

台灣地區生育率模型之研究

A Study of Fertility Projection Model in Taiwan Area

余清祥*

Jack C. Yue*

藍銘偉**

Ming-Wei Lan**

摘 要

近年來台灣地區由於婦女受高等教育比例、就業率的增加等因素，使得育齡婦女的生育時間延後、生育總胎次減少，因此總生育率也逐年降低，在 2001 年時降至 1.4 的歷史新低，加速台灣地區人口老化的腳步。台灣地區婦女生育率的下降，在各年齡層有不同的趨勢，這種各年齡層生育率變化不同步的現象，使得許多常見的生育率模型（如 Gamma 及 Lee-Carter 模型）套用於台灣地區的生育率時產生不小的誤差。為尋求更精確的台灣地區生育率模型，除了考量直接預測年齡別生育率的各種模型外，本文也嘗試以世代（Cohort）的觀點預測未來的生育率，包含世代生育率模型，以及研究胎次別生育率與年齡別生育率之間的關係，再套用國外近年提出的擴散（Diffusion）模型。本文的實證研究以西元 1949 至 1996 年（或 1975 至 1996 年，對胎次別資料

* 政治大學統計系副教授

Associate Professor, Department of Statistics, National Chengchi University

** 政治大學統計系碩士

Master, Department of Statistics, National Chengchi University

收稿日期 2003/05/23，接受刊登 2003/09/10

而言)的資料為基礎,1997至2001年資料為檢測樣本,藉由交叉驗證比較上述方法的優劣;此外,本文也考慮中國習俗對十二生肖的好惡,將龍虎年的效應列入預測方法的修正。研究發現世代模型可能因資料分析及估計的程序較為複雜,預測誤差反而高於單一年齡組個別估計法這類模型較為簡單的方法。

關鍵字：年齡別生育率、胎次別生育率、擴散模型、世代、交叉驗證

Abstract

The total fertility rates in Taiwan area have been falling rapidly since 1960's, which reached a record low 1.4 in 2001, and trigger the population aging in Taiwan area. Possible causes attributing to this rapid decrease include more opportunities of higher education and employment for the female, which generally are treated as two of the main reasons for the delay of childbearing. However, the falling of fertility rates vary quite a lot in different age groups, and thus most well-known fertility models (such as Gamma and Lee-Carter models) fail to produce accurate predictions.

In this paper, we shall extend the study of Huang and Yue (2002) and continue looking for acceptable fertility projection models. The models considered in this study include, in addition to traditional models for age-specific fertility rates, models for cohort fertility rate (CFR), age-and-parity specific birth rate, and diffusion. Data from 1949 to 1996 (or 1975 to 1996 for age-and-parity specific birth rate) are used as a pilot for verifying the model which has the best fit for data gathered from 1997 to 2001. We also consider the effect of Chinese zodiac to adjust the prediction values. We find that simpler models, such as individual group estimation, have smaller prediction errors comparing the complicated models such as the CFR model.

Key Words: age-specific fertility rate, age-and-parity specific birth rate, diffusion model, cohort, cross validation

壹、前言

近年來，台灣由於教育普及、經濟發展等各種因素，已由開發中國家邁入已開發國家的行列。而與世界大多數已開發國家「少產少死」的特性類似，台灣地區居民生育率及死亡率的數值均較 1960~70 年代經濟開始起飛時低了許多，其中生育率的降低尤為顯著，2000 年的總生育率僅達 1960 年的 29%（參考表 1），近幾年來生育率下降更加明顯，屢創歷史新低。這種生育率逐年下降、屢創新低的現象並非台灣專有，由表一可知，除了美國總生育率在 2000 年前後微幅升高外，亞洲、美洲、歐洲、澳洲等地的先進國家的總生育率都比 20 年前低了不少，因之我們認為台灣生育率的下降應是不可避免的趨勢。

表 1 世界各先進國家歷年之總生育率

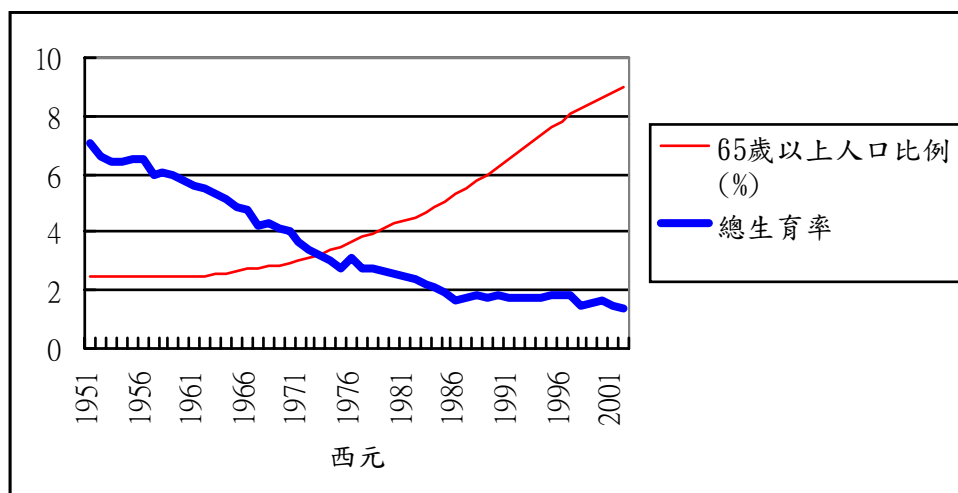
西元	台灣	日本	美國	加拿大	瑞典	義大利	澳大利亞	紐西蘭
1960	5.75 [*]	2.01 ¹	3.64 ¹	3.81 ¹	2.17 ¹	2.31 ¹	3.45 ¹	4.03 ¹
1980	2.51 [*]	1.78 ¹	1.87 ¹	1.75 ¹	1.68 ¹	1.66 ¹	1.89 ¹	2.04 ¹
2000	1.67 [*]	1.74 ^c	2.06 ^c	1.64 ^c	1.53 ^c	1.18 ^c	1.79 ^c	1.80 ^c
2001	1.36 [*]	1.41 ^c	2.06 ^c	1.60 ^c	1.53 ^c	1.18 ^c	1.77 ^c	1.80 ^c
2002	1.30 [*]	1.42 ^c	2.07 ^c	1.60 ^c	1.54 ^c	1.19 ^c	1.77 ^c	1.80 ^c

資料來源：1. ^{*}內政部統計處；

2. ^c Central Intelligence Agency—The World Factbook；

3. ¹ Long-range World Population Projections, United Nations。

台灣地區的人口老化近年因「少產少死」而有加速的傾向。例如：1993 年年底台灣地區 65 歲以上的老年人方纔突破 7%，成為聯合國定義的高齡化社會，不到十年的時間，在 2002 年年底又突破 9%。這個快速的人口老化趨勢與生育率的降低似乎存有同步關係，圖 1 是 1951 年以來 65 歲以上老年人口佔總人口數的比例及總生育率的趨勢圖，可以發現當總生育率下降的同時，65 歲以上老年人口比例也持續上升，尤其是自 1980 年代中期總生育率低於替代水準後，65 歲以上人口比例的上升更為明顯。



資料來源：1. 行政院經建會人力規劃處；
2. 內政部統計處。

圖 1 65 歲以上老年人口比例及總生育率趨勢圖

雖然台灣總生育率近幾年來持續地下降，但是由於不同年齡組的生育率變化方向不盡相同，例如：20 至 24 歲的年齡別生育率逐年下降，但 30 至 34 歲則自 1990 年起反而呈現緩慢攀升。這種各年齡層生育率不同步的現象，使得許多常見的生育率模型（如 Gamma 及 Lee-Carter 模型，參考黃意萍與余清祥（2002）的討論）套用於台灣地區的生育率時產生不小的誤差。為尋求更精確的台灣地區生育率預測模型，本文將以統計分析的觀點探討未來趨勢，除了使用主成份分析法預測年齡別生育率外，也將引進胎次別生育率，與之前有較佳預測結果的單一年齡組個別估計法比較。此外，除了以上提及較常見的橫斷面（Cross-sectional）模型，本文也考慮從縱斷面（Longitudinal 或 Cohort）的角度來預測生育率，例如：Zeng and Land（2002）建議的擴散（Diffusion）模型，與以整個世代為考量的世代生育率。

本文的安排如下：第二節為文獻探討及模型介紹，介紹本文使用的各種預

測台灣生育率的模型；第三節為實證分析，透過台灣歷年生育率資料，藉由交叉驗證比較各種生育率模型的優劣；第四節則加入中國習俗中的龍虎年效應修正生育率預測值；第五節為結論與討論，探討各生育率模型應用在台灣生育率的預測時，需要考慮的事項。

貳、文獻探討及模型介紹

世代生育率（Completed Cohort Fertility Rate，簡稱 CFR 或 CCFR）以每一世代（即縱斷面）各年齡層的方式，紀錄一個婦女一生中生育的子女總數，較能反映各世代實際的生育率變化。但因婦女的生育年齡涵蓋時間長達三、四十年（我國記錄以 15 至 49 歲為原則），若欲比較兩個不同世代的生育行為，需要等至婦女都過了生育年齡方能進行。這樣的方式通常不能符合實際需要，因此才採用總生育率（Total Fertility Rate，簡稱 TFR），折衷地以現行同一年度（即橫斷面）的方式記錄生育率。其中 TFR 的定義為

$$TFR^z = \sum_{x=\alpha}^{\beta} {}_n f_x^z = \sum_{x=\alpha}^{\beta} \frac{{}_n B_x^z}{{}_n F_x(z)},$$

其中 TFR^z 代表第 z 年的總生育率， ${}_n f_x^z$ 為第 z 年 x 歲至 $x+n-1$ 歲的婦女的生育率， ${}_n B_x^z$ 是第 z 年 x 歲至 $x+n-1$ 歲的婦女所生之嬰兒數， ${}_n F_x(z)$ 是第 z 年 x 歲至 $x+n-1$ 歲的年中婦女數， α 及 β 即代表婦女生育年齡的上界及下界。

CFR 與 TFR 的比較如圖 2，由左上至右下的直線為 CFR 的計算方式，由上垂直而下的直線為 TFR。圖 2 中的 CFR 代表的是 1965 年時年紀為 15 歲的婦女，當她們到 50 歲時累積一生的生育數，期間經歷了 35 年；TFR 則代表 1990 年時從 15 歲（在 1975 年出生）至 50 歲（在 1940 年出生）的婦女生育率。雖然 TFR 可能將不同生育行為的婦女放在一起計算，但因不需要等待 35 年，使用上較為便利，因此生育率比較大多仍以 TFR 的格式進行，而非原先認為較為合理的 CFR。

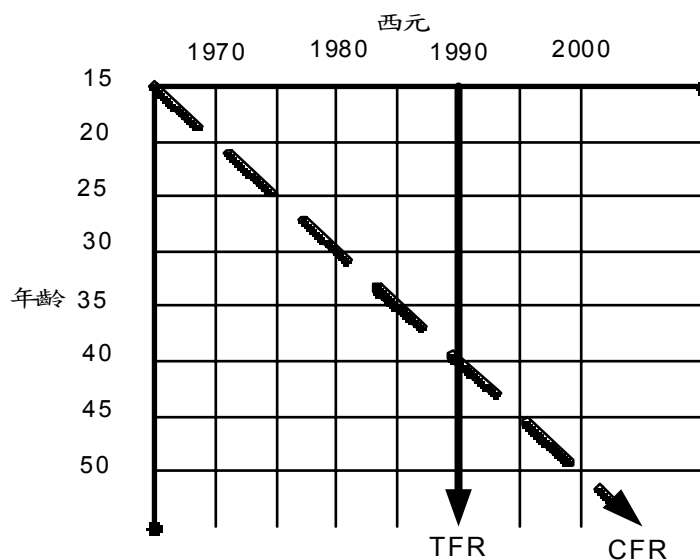


圖 2 TFR 與 CFR 比較圖

採用 CFR 與 TFR 的分別，本文考慮的模型可依橫斷面、縱斷面的分成兩類。屬於橫斷面的方法有：一、年齡別生育率的主成份分析法，二、胎次別的主成份分析法，三、單一年齡組個別估計法；屬於縱斷面的方法有：一、擴散模型，二、CFR 法。以下分別介紹各種方法：

一、年齡別生育率主成份分析法

主成份分析法用於生育率首先由 Bozic and Bell 於 1987 年提出，Lee and Carter (1992) 也使用類似方法，應用在死亡率的預測上。主成份分析法的基本想法為將所有的年齡別生育率（或生育率的對數）視為多個變量，透過主成份分析縮減維度達到資料縮減 (Data Reduction) 的目的。因為各主成份之間互相獨立，可使用迴歸分析或時間序列方法分別對每一主成份預測未來走勢，將原先多變量的預測問題簡化為幾個單變量的預測。

Lee and Carter 於 1992 年所提出的 Lee-Carter 法可視為年齡別主成份法的特例，只是 Lee-Carter 法是先將年齡別生育率取對數。Bell (1997) 以美國白人的死亡率比較 Lee-Carter 模型與主成份分析法的差異，發現 Lee-Carter 模型的估計誤差介於一個主成份與兩個主成份之間，這個結果與曾奕翔及余清祥 (2002) 使用台灣地區死亡率資料分析的結果類似。本文仿照死亡率的研究，嘗試將年齡別生育率、年齡別生育率的對數套用主成份方法 (包括 Lee-Carter 模型)，預測未來的年齡別生育率。

二、胎次別生育率主成份分析法

胎次別生育率 (Parity Specific Fertility Rate) 的觀點類似隨機變數中機率密度函數 (Probability Distribution Function, 記為 $f(x)$)，因為

$$E(X) = \sum x \cdot f(x),$$

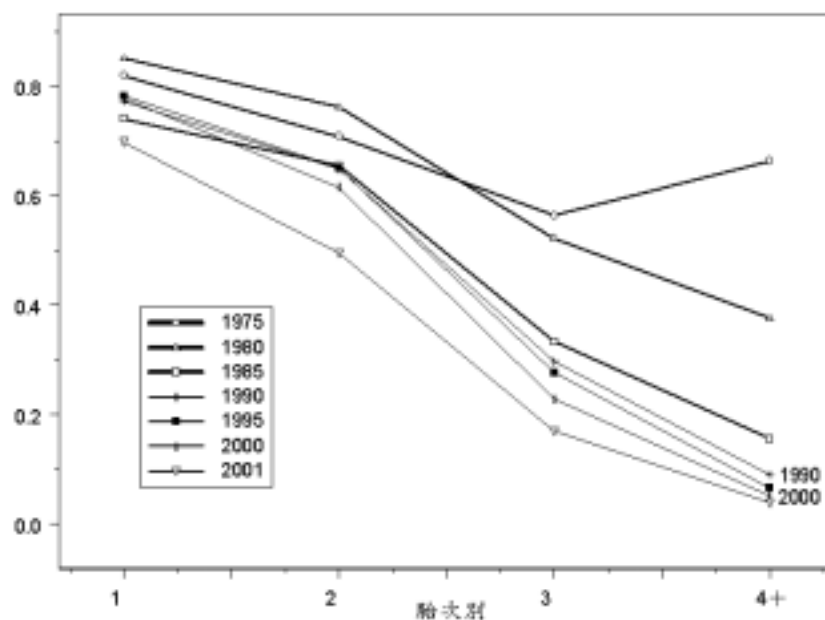
只要知道婦女生一胎、兩胎、... 的機率，總生育率即可代入期望值的公式算出。事實上，如果對所有 $x \geq 3$ 都滿足 $f(x) = 0$ ，總生育率一定不大於 2。胎次別生育率的定義如下：

$$TFR^z(i) = \sum_{x=\alpha}^{\beta} f_x^z(i),$$

其中 $TFR^z(i)$ 代表在第 z 年第 i 胎的胎次別生育率，而 $f_x^z(i)$ 為在第 z 年 x 歲婦女第 i 胎的年齡別胎次別生育率 (Age-and-Parity Specific Birth Rate)， $\alpha (=15)$ 及 $\beta (=49)$ 分別代表生育年齡的下界與上界。

台灣地區自 1975 年 (民國 64 年) 起開始紀錄各胎次別生育率，圖 3 為台灣 1975 年至 2001 年的各胎次別生育率，4 胎與 4 胎以上綜合成一組來表示。個別胎次別生育率與總生育率類似，大體上近年來都呈現下降的趨勢。胎次別生

育率的模型配適也將採主成份分析法，希望可由較少個數的主成份達到資料縮減的目的，便利未來的生育率預測。



資料來源：內政部統計處。

圖 3 1975 年~2001 年 1~4 胎以上合併之胎次別生育率

三、單一年齡組個別估計法

前兩種方法本質上假設各年齡組間的生育率相互關聯，在套用主成份分析方法後分離出互相獨立的幾個主成份，可對每個主成份個別找出最合適的預測模型，換言之，即是將多變量問題簡化為幾個單變量的問題。單一年齡組個別估計法則是直接將每個年齡組視為獨立，然後分別預測每個年齡別生育率的變動趨勢。單一年齡組個別估計法的使用細節及說明參考黃意萍與余清祥(2002)。

四、擴散模型

總生育率(TFR)有時無法衡量出真正生育率的變化情形，例如：如果從今年起婦女一生的生育總數不變、平均延後兩年，則從現在起的十年內總生育率應會下降，但不代表婦女一生的生育總數下降。Bongaarts and Feeney (1998) 認為生育率的計算需將生育時機 (Tempo Component) 與生育數量 (Quantum Component) 同時列入考慮，以去除因生育時間的延後引起總生育率下降的假象。他們考慮『時間對世代的生育數量等式』(Period-Cohort Quantum Equation, 簡稱 B-F 法)，公式為

$$TFR'^t = CFR^t = TFR^t / (1 - r_p(t))$$

上式中 TFR' 以及 CFR 代表的是生育數量，也就是在去掉生育時機影響後的生育數量，而 $r_p(t)$ 是『第 t 年觀察到的生育時機改變率』：

$$r_p(t) = 0.5 \times (MAC(t+1) - MAC(t-1))$$

其中 $MAC(t)$ 為第 t 年的平均生育年齡 (Mean Age at Childbearing)。

2002 年 Zeng and Land 提出以 $r_c(t)$ 修正 $r_p(t)$ ：

$$r_c(t) = \frac{r_p(t)}{1 - r_p(t)}$$

代回原先的公式可推得：

$$TFR'^t = CFR^t = TFR^t / (1 - r_c(t))。$$

本文在此引用擴散模型的概念，將胎次別生育率分解為生育時機與生育數量兩個分量，以求出真正的生育率變化。進行步驟為藉由每個胎次的年齡別胎

次別生育率，求出第 t 年第 i 胎的平均生育年齡（即 $MAC_i(t)$ ），再透過主成份分析法估計生育年齡、再修正胎次別生育率，最後轉換回年齡別生育率。

五、CFR 法

本文假設 CFR 滿足 Gamma 模型，也就是：

$$f(x_t) = \frac{\beta_t^{\alpha_t}}{\Gamma(\alpha_t)} x_t^{\alpha_t-1} e^{-\beta_t x_t} = C_t x_t^{\alpha_t-1} e^{-\beta_t x_t}$$

其中 x_t 代表在第 t 年時為 x 歲該世代的年齡， $f(x_t)$ 代表這個世代在第 t 年 x 歲時的單一年齡生育率， C_t 代替 Gamma 模型的調整係數。將上式等式兩邊同取對數：

$$\log f(x_t) = \log C_t + (\alpha_t - 1) \log x_t + (-\beta_t) x_t + \varepsilon_t,$$

可套用迴歸方法找出模型的參數。

參、實證分析

本節使用的資料為將內政部統計處 1949 年至 2001 年（民國 38~90 年）台灣地區 15 歲至 49 歲的五歲一組婦女生育率、以及年齡別胎次別生育率（僅有 1975 至 2001 年），分別考慮上述介紹之五種生育率模型，先以 1975 至 1996 年¹的資料估計模型參數，再以 1997 至 2001 年資料當成檢測樣本，計算誤差大小，並以此誤差比較模型的優劣。以下分別說明資料分析細節及預測結果：

一、年齡別主成份分析法

使用主成份分析法發現年齡別生育率、年齡別生育率的對數兩者的 3 個主

¹ 因為胎次別生育率完整的資料從民國 64 年開始，為求資料使用一致，因此除了 CFR 法外，其餘方法皆使用民國 64 年以後的資料。

成份皆有接近 0.99 的解釋能力(附錄 1)，取對數後的解釋能力都比原始資料高，因此本文以下只考慮年齡別生育率的對數。分析發現第 1 個主成份對時間作迴歸分析的解釋能力(R^2)有 65%以上，第 2 個及第 3 個主成份則以時間序列模型進行預測。求出 3 個主成份的預測值後，再代入負荷便可求出 1997 年至 2001 年之年齡別生育率及總生育率的預測值。由於兩個主成份（解釋能力 0.925）的預測誤差在主成份方法中最小，以下的討論將以兩個主成份的模型代表年齡別主成份法。

二、胎次別生育率主成份分析法

由主成份分析發現胎次別生育率或胎次別生育率的對數可篩選出兩個主成份（附錄 2），都有約 0.985 的解釋能力，與年齡別生育率類似。預測出胎次別 2 個主成份的未來趨勢後，再代回求出年齡別生育率的預測值，進而求出總生育率之預測值。

三、個別估計法

本文考慮以 7 個年齡組的（對數）年齡別生育率當應變數，以時間當自變數來配適迴歸模型，殘差則以時間序列模型進行預測。預測結果發現取對數後的年齡別預測結果較理想，推測其原因可能為近年來台灣各年齡別生育率變動不穩定，因此取對數後可將其變動減小，而有較佳的預測結果。

四、擴散模型

首先分別求出 5 個胎次的平均生育年齡，再透過擴散模型理論將 5 個胎次的胎次別生育率分解為平均生育年齡及修正後的胎次別生育率兩個變數。分析發現第 1 至 5 胎的平均生育年齡從 1975 年至 1996 年大致都有明顯上升的趨勢，表示台灣婦女近年來生育年齡的延遲。其中第 1 胎的平均生育年齡由 1975 年的 23.71 歲延後至 1996 年的 26.15 歲，而生育第 2 胎的平均年齡則從 25.29 歲延後至 28.15 歲；修正過後的第 1 胎及第 2 胎的胎次別生育率下降的幅度並不明顯，

但修正過後的第 3 胎、第 4 胎及第 5 胎的胎次別生育率快速的下降，修正後的第 3 胎胎次別生育率從 1976 年的 0.65 下降至民國 1995 年的 0.30，這個結果說明了台灣現代婦女生育超過 2 胎的比例明顯下降。

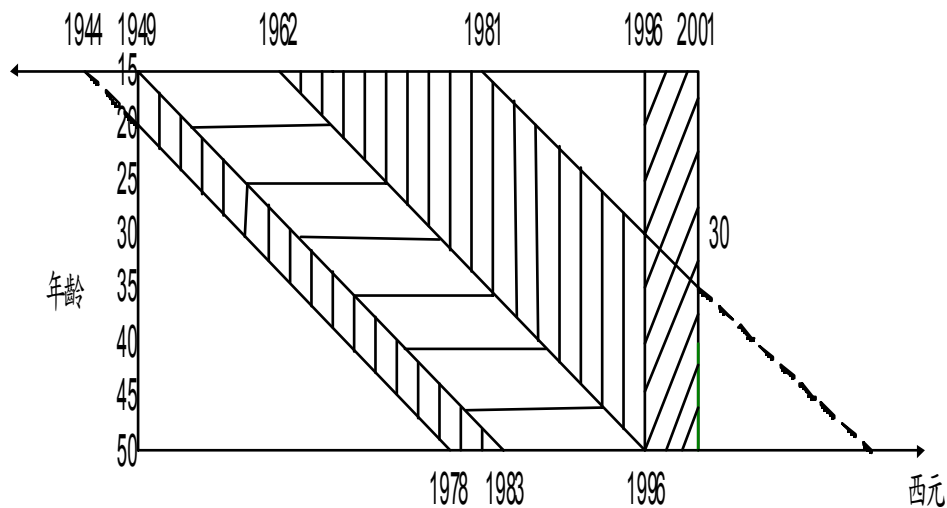
接著以主成份法分別對 5 個胎次的平均年齡、修正後的胎次別生育率進行分析，發現 3 個主成份的解釋能力都在 0.985 以上。之後的步驟與胎次別生育率模型類似，考慮對胎次別生育率（或其對數值），以主成份分析法篩選出 2 個主成份。

本方法與胎次別生育率模型頗為類似，不同之處在於擴散模型之前半部是以類似世代（即縱斷面）的觀點去估計胎次別生育率，而後半部則與胎次別生育率模型相同，是以綜合不同世代（即橫斷面）的觀點預測年齡別生育率。分析結果顯示，不管平均生育年齡與修正過後的胎次別生育率取 1 個、2 個或 3 個主成份估計胎次別生育率，30 至 34 歲這一組的年齡別生育率的配適誤差都不小。

五、CFR 法

由於計算一個世代生育率需要 35 年的觀察，因此這部份的資料配適使用 1949 年至 1996 年的資料，與其他四種方法所使用資料範圍不同。在這些資料中，有完整世代生育率的部份計有 1949 年時 15 歲的世代到 1962 年時 15 歲共有 14 個世代，其他世代或是僅有年輕（接近 15 歲）、或是僅有年紀較大（接近 50 歲）的不完整資料。為求使用最多的資訊，除了使用 Gamma 參數模型，我們也將包含完整的 20 至 30 歲資料的世代列入考慮。（考慮 20 至 30 歲資料的理由有二：以橫斷面來看，這 3 組的年齡別生育率佔了總生育率的 55% 以上；以縱斷面來看，Gamma 函數最重要的高峰部分，也就是模型最具特色之處就在 20 至 30 歲間。）因此，除了之前的完整的 14 組外，再加上 1944 年至 1948 年 15 歲的世代，以及 1963 年 1981 年 15 歲的世代，一共 24 個不完整的世代，共計 38 個世代進行資料分析。以圖 4 表示 CFR 的預測方法。橫軸為時間、縱軸為年齡，

實線代表已有生育率資料，虛線代表沒有的生育率資料，水平線部分為完整世代，垂直線部分為不完整世代，斜線部分是所要預測的年度。



註：橫線部分為完整世代。
直線部分為不完整世代。
斜線部分為所要預測的年度。

圖 4 CFR 法的資料使用說明

將 38 個世代的生育率資料分別以最小平方方法配適 Gamma 函數，求出參數估計值，發現每個世代配適出的結果其解釋能力都有 95% 以上（附錄 3）。接著用迴歸分析加時間序列方法，以時間為自變數，分別配適 3 個 Gamma 模型的參數，進而求出 1982 年之後所有 15 歲世代 Gamma 參數預測值，進而求出 1997 至 2001 年生育率的預測值。另外，1982 年 15 歲以後的世代其 29 歲前的單一年齡生育率已有觀察值，可採用等比例修正來調整單一年齡生育率的估計值。

以上述五種方法得出 1997 年至 2001 年之年齡別生育率與總生育率之預測值後，以下列之誤差標準作為本文交叉驗證中判斷模型優劣：平均絕對誤差

(Mean Absolute Percentage Error, 簡稱 MAPE) :

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\varepsilon_i|}{Y_i} \times 100\%$$

及根均平方誤差(Root Mean Square Percentage Error, 簡稱 RMSPE) :

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{\varepsilon_i}{Y_i}\right)^2} \times 100\%$$

其中 Y_i 是第 i 個觀察值, 也就是第 i 個預測年度的年齡別生育率 (或總生育率); $\varepsilon_i = Y_i - \hat{Y}_i$, \hat{Y}_i 是 Y_i 的預測值。

附錄 4 (年齡別生育率) 及附錄 5 (總生育率) 列出本研究考慮的五種模型的預測誤差。以下我們將依年齡別生育率、總生育率的順序說明實證發現:

1. 各年齡別生育率

五種方法在 7 個年齡組的預測誤差各有其優缺點 (附錄 4)。整體來說最後兩組 (40 至 49 歲) 年齡別生育率在各種方法下的預測皆不理想, 推測其原因, 可能是因為這兩組的生育率非常低, 只要預測稍有偏誤, 即有不小的誤差值。另外, 雖然擴散模型及 CFR 模型都是在世代理論架構下推導出的方法, 但模型較為複雜, 無法只藉由一次估計 (至少需要兩次估計) 獲得生育率的估計值, 因為愈多的估計誤差愈大, 使用較多參數的模型 (包括擴散模型及 CFR 法) 預測結果反而都比較不理想。

除了 40 至 49 歲這兩組外, 單一年齡組個別估計法有最小的產生的預測誤差, 其誤差在 10% 以下, 符合 Lewis (1982) 定義的高精確度的預測能力, 但在最後一組的預測誤差大於 40%, 這個結果與黃意萍與余清祥 (2002) 預測 1996

至 2000 年的生育率結果相同，但本文考慮的是個別估計法取對數。年齡別主成份分析法也有不錯的預測表現，15 至 39 歲的預測誤差大致也可視為具有高精確度，在 40 至 49 歲這兩組的預測誤差僅略大於 30%，尚稱合理，比其他方法穩定。（第一組、最後一組的預測誤差比單一年齡組個別估計法小。）

Lee-Carter 模型也有不錯的預測結果，在各年齡組的預測誤差最為平均，每組最高的誤差皆不超過 30%，未必只在高年齡組有較高的誤差；整體而言，誤差都在合理的範圍內，但某些年齡組的誤差仍與 2 個主成份的模型有不小的差距，與之前的研究結果類似。另外，除了 45 至 49 歲這一組明顯誤差偏高外，胎次別生育率模型的預測都不錯，35 至 44 歲的誤差甚至是各方法中最小者，未來若能降低最後一組的誤差，胎次別也是可能使用的預測模型之一。

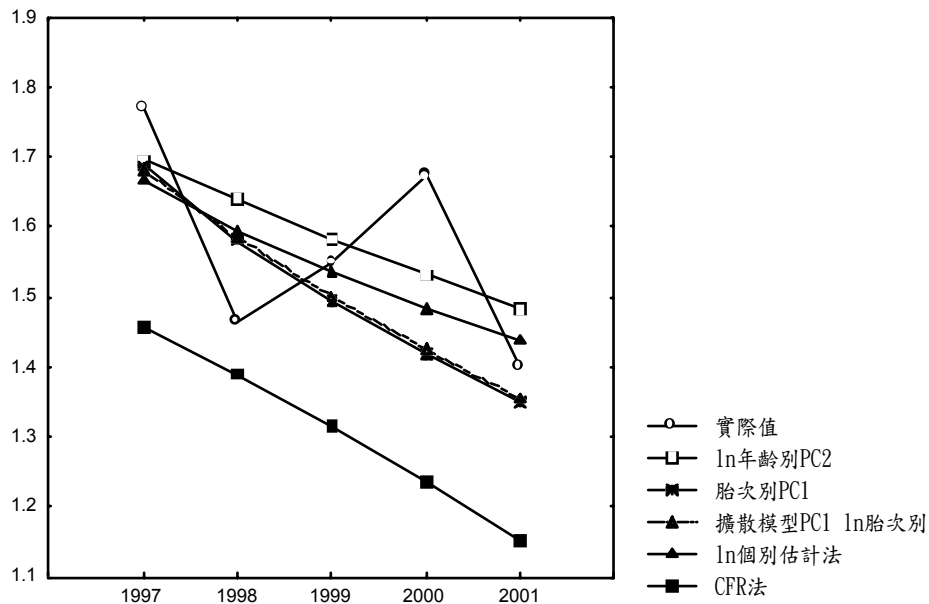


圖 5 實際總生育率與五種方法預測之總生育率

2. 總生育率

五種方法在預測總生育率方面皆有不錯的效果，除了 CFR 法外，其他四種方法（包含 Lee-Carter 法）的 MAPE 大概都不大於 10%（附錄 5）。由圖 5 可看出各年度總生育率的預測細節，CFR 法確實存有系統誤差，明顯低估總生育率的趨勢。其他方法則大致與實際總生育率相差不大，除了 1998 年因孤鸞年（虎年）生育率較低，2000 年因龍年效應而明顯上升，其他 3 個年度的預測值都非常接近實際值。（尤其是單一年齡組個別估計法，在 1997、1999、2001 年幾乎與實際值完全一致！）Lee-Carter 模型因為與主成份法的圖形類似，因此不列入圖 5。

有鑑於國人因各十二生肖年度而有不同的生育行為，在下一節我們將龍虎年（讀者可參考 Goodkind, 1993）的效應列入考慮，修正本文的各模型。

肆、龍虎年效應之修正

關於龍虎年效應的修正，因為 1973 年（牛年）至 1996 年（鼠年）共 24 年正好涵蓋兩整輪的十二生肖年份，時間又接近本文預測的 1997 至 2001 年，因此我們將這 24 年的總生育率與年齡別生育率資料取對數後作為應變數，令時間為自變數配適迴歸模型。求出估計值後再與實際值比較，將第 t 年及第 $t+12$ 年（ $t = 1973, 1974, \dots, 1984$ ）誤差平均後作為修正未來第 $t+24$ 年生育率的依據，修正公式為

$$TFR^{*z} = \exp\{\log TFR^z + \varepsilon_z\}, \quad z = 1997, 1998, \dots$$

其中 ε_z 是上述的誤差平均。

以龍虎年效應修正年齡別生育率與總生育率後的結果如下：

1. 年齡別生育率

大體而言，在 25 至 29 歲、30 至 34 歲（生育率最高的兩組）、40 至 44 歲這三組的年齡別生育率，各模型修正後的誤差都較小（詳見附錄 6）；但其他年齡組的生育率修正前後並無明顯差異或大小關係。

2. 總生育率

除了擴散模型、CFR 法外，其他模型對總生育率的預測誤差在修正後都降低不少（大約都有 30%，詳見附錄 7）。圖 6 為修正後各年度總生育率的詳細預測趨勢圖，與圖 5 為修正前比較，幾乎所有方法在各年度的誤差都是修正後較小，唯有 1999 年（兔年）的誤差並無明顯改善。原因可能因為兔年在虎年與龍年之間，可能受到虎年少生子或龍年多生子正負抵銷，不容易獲得精確的修正值。

整體而言，分析發現以龍虎年效應修正總生育率，獲得某些程度的改善，尤其是單一年齡組個別估計法以及年齡別生育率主成份分析法，但 CFR 之總生育率預測值仍與實際值有一段差距。

伍、結論與建議

本文以五種方法配適台灣生育率模型，而五種方法從衡量生育率的角度可分為兩大類：一類為非世代、另一類為世代的觀點，其中擴散模型與 CFR 法都牽涉到世代的觀念。以世代分析的優點為可看出個別世代的特性，不會混合不同世代（及其不同生育特性）的狀況，若能以世代的角度的去分析，應比非世代角度更適當。像擴散模型就是將不同世代的生育率轉為同一世代再進行分析，CFR 法也是以採同一世代的角度的，理論上這兩個模型應該有較佳的結果，但我們以台灣生育率資料為基礎，分析發現這兩種方法並未明顯優於其他三種方法，CFR 法反而有最大的預測誤差。

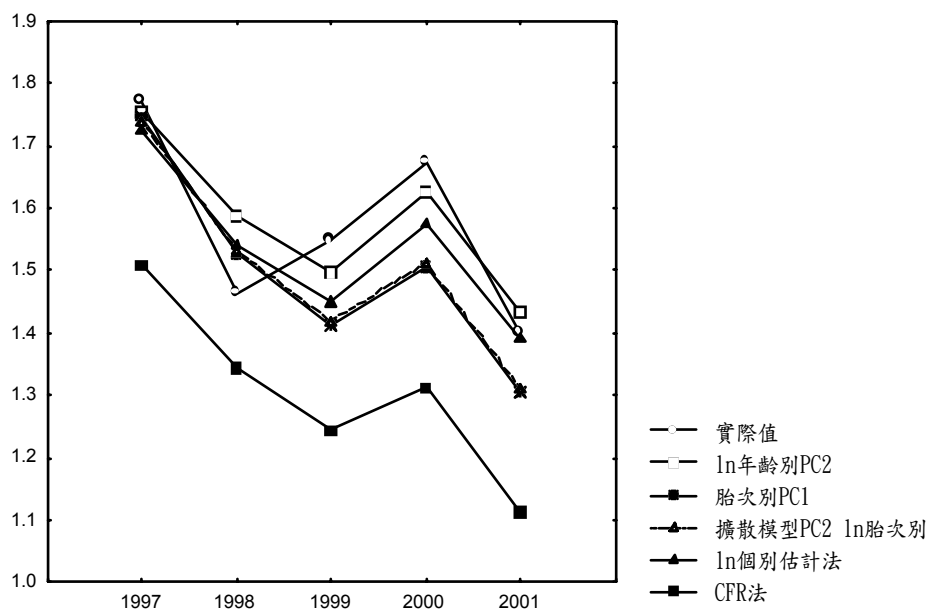


圖 6 實際總生育率與五種方法預測之後總生育率（以龍虎年效應修正）

以統計分析的觀點來看，有可能這兩種方法所需要估計的程序太多，在估計的過程中反而失去了原先理論架構的優勢。例如：擴散模型與胎次別生育率主成份法的差別，在於擴散模型必須先估計平均生育年齡與修正的胎次別生育率，如此才可估計出胎次別生育率，之後的步驟與胎次別生育率主成份法相同，因此擴散模型需要三次估計，胎次別生育率主成份法則用了兩次。CFR 法則更是預測過程繁雜，需要插入、預估的項目更多。再與年齡別主成份法及個別估計法比較，此兩種方法想法簡單，因此整個過程都只使用了一次估計，而整體預測結果也都不錯。另一個可能原因是或許台灣生育率近年來變化很大，使用過於複雜的模型描述台灣生育率，效果恰好適得其反。當然，我們也不能排除預測誤差較大的另一種可能——CFR 與擴散模型不適用於描述台灣地區婦女的生育率。

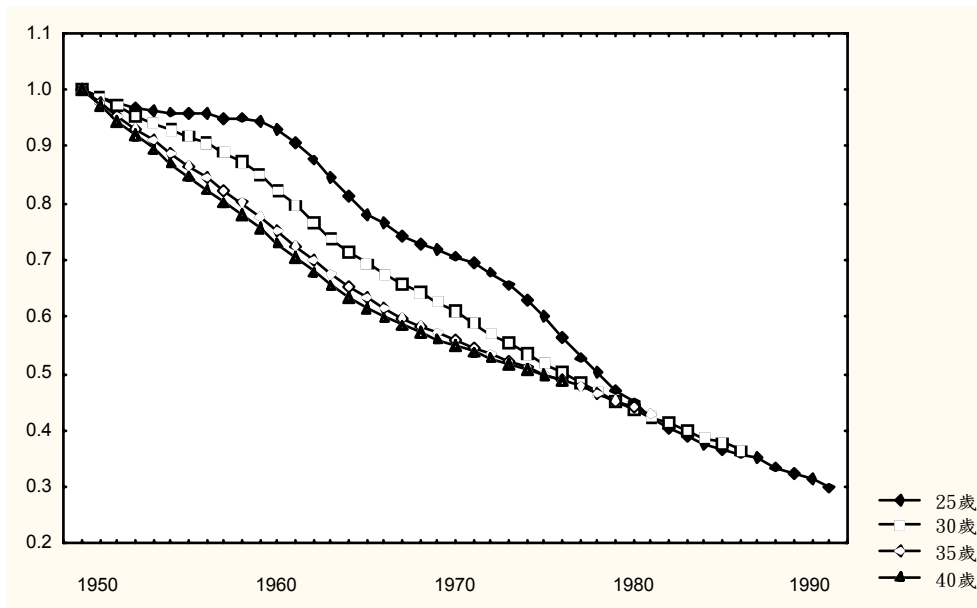


圖 7 台灣 CFR 累積生育率

雖然 CFR 不見得適用於預測未來台灣的婦女生育率，但它彌補了總生育率這類以橫斷面描述某一年度生育行為、而無法兼顧世代間變化的不足。圖 7 為台灣自 1949 年（1949 年時年紀為 15 歲的世代）以來 25、30、35 及 40 歲的累積（Cumulative）CFR，並以 1949 年為基礎（即除以 1949 年數值得出的比值）所計算出的生育率變化趨勢。由圖中可看出台灣自 1949 年以來，各世代婦女的生育數逐年下降，表示近年來台灣生育率下降的趨勢實際上在五十年前早已開始。與圖 8 美國歷年 25、30、35 及 40 歲的累積 CFR 趨勢比較，不難看出我國與美國的差異。美國的生育率變化較符合人口學家 Easterlin 的理論，呈現波浪狀的起伏，兩個波峰（或波谷）的間隔約為兩個世代，約為 40 年至 50 年；而台灣的生育率變化則較接近人口學家 Ermisch 的理論，未來上升的可能性不高（Brown，1997）。

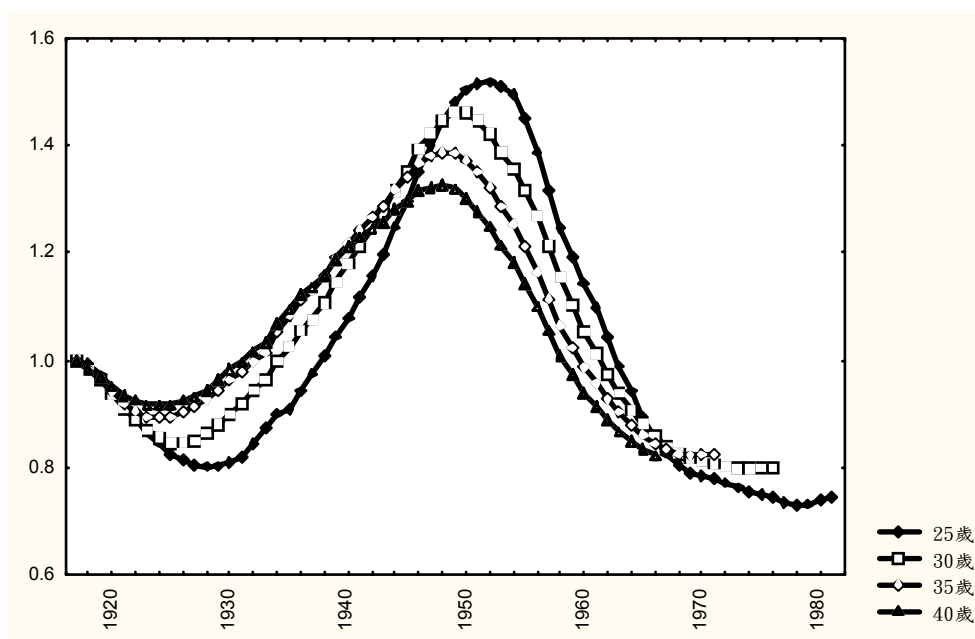


圖 8 美國 CFR 累積生育率

本文另一重點為探討一系列的主成份分析法。本文發現年齡別生育率資料取對數後的結果較佳，年齡別生育率取對數後使用 1 個主成份的誤差通常遠大於 2 個主成份，但 3 個主成份的預測誤差（與估計誤差）有時還比 2 個主成份大，這表示第 3 個主成份可能只反映雜訊。這個結果與 Bell (1997) 略有不同，但基本上我們同意 Bell 認為 Lee-Carter 模型的預測效果不如兩個主成份，加上 Lee-Carter 模型在估計時較為繁瑣，因此若考慮使用主成份分析法預測生育率時，我們不推薦讀者使用 Lee-Carter 模型。

至於 2 個主成份的實質解釋，因為 2 個主成份代表兩個（獨立的）生育率趨勢，其中第一個主成份（直線下降的變化）可詮釋為教育普及、經濟發展等因素使得現代婦女晚婚比例明顯提高、平均生育胎數減少，因此整體的生育率

逐年下降；第二個主成份則可詮釋為婦女晚婚使得原先的生育高峰向後延遲，越來越多人在 30 歲以後才開始生育，提高了近年來 30 至 39 歲的生育率，而原先最高的 20 至 24 歲年齡別生育率已漸漸遞延至 30 至 34 歲。這兩個主成份應可涵蓋台灣大多數的生育率變化趨勢。

關於龍虎年效應，本文發現即使是對成爲已開發國家的今日台灣民眾而言，傳統十二生肖的觀念對生育率的影響依舊存在，因此在預測總生育率時需把此類因素考慮進去；修正生育率時可將重點集中在 20 至 24 歲、25 至 29 歲及 30 至 34 歲這 3 個年齡組，因爲這 3 組對總生率的貢獻最大，且本文發現 25 至 29 歲、30 至 34 歲兩組在修正後的預測誤差降低最多。我們推薦讀者使用龍虎年修正後的個別估計法，或是修正後的年齡別生育率（2 個）主成份分析法。

謝 誌

本文作者感謝涂肇慶教授、林正祥教授、以及兩位匿名評審的寶貴建議；也感謝國科會對本計畫的部份補助，計畫編號 92-2412-H-004-002。

參考文獻

中文部分

中華民國內政部統計資訊網，<http://www.moi.gov.tw/W3/stat/>。

內政部（1949～2001）中華民國台閩地區人口統計，內政部編印。

行政院經濟建設委員會人力規劃處（2002），中華民國台灣地區民國 91 年至 140 年人口推計，編號(91)027.805。

黃意萍、余清祥（2002）台灣地區人口推估研究，*人口學刊*，25：145-171。

曾奕翔、余清祥（2002）台灣地區死亡率推估的實證方法之研究，台灣人口學年會學術研討會。

英文部分

Bell, W. R. (1997) Comparing and assessing time series methods for forecasting age-specific fertility and mortality rates, *Journal of Official Statistics*, 13(3): 279-303.

Bongaarts, J. and G. Feeney (1998) On the quantum and tempo of fertility, *Population and Development Review*, 24(2): 271-291.

Bozik, J. E. and W. R. Bell (1987) Forecasting age specific fertility using principle components, *Proceedings of the American Statistical Association*, Social Statistics Section, 396-401.

Brown, R. L. (1997) *Introduction to the Mathematics of Demography*, ACTEX Publications, Inc.

Central Intelligence Agency, <http://www.cia.gov/>

Goodkind, D. M. (1993) New zodiacal influences on Chinese family formation: Taiwan, 1976, *Demography*, 30: 27-142.

Lee, R. D. and L. R. Carter (1992) Modeling and forecasting U. S. mortality, *Journal of the American Statistical Association*, 87(419): 659-675.

- Lewis, C. D. (1982) *Industrial and Business Forecasting Methods: A practical Guide to Exponential Smoothing and Curve Fitting*, London: Butterworth Scientific.
- Zeng, Y. and K. C. Land (2002) Adjusting period tempo changes with an extension of Ryder's basic translation equation, *Demography*, 39(2): 269-285.
- United Nations (1992) *Long-Range World Population Projection – Two Centuries of Population Growth 1950-2150*, New York: United Nations Publication.

附 錄

附錄 1 (對數)年齡別生育率主成份解釋能力

	PC1	PC2	PC3
年齡別主成份	0.736	0.925	0.985
對數年齡別主成份	0.742	0.961	0.988

附錄 2 (對數)胎次別生育率主成份解釋能力

	PC1	PC2
胎次別主成份	0.917	0.982
對數胎次別主成份	0.920	0.985

附錄 3 各世代 Gamma 參數估計值

世代(西元/15 歲)	$\log C_t$	$\alpha_t - 1$	β_t	R^2
1944	-75.2213	32.3180	-1.1893	0.9777
1945	-72.2195	31.2524	-1.1701	0.9777
1946	-68.5528	29.8602	-1.1369	0.9811
1947	-66.0503	28.9152	-1.1155	0.9839
1948	-63.3857	27.8715	-1.0887	0.9878
1949	-62.2610	27.4595	-1.0818	0.9895
1950	-62.6658	27.6639	-1.0932	0.9909
1951	-63.3975	28.0067	-1.1093	0.9918
1952	-64.0909	28.3087	-1.1220	0.9951
1953	-64.6106	28.5713	-1.1361	0.9952
1954	-64.2165	28.4406	-1.1365	0.9969
1955	-64.0537	28.4098	-1.1405	0.9968
1956	-63.5225	28.1967	-1.1363	0.9977
1957	-63.0952	28.0236	-1.1329	0.9976
1958	-62.0888	27.6038	-1.1209	0.9967
1959	-60.8810	27.0879	-1.1046	0.9956
1960	-60.2530	26.8099	-1.0955	0.9942
1961	-59.5688	26.4786	-1.0822	0.9941
1962	-58.9883	26.1862	-1.0696	0.9934
1963	-58.0285	25.7081	-1.0487	0.9920
1964	-57.0086	25.2201	-1.0284	0.9895
1965	-54.9172	24.2617	-0.9910	0.9880

1966	-54.4116	24.0339	-0.9827	0.9883
1967	-54.7028	24.1483	-0.9862	0.9881
1968	-54.7060	24.1334	-0.9848	0.9870
1969	-54.4553	23.9799	-0.9759	0.9819
1970	-53.2458	23.3729	-0.9476	0.9761
1971	-53.8207	23.5944	-0.9534	0.9722
1972	-54.2632	23.7806	-0.9599	0.9710
1973	-54.2093	23.7466	-0.9581	0.9714
1974	-53.2306	23.2828	-0.9385	0.9717
1975	-51.1349	22.2808	-0.8948	0.9658
1976	-49.6199	21.5475	-0.8625	0.9661
1977	-49.2791	21.3367	-0.8503	0.9709
1978	-49.0087	21.1545	-0.8389	0.9758
1979	-48.4218	20.8132	-0.8199	0.9843
1980	-47.8834	20.4766	-0.7996	0.9902
1981	-47.4480	20.1630	-0.7779	0.9922

附錄 4 各年齡別生育率預測誤差(5 個年度加總)

模型	準則	15~19	20~24	25~29	30~34	35~39	40~44	45~49
Lee-Carter	MAPE	11.05	16.24	14.92	7.66	11.12	26.19	23.36
	RMSPE	12.95	18.09	18.10	8.99	12.90	29.69	26.22
log 年齡別 PC2	MAPE	5.86	5.89	10.18	7.12	15.79	30.11	31.99
	RMSPE	6.79	6.64	13.22	8.36	17.07	32.39	36.67
胎次別 PC1	MAPE	18.50	20.73	7.36	8.09	5.01	18.70	60.92
	RMSPE	24.09	24.00	8.98	9.47	7.14	19.45	66.34
擴散模型 PC2	MAPE	25.26	17.62	7.56	8.09	5.01	53.78	60.92
	RMSPE	29.76	22.38	8.45	9.47	7.14	61.05	66.34
log 胎次別 PC1	MAPE	6.59	5.56	8.20	6.80	13.35	26.07	43.24
	RMSPE	7.00	6.60	9.47	7.99	14.76	31.03	47.15
CFR 法	MAPE	16.81	26.56	26.96	8.08	23.68	50.35	129.56
	RMSPE	22.17	27.93	27.87	9.09	25.67	51.35	130.47

註：網底者為該年齡層各模型中最小者。

附錄 5 總生育率預測誤差(5 個年度加總)

模型	MAPE	RMSPE
Lee-Carter	8.81	10.96
log 年齡別 PC2	6.60	7.44
胎次別 PC1	6.99	8.28
擴散模型 PC2, log 胎次別 PC1	5.94	7.70
log 個別估計法	5.92	7.06
CFR 法	16.32	17.67

註：網底者為各模型中最小者。

附錄 6 各年齡別生育率以龍虎年效應修正後預測誤差(5 個年度加總)

模型	準則	15~19	20~24	25~29	30~34	35~39	40~44	45~49
Lee-Carter	MAPE	9.52	15.41	13.71	8.38	10.18	25.52	27.68
	RMSPE	10.16	16.27	15.20	9.20	12.66	30.86	32.38
log 年齡別 PC2	MAPE	3.78	2.78	7.75	6.66	14.75	29.49	30.08
	RMSPE	5.06	3.26	10.01	7.69	16.58	33.11	36.89
胎次別 PC1	MAPE	19.80	21.34	4.34	5.85	6.12	18.50	57.30
	RMSPE	23.98	23.86	5.27	7.20	6.66	19.45	64.43
擴散模型 PC2, log 胎次別 PC1	MAPE	26.49	18.28	4.79	5.81	6.10	53.41	57.31
	RMSPE	29.96	22.00	5.01	7.18	6.63	59.43	64.44
log 個別估計法	MAPE	3.64	3.81	5.26	6.63	12.29	26.44	39.42
	RMSPE	4.51	4.15	6.05	7.38	14.34	32.28	45.95
CFR 法	MAPE	18.21	27.19	27.77	7.88	24.87	50.62	138.96
	RMSPE	21.98	28.03	28.14	8.78	25.98	52.04	142.31

註：網底者為該年齡層各模型中最小者。

附錄 7 總生育率龍虎年效應修正前後誤差比較(5 個年度加總)

準則	修正前		修正後	
	MAPE	RMSPE	MAPE	RMSPE
Lee-Carter	8.81	10.96	6.96	8.32
log 年齡別 PC2	6.60	7.44	3.59	4.40
胎次別 PC1	6.99	8.28	6.26	6.98
擴散模型 PC2, log 胎次別 PC1	5.94	7.70	5.21	5.92
log 個別估計法	5.92	7.06	4.10	4.64
CFR 法	16.32	17.67	16.93	17.64

註：網底者為各模型修正前後較小者。