

## 第四章、研究設計

根據前文所述，本文之研究目的在於考量文化因素對生育行為的影響後，探討台灣婦女教育程度與生育率之 Granger 因果關係。倘若加入文化因素考量後，與 Cheng (1999) 之研究結果不一致，則必須檢驗文化因素對生育率之影響力是否顯著。另外，將 Cheng (1999) 忽略的經濟景氣亦視為內生變數之一後，其婦女教育程度與生育率之因果關係亦為本文所欲關注的。最後，再依據實證結果比較分析，提出政策建議。為達上述目的，在進行實證研究前，先於本章節詳細說明本文擬用之因果關係檢定法與實證模型設定。

### 第一節、研究方法說明

如前所述，本文之研究目的在於確認婦女教育程度的提昇是否能 Granger 影響育率的下降。然而生育行為與許多重要因素之間的關係是交互影響，並非為單一影響方向。又根據前文之文獻回顧，生育率、婦女教育程度、女性勞動參與率、經濟景氣之間息息相關，並且交互影響。因此，係根據變數之間交互影響的特性，本文將設定數個模型包含不同變數，並比較數個模型之實證結果。各實證模型設定之變數定義，將於下節中詳細說明。本文基於各變數之間有交互影響，因此採用 VAR 模型進行估計。VAR 模型的特點，即在於其將內生變數的動態結構（亦即交互影響作用）納入考量後，再進行模型估計。VAR 模型與聯立方程模型（simultaneous equation system）相比，外生變數在 VAR 模型中只居於次要地位。

本文在 VAR 模型架構下，為審視台灣婦女教育程度與生育率之 Granger 因果關係，可透過對婦女教育程度與生育率兩個內生變數，進行因果關係檢定。Granger 因果關係檢定的概念最早由 Granger (1969) 提出，

往後以此為基礎，發展出多種因果關係檢定方法。為了改進傳統因果檢定之缺失，本文將採用 Toda and Yamamoto (1995) 所提出的因果關係檢定。本節將對傳統因果關係檢定，以及 Toda and Yamamoto 之因果檢定作描述。

### 一、傳統因果關係檢定

由 Granger (1969) 提出，Sims (1972) 推廣的因果關係檢定，其基本概念在探討兩變數間預測關係為何。若在時間序列資料中，變數 Y 與其歷史資訊 ( $y_t, y_{t-1}, y_{t-2}, \dots$ )，對於預測變數 X 無幫助，則定義 Y 不為 X 的因 (Y does not Granger causes X)。更精確地以兩變數的分配函數表示為  $f(x_t | x_{t-1}, x_{t-2}, \dots) = f(x_t | x_t, x_{t-1}, \dots, y_t, y_{t-1}, \dots)$ 。反之，則 X 不為 Y 的因 (X does not Granger causes Y)。以 X 和 Y 的二元 VAR 模型為例，

$$X_t = \sum_{i=1}^n \alpha_i Y_{t-i} + \sum_{j=1}^n \beta_j X_{t-j} + \mu_{1t} \quad (8)$$

$$Y_t = \sum_{i=1}^m \lambda_i Y_{t-i} + \sum_{j=1}^m \delta_j X_{t-j} + \mu_{2t} \quad (9)$$

其中，假設干擾項  $\mu_{1t}$  與  $\mu_{2t}$  為無相關。若 Granger 因果關係存在，則式 (8) 中 Y 的延遲項 (the lagged Y) 顯著地影響 X，式 (9) 中 X 的延遲項顯著地影響 Y。因此，因果關係檢定為使用 F 統計量，聯合檢定延遲項的估計係數  $\sum \alpha_i$  與  $\sum \lambda_i$  是否顯著異於零。

傳統因果關係檢定提出後，在實證研究上應用很廣。然而，傳統因果關係檢定存在一些缺失與問題。首先，時間序列資料多呈現非定態 (non-stationary)，若使用非定態的資料進行迴歸估計會產生偽迴歸 (spurious regression)，而使統計量不具可信度，造成無效率的檢定。Granger and Newbold (1974) 認為，當經濟變數為非定態的時間序列資料時，若是使用傳統計量方法進行實證分析，會產生偽迴歸的問題。如此一

來，便無法正確瞭解變數間的關係，估計結果不具任何經濟上的意義，且會造成無效率的檢定。

考慮序列的定態下，發展出差分 VAR 因果關係檢定。然而，在序列經過差分處理後，可能會導致長期資訊的缺乏，忽略序列間的長期關係。因此，隨後發展出考慮長期關係的共整合概念，以修正差分後的缺失。共整合過程隱含於 Davidson et al. (1978) 提出的誤差校正模型 (ECM, error corrected model)，其中主要的概念直到 Granger (1983) 以及 Engle and Granger (1987) 才有較完整的發展。Engle and Granger (1987) 提出兩階段最小平方法進行共整合檢定，隨後 Johansen (1988) 提出檢定力較佳的跡檢定 (trace test) 與最大特性根檢定 (maximum eigenvalue test)，使用最大概似估計法進行共整合檢定。

ECM 架構下的因果關係檢定，雖然已考慮序列定態以及長期均衡關係，仍然存在一些問題。Toda and Yamamoto (1995) 指出此方法的缺失，在於因果關係檢定之前必須的單根檢定與共整合檢定，其檢定過程可能產生其他問題。Rambaldi and Doran (1996) 亦指出其過程複雜繁瑣，缺乏簡單及容易應用的特性。單根檢定的檢定力不足，以及共整合檢定對於其整合階數的不穩定性等問題，使因果關係檢定過程繁瑣，可能衍生更多複雜的問題。由於本文的研究目的在於探討台灣婦女教育程度與生育率之因果關係，並不關注其是否存在共整合關係。因此，本文將使用 Toda and Yamamoto (1995) 所提出之因果關係檢定，不需考慮變數間是否存在共整合關係，直接針對變數間的因果關係進行檢定。

## 二、Toda and Yamamoto 之因果關係檢定

如前所述，為了避免因 ECM 架構下繁複的共整合檢定所產生的缺失，本文將採用 Toda and Yamamoto (1995) 所提出之因果關係檢定，以免除

共整合檢定的不穩定性。Toda and Yamamoto 之因果關係檢定，最大的特點在於可直接探討變數間的因果關係，而不需考慮 VAR 系統是否具有共整合特性。以下以一  $k$  階 VAR 模型如下式 (10)，說明 Toda and Yamamoto (1995) 所提出之因果關係檢定。

$$X_t = \alpha + \beta_1 t + \beta_2 t^2 + \cdots + \beta_q t^q + J_1 X_{t-1} + J_2 X_{t-2} + \cdots + J_k X_{t-k} + u_t \quad (10)$$

其中， $X_t$  為一包含  $n$  個內生變數的向量； $\alpha$  為常數項； $\beta_1, \beta_2, \dots, \beta_q$  為對應時間趨勢  $t$  之係數向量； $J_1, J_2, \dots, J_k$  為對應落後期數  $X$  的係數向量； $u_t$  為殘差向量。

Toda and Yamamoto 之因果關係檢定只關心式 (10) 中  $X$  變數各落後期數之係數是否顯著，因此其虛無假設為聯合檢定  $J$  向量如下式 (11)：

$$H_0 : J_1 = J_2 = \cdots = J_k = 0 \quad (11)$$

接著以最小平方法估計 VAR 模型，如下式 (12)：

$$Y_t = \hat{\alpha} + \hat{\beta}_1 t + \cdots + \hat{\beta}_q t^q + \hat{J}_1 X_{t-1} + \cdots + \hat{J}_k X_{t-k} + \cdots + \hat{J}_p X_{t-p} + \hat{u}_t \quad (12)$$

其中， $\hat{\alpha}$ 、 $\hat{\beta}_1$ 、 $\hat{\beta}_q$ 、 $\hat{J}_1$ 、 $\hat{J}_k$ 、 $\hat{J}_p$ 、 $\hat{u}_t$  分別表示其估計值； $p \geq k + d_{max}$ ， $d_{max}$  為各變數的整合階數中之最大值。Toda and Yamamoto (1995) 證明其 *MWALD* 統計量服從卡方分配 ( $\chi^2$  distribution)；在進行因果檢定時，只需要檢定 VAR ( $k + d_{max}$ ) 模型中  $k$  階的估計參數即可。因此，Toda and Yamamoto 因果關係檢定之虛無假設如上式 (11)。

本文採用 Toda and Yamamoto 之因果關係檢定最大的優點在於免除繁複的共整合檢定，亦無需將 VAR 模型轉成 ECM 模型。Rambaldi and Doran (1996) 更證明可利用 SUR (seemingly unrelated regression) 型式，進行 Toda and Yamamoto 之因果關係檢定。以下，利用三個內生變數之 VAR 模型如下式 (13) 說明之。

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = A_0 + B_1 \begin{bmatrix} X_{t-1} \\ Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \cdots + B_k \begin{bmatrix} X_{t-k} \\ Y_{t-k} \\ Z_{t-k} \end{bmatrix} + B_{k+1} \begin{bmatrix} X_{t-k-1} \\ Y_{t-k-1} \\ Z_{t-k-1} \end{bmatrix} + \cdots + B_{k+d} \begin{bmatrix} X_{t-k-d} \\ Y_{t-k-d} \\ Z_{t-k-d} \end{bmatrix} \quad (13)$$

式 (13) 為以 SUR 型式表示之三個內生變數 VAR 模型，可對任意兩變數進行因果關係檢定。

欲瞭解預測 X 時，Y 的資訊是否有幫助，可進行虛無假設為下式 (14) 的卡方檢定：

$$H_0 : \beta_1^{(12)} = \beta_2^{(12)} = \cdots = \beta_k^{(12)} = 0 \quad (14)$$

其中， $\beta^{(12)}$  為 Y 的係數。若檢定結果拒絕虛無假設，則代表 Y 是 X 的因 (Y Granger causes X)。反之，欲瞭解預測 Y 時，X 的資訊是否有幫助，可進行虛無假設為下式 (15) 的卡方檢定：

$$H_0 : \beta_1^{(21)} = \beta_2^{(21)} = \cdots = \beta_k^{(21)} = 0 \quad (15)$$

其中， $\beta^{(21)}$  為 X 的係數。若檢定結果拒絕虛無假設，則代表 X 是 Y 的因 (X Granger causes Y)。

使用 Toda and Yamamoto (1995) 模型的好處，在於可以不需要考慮 VAR 模型中，內生變數之間的共整合關係 (cointegration)；亦不需考慮各內生變數是否為定態 (stationary) 序列，但可以正確得知其變數間的因果關係。而 Zapata and Rambaldi (1997) 根據 Monte Carlo 實驗結果亦驗證，Toda and Yamamoto (1995) 的因果檢定方法，並不需要事先知道 VAR 模型中的變數是否具有共整合的特性，只要變數的定態階數不大於模型中所選定的落後期數  $k$  (亦即  $k \geq d$ )，即可進行各變數間的因果檢定。

此外，Rambaldi and Doran (1996) 提出以 SUR (seeming unrelated regression) 的形式進行估計，使 Toda and Yamamoto (1995) 之因果關係檢定之實證應用更為普遍。例如：Shan and Sun (1998) 利用該方法檢驗澳

洲國內存款與國外投資之因果關係。Shan and Wilson (2001) 藉此研究中國貿易與觀光之因果關係。Huang et al. (2006) 利用該方法探討台灣勞動參與率與青少年犯罪率之因果關係。

## 第二節、實證模型與變數定義

### 一、實證模型

本文研究目的在於考量文化因素後，重新審視台灣婦女教育程度與生育率之 Granger 因果關係。為了與 Cheng (1999) 之實證結果相比較，本文將利用與 Cheng (1999) 相同的資料來源及變數定義，將資料期間由 Cheng (1999) 使用的 1952 年至 1996 年，延伸至 1952 年至 2005 年。又根據 Cheng (1999) 的實證模型設定，以其所使用的三個內生變數之 VAR 模型，建立本文的基本實證模型。再將前文提及的文化因素（龍年效果與虎年效果），以虛擬變數納入實證模型。此外，本文再依據 Mocan (1990) 與 Huang (2006) 將經濟景氣對生育行為的影響考慮在內，亦將實質經濟成長率亦視為內生變數。建立一包含四個內生變數及兩個虛擬變數的 VAR 模型如下式 (16)。

$$\begin{bmatrix} BR_t \\ HED_t \\ LFPY_t \\ GDPY_t \end{bmatrix} = A_0 + B_1 \times \begin{bmatrix} BR_{t-1} \\ HED_{t-1} \\ LFPY_{t-1} \\ GDPY_{t-1} \end{bmatrix} + \dots + B_{k+d_{\max}} \begin{bmatrix} BR_{t-k-d_{\max}} \\ HED_{t-k-d_{\max}} \\ LFPY_{t-k-d_{\max}} \\ GDPY_{t-k-d_{\max}} \end{bmatrix} + \Pi \times \begin{bmatrix} D \\ T \end{bmatrix} + \begin{bmatrix} e_{BR_t} \\ e_{HED_t} \\ e_{LFPY_t} \\ e_{GDPY_t} \end{bmatrix} \quad (16)$$

本文為瞭解台灣婦女教育程度與生育率之因果關係，將採用 Toda and Yamamoto (1995) 所提出的因果關係檢定 (Granger causality test)，針對婦女教育程度與生育率兩個內生變數，進行因果檢定。利用 Wald test 檢定虛無假設  $H_0: \beta_1^{(12)} = \beta_2^{(12)} = \dots = \beta_k^{(12)} = 0$ ，其中， $\beta_i^{(12)}$  分別為 VAR ( $k+d_{\max}$ ) 模型中第一個等式包含的解釋變數  $HED_i$  之係數。根據 modified Wald test 之 MWALD 統計量，若檢定結果為拒絕虛無假設，則表示婦女教育程度為生育率的重要解釋因素之一。

根據上述研究方法，進行實證估計，據以觀察台灣生育率與婦女教育

程度之間的因果關係。首先確認模型中各變數的定態階數  $d$ ，並令其中最高階數為  $d_{max}$ 。此外，根據最小 AIC (Akaike Information Criteria) 法找出將此一 VAR 模型的最適落後期數  $k$ 。<sup>27</sup> 接著，再估計  $k+d_{max}$  階的 VAR，亦即 VAR ( $k+d_{max}$ ) 模型。Toda and Yamamoto (1995) 證明，VAR ( $k+d_{max}$ ) 模型中  $k$  階的估計參數，其 MWALD 統計量服從卡方分配 ( $\chi^2$  distribution)；在進行因果檢定時，只需要檢定 VAR ( $k+d_{max}$ ) 模型中  $k$  階估計參數之聯合檢定即可。

使用 Toda and Yamamoto (1995) 的因果關係檢定，首先必須先透過單根檢定確定各變數的定態階數，據以找出  $d_{max}$ 。近二十年來，關於時間序列資料是否為定態的討論相當熱絡。Granger and Newbold (1974) 為首篇提出，當經濟變數為非定態的時間序列資料時，若是使用傳統計量方法進行實證分析，會產生偽迴歸 (spurious regression) 的問題。如此一來，便無法正確瞭解變數間的關係，估計結果不具任何經濟上的意義，且會造成無效率的檢定。因此，使用時間序列的資料進行實證分析時，必須先以單根檢定確認變數是否為定態序列。

單根檢定首先由 Dickey and Fuller (1979) 提出之 DF (Dickey-Fuller) 單根檢定。此後，許多學者致力於此，紛紛提出許多單根檢定法。Dickey and Fuller (1981) 亦提出修正 DF 單根檢定之 ADF (Augmented Dickey-Fuller) 單根檢定法，其檢定方法如下：

對於一時間序列資料  $Y$ ，考慮一  $Y$  為應變數之模型，如下式 (17)

$$Y_t = a + b_1 t + b_2 y_{t-1} + c_1 \Delta y_{t-1} + c_2 \Delta y_{t-2} + \dots + c_q \Delta y_{t-q} + \varepsilon_t \quad (17)$$

其中， $\Delta$  代表一階差分運算因子， $t$  為時間趨勢項， $\varepsilon$  為白噪音過程

<sup>27</sup> 決定 VAR 中最適落後階數  $k$ ，可根據 FPE、AIC、HQIC、BIC 等統計量，各個統計量所選的最佳階數並不一致，本文乃採用 AIC 統計量。AIC 統計量的計算方式為  $AIC = T \ln \sum (Y - \hat{Y})^2 / 2n$ ；其中  $n$  為估計參數的個數， $T$  為觀察樣本數。

(white-noise process)， $q$ 為所選定的落後期數，以最小AIC法選出最適落後期數。<sup>28</sup>對式(17)等式兩邊各減去 $y_{t-1}$ ，可得如下式(18)

$$\Delta Y_t = a + b_1 t + b_2 Y_{t-1} + \sum_{j=1}^q c_j \Delta Y_{t-j} + \varepsilon_t, \quad q = 1, 2, \dots \quad (18)$$

其中， $c_j = -\sum_{k=j+1}^q b_k$ ， $b_2 = \left( \sum_{i=1}^q b_i \right) - 1$ 。單根檢定之虛無假設為 $H_0: b_2 = 0$ ，表示

$Y$ 序列資料存在單根。若拒絕虛無假設，則表示資料為定態資料。若對非定態的序列取 $d$ 階差分，使其成為定態序列，則代表變數的定態階數為 $d$ ，以 $I(d)$ 表示之。

本文是以 Cheng (1999) 的實證為基礎，並加入文化因素的考量，重新審視台灣婦女教育程度與生育率之因果關係。因此，本文首先以 Cheng (1999) 的資料來源與變數定義，利用台灣 1952 年至 2005 年的年資料，進行實證模型估計。藉此檢驗 Cheng (1999) 之結論是否與本文實證結果一致。倘若一致，則根據前文提及之龍年效果與虎年效果，將虛擬變數納入模型後，其結果為何。又根據 Mocan (1990) 認為經濟景氣亦是解釋生育率的重要因素，另外加入經濟景氣之內生變數後，其結果又為何。

利用與 Cheng (1999) 的資料來源與變數定義皆相同的年資料建立實證模型後，本文將據此延伸其他實證模型，以進一步探討台灣生育率與婦女教育程度之因果關係。如前文所述，本文擬將 Cheng (1999) 使用的年資料改用季資料，增加樣本數以增進估計的有效性。Cheng (1999) 在探討婦女教育程度與生育率的因果關係時，使用三個內生變數的 VAR 模型，其內生變數的資料分別使用粗出生率、全國人口教育程度，以及女性勞動

<sup>28</sup> 根據 Greene (1999)，ADF 單根檢定中，AIC (Akaike information criteria) 值計算方式為， $AIC(p) = \ln\left(\frac{e'e}{T - p_{max} - K^*}\right) + (p + K^*)\left(\frac{2}{T - p_{max} - K^*}\right)$ ，其中有隨機項 (random walk) 時  $K^*=1$ ；有隨機項與漂浮項 (random walk with drift) 時  $K^*=2$ ；有時間趨勢穩態時 (trend stationary)  $K^*=3$ ； $p_{max}$ =最大選擇的落後期數。

參與率。<sup>29</sup>根據 Whittington et al. (1990) 指出，使用一般生育率來反映生育水準較為適當，其好處在於相較於粗出生率，一般生育率較不容易受到人口年齡組成的影響。據此，本文乃採用一般生育率做為衡量生育水準的變數，建立更進一步的實證模型。

又根據 Huang (2002, 2006) 認為在探討生育率與相關經濟變數之間的關係時，必須考慮生育行為的特性。由於懷胎期間為 280 天，約九個月的時間。因此，目前的生育決策（受孕）所產生的結果，應當反映在 3 季之後的出生率 ( $GFR_{t+3}$ )。是故，本研究將 3 季之後的生育率視為當期的受孕率。<sup>30</sup>這樣的設定，亦是 Cheng (1999) 使用年資料進行估計時，無法考量的因素。此外，龍年效果與虎年效果的虛擬變數 (dummy variable)，<sup>31</sup>仍隨著受孕的時間而設定。若預期在龍年生育，則受孕期為兔年的第二季至龍年的第一季。反之，若要避免在虎年生育，則應避免於牛年的第二季至虎年的第一季期間受孕。<sup>32</sup>

據此，本文將延伸 (15) 式的實證模型設定，以一般生育率替換粗出生率，並考慮受孕與生育之時間不一致，以受孕率 ( $GFR_{t+3}$ ) 替換當期生育率 ( $GFR_t$ )，並且以育齡婦女教育程度替換全國教育程度，建立一包含四個內生變數及兩個虛擬變數的 VAR 模型如下式 (19)：

<sup>29</sup> Cheng (1999) 所使用的資料來源為 *Taiwan Statistical Data Book* (1989, 1996)。教育程度衡量的指標有二，一是以教育程度大專以上人口數與適齡 (15 歲至 64 歲) 工作人口數的比率來衡量；另一是以教育程度國中以上人口數與適齡工作人口數的比率來衡量。

<sup>30</sup> 本文假設懷胎期間的意外，例如流產、死胎、早產、過期妊娠等情況的發生是隨機的，期望值為零，因此可以排除考慮。並且假設受孕即會生育，將複雜的生育情況簡化。

<sup>31</sup> 本文為實際反應龍年與虎年效果對生育行為實際的影響狀況，乃將兩個虛擬變數分別設為，在龍年生育設為 1，其餘為 0；以及在虎年生育為 1，其餘為 0。

<sup>32</sup> 本文使用的資料期間，1988 年與 2000 年的龍年，其農曆春節為 2 月 17 日與 2 月 5 日；1986 年與 1998 年的虎年，其農曆春節為 2 月 9 日與 1 月 28 日。為考慮農曆與國曆的差異，本文以農曆的月份為基礎，設立龍年與虎年的虛擬變數。由於龍 (虎) 年的農曆過年後出生的小孩，才真正屬於龍 (虎) 年出生，而農曆的過年多在一月或二月。因此，本文的虛擬變數乃是自龍 (虎) 年的第二季起至隔年的第一季止設定為一，其餘其間設定為零。

$$\begin{bmatrix} GFR_{t+3} \\ HED_t \\ LFP_t \\ GDP_t \end{bmatrix} = A_0 + B_1 \times \begin{bmatrix} GFR_{t+3-1} \\ HED_{t-1} \\ LFP_{t-1} \\ GDP_{t-1} \end{bmatrix} + \dots + B_{k+d_{\max}} \begin{bmatrix} GFR_{t+3-k-d_{\max}} \\ HED_{t-k-d_{\max}} \\ LFP_{t-k-d_{\max}} \\ GDP_{t-k-d_{\max}} \end{bmatrix} + \Pi \times \begin{bmatrix} D \\ T \end{bmatrix} + \begin{bmatrix} e_{GFR_{t+3}} \\ e_{HED_t} \\ e_{LFP_t} \\ e_{GDP_t} \end{bmatrix} \quad (19)$$

其中， $GFR_t$  為  $t$  期一般生育率； $GFR_{t+3}$  為  $t$  期受孕率； $HED_t$  代表  $t$  期育齡婦女教育程度大專以上的比率； $LFP_t$  為  $t$  期女性勞動參與率； $D$  代表在龍年生育為 1，其餘為 0 的虛擬變數； $T$  則代表在虎年生育為 1 其餘為 0 的虛擬變數。 $A_0$  為  $4 \times 1$  係數矩陣， $B_s$  為  $4 \times 4$  係數矩陣， $\Pi$  為  $4 \times 2$  係數矩陣。

## 二、變數定義

首先使用與 Cheng (1999) 相同來源之資料，利用台灣 1952 年至 2004 年之年資料進行實證估計。粗出生率、全國人口大專以上之比例、女性勞動參與率之數據，來自行政院經濟建設委員會編印之 *Taiwan Statistical Data Book*；實質經濟成長率之數據來自《中華民國統計月報》。文化因素之虛擬變數設定，龍年效果之虛擬變數為 1952 年、1964 年、1976 年、1988 年、2000 年共五年設為 1，其餘為零；虎年效果之虛擬變數為 1962、1974 年、1986 年、1998 年共四年設為 1，其餘為零。

本文延伸之實證模型，在變數選擇上與 Cheng (1999) 有所不同。延伸之實證模型以一般生育率衡量生育水準，並用育齡婦女教育程度大專以上的比率來衡量教育水準。如此一來，可以更直接地探討生育率與婦女教育程度之間的因果關係。採用育齡婦女教育程度大專以上的比率來衡量婦女教育程度的水準，理由在於與一般生育率的衡量基準具有一致性，兩者分別是描述育齡婦女教育程度與育齡婦女生育情況的變數。本文所使用的變數定義與基本統計量，呈現於表 2。

表 2：相關變數之定義與基本統計量

變數名稱	符號	變數定義說明	平均值	標準差
<b>A、年資料</b>				
粗出生率	$BR_t$	每年出生活產數÷ 總人口數 (單位：‰)	24.80	11.04
全國教育程度 大專以上比率	$HED_t$	全國教育程度大專以上人口數÷ 總人口數 (單位：%)	10.58	8.60
女性勞動參與率	$LFPY_t$	女性勞動力人口÷ 女性適齡工作人口 (單位：%)	40.71	21.17
實質經濟成長率	$GDPY_t$	每年實質經濟成長率 (單位：%，以 2001 年價格為基準)	7.97	10.46
龍年效果	$D$	龍年為 1， 其餘為 0 的虛擬變數	0.09	0.29
虎年效果	$T$	虎年為 1， 其餘為 0 的虛擬變數	0.07	0.26
樣本數		1952 年至 2005 年	54	
<b>B、季資料</b>				
一般生育率	$GFR_t$	每季出生活產數÷ 年中育齡婦女人口數 (單位：‰)	15.36	4.49
受孕率	$GFR_{t+3}$	當季受孕率=後三季的一般生育率 (單位：‰)	15.12	4.32
育齡婦女教育程 度大專以上比率	$FHED_t$	育齡婦女教育程度大專以上÷ 育齡婦女人口數 (單位：%)	17.28	8.27
女性勞動參與率	$LFPQ_t$	女性勞動力人口÷ 女性適齡工作人口 (單位：%)	44.20	2.72
實質經濟成長率	$GDPQ_t$	每季實質經濟成長率 (單位：%，以 2001 年價格為基準)	6.46	3.21
龍年效果	$D$	在龍年生育為 1， 其餘為 0 的虛擬變數	0.07	0.26
虎年效果	$T$	在虎年生育為 1， 其餘為 0 的虛擬變數	0.07	0.26
樣本數		1978 第一季至 2005 年第四季	112	

資料來源：年資料來自歷年 *Taiwan Statistical Data Book*；季資料來自歷年《臺閩地區人口統計》、  
歷年《中華民國統計月報》、1987 年至 2005 年《人力資源統計月報》、1978 年至 1987 年《勞工統  
計月報》。

實證模型進行估計。又使用季資料時，根據 Huang (2002) 考慮生育時間與受孕時間上的不一致，將以前文所提及之受孕率 ( $GFR_{t+3}$ ) 替換當期生育率 ( $GFR_t$ )，<sup>33</sup>再進行實證估計。最後，根據 Mocan (1990) 將實質經濟成長率亦視為內生變數之一，估計本文式 (18) 所建立的 VAR 模型，並利用 Toda and Yamamoto (1995) 提出的方法，檢驗婦女教育程度與生育率之因果關係。

基於上述幾點，本文將依序進行七個模型的實證估計，分別為模型一、模型二、模型三、模型四、模型五、模型六與模型七。首先利用 Cheng (1999) 的變數定義與相同來源的資料，使用 1952 年至 2005 年的年資料估計模型一、模型二與模型三。模型一為包含三個內生變數 ( $BR_t$ 、 $HED_t$  與  $LFPY_t$ ) 之 VAR 模型；模型二為模型一再加入龍年與虎年效果 ( $D$ 、 $T$ ) 之 VAR 模型；模型三為模型二中加入經濟景氣  $GDPY_t$  的包含四個內生變數之 VAR 模型。

再者，係根據前文所述，將粗出生率替換成一般生育率，全國人口教育程度替換成育齡婦女教育程度，並使用 1978 年第一季至 2005 年第四季的季資料，估計模型四、模型五、模型六與模型七。模型四為包含三個內生變數 ( $GFR_t$ 、 $FHED_t$  與  $LFPQ_t$ ) 之 VAR 模型；模型五為模型四再加入龍年與虎年效果 ( $D$ 、 $T$ ) 之 VAR 模型；模型六為模型五中，以受孕率  $GFR_{t+3}$  替代生育率  $GFR_t$ ；模型七為模型六再加上代表經濟景氣之  $GDPQ_t$  內生變數。茲將七個實證模型之變數整理於表 3。

---

<sup>33</sup> 本文將以  $GFR_t$  表示當期生育率，當期受孕率則以  $GFR_{t+3}$  表示之。

表 3：實證模型之設定

	A、年資料			B、季資料			
	模型一	模型二	模型三	模型四	模型五	模型六	模型七
內生變數	$BR_t$	$BR_t$	$BR_t$	$GFR_t$	$GFR_t$	<b><math>GFR_{t+3}</math></b>	$GFR_{t+3}$
	$HED_t$	$HED_t$	$HED_t$	$FHED_t$	$FHED_t$	$FHED_t$	$FHED_t$
	$LFPY_t$	$LFPY_t$	$LFPY_t$	$LFPQ_t$	$LFPQ_t$	$LFPQ_t$	$LFPQ_t$
			<b><math>GDPY_t</math></b>				<b><math>GDPQ_t</math></b>
外生變數	—	$D、T$	$D、T$	—	$D、T$	$D、T$	$D、T$

註：1.  $BR_t$ 、 $HED_t$ 、 $LFPY_t$ 與  $GDPY_t$ 分別代表年資料的粗出生率、全體教育程度大專以上比率、婦女勞動參與率與實質經濟成長率。 $GFR_t$ 、 $GFR_{t+3}$ 、 $FHED_t$ 、 $LFPQ_t$ 與  $GDPQ_t$ 分別代表季資料的一般生育率、受孕率、育齡婦女教育程度大專以上比率、婦女勞動參與率與實質經濟成長率。

2.  $D$ 為在龍年生育為1，其餘為0的虛擬變數； $T$ 為在虎年生育為1，其餘為0的虛擬變數。

3. 粗體字表示相較於上一個模型，該模型增加或改變的變數。

### 第三節、研究範圍與研究限制

#### 一、研究範圍

本文為了與 Cheng (1999) 之實證結果比較，首先採用 1952 年至 2005 年之年資料進行模型一、模型二與模型三之估計。使用年資料共有 54 個樣本數，然 Yamamoto and Toda (1998) 指出若是樣本數少餘五十，Toda and Yamamoto 之因果關係檢定之檢定力會不足。為增進模型估計之有效性，本文採用 1978 年第一季至 2005 年第四季之季資料，共 112 個樣本數，以解決樣本數不足的問題。此外，Cheng (1999) 之實證結果顯示全國教育程度對粗出生率之因果關係不顯著。然而，如前所述基於粗出生率易受人口年齡組成之影響，以及生育率與婦女教育程度之關係較為直接等因素，本文將檢定婦女教育程度對一般生育率之因果關係。

#### 二、研究限制

由於本論文採取的是實證研究途徑，主要限制乃在於研究設計與資料蒐集上，茲說明如下：

第一，在對文化因素之虛擬變數設定時，由於龍年與虎年以農民曆計算，若要精確衡量龍年效果與虎年效果有其困難性。以本文使用的季資料為例，<sup>34</sup>由於皆始於當年一月底二月初，結束於隔年一月底二月初。又由於其餘資料皆是以西曆計算，因此本文進行實證估計時，將龍年效果與虎年效果設為當年第一季至第四季，並不考慮隔年第一季。處理虛擬變數時，當年第一季與隔年第一季之效果的精確性，為本文的研究限制。

第二，本文於第一節設定實證模型時，礙於現有資料，受孕率的定義

---

<sup>34</sup> 1988 年之龍年始於二月十七日，結束於 1989 年二月五日；2000 年之龍年始於二月五日，結束於 2001 年一月二十三日；1986 年之虎年始於二月九日，結束 1987 年一月二十八日；1998 年之虎年始於一月二十八日，結束於 1999 年二月十五日。

上有諸多限制。本文假設懷胎期間的意外，例如流產、死胎、早產、過期妊娠等情況的發生是隨機的，因此可以排除考慮，將複雜的生育情況簡化。並且假設受孕即會生育，亦即估計育齡婦女教育程度對受孕率的影響，等同於育齡婦女教育程度對生育率的影響。因此，本文先將生育行為簡單化後，再探討生育率與婦女教育程度之因果關係。

第三，本文主要係依據 Cheng (1999) 之實證結果，利用總體資料觀察台灣生育率與婦女教育程度之因果關係。因此，僅能證實婦女教育程度的提昇是否為生育率下降的重要解釋變數之一。故深入之相關議題，例如，教育程度愈高之婦女其生育率是否愈低等，則有待進一步更詳盡的個體資料，進行深入的研究調查才可證實。

## 第四節、本章小結

本文的研究目的在於考慮文化因素對生育率的影響後，台灣婦女教育程度與生育率之 Granger 因果關係為何。為修正傳統因果關係檢定的缺失，以及 VECM 模型的繁複模型轉換與因果檢定過程，容易導致因果關係檢定力不足。故本文故決定採用 Toda and Yamamoto (1995) 之因果關係檢定，以避免繁複過程中產生其他問題。本章首先介紹本文採用的研究方法；其次為建立實證模型，並說明各變數之定義及資料來源；最後再對本文之研究範圍及研究限制加以說明。本章重點整理如下：

### 一、研究方法說明

本文為了避免 VECM 模型中因果關係檢定的繁複過程，將採用 Toda and Yamamoto (1995) 之因果關係檢定，以避免共整合檢定之不穩定性，及繁複的模型轉換等因素產生其他問題。Rambaldi and Doran (1996) 更提出以 SUR (seemingly unrelated regression) 型式，進行 Toda and Yamamoto 之因果關係檢定。首先，考慮各變數的定態階數  $d$ ，令其最大者為  $d_{max}$ ；再以 AIC 法則決定 VAR 模型最適階數  $k$ ；隨後估計 VAR ( $k+d_{max}$ ) 模型。最後，在進行因果檢定時，只需要檢定 VAR ( $k+d_{max}$ ) 模型中  $k$  階估計參數之聯合檢定即可。

### 二、實證模型與變數定義

本文首先根據 Cheng (1999) 的實證模型設定，再將文化因素（龍年效果與虎年效果）納入考量，並根據 Mocan (1990) 將經濟景氣亦視為內生變數之一，依序建立本文模型一、模型二與模型三的基本實證模型。本文為了進一步探討台灣生育率與婦女教育程度之因果關係，利用 1978 年第一季至 2005 年第四季之季資料，估計延伸的模型四、模型五、模型六與模型七等實證模型。本文藉由不同的變數定義，重新審視婦女教育程度

與生育率之因果關係。衡量生育水準時，以一般生育率替換粗出生率；衡量教育程度時，以育齡婦女大專以上比例替換全國人口大專以上比例。考慮並將 1952 年至 2005 年之年資料，改為 1978 年至 2005 年之季資料，增加樣本數以增進估計的有效性。除了估計 Cheng (1999) 所建立的三個內生變數之 VAR 模型外，並估計加上文化因素之虛擬變數後之實證模型。除此之外，由於使用季資料估計模型，因此考慮受孕與生育之時間不一致，最後加入經濟景氣之內生變數。是故，本文依序建立七個實證模型，各模型之變數說明可參照表 2。

上述變數 1952 年至 2005 年之年資料來源為歷年之 *Taiwan Statistical Data Book* 與歷年《中華民國統計月報》；1978 年至 2005 年之季資料來源為各季《臺閩地區人口統計》、各期《中華民國統計月報》、1987 年至 2005 年各期《人力資源統計月報》，及 1978 年至 1987 年各期《勞工統計月報》。

### 三、研究範圍與研究限制

本文的研究範圍首先採用 1952 年至 2005 年之年資料，共 54 個樣本數。為增進模型估計之有效性，乃採用 1978 年第一季至 2005 年第四季之季資料，共 112 個樣本數，解決樣本數不足的問題。本文的研究限制如下：第一，文化因素虛擬變數之設定，由於龍年與虎年以農民曆計算，若要精確衡量龍年效果與虎年效果有其困難性，本文僅以季為單位計算之。第二，本文假設懷胎期間的意外事件是隨機的，因此可以排除考慮。本文先將生育行為簡單化後，並假設受孕即會生育，再探討受孕率（亦即生育率）與婦女教育程度之因果關係，其與實際生育情形有差異。第三，本文利用總體資料觀察台灣生育率與婦女教育程度之因果關係，僅能證實婦女教育程度的提昇是否為生育率下降的重要解釋變數之一。深入之相關議題，則有待進一步進行深入的個體調查才可證實。