (0.087) ***
對訪問負面的態度		
訪問太長	0.103 (0.172)	
覺得困惑	-0.187 (0.200)	
想要終止訪問	-0.509 (0.250) *	
訪問太複雜	-0.461 (0.280)	
訪問無趣或乏味	-0.071 (0.277)	
有負面的反應	-0.219 (0.296)	
對訪問負面態度的總和		-0.203 (0.076) **
常數項	-0.096 (0.227)	-0.105 (0.226)
樣本數	1,949	1,949
模型卡方值	64.995	59.615
自由度	12	7
正確預測率	69.47%	69.78%

說明：1. 依變數：參與1992年續訪為「1」，未參加1992年續訪為「0」。其他解釋變數的處理方式，請參與前述各表以及文內說明。

2. * : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$ 。

資料來源：美國全國選舉研究 (ANES)，1990年。

假設其他變數值為其平均值時，當受訪者的對訪問興趣由「低」(參加訪問的機率为53.3%)提高為「一般」程度(機率为60.6%)或是由「一般」提高到「高」程度(機率为67.4%)時，其參與續訪的機率會提高七個百分點(計算方式請參閱本文附錄一，附表1有各變數的平均值)。換言之，一個對訪問興趣「高」的受訪者參加續訪的機率比一個對訪問興趣「低」的受訪者，高出十四個百分點。若是我們要看「想要終止訪問」對參與續訪機率的影響，可以將其他變數設為其平均值，我們發現一旦受訪者在

1990年的訪問中提到他想終止訪問時，他參加1992年的續訪的機率下降了11.6個百分點（從70.1%降低為58.5%）。此外，表7中的第二個模型，是按照表6建立「負面態度總和」指標的方式，建立一個從「0」到「3」四分類的指標。我們發現，所有變數的參數估計的正負值也與我們的理論假設方向相符，而且大多的估計係數與模型一幾乎相同。同樣地，我們按照上述的機率計算方法可以發現，當受訪者對訪問有一個負面的評價時，他參加續訪的機率將降低四點五個百分點（從69.2%下降到64.7%）。整體而言，表7告訴我們，受訪者對訪問的興趣以及對訪問的負面印象，影響他參加續訪的機率。

正如 Brehm (1993: 92) 所言：「我們得想一個辦法彌補受訪者所花的時間以及縮短我們的訪問。…調查訪問可以變得好玩、有趣以及具有知性 (informative)，或者是痛苦與無聊，而此選擇權是操在調查研究者的手上。」調查研究工作者面對的兩難是，想要將有限的研究經費做極大化的發揮，因此，蒐集愈多的資訊愈好。不過，問卷過長的結果是，成功的受訪者並不能代表我們要推論的一般大眾，使我們研究的焦距失真，結果的效度受到質疑。

肆、樣本流失與研究效度：兩個投票模型的測試

社會科學研究的目標之一是在解釋變數之間的因果關係。因此，我將繼續探討流失樣本對我們研究「內在效度」的影響。對於研究選民投票行為的學者而言，我們關心的不外以下兩個問題：「選民會不會去投票？」以及「他／她會投給哪一位候選人？」。我將運用 Teixeira 的選民投票參與的模型以及 Jacobson 的國會選舉模型來檢視樣本流失對研究效度的影響。

一、Teixeira 的投票參與模型

「如果在二十世紀晚期，美國的投票參與有一個明確事實的話，那就是愈來愈少美國人去投票了」（Niemi and Weisberg, 1993: 14）。對於美國選民投票率自1960年以來降低的原因，學者們提出不同的解釋（Abramson et al., 1998: 79）。有些學者以選民的教育程度、選民總體的平均年齡或是政治世代交替（generational replacement）等社會因素提出解釋。有些學者著重在一些像是政黨認同以及政治功效感（political efficacy）等政治態度強度的變化。也有學者討論一些制度的因素，著眼於諸如選民選舉的登記制度。有些則分析政治人物對選民的動員能力，對選民投票與否的影響。本文採用 Teixeira 在1987年提出較廣泛的模型，探討樣本流失對選民選舉參與研

究效度的影響。

Teixeira 提出的模型包括了社會因素、政黨認同的強弱、政治功效感以及閱讀報紙等解釋因素，來說明美國選民投票率降低的原因（各變數的處理方式，請見附錄二）。我們的假設是：教育程度低者、年輕者、未婚或是配偶過世者、經常搬遷者、藍領階級、低收入者、女性、非白種人、居住在南方者、政黨認同的強度弱者、政治外在功效感低者以及不太注意選舉消息者，較不會去投票。在表 8 中我們可以發現，以全體受訪樣本為分析對象時，模型中絕大多數變數的參數估計值，與我們理論預期的方向相符合。就受訪者的人口學背景而言，低教育程度者者（少於12年）、年輕選民、在現址居住未滿兩年者、收入較低者、女性、非白種人、以及居住在南方的選民較少投票。其中，受訪者的教育程度、年齡（37歲及以上者）、居住狀態和家庭收入等變數，都達到統計上的顯著水準。模型中也顯示：已婚且與配偶住在一起者（包括未婚但是與 partner 同住者）的投票率比其他類別低，這是出乎我們意料之外的。不過，該參數估計值未達統計上的顯著水準。同時，在教育程度上，我們也發現受過一些大學教育者（13至15年）的投票率最高，其次是高中畢業者，接下來才是拿到學士文憑或是受過更高教育者。不過，這三個類別受訪者的投票率，都比教育程度較低（少於12年者）為高。

除了受訪者的一些人口學背景之外，研究投票參與的學者，也試圖用一些政治態度以及選民對資訊的吸收情形，來解釋選民的投票參與。在美國政治中，最受到重視的一個問題之一，就是在1960年代中期之後，美國選民中，兩大政黨認同者的比列開始下滑，這個現象直到1970年代末以及1980年初其才趨緩和（Wattenberg, 1994）。我們從表 8 中可以發現，選民對政黨認同的強弱，對其投票與否，仍然有顯著的影響。選民政黨認同的強度愈強，他去投票的機率愈高。不過，對於政治功效感這個變數而言，則與我們的理論預期不同。表 8 中可以發現，政治功效愈高的選民，愈不傾向投票，不過，此一變數並未達統計上的顯著水準。另一個關於選民政治興趣的變數，則是受訪者對報上選舉消息的注意程度。我們發現，對選舉消息愈注意的受訪者，投票的機率愈高，該兩個虛擬變數皆達統計上的顯著水準。

我們當然也關心樣本流失的問題，對於我們投票參與的模型中，各變數的參數估計值是否會產生影響。從表中可以發現，續訪樣本中絕大多數的參數估計與所有樣本的估計值相當接近，不過，有關受訪者的年齡以及對報上選舉消息的注意情況這兩類變數，卻出現低估的傾向。年齡的幾個虛擬變數的估計值差異不小，不但低估原來各變數的參數，且原先顯著的一個變數（37歲及以上）不再顯著了。我們可以估計年齡37歲及以上的受訪者，他們在期中選舉投票的機率。以所有樣本為例，當一個受訪者在37歲以

表 8 Teixeira 的投票參與模型

	所有樣本 係數(標準誤)	續訪樣本 係數(標準誤)	流失樣本 係數(標準誤)
教育程度			
12 年	0.707(0.243)**	0.658(0.296)*	0.648(0.456)
13-15 年	1.128(0.285)***	1.273(0.358)***	0.705(0.519)
16 年及以上	0.673(0.280)*	0.611(0.336)\$	0.523(0.558)
年齡			
25-28 歲	-0.305(0.420)	-0.818(0.521)	1.029(0.870)
29-32 歲	-0.103(0.408)	-0.838(0.514)	1.863(0.849)*
33-36 歲	0.248(0.398)	-0.257(0.521)	1.608(0.765)*
37 歲幾以上	0.761(0.336)*	0.161(0.433)	2.356(0.696)***
婚姻狀態			
已婚且配偶健在	-0.129(0.178)	0.011(0.216)	-0.379(0.343)
居住狀態			
居住兩年及以上	0.788(0.287)**	0.756(0.358)*	1.018(0.534)\$
職業			
藍領工人	-0.221(0.257)	-0.226(0.319)	-0.290(0.467)
白領階級	0.052(0.201)	0.113(0.246)	-0.010(0.386)
家庭收入			
年薪高於 \$ 10,000	0.725(0.268)**	0.764(0.327)*	0.599(0.505)
性別			
男性	0.137(0.176)	0.170(0.212)	-0.003(0.344)
種族			
白人	0.508(0.232)*	0.577(0.283)*	0.253(0.449)
地區			
非南方	0.222(0.183)	0.202(0.227)	0.370(0.340)
政黨認同強度			
三刻度量表	-0.355(0.142)*	0.342(0.174)*	0.393(0.261)
政治功效感			
三刻度量表	-0.074(0.103)	-0.148(0.121)	0.237(0.223)
注意報上選舉消息程度			
有一點注意	0.847(0.208)***	0.637(0.252)*	1.442(0.405)***
很注意	1.183(0.219)***	1.083(0.267)***	1.532(0.414)***
常數項	-3.248(0.565)***	-2.568(0.693)***	-5.122(1.123)***
樣本數	935	671	264
模型卡方值	120.486	81.617	54.906
自由度	19	19	19
正確預測率	76.36%	76.90%	75.00%

說明：1. 依變數：1990年投票為「1」，未投票為「0」。

2. \$; $p < .10$; * : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$ 。

資料來源：美國全國選舉研究 (ANES) ，1990年。

及以上時，他投票的機率增加了14.8個百分點（由65.3%增加為80.1%）。而以續訪樣本為分析對象時，年齡為37歲及以上者，其投票率僅增加為2.8個百分點（從76.3%增加到79.1%）。所以，續訪問本低估了年齡（37歲及以上者）對受訪者投票機率的重要性達12個百分點。此外，在續訪樣本中，受訪者對報上選舉消息的注意程度的重要性也被低估。以受訪者對報上的消息「有一點注意」者為例，在全體樣本中，他們的投票率會增加14.8個百分點（從69.2%增加為84.0%），不過，在續訪樣本中，他們的投票率僅增加10.5個百分點（從73.5%增加為84.0%）。兩個樣本對這個變數重要性的估計差異達4.2個百分點。整體而言，「樣本流失」的問題，對我們投票參與的模型中，其他各變數的參數估計值的影響並不大，不過，卻低估了受訪者年齡以及對報紙選舉消息注意程度的重要性。因此，對研究的「內在效度」有相當的影響。

二、Jacobson 的國會選舉的研究模型

對於美國國會議員選舉的研究，學者們或者著重一些諸如總統的聲望、整體的經濟情況等全國性的因素（national force），或者探討一些像是選民服務、參選人的知名度、現任者優勢（incumbent advantage）與競選經費花費等個別選區因素（local force）（Niemi and Weisberg, 1993b: ch. 13）。由於本文使用的是1990年國會選舉的切面（cross-sectional）的資料，因此，Jacobson 探討個別選區因素對國會選舉影響的模型較適宜本文之分析。

Jacobson（1997）的國會選舉的模型中，除了探討選民對眾院議員的熟悉程度（name recognition）對其投票行為的影響之外，他也分析現任者（incumbent）所具有的選舉優勢。我採用他在1997年著作提出的一個模型，來展現樣本流失對國會選舉研究效度的影響（關於各變數的處理方式請參閱附錄三）。我們可以預期，現任議員以及選民對其熟悉的候選人，在競選時應該享有相當的優勢。我的分析結果列於表9。

從表9的分析結果可以得知：當包括所有樣本時，所有解釋變數均達統計上的顯著水準，而且變數的正負值與我的理論預期相符。在五個解釋變數中，政黨認同尤其是解釋期中選舉的重要因素。而其他四個關於當「現任者」以及「熟悉程度」的變數中，以「民主黨為現任者」最為重要（註三）。當其他變數設定為其平均值時，一個選民在民主黨為現任者的選區中，投給民主黨候選人的機率會增加35個百分點（從48.4%增加到83.7%）。相對地，一個尋求連任的現任共和黨眾議員，其所獲得的「現任者優勢」只有增加27個百分點的當選機率（從21.2%增加為48.3%）。所以，民主黨的現任者，比他們國會中共和黨的同僚，享有更大的競選連任優勢。

表9 Jacobson 的國會選舉模型

	所有樣本 係數 (標準誤)	續訪樣本 係數 (標準誤)	流失樣本 係數 (標準誤)
政黨認同	1.857 (0.171) ***	1.842 (0.196) ***	1.989 (0.369) ***
共和黨議員為現任者		-0.886 (0.420) *	
民主黨議員為現任者	1.698 (0.336) ***	1.851 (0.405) ***	1.585 (0.628) *
可指出共和黨議員名字		-1.057 (0.386) **	-1.589 (0.811) \$
可指出民主黨議員名字	0.772 (0.314) *	0.746 (0.344) *	0.959 (0.782)
常數項	-0.021 (0.299)	-0.351 (0.365)	0.720 (0.550)
樣本數	801	587	214
模型卡方值	520.256***	366.77***	159.788***
自由度	5	5	5
正確預測率	86.52%	85.86%	90.65%

說明：1. 依變數：1990年投票給民主黨眾議員候選人為「1」，投給其他候選人為「0」。

2. \$: $p < .10$; * : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$ 。

資料來源：美國全國選舉研究 (ANES)，1990年。

不過，當我們只用「續訪樣本」做推論時，我們得到的結果就會有一些出入。從估計的參數來看，續訪樣本 (1.851) 比所有樣本 (1.698) 高估 (overestimate) 了民主黨現任者的優勢。同樣地，我們將其他變數代入其平均值，我們可以發現當一個選民面對一個民主黨的現任者時，他投給這位候選人的機率增加了39個百分點 (從44.2%增加到83.5%)。所以，續訪樣本中高估了民主黨現任議員的連任機率，達四個百分點。此外，對照表9內的兩個樣本，我們也可以發現，續訪樣本中低估了共和黨現任者的優勢，也同時低估了選民熟悉共和黨的候選人時，共和黨候選人的當選機率。我們援用上述的計算方式，將其他變數代入其平均值，計算得出：在全體樣本中，當共和黨候選人為現任者，其當選機會提高了27個百分點 (選民投給民主黨議員的機率從78.8%降低為51.7%)；而在續訪樣本中，選民投給共和黨現任議員的機率則提高了19個百分點 (選民投給民主黨議員的機率從76.5%降低到57.3%)。因此，續訪樣本低估了共和黨現任者的當選機率，高達八個百分點 (註四)。

從上述的討論中可以發現，樣本流失對我們不同研究的「內在效度」產生不同的影響。在投票參與模型方面，大多數變數的參數估計相當穩定，不過，在年齡的變數上則出現低估25歲以上各年齡層受訪者投票機率的情況 (對照組為24歲以及以下的受訪者)，而且，原先顯著的一個變數 (37歲及以上) 變得並不顯著。此外，續訪樣本也

低估了選民對報上選舉消息的注意程度，對其投票機率的影響。至於國會投票模型方面，樣本流失則對我們的參數估計也產生了影響。特別是我們要以訪問結果來解釋不同變數對選舉結果的影響時，我們可能會因樣本的流失，而高估了民主黨現任者的選舉優勢，或是錯估共和黨現任議員的劣勢。

伍、結論

「固定樣本連續訪談法」雖然是研究個體政治態度持續與變遷一種很好的研究設計，不過，當研究中存在著特定背景的樣本，於續訪中流失的情況時，我們的研究效度就會受到影響。本文分析美國全國選舉研究1990年到1992年的固定樣本連續訪談的資料，研究結果發現續訪樣本與流失樣本之間，在政治興趣、政治知識與政治參與上，存在著顯著的差異。因此，當研究的目的是在描述母體在政治興趣、政治知識與政治參與的分佈時，樣本流失的問題會影響研究的「外在效度」。

當選民決定參與續訪與否的原因，與我們研究的變數互相關連時，樣本流失的問題也影響了研究的「內在效度」。從本文使用的投票參與模型中可以發現，樣本流失低估了受訪者年齡以及對報上選舉消息注意程度這兩個變數，對其投票與否的影響。當我們要解釋國會選舉中現任者的優勢時，樣本流失的問題高估了民主黨議員現任者優勢的重要性，也低估了共和黨現任議員的重要性。

本研究另外一項重要發現，則是續訪樣本與流失樣本之間，在許多人口學的變項上，並不存在著顯著的差異。對許多調查研究的實務工作者而言，這些人口學的變數，常常是決定該樣本是否足以代表母體的一個重要或是唯一的判斷標準。本研究結果則發現，這兩群看似相同的樣本，其實卻具有不同的政治態度。因此，以人口學的變數來作為推論樣本代表性的檢定以及加權的處理方式，必須審慎為之。

從本研究中也得知，參與固定樣本連續訪問與否，取決於受訪者對訪問本身的興趣高低以及是否對訪問存有負面態度。由於受訪者合作意願的高低，是調查研究成功的重要關鍵因素。因此，如何增加訪問過程的順暢與減低受訪者的負擔和不必要的尷尬，是從事實際調查研究工作者無可避免的責任。當然，愉快的訪談過程以及成功率高的訪問，最後受惠的，還是實際的調查研究工作者，以及使用資料的所有人員。

註 釋

- 註 一：有關加權方式的討論與運用，請參考洪永泰（1996：107－116）。
- 註 二：對於「對數成敗比模型」估計係數的解釋方式有很多種，請參見 Hamilton 1992：ch. 7。不過，由於「對數成敗比模型」對機率分佈的假設是成 S 型，所以，一個解釋變數對機率的影響大小不單其本身估計的係數大小是觀察重點之一，也必須視其他變數的大小而定（Hanushek and Jackson 1977：188－189）。因此，將其他變數代入其平均值後再操縱我們感興趣的變數，是看出該變數對參與續訪機率的影響力之較佳方式。請參閱本文附錄一之說明。
- 註 三：由於四個變數的 range 皆在 0 到 1 之間，所以，可以比較他們之間的理論上的重要性。對於一個解釋變數重要程度（theoretical, level, and dispersion importance）的討論，請參考 Achen 1982：68－77。
- 註 四：盛杏媛（1994）在討論 selection bias 時，也測試過 Jacobson 的國會選舉模型，她發現對「現任者」解釋變數的估計值，會受到 selection bias 而影響。

附錄與附表

附錄一：對數成敗比模型的解釋方式

爲了方便說明起見，我們假設我們模型中有兩個解釋變數以及一個常數項，此模型爲：

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + u_i \dots\dots\dots (1)$$

其中我們的依變數， Y_i ，是一個二分類的變數，假設「1」代表參加續訪，「0」代表流失樣本。我們要預測受訪者參加續訪的機率。因此，我們的模型可以寫成：

$$P_i = \Pr(Y_i = 1) = F(\beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3}) = F(X\beta) \dots\dots\dots (2)$$

而

$$1 - P_i = \Pr(Y_i = 0) = 1 - F(\beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3}) = 1 - F(X\beta) \dots\dots (3)$$

其中 $F(X\beta)$ 是描述機率與解釋變數之間如何相關的累積分配函數（cumulative distribution function）。我們使用的對數成敗比分佈可以定義爲：

「固定樣本連續訪談法」(panel studies) 樣本流失問題之探討

$$P_i = \frac{1}{(1 + e^{-X\beta})} \dots\dots\dots (4)$$

當 $X\beta$ 的大小從 $-\infty$ 到 $+\infty$ ，其機率 (P_i) 分佈的範圍 (range) 是從0到1之間。同樣地，我們可以計算受訪者不參加續訪的機率是：

$$1 - P_i = \frac{e^{-X\beta}}{(1 + e^{-X\beta})} = \frac{1}{(1 + e^{X\beta})} \dots\dots\dots (5)$$

將公式(4)與(5)帶入以下公式，我們可以得到：

$$L = \log\left(\frac{P_i}{1 - P_i}\right) = \log\left[\frac{1/(1 + e^{-X\beta})}{e^{-X\beta}/(1 + e^{-X\beta})}\right] = \log\left(\frac{1}{e^{-X\beta}}\right) = \log(e^{X\beta}) = X\beta \dots\dots(6)$$

其中 L 稱為 **logit** 或是 **the log of the odd ratio**，而我們一般以公式(6)所進行的分析，就成為「對數成敗比的分析」(**logit analysis**)。

對數成敗比的分析具有兩個特性。首先，「對數成敗比」的分佈與解釋變數之間，是呈線性函數關係。不過，依變數的機率分佈 (P_i) 與解釋變數之間的關係，就不同於一般迴歸分析中，解釋變數與依變數之間的線性關係。我們可以用偏微分來證明機率與解釋變數之間的關係 (Hanushek and Jackson, 1977 : 189)：

$$\frac{\partial P_i}{\partial X_{ik}} = \frac{\partial [1/(1 + e^{-X\beta})]}{\partial X_{ik}},$$

因為

$$\frac{\partial (1/z)}{\partial x} = -\frac{1}{z^2} \frac{\partial z}{\partial x},$$

而且

$$\frac{\partial e^w}{\partial x} = e^w \frac{\partial w}{\partial x}.$$

我們可以令 $(1 + e^{-X\beta}) = z$ ，以及 $-X\beta = w$ 。

所以，

$$\begin{aligned} &= \frac{\partial [1/(1 + e^{-X\beta})]}{\partial X_{ik}} \\ &= -\frac{1}{(1 + e^{-X\beta})^2} \frac{\partial (1 + e^{-X\beta})}{\partial x} \\ &= -\frac{1}{(1 + e^{-X\beta})^2} \frac{\partial (e^{-X\beta})}{\partial x} \\ &= -\frac{1}{(1 + e^{-X\beta})^2} (e^{-X\beta}) \frac{\partial (-X\beta)}{\partial x} \end{aligned}$$

$$\begin{aligned}
 &= - \frac{1}{(1 + e^{-X\beta})^2} (e^{-X\beta}) (-\beta_{tk}) \\
 &= \frac{1}{(1 + e^{-X\beta})^2} \beta_{tk} e^{-X\beta} \\
 &= \beta_{tk} \left[\frac{1}{(1 + e^{-X\beta})} \right] \left[\frac{e^{-X\beta}}{(1 + e^{-X\beta})} \right] \\
 &= \beta_{tk} P_t (1 - P_t) \dots\dots\dots (7)
 \end{aligned}$$

其次，我們在說明解釋某一個特定變數對依變數的影響時，必須同時考慮到其他解釋變數的值。依變數機率的分布，同時受某特定解釋變數以及其他解釋變數的影響。所以，本文作者在解釋對數成敗比模型時，並不採用一般常用的「對數成敗比」(logit 或是稱之為 log odds) 以及「成敗比」(odds 或是稱之為 ratios of odds) 的解釋方式，而著眼於機率的解釋，正是考慮到對數成敗比分析的上述兩個特性。對於相關的討論有興趣的讀者，可以參閱 Hanushek and Jackson, 1977: ch. 7, 而 Gujarati, 1995: ch. 16; Hamilton, 1992: ch. 7; 以及 Kennedy, 1992: ch. 15 對於對數成敗比模型也有相關的討論。

附錄二、Teixeira 模型中變數的處理方式

本模型中的依變數是一個虛擬變數 (dummy variable)。選民在1990年投票者編碼為「1」，未投票者為「0」。教育程度為三個虛擬變數，分別依受訪者接受教育年數為「12年」、「13至15年」以及「16年及以上」編碼為三個虛擬變數，對照組為受教育程度為「11年及以下」者。年齡則是四個虛擬變數，將受訪者分為「25至28歲」、「29至32歲」、「33至36歲」和「37歲及以上」四組。對照組是受訪者年齡為「24歲及以下」者。婚姻狀態是一個虛擬變數，受訪者已婚且配偶健在者以及未婚但與 partner 同住者為「1」，其餘為「0」。居住狀況是將居住在目前城市的時間達兩年及以上者編碼為「1」，低於兩年為「0」。職業為兩個虛擬變數，一個為「藍領階級」，包括一些以勞力為主的工作 (precision, craft, repair, machine operator, assembler, transportation and material moving, handler, equipment cleaner, and helper)。「而白領階級」則為擔任執行、行政、銷售代表和專業工作等受訪者。家庭收入則將年收入為10,000元及以上者編碼為「1」，其他為「0」。性別則男性為「1」，女性為「0」。種族為白人為「1」，其他為「0」。地區以南方為「0」，非南方為「1」。政黨認同的分布則從「0」到「2」。「獨立選民以及對政治冷漠者 (apolitical)」為「0」，「偏向某一政黨或是輕度認同政黨者」為「1」，「強烈政黨認同者」為「2」。政治功效感主要是由兩個測量外在政治功效的變數組成的。兩個問題是

「我不相信政府官員關心像我這樣的人的想法」以及「像我這樣的人，對於政府所做的事情沒有什麼說話的餘地」。受訪者不同意這兩個問題者（包括「不同意」與「很不同意」），政治功效感為「高」（編碼為「2」），受訪者同意這兩個問題者（包括「同意」與「很同意」）政治功效感為「低」（編碼為「0」），其餘歸類為「中」等程度政治功效感（編碼為「1」）。不過，回答「不知道」者不納入分析。在受訪者對報上選舉消息的注意程度上，對照組是對這些消息「不太注意」或是「根本不注意者」。兩個虛擬變數一個是對選舉消息「有一點注意」（some）者，另一個是對選舉消息「很注意」者（包括“a great deal”與“quite a bit”）。

附錄三、Jacobson 模型中變數的處理方式

本模型中，依變數為一個虛擬變數，選民投票支持民主黨眾議員為「1」，支持共和黨議員為「0」，而其他選項的樣本則不包括在本文的分析之內。對於解釋變數的處理上，政黨認同原來是一個七個類別的量表（從「強烈認同民主黨」到「強烈認同共和黨」），我將之編碼如下：「-1」為「強烈共和黨認同者」、「-0.67」為「弱共和黨認同者」、「-0.33」為「傾向共和黨者」、「0」為「獨立選民」（independent）、「0.33」為「傾向民主黨者」、「0.67」為「弱民主黨認同者」而「1」為「強烈民主黨認同者」。關於「現任者」的解釋變項，我用「共和黨為現任者」以及「民主黨為現任者」兩個虛擬變數，而以「1」代表兩個變數的個別政黨現任者，「0」為其他。另外對候選人熟悉程度的變數則為「熟悉共和黨候選人」以及「熟悉民主黨候選人」等兩個變數。當詢問受訪者特定政黨提名的候選人時，若是受訪者可以指出該候選人以及他所屬的政黨，則編碼為「1」，若只能指出名字而無法說出政黨為「0.5」，而名字及政黨都無法認出時則為「0」。

附表 1 受訪者參與1992年續訪的預測模型，1990 (ANES)

	模型一	模型二
政治興趣		
對選舉的興趣	1.879	1.879
談論政治	0.692	0.692
政治知識	1.426	1.426
政治參與的總和	0.605	0.605
對陌生者的態度		
疑慮	0.182	0.182
對訪問的態度		
對訪問感興趣	2.334	2.334
對訪問負面的態度		
訪問太長	0.118	
覺得困惑	0.071	
想要終止訪問	0.048	
訪問太複雜	0.036	
訪問無趣或乏味	0.034	
有負面的反應	0.031	
對訪問負面態度的總和		0.317
樣本數	1,949	1,949

說明：1.表內數值為各變數的平均值 (mean)。

2.各模型的估計參數請參考表 7，變數的編碼方式請見內文說明。

資料來源：美國全國選舉研究 (ANES)，1990年。

附表 2 Teixeira 的投票參與模型

	所有樣本	續訪樣本	流失樣本
教育程度			
12年	0.324	0.310	0.360
13 - 15年	0.227	0.228	0.224
16年及以上	0.300	0.315	0.261
年齡			
25 - 28歲	0.062	0.064	0.067
29 - 32歲	0.081	0.082	0.080
33 - 36歲	0.101	0.088	0.133
37歲幾以上	0.703	0.714	0.674
婚姻狀態			
已婚且配偶健在	0.625	0.630	0.610
居住狀態			
居住兩年及以上	0.927	0.930	0.921
職業			
藍領工人	0.154	0.152	0.159
白領階級	0.423	0.431	0.421
家庭收入			
年薪高於 \$ 10,000	0.889	0.890	0.886
性別			
男性	0.480	0.481	0.477
種族			
白人	0.862	0.864	0.856
地區			
非南方	0.706	0.714	0.686
政黨認同強度			
三刻度量表	1.302	1.315	1.269
政治功效感			
三刻度量表	0.627	0.644	0.583
注意報上選舉消息程度			
有一點注意	0.419	0.414	0.432
很注意	0.406	0.414	0.386
樣本數	935	671	264

說明：1. 表內數值為各變數的平均值 (mean)。

2. 各模型的估計參數請參考表 8，變數的編碼方式請見附錄二說明。

資料來源：美國全國選舉研究 (ANES)，1990年。

附表3 Jacobson 的國會選舉模型

	所有樣本	續訪樣本	流失樣本
政黨認同	0.129	0.189	0.129
共和黨議員為現任者	0.257	0.285	0.257
民主黨議員為現任者	0.622	0.627	0.622
可指出共和黨議員名字	0.138	0.193	0.138
可指出民主黨議員名字	0.255	0.303	0.255
樣本數	801	587	214

說明：1.表內數值為各變數的平均值（mean）。

2.各模型的估計參數請參考表9，變數的編碼方式請見附錄三說明。

資料來源：美國全國選舉研究（ANES），1990年。

參考書目

I. 中文部分

洪永泰

1996 戶中選樣之研究，台北：時英出版社。

洪永泰、陳陸輝

1995 「研究方法」，陳義彥等著，選舉行為與台灣地區的政治民主化—從八十三年省市長選舉探討，台北：國立政治大學選舉研究中心。

II. 英文部分

Abramson, Paul R., John H. Aldrich, and David W. Rohde.

1998 *Change and Continuity in the 1996 Elections*. Washington D.C. : CQ Press.

Achen, Christopher H.

1975 “Mass Political Attitudes and the Survey Response,” *American Political Science Review* 69 : 1218 – 31.

1982 *Interpreting and Using Regression*. Beverly Hills : Sage.

Aldrich, John R., and Forrest D. Nelson.

1984 *Linear Probability, Logit, and Probit Models*. Beverly Hills : Sage.

Brady, Henry E., Sidney Verba, and Kay Lehman Schlozman.

1995 “Beyond SES : A Resource Model of Political Participation,” *American Political Science Review* 89 : 271 – 294.

Brehm, John.

1993 *The Phantom Respondents*. Ann Arbor : University of Michigan Press.

Campbell, Donald T. and Julian C. Stanley.

1966 *Experimental and Quasi-Experimental Designs for Research*. Chicago : Rand-McNally.

Converse, Philip E.

1964 “The Nature of Belief Systems in Mass Publics,” In David Apter, eds. *Ideology and Discontent*. New York : Free Press.

Converse, Philip E. and Gregory B. Markus

1979 “ ‘ Plus ca Change... ’ The New CPS Election Study Panel, ” *American Political Science Review* 73 : 2 – 49.

Cook, Thomas D. and Donald T. Campbell

1979 *Quasi-Experimentation : Design and Analysis Issues for Field Settings* . Chicago : Rand-McNally.

Green, Donald Philip, and Bradley Palmquist

1990 “ Of Artifacts and Partisan Instability, ” *American Journal of Political Science* 34 : 872 – 901.

Gujarati, Damodar N.

1995 *Basic Econometrics* . 3rd Edition. New York : McGraw-Hill, Inc.

Hamilton, Lawrence C.

1992 *Regression with Graphics : A Second Course in Applied Statistics* . California : Brooks /Cole publishing Company.

Hanushek and Jackson, 1977

Herbst, Susan

1993 *Numbered Voices : How Opinion Polling Has Shaped American Politics* . Chicago : The University of Chicago Press.

Jacobson, Gary C.

1997 *The Politics of Congressional Elections* . 4th Edition. New York : Addison Wesley Longman, Inc.

Jennings, M. Kent, and Gregory B. Markus

1984 “ Partisan Orientations over the Long Hall : Results from the Three-wave Political Socialization Panel Study, ” *American Political Science Review* 78 : 1000 – 18.

Jennings, M. Kent, and Richard G. Niemi

1981 *Generations and Politics* . New Jersey : Princeton University Press.

Johnston, Jack, and John DiNardo

1997 *Econometric Methods* . 4th Edition. New York : the McGraw-Hill Companies, Inc.

Kennedy 1992

Kidder, Louise H., Charles M. Judd, and Eliot R. Smith

- 1986 *Research Methods in Social Relations*. 5th Edition. New York : CBS College Publishing.
- King, Gary, Robert O. Keohane, and Sidney Verba
1994 *Designing Social Inquiry*. New Jersey : Princeton University Press.
- Kish, Leslie
1965 *Survey Sampling*. New York : John Wiley and Sons.
- Manheim, Jarol B. and Richard C. Rich
1995 *Empirical Political Analysis : Research Methods in Political Science*. 4th Edition. New York : Longman.
- Niemi, G. Richard, and Herbert F. Weisberg
1993 *Controversies in Voting Behavior*. Washington D. C. : Congressional Quarterly Inc.
- Sheng, Shing-yuan (盛杏媛)
1994 “ Selection Bias in Vote Choice Models, ” *Journal of Electoral Studies* (選舉研究) 1 (2) : 221 - 250.
- Smith, Eric R. A. N.
1989 *The Unchanging American Voter*. Berkeley : University of California Press.
- Teixeira, Ruy A.
1987 *Why Americans Don't Vote : Turnout Decline in United States 1960 - 1984*. New York : Greenwood Press.
- Verba, Sidney
1996 “ The Citizen Respondent : Sample Surveys and American Democracy, ” *American Political Science Review* 90 : 1 - 7.
- Verba, Sidney, and Norman H. Nie
1972 *Participation in America : Political Democracy and Social Equality*. New York : Harper and Row.
- Wattenberg, Martin P.
1994 *The Decline of American Political Parties 1952 - 1992*. Cambridge, Mass : Harvard University Press.

Mortality: The Threat to the Validity of Panel Studies

Lu-huei Chen

Abstract

Mortality is a problem when it exits differential loss of respondents during our panel studies. By examining the 1990 – 1992 NES panel studies, I demonstrate that there were significant differences between the panel sample and the mortality sample on their political interest, political knowledge, and political participation. Mortality undermines the external validity when our research interest is to describe the political interest, political knowledge, and political participation of the mass public.

When the selection process correlates with some variables in our model, mortality also undermines the internal validity of our research. I demonstrate that we will underestimate the importance of the respondent's age and his /her interest on newspaper on explaining his /her voting participation when we have the mortality problem in our panel studies. When we want to explain the importance of incumbent advantage, mortality makes us overestimate the importance of Democrat's incumbency and underestimate the importance of Republican incumbency.

Participating in panel studies is a function of the respondent's interest and attitudes toward surveys. We shall always keep this in mind since the respondent's cooperation is the key point for the success of survey research. As the nonresponse rate increases in most surveys, social scientists have the obligation to make our interview more pleasant for our respondents.

Keywords: internal validity, external validity, mortality, panel studies, political participation, survey research.

審查意見答覆

審委意見(一)

我建議作者可以將研究的年度與資料加以延伸。比方說，Jennings 等人在1965、1973以及1982所進行的 student-parent study 以及 SRC 在1956-58-60以及1972-74-762的 panel studies，都可以運用類似的模型加以探討。此外，也可以考慮運用流失樣本的相關資訊，對這些 panel studies 給予相當的加權處理，看一看「selection bias」對研究的影響（可以參考盛杏媛1994：221－250）。

審委意見(二)

1. Panel studies 中樣本往往不是以隨機的方式流失，為眾所周知，但要明確找出原因，也需要嚴謹的研究設計。本論文撰稿人雖引用了許多研究設計的文獻和語彙，卻始終未將其因果推論之研究設計具體呈現出來。究竟這是屬於准實驗還是非實驗設計？是其中哪一型？若是屬於非實驗設計，則其第四節中「內在效度」之判定基礎為何？均應詳加說明。
2. 撰稿人雖在第三頁聲稱其研究重點只在比較續訪樣本及流失樣本，但是「外在效度」關心的既是各樣本結果推論母體的有效性，則比較重心便應該在母體資料和1990NES 所有樣本、1992續訪樣本、流失樣本之間的對照，而不只是後兩者之對照而已。因此至少表 1 中必須也將1990之 census 結果列出，並考驗三組樣本與其有無顯著差異。
3. 在1992年流失的621個樣本中，有多少無法追蹤？有多少是追蹤到但是拒訪？應加以說明，並納入分析之考量。

論文評審意見的答覆

- 一、對於第一位評審提出的建議，作者認為本文尚能回答「樣本流失對『固定樣本連續訪談法』研究效度影響」這個問題。日後若繼續延伸，當會涵蓋更多美國與國內「固定樣本連續訪談」的資料。至於運用加權的方式，探討「selection bias」對研究的影響這一部份，作者認為是相當有價值的建議，日後當繼續加以研究。
- 二、對於第二位評審提出的指正與建議。作者除針對若干缺失據以修正之外，提出如下的答覆。

1. 本研究係針對既有的調查訪問資料所做的二手分析，因此，僅能運用研究資料中既有的研究變數，來測試本研究的假設。至於文中第四節探討的「內在效度」，係指樣本流失對我們研究中解釋變數的參數估計的影響。也就是討論：相較於1990年所有的成功樣本，當我們僅運用「連續樣本」作變數間的因果推論時，是否會錯估（高估或是低估）解釋變數的影響力。
2. 作者曾考慮增列母體的基本分佈，不過，一般普查資料僅包括母體的人口學變數的分佈，而無選民政治態度與政治參與的相關資訊。從表1中可以發現，續訪樣本與流失樣本在人口學的變項上並不存在著顯著差異。換言之，若是1990年的訪談樣本（包括續訪樣本與流失樣本）是具有代表性的樣本的話，續訪樣本與流失樣本皆是具有代表性的樣本。但是，從表2至表6中卻可以發現，續訪樣本與流失樣本之間，在政治興趣、政治知識、政治參與以及對訪問的態度上是具有差異的。因此，在母體的相關普查資料不包括上述選民的政治態度與政治參與的資訊的情況下，本文只能針對續訪樣本與流失樣本之間的差異作探討。
3. 至於「在1992年流失的621個樣本中，有多少無法追蹤」這個有關訪問執行程序的問題，並非本研究的重點，因此，作者並未加以探究。