

從美國經驗探討抽樣在普查的角色¹

余清祥* 胡玉蕙**

摘要

戶口普查的主旨在於調查一個國家或地區的人口及其相關特性，其資料對國家的人力政策規劃有重要的指標作用。隨著社會的變遷與調查方法的改變，普查費用持續增加，但資料的正確性卻因完訪率的下降與調查時間的限制，而受到空前的質疑。為提高資料品質與降低調查成本，美國引進一些與抽樣有關的方法，包括事後複查法(Post Enumeration Survey; PES)、未完訪者的追蹤調查(Non-response follow-up; NRFU)、整合抽樣與普查的涵蓋測量值(Integrated coverage measurement; ICM)等方法。本文以美國的經驗為背景介紹這些方法，並探討我國在公元 2000 年戶口普查中引進這些方法的可能性。

關鍵字：戶口普查；事後複查法；未完訪者的追蹤調查；涵蓋測量值。

¹ 筆者感謝政治大學洪永泰教授、主計處龍文彬專員與鄭敏祿科長的指正與建議。

* 余清祥(Jack C. Yue)；國立政治大學統計系副教授

** 胡玉蕙(Debbie Y. Hu)；國立台灣大學經濟系助理教授

ABSTRACT

The objective of census is to explore demography related characteristics of a country or an area, and the census data can be used to provide important indicators for constructing national manpower policy. However, the cost of conducting a census has been continuously increasing, but the data accuracy has been questioned because of low completion rate. To improve the data quality and reduce the cost, the United States plan to adopt new sampling methods, such as Post Enumeration Survey (PES), Non-Response Follow-Up (NRFU), and Integrated Coverage Measurement (ICM) in year 2000. In this study, we introduce these sampling methods used in the United States, and discuss the feasibility of applying them to Census 2000 in Taiwan.

Key Words: Census ; Post enumeration survey ; Non-response follow-up survey ;
Integrated coverage measurement

一、前言

人口資料的來源一般可分為戶口普查(Census)、抽樣調查(Survey)、以及生命統計登記系統(Vital Statistics Registration System)三種。其中戶口普查是在某一特定時間調查母體內的所有個體，如果資料可靠且所有個體都被數到(Count)，則戶口普查是最可靠的人口資料蒐集方式，可提供精確與詳盡的母體資訊。然而戶口普查所需的人力與物力非常龐大，又受限於特定時間內完成，通常只有國家層級的政府組織有能力完成；世界各國的戶口普查大多每十年實施一次，少數的國家(如加拿大)也有每五年一次。

近年來因為人口流動的加速、民眾對普查配合度的降低、以及人口資料需求的多樣化，使得戶口普查的精確度受到存疑，所提供的資訊也不敷各界需要，如何因應這些問題而提出改善的辦法，是歐美各國近來相當熱門的話題。其中抽樣因具有多樣性、經濟、與即時更新資訊的優點，除了可偵錯戶口普查的資料正確性，更可提供兩次普查間的資料更新與普查中沒有包含的資料，在今日世界各國節省公共支出的潮流中受到青睞，因而抽樣在普查中的角色也日趨重要，不再侷限在戶口普查資料的偵錯上。

本文將介紹美國普查中有關抽樣的理念與作法，探討抽樣在普查的角色演變，下一節先回顧美國上次的戶口普查中有關抽樣的觀念與方法，並概述其操作時所遭遇的問題；第三節則介紹美國近十年來在降低費用及提高資料精確度的目標下，為公元兩千年的戶口普查所引進更多、更廣範的抽樣研究，簡述其初步結果並討論其優缺點；最後一節則以美國的抽樣設計為參考，討論在我國戶口普查中引進類似抽樣方法的可行性。

二、美國 1990 年的普查經驗

美國的戶口普查結果用於決定國會(Congress)議員的席次(共 435 個席次)以及聯邦政府的預算，人口數的精確與否牽涉到地區的實質利益。由於過去普查經驗大多認定人口數有低估的傾向，尤其是在都會區及少數族群(如黑人)，各界對是否調整普查結果鬧得不可開交，報章雜誌及傳播媒體上經常可見學者專家對是否調整各抒己見、口誅筆伐，甚至需要大法官來裁決。1990 年的辯論尤其激烈，不但延續了 1980 年普查的爭議，更因統計方法與抽樣理論的發展，克服許多過去無法解決的問題，使得是否調整

普查結果與如何調整再度成為熱門話題。

雖然在 1990 年戶口普查前，美國商業部(Department of Commerce)公佈了是否調整的原則²(1989)，但大家對普查資料的正確性及用什麼方法調整仍舊意見不一。即使商業部長在 1991 年 7 月 15 日宣佈不調整普查結果後，仍有不少人無法認同，與商業部對簿公堂，纏訟近兩年後(1993 年 4 月 13 日)因法院宣判商業部勝訴才告一段落。之後美國統計學界對此議題繼續探討，但依然無法在統計學界獲得共識，對於普查結果是否因資料品質(Data Quality)而產生低估的現象並未獲得結論。(對此議題有興趣的讀者可參考 1994 年第四期的 *American Statistician* 中有修正普查結果的詳細討論。)

法庭上的敗訴並未減低普查局(Census Bureau)對抽樣的興趣，反而因為需要控制日益高漲的調查費用及個人隱私權的注重，抽樣的角色更加突出，在普查中增加抽樣的比重成為 1990 年普查以來的研究趨勢，近年來經常可見新抽樣方法的理論及實證研究結果的發表。本節以下先介紹 1990 年普查中使用的事後複查法(Post Enumeration Survey; PES)以及引起許多爭議的調整方法：二重系統估計(Dual-system Estimate; DSE)，而美國預定於公元 2000 年普查中新增的抽樣方法則於下一節詳細介紹。

(一)事後複查法(PES)：

PES 在美國向來用於評估戶口普查中涵蓋人口數的精確性，但由於人口數的低估有愈來愈嚴重的傾向，1980 年普查後與 1990 年普查前這段時間，有人提出以 PES 修正普查結果的想法，受到熱烈迴響，普查局也計劃實施 PES 且依照其結果調整普查數字。商業部卻持較保守的態度，對調整普查不預設立場，但對是否調整訂定了幾個參考原則，因此 1990 年 PES 的抽樣設計特別引人注目。

1990 年 PES 的抽樣單位(Sampling Unit)是由一個或多個區集(Block)所組成的區集叢聚(Block Cluster)，全國抽出 5,290 個區集叢聚(依分層隨機抽樣; Stratified Random Sampling)，最後共抽出 169,000 戶，或是 380,000 人。這些樣本再依年齡、性別、種族、是否租屋、居住城市型態、地區別(全國分為九區)分為 1,392 個事後層別(Poststrata)處理，作為調整普查結果的基

² Guideline 1: "The Census shall be considered the most accurate count of the population of the United States, at the national, state, and local level, unless an adjusted count is shown to be more accurate."

礎；其他關於 PES 的作業細節可參考 Hogan(1993)的說明。

(二) 二重系統估計(DSE)：

藉由比對(Match)普查與 PES 結果、以及透過重複捕獲模式(Capture-recapture Model)，可修正普查可能有的偏差。重複捕獲模式原先用於估計野生動物的數量，例如先捕獲 10,000 條鮭魚，每條鮭魚身上做好標示後再全數放回；之後再捕獲 100 條鮭魚檢查原先是否已點算到所有鮭魚。如果後來的 100 條鮭魚中 98 條有標示，表示鮭魚全體的(不偏)估計值是 $10,204(=10,000 \cdot 100 / 98)$ ，低估的比率約為 2.0%，這樣算出的估計值就稱為重複捕獲估計值。

每個人依照其是否出現在普查及 PES 中，可作以下的分類：

Census Enumeration			
PES	In	Out	Total
In	N_{11}	N_{12}	N_{1+}
Out	N_{21}	N_{22}	N_{2+}
Total	N_{+1}	N_{+2}	N_{++}

其中只有 N_{22} 觀察不到。在普查與 PES 結果互相獨立的假設下，所有人口數 N_{++} 等於：

$$N_{++} = (N_{+1})(N_{1+}) / N_{11}$$

DSE 的原理雖然簡單明瞭，但其假設條件是否成立與適用性引起不同的反應，1990 年代初期統計學家間的爭論即肇因於此，讀者可參考 Breiman (1994)與 Freeman and Wachter (1994)等人的討論。

三、 美國 2000 年的普查

為了解決不斷上升的普查費用與人口數低估比例提高、以及問卷回收率的持續下降等問題，美國預計在 2000 年的普查中引進(革命性的)兩個抽樣計劃：未完訪者的追蹤調查 (Nonresponse Follow-up; NRFU)與整合涵蓋估測 (Integrated Coverage Measurement; ICM)。其中 NRFU 屬於支出控制，針對未回函者作隨機抽樣；ICM 則為提高資料品質，主旨在於修正人口數低估比例。本節依此順序，先後介紹 NRFU 與 ICM 的操作方式與原理。

(一)未完訪者的追蹤調查(NRFU)：

美國在 1990 年普查郵寄問卷中約有 25.9% 的住戶未寄回，針對這些未回函者(Nonrespondent)雇用了 500,000 名訪員，以面訪的方式蒐集未回函者的資料；而普查總預算 26 億美元中約有 20% 用於支付這些未回函者的追蹤調查，佔了非常高的比例。然而以面訪覆查非常費時，資料品質也難有保障，某些未回函者的資料甚至是透過鄰居代為回答獲得。為解決此一問題，美國普查局預計在 2000 年普查引進 NRFU，以抽樣覆查代替全面覆查，希冀同時兼顧時效、成本控制兩個目標。預計 NRFU 可能採用以下三種方法之一(參考 Farber and Navarro, 1997)：

1. 直接抽樣(Direct Sampling)：原回函率不到 90% 的地區必須抽取樣本，使得最後的回函率至少可達 90%；而原回函率達到 90% 的地區則採十戶取一戶的抽樣；未回函者的資料將由其鄰居、及原先未回函但覆查時回覆者，依統計估計方法代入。
2. 90% 截略抽樣(90 Percent Truncation Sampling)：原回函率不到 90% 的地區由訪員逐一面訪(直接抽樣為抽取部分樣本)，使得最後的回函率至少可達 90%；未回函者的訊息則由十戶取一戶的樣本資料中獲得。
3. 時間截略抽樣(Time Truncation Sampling)：原始普查問卷回收結束後，立即以面訪的方式調查所有未回函者，訪員必須拜訪到所有未回函者(不論是否可找到)。面訪一段時間後，對於若回函率仍不足 90% 的地區，再抽取足夠的樣本使得最後的回函率至少可達 90%；而回函率達到 90% 的地區則採十戶取一戶的抽樣。

以上各方法均要求隨機抽樣，使未回函者被抽出的機會相等，初步的實證研究發現直接抽樣較為理想。

(二)整合涵蓋估測 (ICM)：

NRFU 針對的是時效與成本，並無法降低涵蓋誤差，美國普查局計劃以 ICM 提昇普查的精確度。ICM 的作法與 1990 年的 PES 非常類似，但樣本數卻更大(ICM 預計抽取 750,000 戶樣本，而 PES 原本要抽取 150,000 戶)。1990 年 PES 受限於樣本數的不足，對普查的修正只能以較大的區域來考慮，全國共分為 13 層。ICM 因抽取戶數的增加使普查修正值以州(State)為單位表示，而每一州的人口總數及其人口特性都可藉由該州的樣本戶求

出，減少估計值的誤差。公元 2000 年普查修正的方法與 1990 年類似，也是使用 DSE，讀者請參考上一節的說明。

Farber and Navarro (1997) 以 1990 年普查的資料為例，模擬 (Simulate) NRFU 與 ICM 的抽樣，研究發現引用抽樣方法得出的結果，在以普查的分區 (Tract)、國會議員選區、州、或全國為比較單位均較理想；只有對於每一區中人數較少的族群，修正普查結果反而較差。

四、 討論與建議

從以上兩節我們看到了美國普查局與學術界的專家學者們為改進普查業務所作的努力，雖然 1990 年以 DSE 修正普查結果的計劃最後作罷，但學界與普查局仍然努力不懈，不斷地測試、研究，繼續發展更節省、更有效率的抽樣方法為公元 2000 年的普查作準備。再回頭看我國政府與學界為公元 2000 年戶口普查所作的準備，起步雖然較晚，所幸國內已有許多先進在這方面作持續性研究，但他山之石可以功錯，其他國家的經驗與研究亦值得我們參考，以下是筆者根據美國經驗在抽樣方法上提出的建議。

台灣因採戶籍登記制度，容易掌握人口總數，美國的人口低估問題應該不會發生，然而台灣地區每十年一次的戶口普查，在投資了大量的人力及物力資源後，最後普查結果若僅作為戶籍資料的參考與編制十年一次的國民生命表 (Complete Life Table) 之用，筆者認為不免有些可惜。若普查目的在於檢驗基礎人口資料，不妨以丹麥、挪威、芬蘭等北歐國家的作法為範本 (參考龍文彬, 1998, 的介紹)，因為這些國家有完備的人口登記制度 (與台灣的現況相近)，基礎人口資料可透過行政紀錄檔案取得；若可加強行政紀錄檔案的連結運用，則可整合不同行政部門的統計檔案，供編製普查結果使用。(但屆時每十年一次的戶口普查是否仍有必要?)

然而誠如筆者前面提及，戶口普查的調查對象其涵蓋面最廣，當資料品質沒問題，所得的母體資訊將最為精確與詳盡；若社會各界需要的資訊 (例如重要的公共議題) 可列入普查問卷的問題中，普查的結果將較能反映全民的需求與意見。英國 1991 年的問卷設計可供我們參考 (Thompson, 1995)：除了歐盟 (European Communities) 各國均認可的基本 (Core) 問題外，其它問題是否可列入問卷中則諮詢政府各部門、地方、學術團體、與工商業界；在評估問題的結果是否必須藉由普查方能得知、該結果的可能

使用者、以及結果對使用者的價值等三項因素後，再經過各界的討論與建議後，決定問題是否可置入。

美國的長短問卷(Long and Short Questionnaires)也可作為我們降低成本與增加資訊蒐集的方法：當問卷問題過多時，可以讓大多數的民眾回答基本問題(短問卷，也是完整問卷的一部份)，僅選擇部份民眾回答完整問卷(長問卷)，美國 1990 年普查中約有 5/6 的住戶回答短問卷、1/6 回答長問卷，而 1990 年的結果也顯示長短問卷的回函率差距並不大(以成本的觀點而言)，僅有約 4.5% 的差別。台灣也可考慮測試長短問卷的成效，依估計精確性及調查成本的條件下選擇部份的民眾回答長問卷；對於填長問卷者中可能會有較多的未回函者可採 NRFU 的作法，使用類似直接抽樣的方式保證樣本資料的代表性。

長短問卷的調查方式也不一定要透過郵寄或面訪，電腦輔助電話訪問(Computer Assisted Telephone Interviewing；CATI)系統之類的電話訪問也是一種替代方案，因為台北市、高雄市與其他都市化程度較高地區的電話普及率也高，使用 CATI 可達到節約與縮短資料搜集時間，僅需對未回函者作追蹤調查。美國普查局在 1995 年對使用 CATI 蒐集長問卷資料作測試，初步的結果尚為樂觀，參考 Dawson 等人(1995)的報告。

另外，以美國自 1980 年代至今由普查局主導的戶口普查業務研究，與我國近年來戶口普查的經驗來比較，在總預算方面，我國在 2000 年普查約有新台幣 18 億元(平均每戶的普查預算不到 300 元)，約只有美國的 50 分之一(美國以 26 億美元計算)；若以平均每人而言，也只有美國的 5 分之一。在執行機構方面，美國的普查由專門機構普查局負責，有較充裕的時間與人力來規劃及研究普查業務；我國則僅有主計處第四局的一個科負責，其職掌業務不只戶口普查，規劃時間也不到三年。在如此懸殊的人力與物力的差距下，戶口普查的品質很難與美國比較，建議政府正視人口資料的重要性，設置專司戶口普查的政府機關(例如主計處增設戶口調查局)。

筆者不揣淺陋，提供以上幾點拙見，供普查設計執行單位與學界先進們參考，以收拋磚引玉之效，作為筆者日後深入研究之開端。

參考文獻

一、中文部份

- 行政院(1990)“中華民國七十九年台閩地區戶口及住宅普查法規彙編”。

- 龍文彬(1998)“公元二千年世界人口普查之趨勢及發展”，調查研究，中央研究院調查研究工作室，第五期，頁 37-74。

二、英文部份

- 1990 Population Census of Japan, Vol. 5, *Results of Detailed Sample Tabulation*.
- Belin, Thomas R. and Rolph, John E.(1994) “Can We Reach Consensus on Census Adjustment?”, *American Statistician*, Vol. 9, No. 4, 486-504.
- Breiman, Leo (1994) “The 1991 Census Adjustment: Undercount or Bad Data?”, *American Statistician*, Vol. 9, No. 4, 458-475.
- Brown, Robert L. (1991) *Introduction to the Mathematics of Demography*, 2nd edition, ACTEX Publications, Winsted, Connecticut.
- Childers, Danny R. (1996) “Integrated Coverage Measurement Processing Evaluations”, *Proceedings on the Section of Survey Research Methods*, American Statistical Association, 842-847.
- Dawson, Kenneth B., Sebold, Janice A., Love, Susan P., and Weidman, Lynn (1995), “Collecting Census Long Form Data Over the Telephone: Operational Results of the 1995 Continuous Measurement CATI Test”, *Proceedings on the Section of Survey Research Methods*, American Statistical Association, 590-595.
- Department of Commerce (1989) *Federal Register*, Vol. 54, 51, 002-51, 004.
- Farber, James and Navarro, Alfredo (1997) “A Comparison of Alternative Sampling Methodologies for Census 2000”, *Proceedings on the Section of Survey Research Methods*, American Statistical Association, 683-688.
- Freedman, D. and Wachter, K.(1994) “Heterogeneity and Census Adjustment for the Intercensal Base”, *American Statistician*, Vol. 9, No. 4, 476-485.
- Hogan, Howard (1992) “The 1990 Post-enumeration Survey: An Overview”, *The American Statistician*, Vol. 46, No. 4, 261-269.
- Hogan, Howard (1993) “The 1990 Post-enumeration Survey: Operations and Results”, *Journal of the American Statistical Association*, Vol. 88, No. 423, 1047-1060.
- Lessler, Judith T. and Kalsbeek, William D. (1992) *Nonsampling error in surveys*, New York, Wiley.
- Mahon, Basil and Pearce, David (1991) “The 1991 census of Great Britain:

Plans for Content and Output”, *Social Trends*, Vol. 21, 13-21.

- Riche, Martha F. (1997) “Should the Census Bureau use ‘Statistical Sampling’ in Census 2000”, *Insight magazine*, August 18, 1997.
- Schindler, Eric and Navarro, Alfredo (1996) “Effect of Sampling for Nonresponse Follow-up on Estimates from Sample Data”, *Proceedings on the Section of Survey Research Methods*, American Statistical Association, 371-376.
- Thompson, Eric. J. (1995) “The 1991 Census of Population in England and Wales”, *Journal of Royal Statistical Society, Series A*, Vol. 158, 203-240.
- Treat, James B. (1996) “The Effect of Questionnaire Length on Response”, *Proceedings on the Section of Survey Research Methods*, American Statistical Association, 734-739.
- Wright, Tommy (1998) “Sampling and census 2000: The Concepts”, *American Scientist*, Vol. 86, 245-253.
- Zaslavsky, Alan M. (1993) “Combining Census, Dual-system, and Evaluation Study Data to Estimate Population Shares”, *Journal of the American Statistical Association*, Vol. 88, No. 423, 1092-1105.