量化專家意見與人口推估
Combining Expert Opinions with Stochastic Forecast

Yu-Ting Cheng[1] 鄭宇庭　　Chen-Yu Lee[2] 李政豫　　Jack C. Yue[3] 余清祥

摘要

人口推估通常使用人口變動要素合成法，通常根據現有人口總數及結構，加上對未來生育、死亡、遷移三要素的預期(或情境)，推估未來的人口數及其結構；三要素中的生育和遷移變動趨勢較大，較不容易以歷史資料找出趨勢，大多倚賴專家訂定情境假設。但專家意見也有蒐集上的疑慮，例如意見沒有固定的形成格式，推估結果也難以機率角度詮釋發生可能性，而且推估數值大多不隨時間波動。為了彌補專家意見的不足，近年不少國家的人口推估使用隨機方法(Stochastic Methods)，這些方法大多與專家意見組合使用，使得推估結果可兼具兩種方法的特色。

專家意見需要適當的量化，方能與隨機方法整合，因此本文即以量化專家意見為探討目標。然而，專家意見大多以非數量化（或質性）格式呈現，整理及分析的方式視情況調整；另一方面，資料格式也會影響結果的精確性，也使制訂專家意見的量化規範更為不易。本文目的探討專家意見格式對推估結果的影響，並提出蒐集專家意見的方法，作為未來人口推估量化專家意見的參考。本文將先介紹常見專家意見的蒐集方法，以及對應的量化專家意見的資料格式，接著以電腦模擬評估資料格式對人口推估的影響，再衡量本文提出的專家意見蒐集是否合適。

關鍵詞：人口推估、隨機推估、專家意見、德菲法、電腦模擬

[1]政治大學統計系副教授，電子信箱：ting@nccu.edu.tw
[2]政治大學統計研究所碩士
[3]政治大學統計系教授，電子信箱：csyue@nccu.edu.tw

Abstract

Population projection is important to the government policy planning in many countries and the cohort component method is a popular method used to forecast the future population. The cohort component method needs the information of future fertility, mortality, and migration to project future population. These three factors are usually decided by expert opinions, as well as the historical data. However, there are concerns in using expert opinions. For example, it is not easy to quantify the expert opinions, the expert opinions do not have the meaning in probability, and the projection results via the expert opinion usually have small variances. In recent years, the stochastic methods combing with the expert opinions have been widely applied in order to improve the quality of population projection.

Quantifying the expert opinion is essential in combining with projection methods and the data format of quantification also plays a crucial role. In this study, we shall first review popular methods for quantifying the expert opinion, including the focused group interview and Delphi method. Then, we use computer simulation to evaluate the influence of data format on the analysis methods. It seems that, if the expert opinion is quantified in the form of payment card or binary data, the results are not very reliable. Therefore, we propose a sequential design to collect expert opinion and the simulation study shows promising results.

Key words: Population Projection, Stochastic Projection, Expert Opinion, Delphi Method, Simulation.

一、前言

人口推估(Population Projection)是預測一個國家或地區未來的人口總數及其結構，進而根據未來需求擬定對應政策，以達防範未然的效果。過去人口推估最常使用人口變動要素合成法(Cohort Component Method)，由基年(Baseline year)人口數及結構，加上規範未來生育(Fertility)、死亡(Death)、遷移(Migration)人口三要素的變動範圍，逐年將各參數變化值累加得到推估結果。目前臺灣官方人口推估亦使用這種方法，由行政院經濟建設委員會[4]人力規劃處負責，每兩年公布一次推估結果，最近一次推估在 2012 年公布。

圖 1、1960-2012 年臺灣總生育率
（資料來源：行政院經濟建設委員會）

由於生育容易受到政策與外在環境等因素影響，單從歷史資料推敲，未來趨勢大多只是複製過去的經驗，無法得到異於過去的結果。例如：根據 1960-2010 年的臺灣總生育率[5](Total Fertility Rate；TFR)資料(圖 1，1960-2012 年臺灣總生育率)，很難預料 2012 年總生育率會回升至 1.27，因此各國在推估生育

[4]因應國家發展委員會（以下簡稱國發會），經濟建設委員會已併入國發會。
[5]總生育率指平均每位婦女（一般是指 15 至 49 歲之間）一生中所生育之子女數，國際間評量及比較生育率即是以總生育率為標準。

或遷移時通常會採用專家意見(Expert opinion)。專家意見的格式大多為質性資料，優點是彈性較大且蘊含歷史資料所沒有的資訊，可由專家依專業知識彌補歷史資料無法反映的趨勢，但在蒐集上也有困難。例如：意見沒有固定蒐集及量化的方式，推估結果難以機率角度詮釋，而且數值大多不隨時間波動，像是經建會 2012 年的人口推估，未來生育率採用專家意見，分成低推計、中推計、高推計三種情境(圖 2)。

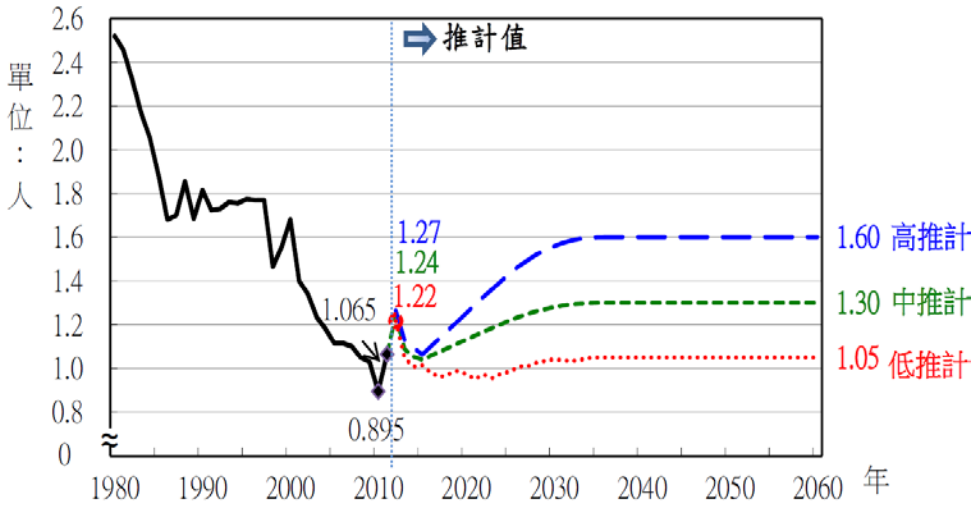

圖 2、採用專家意見的生育率推估

（資料來源：行政院經濟建設委員會「中華民國 2012 年至 2060 年人口推計」）

考慮專家意見的不足，近年來有學者提出隨機推估(Stochastic Projection)，以過去人口推估或是生命統計(Vital Statistics)數值代入數量模型，預測數值的未來變化，這些推估大多藉由電腦模擬進行，可以模擬出數值未來的波動變化，並給予專家意見在機率上的詮釋，例如：Stoto(1983)提出推估誤差法，專家意見的高、低推計值與推估誤差的 68%預測區間接近，約為一倍標準差的預測區間；余清祥與郭孟坤(2008)使用臺灣資料也得到相同結論。另外，有學者提出將專家意見視為一種情境，再與隨機推估模型合併，例如 Lutz et al.(1996)將專家意見當作情境，以時間序列模型進行推估；余清祥與郭孟坤(2008)使用評估電腦模擬應用於人口推估的可能，當未來趨勢與過去類似時，區塊拔靴法能提供可靠及穩定的推估結果，反之，則不推薦區塊拔靴法。由於隨機推估使用方

便，而且改善部份過去推估時所遇到的問題，隨機方法逐漸在人口推估上受到青睞。

專家意見、隨機推估各有優缺點，近年不少國家將兩種方法合併使用，研究如何來訂定生育、死亡、遷移三要素，以解決或降低推估誤差引發的問題，同時又可兼顧專家知識的推估方法。例如：Betz and Lipps(2004)推估德國 2002 年至 2050 年的人口，其中生育方面，以 1973 年至 2000 年的年齡別生育率資料配適常態分配曲線(Gaussian Curve)模型，至於遷徒方面，假設遵循一階自迴歸模型(AR(1) Process)；Betz and Lipps 發現他們的推估值高於德國官方推估值，與政府高、低推計的差距約為一個標準差，因此 Betz and Lipps 認為德國政府應再增大高、低推估之間的差距。

在結合專家意見與隨機推估時，通常量化專家意見較為困難，大多根據各國國情訂定專家意見的蒐集、整理方式，與主事者、推估當時的狀況有關，比較缺乏整體的標準作業程序。例如：台灣人口推估中的未來生育率假設，雖然歷年大多將專家意見分為高推計、中推計、低推計等三種情境，且以中推計為主要依據推估未來人口，但三種情境並非固定，在 2004 年有七種情境、2008 年有五種情境。另外，量化的資料格式也影響專家意見的呈現，甚至意見的估計結果，直觀而言，專家提供的意見愈籠統(或是可能範圍愈大)、估計結果愈不精確，因此資料格式也是設計專家問卷時的必要考量因素。

本文以專家意見的量化為研究目標，除了探討專家意見格式與估計精確性的關係，本文也將提出新的量化方式，可獨立操作、亦可結合德菲法等質性研究方法，將專家意見轉化為數據，提供更有效率的統計估計。以下將先回顧量化專家意見的相關研究，接著再提出本文的研究方法，並以電腦模擬比較不同資料格式與估計方法的關係。

二、文獻回顧

在引進本文的專家意見量化方法之前，本節先整理蒐集專家意見的方法，以及過去與人口推估（專家意見）有關的相關研究。與蒐集專家意見相關的方法大致可歸類質性研究，比較常見的方法包括焦點團體訪談、德菲法，或是這

兩種方法的混合，以綜合兩者的優點。以下先整理焦點團體訪談、德菲法的想法及操作方式：

(一) 焦點團體訪談(Focus Groups Interview)

二次大戰後由於商務蓬勃發展，市場調查員為了瞭解消費者對產品的看法，以促進商品的銷售量，集合與該商品有關的消費者(焦點團體)，蒐集這些參與者的意見及感受，這種意見蒐集方式稱為焦點團體訪談。焦點團體訪談是一個謹慎規劃的系列討論，目的在於蒐集質性資料(Qualitative Data)，以商品開發為例，根據意見在商品的設計、包裝或廣告期間反覆修改，突出消費者認為該產品認為最正向的特質。這種方法始於 1950 年代的美國市場調查研究公司，學術界在 1980 年代正式引進，擷取商界的策略並加以修改，使焦點團體訪談適用於其他領域(Krueger and Casey, 2000)。

每個焦點團體由一個主持人(moderator)帶領，通常有 4 到 12 個參與者，而參與者均具有與該焦點團體的主題有關的某些特質，訪談時透過團體成員的互動，使成員表達、分享個人經驗及想法，亦即藉由團體互動刺激思考(周雅容，1997)。焦點團體訪談為質性研究中最常用的方法，比個別訪談更容易發現新概念、新創意而且快速，能節省大量時間(Morgan, 1988)。此外，焦點團體的特色也包括在有共同聚焦的討論主題下，由於沒有必須達到某個共識的壓力，或是以檢驗某些假設或理論為前提，採取對參與者的發言採取開放性、批判式的評論，因此能蒐集更為多樣的資訊。

訪談主持人是焦點訪談的靈魂人物，通常同時扮演傾聽者、提問者的角色，又需維持討論不離題、平衡參與者的發言時間，使得最後蒐集的意見更能聚焦。此外，焦點團體的成員選擇不易，類似抽樣調查的樣本選取，需考量各種人口面向(例如：性別、年齡、居住地區、人格特質)，以期反映整個母體。例如：政府想要瞭解洗腎病患及其家屬對健保服務的觀感，對焦點訪談參與者的篩選因素宜加入家庭人口結構、經濟收入、就業狀態等，實際操作甚至會比醫學的病例對照研究(Case-Control Study)更複雜，很難面面俱到。意見的量化也是關鍵之一，因為焦點訪談蒐集質性意見，記錄上是以文字敘述為原則，但文字的數量化(或是編碼)並無固定規範，會因人、因時、或視研究需要而調整，不易與量化資料的結果結合。

(二) 德菲法(Delphi Method)

德菲法最早是在 1964 年由 Gordon and Helmer 提出，開始時用於解決複雜的軍事問題及軍事預測，後來逐漸被應用於任何領域的預測，如人口預測、醫療保健預測、經營和需求預測、教育預測等。德菲法(Delphi Method；或譯為疊慧法)是常見的專家意見蒐集方法，以不記名的書面通訊問卷，經過多次、來回地意見回饋及彙整，根據最後一次問卷的統計結果，概略估算出專家們預期的未來趨勢(王美鴻，2008)。因為需要多次的意見交換，規劃德菲法非常不容易，加上問卷格式、人力等配合，執行的難度相當高。

Dalkey (1969)認為，德菲法的理論有幾個假設：

1. 團體比個人擁有更多資訊；
2. 專家擁有專業知識，因此以專家進行預測或判斷堪稱合理；
3. 一群專家比其他群體更能提供正確資訊；
4. 匿名式訪問或問卷可以降低人際關係的負面影響，有利呈現真實意見，使少數意見得到尊重。

與焦點團體訪談不同，德菲法是匿名式的專家團體意見判斷方法，具有評估現況、預測未來的功能，主要以問卷蒐集專家的團體意見，問卷發放方式可依受訪專家方便而調整(郵寄或電子郵件)。因為匿名參與討論，原則上專家之間無法互相討論，較能保留個人的想法，降低因為意見相左、或是礙於情面不發表各人意見等之人際衝突，使得意見的面向更廣泛、更多元。另外，德菲法不侷限在質性資訊，亦能蒐集量化資料(例如：專家對於明年經濟成率的看法)，使用時更有彈性，且因其集思廣益、維持專家獨立想法、打破時空距離等優點，面對不明確性度高、具爭論性之議題更能發揮效果，像是企業人力資源的需求預測，以及新型態商品銷售量預測。

德菲法的第一次問卷的為開放式問卷，再整理所有意見後在回饋給所有專家，專家們再根據綜合意見修正原先的意見，接著再彙整、整理，如此反覆討論，逐步達到一致結論。然而，由於德菲法類似召開專家座談會，其過程通常比較複雜、花費時間較長(事前很難評估)，不像焦點團體訪談大約在 2 至 4 小時結束。另外，類似訪談主持人是焦點團體的靈魂人物，德菲法的問卷設計也很重要，像是提出簡單明瞭的問題，使得專家能夠在相同觀念及基礎上反映意

見；由於最後需達到一致結論，如果專家之間的差異性較大(例如：對未來看法、或是風險取向不同)，屆時如何彙整所有意見將是挑戰。

由於德菲法費時較久，近年也有學者提出修正式德菲法(Modified Delphi Method)為近年簡化德菲法的修正方法，結合焦點團體訪談(Focus Groups Interview)與德菲法(Delphi Method)兩者的優點，除了可節省時間外，並可提高問卷回收率(Murry and Hommons, 1995)。修正式德菲法的操作、統計方式與傳統德菲法大致相同，相異之處在於省略第一回合開放式問卷的繁複步驟，改以參考過去文獻或是採取焦點團體訪談設計第一回合問卷；而且對於達到一致結論的要求，也有更為具體的設定。(Holden and Wedman, 1993；林倫豪與徐昊杲，2012)

隨機方法近年在人口推估中漸受重視，這些方法大致可分為三種類型：一為推估誤差(ex post Method)、一為模擬情境(Random Scenario)、一為隨機推估(Stochastic Forecast)。第一類方法以過去推估誤差提供可能的誤差範圍，進而求出未來人口推估值(人口總數、年齡結構等)的預測區間；第二類方法由專家決定某些特定年度(例如：每隔十年)的人口總數(或生育率、死亡率和遷移)數值及其分配，中間年度的預測值再用平滑曲線補齊，進而得到總人口數的隨機分配。第三類方法假設生育率、死亡率和遷移遵循隨機模型，再利用過去資料去估計這些隨機模型的參數，進而推估未來總人口數的隨機分配。

第二類模擬情境法與專家意見較為有關，通常由專家決定幾個可能性，再根據電腦模擬等方法、結合這幾個可能性建立三個要素的分配函數，據此建立相關的預測區間。模擬情境法最早由人口學家 Lutz 和她的同事提出(Lutz et al., 2004; De Beer and Alders, 1999)，基本想法是先由專家提供生命參數值(Vital Rates)在幾個未來目標年度的高、中、低三個數值，再將此高、中、低值對應到某個常態分配的百分之五、百分之五十及百分之九十五的位置，而生命參數值在目標年度即假設遵循此常態分配；至於非目標年度，則用(線性)差補法或平滑的曲線填補，最後再用蒙地卡羅模擬求得這些生命參數值平均推估值及預測區間。

將專家意見加入隨機性的作法，與現行作法最為接近，經過適度修正後即可使用。模擬情境法的缺失則是在目標年度的生命參數值決定後，中間年度即

跟著確定並成為一直線或平滑曲線，缺少隨機上下震盪的現象(Lee, 1998)，例如：Lutz et al.(1998)利用奧地利資料做敏感度分析(Sensitivity Analysis)，發現此方法在短期時間內的確會呈現變異過小的現象，但在中長期時，則和自迴歸模型(Auto-regressive Model)不相上下。另外，大部分專家會偏向樂觀，產生了過窄的預測區間(Lutz et al., 1998)；生育率和死亡率間一般會有強烈的關聯性(Galloway, 1988)，例如：生育率和死亡率同時都受到社會經濟或醫療技術的影響。

上述與模擬情境的相關研究中，雖然提到在隨機方法中加入專家意見，其中意見均已量化，甚至可根據發生機率找出對應數值，但對於如何蒐集、量化專家意見並未詳細說明，至多只提出專家人數的建議。例如：可接受的專家人數範圍為 10 到 50 人 (Jones and Twiss, 1978)，20 人左右最適合，持續參加的專家至少需有 10 人以上(Couper, 1984)。根據這些專家人數的建議，我們將在下一節提出新的量化專家意見的方法，希冀能提供有別於歷史資料的相關領域知識，作為臺灣人口推估的參考。

除了專家意見的蒐集方法外，意見的量化格式也扮演重要角色，

我們在下一節透過電腦模擬，探討專家意見的格式與估計結果間的關連，原則上將意見的資料格式分為四類，並提出對應的分析方式，第四節則以電腦模擬驗證本研究的方法。

三、研究方法

本節根據問卷格式(或是專家提供的意見格式)，說明資料格式與分析方法間的關連。以數量化的程度來看，若能直接將專家意見以數值呈現當然最理想，但想法與觀念未必能轉化為文字，更何況進一步的量化，比較實際的作法或許是列出幾種可能、某些範圍，再由專家從中挑選最接近他們想法的選項。上述的資料格式與統計分析的資料格式(Data Format)類似，大致可將專家意見分為四個類型：比值(Ratio)、區間(Interval)、願付卡(Payment Card)、二元(Binary)，以下先定義這些類型的資料：

- 比值：假設每位專家可參考過去資料和自己想法，具體表達出未來生命統計(生育、死亡、或遷移)的數值。

- 區間：假設專家也能具體說出自己的看法，但意見並非以單一數值呈現，而是落至某個可能的範圍(亦即區間)之內。
- 願付卡：在問卷中列出幾個連續數值(例如：0-100 元、101-200 元、201-300 元、…等)，由專家從中找出最合乎想法的選項，此種方法可改善拒答率過高之缺點(或無法給予確切數值的窘境)。
- 二元：給予專家一個預設數值，由專家回答這個數值是否高於、或不高於心中預期(答案只有「是」、「否」)，代表其認為的數值是否高於詢問值。

對應於上述四種資料，以下介紹可行的統計分析，大致有三種類型的方法：動差估計法、最小平方估計量、Trimean 法。比值資料的分析最為單純，一般的分析方法大多可直接套用，如果專家意見服從常態分配 N(μ,σ^2)，則樣本平均數、樣本變異數可用於估計常態分配的期望值 μ、變異數σ^2，這個方法也稱為動差估計法(Method of Moment Estimator, MME)。另外，最小平方估計量(Least Squares Estimator, LSE)亦可使用，需先將觀察值排序，由 $x_1,x_2,\ldots,x_n$ 由小到大排序成 $x_{(1)},x_{(2)},\ldots,x_{(n)}$，再參考常態分配的百分位數(Percentile)轉換觀察值，作法類似 Z 分數(Z-Score)的作法，最後再仿造迴歸分析的最小平方法：

$$x_{(1)},x_{(2)},\ldots,x_{(n)} \to \mu+\Phi(0.05)\sigma,\mu+\Phi(0.05+\frac{1}{n})\sigma,\ldots,\mu+\Phi(0.05+\frac{n-1}{n})\sigma,$$

而(μ,σ^2)的估計值為最小化以下方程式的解：

$$\underset{\hat{\mu},\hat{\sigma}}{Min}\sum_{i=1}^{n}\left(x_i-\mu-\Phi(0.05+\frac{n+1-i}{n})\sigma\right)^2. \qquad (1)$$

另一種可行的方法為 Trimean 法(Tukey, 1977)，也是透過百分位數估計參數，但只使用四分位數(Quantile)、最小值、最大值，若以Q_1、Q_2、Q_3代表第一至第三個四分位數，則期望值估計值可表為：

$$\hat{\mu}_T=\frac{Q_1+2Q_2+Q_3}{4}\ ; \qquad (2)$$

仿造 Tukey 的想法，藉由電腦模擬找出合適c_1及c_2值，我們也提出依照百分數得出的變異數估計值：

$$\hat{\sigma}_T = \frac{1}{2}\left[\frac{x_{(n)} - x_{(1)}}{c_1} + \frac{Q_3 - Q_1}{c_2}\right] 。 \qquad (3)$$

如果區間中點(或接近中點)即為專家的想法，則區間資料無異於比值資料，估計方法當然可直接引上述三種方法：MME、LSE、Trimean；如果專家意見偏向於區間左側、或是區間右側的機會相當，預期估計效果也會與比值資料接近。但若所有專家都傾向於選擇區間的某一側，以上述處理比值資料的方式面對區間資料，勢必產生某種程度的偏誤，然而在缺乏專家偏好的相關資訊，上述三種方法仍屬可行方法，我們將在下一節以電腦模擬檢驗，區間資料因為專家偏好而產生的影響。

以願付卡的方式蒐集專家意見，有兩種可能結果：專家選擇某個區間的數值、專家選擇不在這些區間內，例如：要求專家提供 10 年以後的臺灣總生育率，願付卡的選項是[1.0, 1.2)、[1.2, 1.4)、[1.4, 1.6)、[1.6, 1.8)，其中[a, b)代表 a ≤ 總生育率< b。如果專家認為總生育率在上述四個區間內，像是[1.2, 1.4)，則仿造區間資料，可認定總生育率是區間端點的平均=(1.2+1.4)/2=1.3；如果專家認為總生育率會小於 1、或是大於 1.8，則資料紀錄會是(–∞, 1.0)或[1.8, ∞)，無法以區間中點的方式代入計算，直接刪除這種觀察值也會造成估計上的偏差。針對願付卡的資料特性，上述分析方法可透過重新加權、調整對應的百分位數等方式修正，操作方式視情況而改變，下一節電腦模擬中再詳細敘述。

二元資料的想法來自於願付價值(Willingness to Pay)，以順序提問法(Sequential Pretest)量化專家意見，專家對問卷題目回答「是」或「否」，或是以二分法(Dichotomous)切割可能性。傳統的願付價值會詢問每個受訪者一個或兩個問題，分別得到單邊區間(Single-bound)或雙邊區間(Double-bound)確定的資料，例如：如果專家認為總生育率會高於 1.0、但低於 2.0（即詢問了兩個問題），則專家意見應該介於[1.0, 2.0)，得到和上述願付卡的結果類似。本文採用修正的順序提問法(余純君，2000)，先依據過去資料或是由第一回合德菲法問卷設計初始詢問值，將這個初始值詢問第一位專家，依據專家的回答設定下一個詢問值，再以這個數值詢問第二位專家；如此反覆修正詢問值(類似德菲法)，直到獲得較為穩定的詢問值。

順序提問法具體的操作過程如下：

1. 由過去資料或第一回合德菲法問卷設計初始詢問值 x_1。

2. 以初始詢問值 x_1 詢問第一位受訪者，如果第一次回答是高於初始值 $(I_1=1)$，則 $x_2=2x_1$；如果回答是低於初始值 $(I_1=0)$，則 $x_2=x_1/2$，接著以 x_2 為詢問值詢問下一個受訪者繼續執行直至設定的最後步驟（或最後一個人）。

3. 跳升步驟：

(1)在詢問第 k 次時，若前一次回答是低於詢問值 $(I_{k-1}=0)$，問第 k 次時回答是高於詢問值 $(I_k=1)$。此時下一次的詢問值為第 k 次詢問值和大於 x_k 且最接近的詢問值平均，亦即 $x_{k+1}=\dfrac{x_k+x_\beta}{2}$，其中 $x_k<x_\beta$，x_β 為大於 x_k 最接近 x_k 的詢問值。

(2)在詢問第 k 次時，若第 $k-1$ 次時回答是高於詢問值 $(I_{k-1}=1)$，第 k 次時也回答高於詢問值 $(I_k=1)$。如果第 k 次詢問值 x_k 是曾經詢問過的值中最大的，也就是 $x_k \geq x_i, \forall i \leq k$，則下一次詢問值為第 k 次詢問值的兩倍，亦即 $x_{k+1}=2x_k$；若不是曾經詢問的值中最大的，則取大於 x_k 且最接近的詢問值 $x_{k+1}=x_\beta$ 其中 $x_k<x_\beta$，x_β 為大於 x_k 最接近 x_k 的詢問值。

4. 跳降步驟：

(1)在詢問第 k 次時，若前一次回答是高於詢問值 $(I_{k-1}=1)$，問第 k 次時回答是低於詢問值 $(I_k=0)$，此時 $x_{k+1}=\dfrac{x_\alpha+x_k}{2}$，其中 $x_\alpha<x_k.$，x_α 為小於 x_k 最接近 x_k 的詢問值。

(2)在詢問第 k 次時，若第 $k-1$ 次時回答是低於詢問值 $(I_{k-1}=0)$，第 k 次時也回答低於詢問值 $(I_k=0)$。如果 x_k 是曾經詢問過的值中最小的，也就是 $x_k \leq x_i, \forall i \leq k$，則 $x_{k+1}=x_k/2$；若不是最小的則取小於 x_k 的最近一個詢問值 $x_{k+1}=x_\alpha$ 其中 $x_\alpha<x_k.$，x_α 為小於 x_k 最接近 x_k 的詢問值。

上述詢問通常 20 次即可得到穩定的結果，如果每位專家詢問兩次，大致需要 10 位專家。另外，因為 20 次即可得到不錯的結果，分析時可使用最後一次詢問值、最後三次詢問值的平均、最後五次詢問值的平均，以這三種方法估計專家的期望值。

四、電腦模擬

本節將以電腦模擬評估上述的資料格式及其分析方法，並且測試需要多少位專家，才能提供比較穩定、可靠的意見期望值之估計。電腦模擬研究的探討順序和上一節相同，先假設專家能夠確切量化心中意見(比值資料)，且其意見為服從期望值 $\mu=1.5$、標準差 $\sigma=0.2$ 的常態分配變數 [6]。我們先以十位專家為例，每次模擬抽出十位專家(亦即 10 筆常態分配亂數)，分別根據這些亂數計算三種分析方法的估計值，這個步驟將重複 1,000 次。除了從 1,000 次模擬結果計算三種方法對於參數 μ、σ 估計值之平均數、變異數外，也計算這兩個參數的涵蓋機率(Coverage Probability)，或是每次模擬估計值$\pm 1.96\sqrt{\text{估計值之變異數}}$(建立參數 μ、σ 的信賴區間)，其中 1.96 對應於 95%信賴水準。

表 1 為十位比值資料的專家意見之電腦模擬結果，其中 Trimean 的 σ 估計值根據(3)式，在十位專家的情形下，以電腦模擬求得 c_1=3.0707，c_2= 1.1694。三種分析方法的結果大致相當，無論是參數 μ 或 σ 的估計值接近不偏，且三種估計值的標準誤大小也很接近。進一步比較三種方法的涵蓋機率，數值介於 0.936 及 0.964 代表大略滿足 95%信心水準的要求，三種方法也大致滿足，除了 Trimean 的 μ 涵蓋機率稍微低了一些，實證上這三種方法都可使用。除了十位專家外，我們也嘗試了五位、二十位等不同專家的情境，結果與表 1、表 2 類似，在此不贅述。

表 1、三種分析方法的估計值（比值資料）

	MME	LSE	Trimean
$\hat{\mu}$	1.501	1.500	1.497
(s.e.)	(0.062)	(0.064)	(0.067)
$\hat{\sigma}$	0.197	0.189	0.201
(s.e.)	(0.047)	(0.046)	(0.052)

表 2、三種分析方法的估計值涵蓋率（比值資料）

[6]這些參數值可視為總生育率相關的數值；另外，我們也嘗試其他分配的電腦模擬，限於篇幅、且模擬結果類似，本文只列出常態分配的結果。

	MME	LSE	Trimean
$\hat{\mu}$	0.946	0.955	0.934
$\hat{\sigma}$	0.950	0.954	0.948

區間資料的模擬設定和比值資料類似，但假設專家意見以 $x+d$ 的格式呈現，其中 x 服從期望值 $\mu=1.5$、標準差 $\sigma=0.2$ 的常態分配變數，d 服從均勻分配 $U(a,\ b)$，亦即意見以區間方式表達。在此也只列出 10 位專家 1,000 次模擬結果，而且為了探討區間寬度的影響，我們嘗試五種不同 $U(a,\ b)$的選擇，區間長度分別為 0.1、0.2、0.4、0.6、0.8。

表 3、三種分析方法的估計值（區間資料）

	參數估計	MME	LSE	Trimean
a=0.1, b=0.2	$\hat{\mu}$ (s.e.)	1.497 (0.065)	1.500 (0.064)	1.497 (0.069)
	$\hat{\sigma}$ (s.e.)	0.194 (0.046)	0.189 (0.046)	0.199 (0.051)
a=0.1, b=0.3	$\hat{\mu}$ (s.e.)	1.499 (0.063)	1.500 (0.064)	1.499 (0.067)
	$\hat{\sigma}$ (s.e.)	0.196 (0.048)	0.188 (0.045)	0.198 (0.054)
a=0.1, b=0.5	$\hat{\mu}$ (s.e.)	1.500 (0.063)	1.500 (0.063)	1.500 (0.068)
	$\hat{\sigma}$ (s.e.)	0.196 (0.046)	0.190 (0.046)	0.200 (0.052)
a=0.1, b=0.7	$\hat{\mu}$ (s.e.)	1.502 (0.064)	1.500 (0.063)	1.501 (0.069)
	$\hat{\sigma}$ (s.e.)	0.193 (0.046)	0.190 (0.045)	0.201 (0.053)
a=0.1, b=0.9	$\hat{\mu}$ (s.e.)	1.499 (0.066)	1.500 (0.062)	1.499 (0.070)
	$\hat{\sigma}$ (s.e.)	0.194 (0.046)	0.190 (0.045)	0.201 (0.053)

表 4、三種分析方法的估計值涵蓋率（區間資料）

	參數	MME	LSE	Trimean
a=0.1, b=0.2	μ	0.954	0. 947	0.952
	σ	0.956	0. 958	0.963
a=0.1, b=0.3	μ	0.964	0.952	0.962
	σ	0.945	0.952	0.937

a=0.1, b=0.5	μ	0.947	0.937	0.942
	σ	0.958	0.961	0.964
a=0.1, b=0.7	μ	0.960	0.946	0.951
	σ	0.950	0.951	0.936
a=0.1, b=0.9	μ	0.955	0.939	0.953
	σ	0.947	0.949	0.944

表 3 及表 4 分別列出參數 μ、σ 的估計值(及其變異數)，以及估計值的涵蓋機率，Trimean 的 σ 估計值根據(3)式，在十位專家的情形設定 c_1=3.0707，c_2 = 1.1694。三種分析方法的參數 μ、σ 估計值同樣都接近不偏，但 Trimean 的參數估計值之變異數稍微高一些，LSE 的 σ 估計值則較低，整體而言，三種方法皆能產生穩定精確的參數估計值，且不受區間寬度 d 的影響。參數的涵蓋機率結果也與比值資料類似，大約都與設定的 95%相近，實證分析時當估計值的變異數未知時，可搭配 Bootstrap 模擬求取估計值的變異數，建立參數 μ 與 σ 的信賴區間。

表 5、三種分析方法的估計值（非中心點的區間資料）

	參數估計	MME	LSE	Trimean
$d_1, d_2 \sim U(\mathrm{a}, \mathrm{b})$ a=0.1, b=0.2	$\hat{\mu}$ (st.d.)	1.498 (0.065)	1.499 (0.063)	1.498 (0.068)
	$\hat{\sigma}$ (st.d.)	0.195 (0.046)	0.190 (0.046)	0.200 (0.052)
$d_1, d_2 \sim U(\mathrm{a}, \mathrm{b})$ a=0.1, b=0.3	$\hat{\mu}$ (st.d.)	1.500 (0.065)	1.501 (0.064)	1.500 (0.069)
	$\hat{\sigma}$ (st.d.)	0.200 (0.048)	0.194 (0.046)	0.205 (0.053)
$d_1, d_2 \sim U(\mathrm{a}, \mathrm{b})$ a=0.1, b=0.5	$\hat{\mu}$ (st.d.)	1.500 (0.067)	1.500 (0.067)	1.500 (0.073)
	$\hat{\sigma}$ (st.d.)	0.212 (0.052)	0.205 (0.049)	0.215 (0.056)
$d_1, d_2 \sim U(\mathrm{a}, \mathrm{b})$ a=0.1, b=0.7	$\hat{\mu}$ (st.d.)	1.498 (0.076)	1.496 (0.077)	1.496 (0.081)
	$\hat{\sigma}$ (st.d.)	0.227 (0.053)	0.226 (0.054)	0.236 (0.063)
$d_1, d_2 \sim U(\mathrm{a}, \mathrm{b})$ a=0.1, b=0.9	$\hat{\mu}$ (st.d.)	1.500 (0.082)	1.498 (0.082)	1.499 (0.086)
	$\hat{\sigma}$ (st.d.)	0.252 (0.060)	0.243 (0.056)	0.259 (0.065)

在專家意見可量化成區間資料時，若真實意見不見得是區間中點，或是以 $(x-d_1, x+d_2)$ 的格式呈現，其中 x 為真實的專家意見，$d_1, d_2 \sim U(a,b)$。我們重複區間資料的電腦模擬，同樣假設 x 服從期望值 $\mu=1.5$、標準差 $\sigma=0.2$ 的常態分配變數，五種不同 $U(a, b)$的選擇，以及選出 10 位專家。限於篇幅，僅列出參數估計值及其變異數(表 5)。雖然 μ 的估計值也接近不偏，但估計值的變異數隨著區間寬度而變大，且與意見為區間中心點的情形不同，當區間寬度大於 0.4 時，σ 估計值明顯異於真實值。根據這些結果大致可推論(實際結果也如此，限於篇幅不列出)，μ 的估計值涵蓋機率仍舊可達 95%，但 σ 的估計值涵蓋機率隨著區間寬度增大而變差。上述區間資料的模擬結果，大略與原先的預期接近，如果專家意見不見得是區間中點時，本文使用的三種分析方法可得出接近不偏的 μ 估計值，但 σ 的估計值會偏大。

願付卡資料假設專家無法主動提供確切建議，仰賴蒐集者提供某些範圍後、再由專家選擇，如何決定這些範圍也是估計專家意見的關鍵之一，但限於篇幅、以及非本文探討目標，在此僅探討設定範圍與估計方法間的關連。為求簡化研究設定，假設願付卡的範圍為(–∞, 1.1)、[1.1, 1.3)、[1.3, 1.5)、[1.5, 1.7)、[1.7, 1.9)、(1.9, ∞)總共六個區間，而專間意見為服從常態分配(期望值 μ、標準差 σ）的變數，專家人數同樣是 10 人、重複電腦模擬 1,000 次。因為選擇兩個最外側的意見區間(–∞, 1.1)、(1.9, ∞)無法以區間中點代入，可以採取刪除意見、或是以端點 1.1、1.9 代入這兩種方式處理，因為刪除意見得出的估計結果較佳，以下只列出刪除離群值的估計結果。

表 6、三種方法的估計值（願付卡資料）

	參數估計	MME	LSE	Trimean
μ=1.1	$\hat{\mu}$ (s.e.)	1.281 (0.054)	1.200 (0.032)	1.104 (0.036)
	$\hat{\sigma}$ (s.e.)	0.125 (0.041)	0.127 (0.065)	0.074 (0.055)
μ=1.3	$\hat{\mu}$ (s.e.)	1.358 (0.054)	1.314 (0.062)	1.309 (0.065)
	$\hat{\sigma}$ (s.e.)	0.153 (0.038)	0.175 (0.056)	0.148 (0.053)
μ=1.5	$\hat{\mu}$ (s.e.)	1.502 (0.059)	1.497 (0.067)	1.502 (0.074)
	$\hat{\sigma}$	0.177	0.177	0.177

	(s.e.)	(0.035)	(0.043)	(0.051)
μ=1.7	$\hat{\mu}$ (s.e.)	1.637 (0.056)	1.685 (0.061)	1.689 (0.079)
	$\hat{\sigma}$ (s.e.)	0.153 (0.039)	0.170 (0.053)	0.146 (0.051)
μ=1.9	$\hat{\mu}$ (s.e.)	1.716 (0.060)	1.799 (0.031)	1.822 (0.049)
	$\hat{\sigma}$ (s.e.)	0.125 (0.037)	0.132 (0.065)	0.073 (0.055)

表 6 為不同期望值的模擬結果，因為提供的願付卡範圍以 1.5 為中心，當期望值 μ 愈遠離 1.5、預期估計結果愈不理想。表 6 結果與預期的想法吻合，當期望值 μ 數值與 1.5 愈遠，μ 與 σ 估計值都愈不理想，其中 Trieman 的 μ 估計值都接近不偏，但 MSE、LSE 兩種方法僅有在 1.5 時接近不偏。另一方面，標準差 σ 估計值有偏低的趨勢，尤其是在期望值 μ 與 1.5 距離愈遠時、σ 估計值的偏差愈大，其中又以 Trimean 的偏差較大。根據上述三種資料的模擬研究，專家若能清楚表達出意見，分析結果較為準確，當專家僅能以區間方式表達意見，則本文提出的三種分析方法都無法同時得出較可靠的 μ 與 σ 估計值。

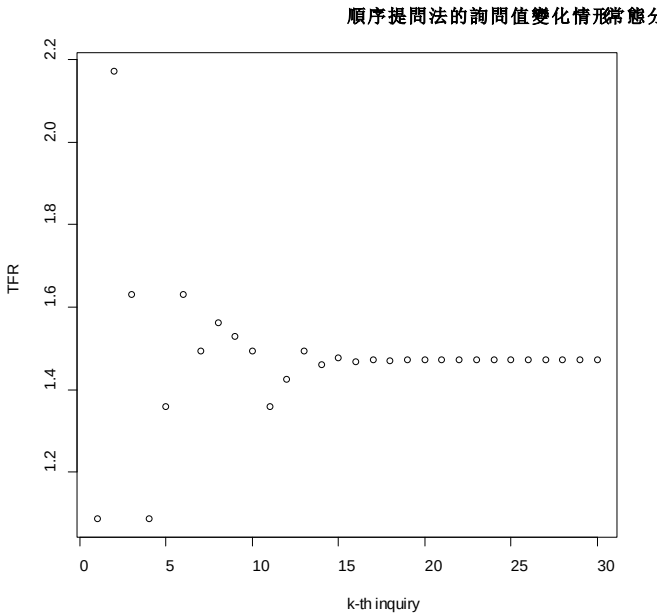

圖 3、順序提問法的詢問值與詢問次數間的變化關係

當專家僅能以「是」、「否」方式表達意見，可預期其估計結果勢必會比區間資料差一些，頂多只能提供百分位數(percentile)的估計值，且需要至少

1,000 位專家才能降低最大誤差至 3%。為求有效率地反映專家意見，本文針對二元資料採用上一節提到的順序提問法，希冀可得出不錯的參數估計值。與上述三種資料的模擬類似，假設專家意見服從常態分配（期望值 $\mu=1.5$ 、標準差 $\sigma=0.2$ ），且以最後一位、三位、五位專家的詢問值平均作為期望值的估計值；為探討所需專家的人數，在此仿造 Jones and Twiss(1978)的做法，嘗試專家人數設為 5 人至 30 人，評估對估計值的影響。另外，不同於願付卡資料，順序提問法不會受到第一次詢問值的干擾，在此假設第一位專家的詢問值為介於 1.0 及 2.0 的均勻分配亂數(其他範圍的模擬結果也類似，在此不贅述)。圖 3 為某次順序提問法的模擬結果，前 10 位專家的詢問值有不小震盪，第 15 位專家以後趨於穩定，這個結果也出現在以下兩個表格中。

表 7、順序提問法的三種估計值

專家人數	參數估計	最後一位	最後三位	最後五位
5	$\hat{\mu}$ (s.e.)	1.609 (0.564)	1.564 (0.227)	1.566 (0.224)
10	$\hat{\mu}$ (s.e.)	1.508 (0.250)	1.518 (0.217)	1.532 (0.183)
15	$\hat{\mu}$ (s.e.)	1.506 (0.157)	1.509 (0.164)	1.513 (0.147)
20	$\hat{\mu}$ (s.e.)	1.500 (0.124)	1.501 (0.123)	1.501 (0.106)
30	$\hat{\mu}$ (s.e.)	1.502 (0.090)	1.498 (0.087)	1.505 (0.086)

表 7 為順序提問法的 1,000 次電腦模擬結果。無論是最後一位、最後三位、或是最後五位專家的平均數，都能提供還算精確的期望值估計，與真實值間的差距在兩倍標準誤之內；另外，標準誤隨著專家人數增加而下降。原則上，大約 15~20 位專家即能提供不錯的 μ 估計值，且三種估計方法的結果都很接近。進一步計算三種估計方法的涵蓋機率(表 8)，也與預設的 95%非常接近，在 20 位及 30 位專家時，三種估計方法的結果都符合 95%信賴區間的要求，建議選用最後三位、或最後五位專家的詢問值平均，避免單一數值造成的震盪。

表 8、順序提問法三種估計值的涵蓋率

專家人數	涵蓋率	最後一位	最後三位	最後五位
n=5	μ	0.991	0.874	0.899
n=10	μ	0.946	0.966	0.947
n=15	μ	0.954	0.957	0.965
n=20	μ	0.963	0.962	0.949
n=30	μ	0.958	0.960	0.961

只要專家人數不少於 15~20 人，上述的結果顯示順序提問法能夠提供可靠的參考期望值，雖然 μ 估計值的標準誤大約為比值資料、區間資料的兩倍，但估計值較願付卡穩定，不會受到提供的參考值(第一次詢問值、或是願付卡範圍)影響。我們也嘗試了專家意見符合其他分配，順序詢問值也能取得可靠的期望值之估計值，無論是極度左偏(或右偏)的指數分配、對數常態(Log-normal)分配等，20 位專家大約都已足夠。因為順序提問法比較不受分配影響，且詢問的答案格式為「是」、「否」，受訪者比較容易回答，在缺乏專家意見的參考範圍時，我們推薦使用這個方法量化專家意見。另一方面，當專家意見服從常態分配時，順序提問法也能提供變異數的估計值，限於篇幅及人口推估大多注重期望值，本文在此省略變異數的討論。

五、結論與建議

人口推估提供一個國家或地區的未來人口總數及結構，政府可根據趨勢提早因應，避免問題一口氣爆發而無法收拾，近年臺灣生育率、死亡率下降而造成快速人口老化，因為政府及早規劃配套措施，像是 1995 年實施全民健保、2005 年勞工退休金條例(勞退新制)、以及規劃中的長期照護保險，都是與老化有關的社會保險制度。人口推估各國作法不一，臺灣大多根據生命統計數值的趨勢變化，再加入專家建議設定未來各數值，這樣的做法可結合主觀的專家意見與客觀的歷史資料，兼具兩者特色與優點。然而專家意見缺乏機率詮釋，且其數值不具有隨機波動，有時難以與量化資料結合，近年不少國家加入隨機推估，以彌補主觀意見的不足。由於生育率與遷移的趨勢變化較大，像是台灣的總生育率從 1960 年代的 6.0 急遽下降至 2010 年的 1.0 左右，單僅憑歷史資料及

隨機推估無法反映這種趨勢，必須仰賴專家提供專業見解，填補隨機推估無法跳脫過去趨勢的缺點。

但在形成專家意見的過程中，過去研究大多只描述意見蒐集步驟，較缺乏如何量化的具體作法，更遑論進一步的整併專家意見及隨機推估。有鑑於此，本文以專家意見的量化為研究目標，除了探討專家意見格式與估計精確性的關係，本文也提出新的量化方式，可獨立操作、亦可結合德菲法等質性研究方法，將專家意見轉化為數據，提供更有效率的統計估計。其中專家意見經過整併後大致可量化為四種資料類型：比值(Ratio)、區間(Interval)、願付卡(Payment Card)、二元(Binary)，再設計不同的分析方法將資料整合，並以電腦模擬比較不同量化方法的優劣。模擬結果顯示，量化的格式影響估計精確性，資料為比值、區間類型時，本文考慮的三種方法之估計結果都不錯，但願付卡容易受到資料範圍干擾，二元資料的估計準確性就更差，需要的專家人數也更多。

本文提出的順序提問法(Sequential Pretest)，專家針對詢問值高於心中預期回答「是」或「否」，適用於專家無法提供類似比值、區間等格式的確切意見時，且不像願付卡會受到詢問值影響，在缺乏過去專家的相關資訊時，可作為蒐集專家意見的替代方法。電腦模擬顯示在專家人數至少 15~20 人時，順序提問法可提供不錯的期望值之估計值，雖然估計值的變異數約為比值、區間資料的四倍，但涵蓋真實期望值的機率滿足預期，不失為蒐集、量化專家意見的可行方法。實際應用時若專家人數不足，也可套用願付價值(Willingness-to-pay)中的雙極限(Double Bound)，每位專家詢問兩次，所需專家人數可減半，電腦模擬顯示和詢問一次(Single Bound)的結果類似。另外，如果專家意見服從常態分配，順序提問法可提供變異數的估計值，但因計算過程較複雜，在此省略。

取得專家意見期望值的估計值後，可仿造 Lutz 等人(Lutz et al., 2004)的模擬情境，結合隨機推估的期望值及變異數，獲得具有機率意涵的人口推估。另一種可能是將蒐集到的專家意見及隨機推估的期望值、變異數，透過可信度理論(Credibility Theory)、貝氏理論(Bayesian Theory)、或是簡單的加權平均，整併專家意見及隨機推估兩種意見。實證上兩種意見差異很大時，權數(Weight)的挑選很重要，單單依賴可信度理論會有偏頗的疑慮，除了由專家決定權數的範圍外，也可仿造 Stoto (1983)的推估誤差法，由過去推估結果給予專家意見、隨機推估的機率詮釋，再轉化為合適的權數。當專家意見為不同資料類型時，或可考慮

整合分析(Meta Analysis)合併，或可將資料統整為較不精確的格式，像是比值資料視為區間資料，方便套入適當的估計方法(傾向於使用 Trimean)。

本文研究專家意見的量化，主要探討資料格式對估計結果的影響，但實際上蒐集專家意見也是關鍵，尤其是挑選專家、蒐集專家意見，對資料品質更是具有決定性的影響。Lutz 等人認為專家應該分為國內、國際兩個來源(Lutz et al, 2004)，其中國際專家可從聯合國或 International Institute for Applied Systems Analysis (IIASA)，或是瞭解台灣情況的外籍學者中挑選。我們建議資料蒐集可採取修正式德菲法，第一回合採取焦點團體訪談，並以訪談結果設計第一回合問卷，再整理第一回合問卷專家填答結果，包括平均數、標準差、眾數、四分位差等，作為設計後續問卷的參考(專家意見的可能區間)，在第二回合可透過願付卡獲得更精確的估計值。

參考文獻

一、中文部份

中華民國 2012 年至 2060 年人口推計 (2012) 行政院經濟建設委員會人力規劃處。

王美鴻 (2008) 疊慧法：以圖書館與資訊科學的應用為例，圖書與資訊學刊，Vol. 23, 45-60。

余純君 (2000) 願付價值及其前測的研究，國立政治大學統計研究所論文。

余清祥與藍銘偉 (2003) 台灣地區生育率模型之研究，人口學刊，Vol. 27，105-131。

余清祥與郭孟坤 (2008) 電腦模擬、隨機方法與人口推估的實證研究，人口學刊，Vol. 36，67-98。

林倫豪與徐昊杲 (2012) 保險技職教育校外實習專業能力指標之評選-修正式德菲法與層級分析法之應用，保險經營與制度，Vol. 11(2)，245-282。

周雅容 (1997) 焦點團體法在調查研究上的應用，中央研究院調查研究，Vol. 3, 51-73。

二、英文部份

Couper, M. R. (1984). “The Delphi technique: characteristics and sequence model.” *Advances in Nursing Science*, Vol. 7 (1), 72-77.

Betz, F. and O. Lipps. (2004). “Stochastic Population Projection for Germany-based on the QS-approach to Modeling Age Specific Fertility Rates.” No 4059, *MEA Discussion Paper Series from Mannheim Research Institute for the Economics of Aging, University of Mannheim.*

Dalkey, N. (1969). The Delphi method：*An experimental study of group opinion*.

CA：Rand.

De Beer, J. and Alders, M. (1999). *Probabilistic population and household forecasts for the Netherlands*. Working Paper No. 45. Paper presented at the Joint ECE-Eurostat Work Session on Demographic Projections, 3–7 May 1999, Perugia, Italy.

Galloway, P. R. (1988). “Basic Patterns in Annual Variation in Fertility, Nuptiality, Mortality, and Prices in Pre-industrial Europe.” *Population Studies*, 42(2), 275-302.

Holden, M. C. and Wedman, J. F. (1993). “Future Issues of Computer-mediated

Communication: The results of a Delphi Study." *Educational Technology Research and Development*, 41(4), 5-24.

Jones, H. and Twiss, B. C. (1978). *Forecasting technology for planning decision*. London, UK: Macmillan Press Ltd.

Krueger, R. A. and Casey, M. A. (2000). *Focus Group: A Practical Guide for Applied Research*, 3rd Edition, Sage Publications.

Lee, R. D. (1998), "Probabilistic Approaches to Population Forecasting", *Population and Development Review*, Vol. 24, Supplement : Frontiers of Population Forecasting. 156-190.

Lutz, W., Sanderson, W. and Scherbov, S. (1996). "Probabilistic Population Projections based on Expert Opinion." In *The Future Population of the World: What Can We Assume Today?* 2nd edition, Lutz ed., Earthscan: London; 397-428.

Lutz, W., Sanderson, W. and Scherbov, S. (1998). "Expert-Based Probabilistic Population Projections." *Population and Development Review*, Vol. 24, 139-155.

Lutz, W., Saariluoma P., Sanderson, W. and Scherbov, S. (2000). New Developments in the Methodology of Expert- and Argument-Based Probabilistic Forecasting. *IIASA Interim Report*, IR-00-020.

Lutz, W., Sanderson, W. and Scherbov, S. (2001). "The End of World Population Growth." *Nature*, Vol. 412, 543–545.

Lutz, W., Sanderson, W. and Scherbov, S. (2004). The End of World Population Growth. In Lutz & Sanderson (Eds.), *The End of World Population Growth in the 21st Century: New Challenges for Human Capital Formation and Dustainable Development* (pp. 17–84). London7 Earthscan.

Lutz, W. and Scherbov, S. (1998). "An Expert-based Framework for Probabilistic National Population Projections: The Example of Austria." *European Journal of Population*, Vol. 14, 1–17.

Morgan, D. L. (1988). *Focus Groups as Qualitative Research*. Newbury Park, *SAGE Publications.*

Murry, J. W. and Hammons, J. O. (1995). "Delphi: A Versatile Methodology for Conducting Qualitative Research." *Review of Higher Education*, Vol. 18, 423-436.

Stoto, M. A. (1983). "The Accuracy of Population Projections." *Journal of the American Statistical Association*, Vol. 78, 13-20.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.