

緩長記憶與狀態變換

謝俊魁·林建甫*

針對眾所周知之財務資料波動程度的高度相關性，辨別這種相關性是否就是「緩長記憶」在經濟政策、財務決策、實證方法、及計量理論上都具有非常重要的意義。本文結合 Hamilton and Susmel (1994) 之狀態變換 (regime switching) SWARCH 模型，以及 Baillie, Bollerslev, and Mikkelsen (1996) 的緩長記憶 FARCH 模型，新創相當一般化之 SW(k)-FIGARCH-L (0,d,0) 模型，同時估計狀態轉變參數及部分差分係數，使我們既可在狀態變換的前提下檢定緩長記憶假說，也能在考慮緩長記憶可能存在的前提下檢定是否有狀態變換發生，並證明波動程度之狀態變換的確會造成相當嚴重的假性緩長記憶問題。此新模型不但可避免以單純緩長記憶模型估計狀態變換母體所導致的假性緩長記憶問題，更解決了單純狀態變換模型殘差項平方值仍具高度跨期相關的窘境，是一個相對優良的模型，在財務資料的估計上十分具有發展潛力。根據實證結果，未考慮狀態變換因素前，台股指數日報酬率的波動程度的確顯現出緩長記憶特徵，然而同時考慮狀態變換因素後，緩長記憶特徵不復存在，僅具中等記憶。

關鍵詞: FIGARCH 模型, Markov-switching 模型, SWARCH 模型,
SW(k)-FIGARCH-L 模型, 緩長記憶
JEL 分類代號: C52

1 前言

「緩長記憶」(long memory) 是近年研究的熱門課題。所謂「緩長記憶」係指某些時間序列具有的長時距依存現象 (long range dependence); 其表現在樣本

*作者分別為國立台灣大學財務金融系博士班學生與國立台灣大學經濟系教授。作者感謝中華經濟研究院八十八年度暑期訪問計劃研究的補助，中華經濟研究院研討會，輔仁大學創校七十週年校慶總體經濟與國際金融學術研討會與會人士討論提供意見。作者並特別感謝兩位匿名評審提供意見。當然，若有任何錯誤，作者應負擔所有的責任。

經濟論文叢刊 (*Taiwan Economic Review*), 32:2 (2004), 193-232。

國立台灣大學經濟學系出版

資料上的特徵，為樣本自我相關函數 (autocorrelation function, 簡稱 ACF) 遞減的速度相當緩慢。許多經濟及財務變數曾被論證或檢定具有緩長記憶性質，或在經過某些變數轉換後具有緩長記憶性質；如股市成交量 (Lobato and Velasco 2000)、股價報酬率 (Sadique and Silvapulle 2001; Greene and Fielitz 1977; Ding, Granger and Engle 1993; Lobato and Savin 1998)、股價波動性 (So 2000) 匯率 (Cheung 1993a)、利率 (Tsay 2000; Backus and Zin 1993)、實質國民生產毛額 (Diebold and Rudebusch 1989; Sowell 1992)、物價指數 (Geweke and Porter-Hudak 1983; Baillie, Chung and Tieslau 1996) 等等。

可是任何針對母體性質所作的統計推論都是根據樣本資料所透露的訊息而得到的。一旦兩種不同的資料產生過程 (data generating process, 簡稱 DGP) 所產生的樣本資料具有極為相似的性質，我們很難透過樣本訊息分辨出真正的 DGP。近年來計量經濟學界中單根 (unit root) 與結構性改變 (structural change) 的爭論即根源於此，因為早期大部分的總體變數往往被認定為趨勢定態；但 Nelson and Plosser (1982) 使用檢定後發現不能推翻單根的性質；Perron (1989) 卻質疑資料所呈現的單根性質可能是結構性改變造成的結果，Zivot and Andrews (1992) 認為總體變數中部分是單根序列，部分則是趨勢定態但具結構改變的序列。目前許多計量學者仍致力於研究更準確分辨這兩類 DGP 的計量方法。同樣的問題也可能發生在緩長記憶性質的估計及檢定上。

Lobato and Savin (1998) 檢定 S&P 500 指數日報酬率絕對值及平方值資料的緩長記憶性質時，曾注意到結構性改變可能造成「假性緩長記憶」於是將資料依可能的結構改變發生點分成兩部分，分別對整筆資料及兩筆子資料 (sub-sample) 進行檢定，結果顯示整筆資料及子資料都有顯著的緩長記憶性質。Baillie (1998) 認為報酬率絕對值及平方值資料的緩長記憶現象可用其條件二階動差的緩長記憶設定來刻劃，於是採用 FIGARCH 模型對 S&P 500 的日報酬率資料進行估計；同樣考慮樣本期間可能發生結構性改變，而將資料分成 5 部分，分別對整筆資料及各子資料進行估計，結果顯示其都有顯著的緩長記憶性質。

然而我們知道，「整筆資料及各子資料都有顯著的緩長記憶性質」並不能證明「即使考慮了結構性改變，緩長記憶性質仍然是真實的」。首先，用恣意認定結構改變點來進行資料劃分並不恰當。再者，即使所選擇的改變點的確發生了結構性改變，我們仍不能肯定子資料所顯示的緩長記憶性質是不是真實的；

倘若樣本期間 DGP 的結構並不穩定, 我們所選擇的結構改變點間可能還存在更多其他的改變點, 那麼子資料的緩長記憶性質當然也可能是結構性改變所造成的假性現象。由圖 1 我們可以發現, 1980 年 1 月至 1999 年 2 月台股指數報酬率。絕對值的 ACF 呈現十分緩慢的下降速度, 在相距 900 期時, 相關係數仍將近 0.06, 顯示可能有緩長記憶現象。然而, 當我們將資料分成 3 個樣本數夠大的子資料後, 子資料的 ACF 即不再那麼「規則且緩慢」地下降, 這顯示樣本的緩長記憶特徵並不穩定。

辨別緩長記憶現象的真偽在實證方法、及計量理論甚至政策分析上都具有重要的意義。經濟變數若具緩長記憶, 表示各期隨機衝擊的影響會持續一段相當長的時間; 若其緩長記憶特徵係結構性改變所造成的假性現象, 則表示僅有某些偶發的重大事件會造成 DGP 的結構性改變, 進而對該經濟變數產生較長遠的影響, 其他絕大部分的隨機衝擊則僅具相對短暫的影響效果。相同的樣本特徵, 卻對應著截然不同的經濟解釋, 若不予以辨別, 不但會因為實證方法的誤用而導致統計推論錯誤, 更可能會對經濟現象的判讀錯誤而制訂出不當的政策, 造成金融商品市場不具效率性。此外, Tsay and Chung (2000) 發現被解釋變數與解釋變數都具有緩長記憶, 雖然兩者都是定態序列, 仍可能會造成假性迴歸 (spurious regression) 現象。因此在進行迴歸分析之前, 先確認資料緩長記憶的存在性是不能忽略的工作。

序列的長時距依存現象會表現在其時間路徑動向的高持續性 (persistence) 上。但持續性可能是虛假的緩長記憶現象。Anderson (2000) 則發現為部份積分模型配上 AR 係數為正且接近 1, 則一定產生緩長記憶現象, 但若 AR 係數為負則高持續性減少, 若僅有 MA 係數則造成緩長記憶現象不容易顯著。所以如何解釋緩長記憶現象引起最近的學界高度重視。一種文獻解釋是水準值的變動, 例如 Bos, Franses and Ooms (1998) 用 G7 的通貨膨脹資料具體檢驗發現緩長記憶現象應是由水準值的變動所引起。Granger and Hyung (1999) 使用偶然水準值的變動也造成緩長記憶效果。但另一種文獻的解釋是 Diebold (1986) 所指出的序列的「狀態變換」(regime switching) 也可能會造成時間路徑動向的高持續性。所謂的狀態變換事實上就是一種結構性改變, 只不過一般所稱的結構性改變通常是指參數結構狀態的 1 次變動, 而狀態變換所描述的則是參數結構在兩個以上的狀態間來回變換的現象。Debold (1986) 並未對狀態變換如何形成假性緩長記憶作更深入的說明; 但 Lamoureux and Lastrapes (1990)

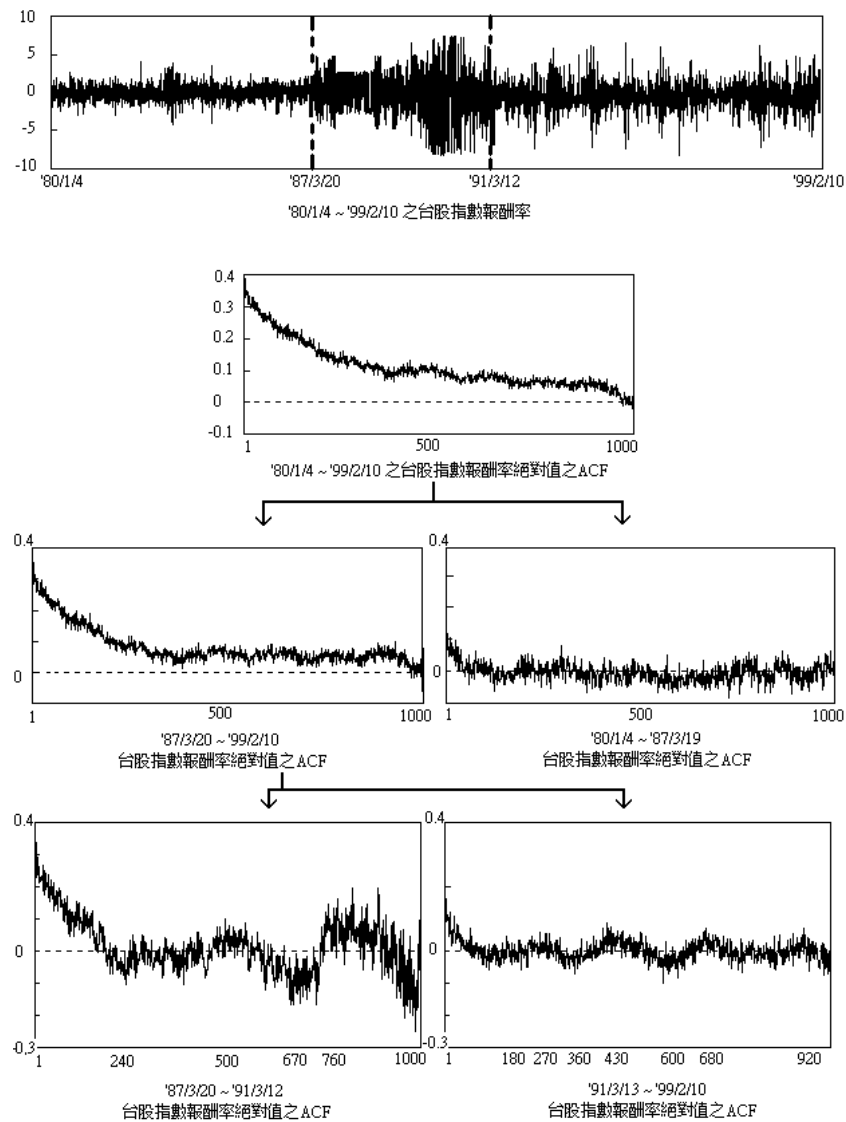


圖 1: 1980/1/4 至 1999/2/10 期間, 台股指數日報酬率資料在取絕對值後呈現出典型緩長記憶型態的 ACF。以 1987/3/20 為界將資料分為前後兩段後, 前段部分呈現出起伏型態的 ACF, 後段部分仍為緩長記憶型態, 但已稍微減弱。再進一步以 1991/3/12 為界將後段資料分為兩部分, 則緩長記憶型態進一步被削弱, 取而代之的是起伏型態的 ACF。

以及 Engle and Mustafa (1992) 也都認為以 ARCH 或 GARCH 模型得到的參數估計值有結構不穩定的現象, 其中 Lamoureux and Lastrapes (1990) 進一步認為 GARCH 模型中, 條件變異數結構一旦不穩定, 可能會形成高持續性的假象。基於辨別真偽緩長記憶現象的重要性, 我們有必要對狀態變換及其形成假性緩長記憶的過程加以探究。

欲辨別狀態變換與緩長記憶, 一個適當的方法是: 建立一套同時包含狀態變換設定與緩長記憶設定的模型, 使我們既可以在狀態變換的前提下檢定緩長記憶假說, 也能在考慮緩長記憶可能存在的前提下檢定是否有狀態變換發生。換言之, 我們可藉由同時估計狀態變換與緩長記憶, 讓資料自行判定緩長記憶的真偽。在緩長記憶方面, 原則上我們採納 Baillie (1998) Baillie, Bollerslev, and Mikkelsen (1996) 及 Bollerslev and Mikkelsen (1996) 的設定方式 (以下簡稱為 BBM/BM), 以 FIGARCH 模型捕捉報酬率絕對值及平方值資料可能存在的緩長記憶現象, 但在設定的細節方面則做了些許更動。在狀態變換方面, 我們參考 Hamilton and Susmel (1994) 所提出的 SWARCH-L 模型, 讓各個狀態依循馬可夫過程 (Markov process) 進行變換, 但在期望條件密度函數的推導方式上則做了必要的更動, 使馬可夫變換 (Markov switching) 的想法得以落實在 FIGARCH 設定上。

綜上所述, 我們將建立一個新模型 SW-FIGARCH-L (Markov switching FIFARCH with leverage effect) 模型。這個模型在某些條件下, 可分別簡化為 SW-GARCH-L 模型或 AR(1)-FIGARCH-L 模型; 換言之, 此模型允許狀態變換與緩長記憶同時存在, 而狀態變換模型與緩長記憶模型都是它的特例。此外我們要特別提出一點: ARCH 或 GARCH 設定下的馬可夫變換模型主要想刻劃的是序列干擾項平方值的動態性質, 因此, 檢查殘差項的平方值, 也就是模型未解釋部分的平方值是否仍具高度跨期相關應是一個相當重要的實證步驟; 然而就我們所見, 以股價指數報酬率為樣本而估計上述馬可夫變換模型的文獻卻都忽略了此一重要步驟。其中, 部分文獻進行了殘差項本身的跨期相關檢定, 卻仍未進一步對殘差項平方值進行檢定。事實上, 由於 SWARCH (或 SWARCH-L) 模型及 SW-GARCH (或 SW-GARCH-L) 模型中 ARMA 形式的條件變異數設定所能捕捉的僅限於干擾項平方值的「短期動態行為」; 倘若干擾項平方值具有顯著的「中長期性質」估計結果的殘差項平方值很可能會反映未被捕捉的中長期性質而呈現出高度跨期相關。換言之, ARCH 或

GARCH 設定下的馬可夫變換模型所要解釋的是序列平方值的動態行爲，但卻可能會因爲未同時捕捉其中長期性質而面臨未解釋部分平方值仍相當不乾淨的窘境。允許同時估計中長期行爲的 SW-FIGARCH-L 模型將會是解決這項問題的關鍵。

文章安排如下：第 2 節回顧緩長記憶模型的設定及其統計特性，並介 Lo-bato and Robinson (1997) 所提出的拉氏乘數檢定 (Lagrange multiplier test, 簡稱 LM 檢定)。第 3 節則檢討 Hamilton and Susmel (1994) 的 SWARCH-L 模型設定，並說明更動其期望條件密度函數的必要性。第 4 節對台股指數報酬率絕對值及平方值資料進行緩長記憶的 LM 檢定，並建立 SW-FIGARCH-L 模型，然後以 1991 年至 1999 年的台股指數日報酬率資料進行估計和檢定，並進行不同模型的比較。第 5 節討論狀態變換形成假性緩長記憶的可能過程，並以模擬方式說明狀態變換的確會形成假性緩長記憶現象。第 6 節則爲結論。

2 緩長記憶模型

第 1 節曾提及，經濟變數若具緩長記憶，表示各期隨機衝擊對該經濟變數的影響都會持續一段相當長的時間。這裏所謂的「一段相當長的時間」係針對自迴歸移動平均模型 (autoregressive moving average 模型，以下簡稱 ARMA 模型) 中所刻劃，相對短暫的持續性而言；若與單位根模型相比，則無其恆久 (permanent) 的影響。換言之，一個成功緩長記憶設定所描述的持續性必須介於 ARMA 模型與單位根模型間；持續性必須遠高於 ARMA 模型，但不能是恆久的。爲了刻劃序列的緩長記憶現象，Granger (1980), Granger and Joyeux (1980) 及 Hosking (1981, 1984) 曾導出「部分整合」(fractionally integrated) 模型。該模型的特色在於 ACF 遞減的速度遠比 ARMA 模型緩慢。

對一個有平均數 μ 的序列而言， y_t 若符合下式即稱其爲 d 階整合序列，(簡稱爲 $I(d)$ 序列)：

$$(1 - B)^d (y_t - \mu) = \varepsilon_t, \quad (2.1)$$

d 即所謂的「部分差分係數」， B 爲落後運算子 (back shift operator)， ε_t 是一個以 0 爲平均數之弱定態 (weak stationary) $I(0)$ 序列。部分差分係數 d 可捕捉序列的長期行爲。爲同時考慮短期行爲，令 ε_t 爲定態且可逆 (invertible) 的 ARMA 序列： $\phi(B)\varepsilon_t$ ，則 (2.1) 式可改寫爲：

$$\phi(B)(1 - B)^d(y_t - \mu) = \theta(B)\varepsilon_t, \quad (2.2)$$

這裡 $\phi(B) = 1 - \sum_{j=1}^p \phi_j B^j$ 為 p 階自迴歸多項式, 而 $\theta(B) = 1 + \sum_{j=1}^q \theta_j B^j$ 為 q 階移動平均多項式, ε_t 是一以 σ^2 為變異數之白干擾 (white noise) 序列。(2.2) 式即自迴歸部分整合移動平均模型, 簡稱 ARFIMA(p,d,q) 模型。 $d > -0.5$ 時 (2.2) 式為一可逆序列, $d < 0.5$ 時 (2.2) 式則為一定態序列, 換言之, $-0.5 < d < 0.5$ 時 ARFIMA(p,d,q) 為定態且可逆的序列。當 $p = q = 0$ 時, (2.2) 式可簡化為單純的 I(d) 形式, 換言之, ARFIMA(0,d,0) 即為單純的 I(d) 模型; 當 $d = 0$ 時, (2.2) 式則簡化為標準的 ARMA(p,q) 模型。

雖然 ARMA(p,q) 與 ARFIMA(p,d,q) 唯一的差別在於 $d = 0$ 與否, 兩者 ACF 的遞減速度卻相去甚遠。ARMA(p,q) 的 ACF 呈快速的幾何率遞減:

$$|\rho_j| \sim C_1 r^j, \quad j \rightarrow \infty, \quad (2.3)$$

ρ_j 表落後 j 期之自我相關係數 C_1 為大於 0 的常數, r 為 0 和 1 間的分數。 $-0.5 < d < 0.5$, 但 $d \neq 0$ 時, ARFIMA(p,d,q) 的 ACF 則呈較緩慢的雙曲線率 (hyperbolic rate) 遞減:

$$|\rho_j| \sim C_2 j^{2d-1}, \quad j \rightarrow \infty, \quad (2.4)$$

ρ_j 表落後 j 期之自我相關係數 C_2 為大於 0 的常數。理論上, 當 $0 < d < 0.5$ 時, 此時 (2.3) 式可寫為 $\rho_j \sim C_2 j^{2d-1}, j \rightarrow \infty$ 這就是緩長記憶序列所表現出的 ACF 緩慢遞減現象。綜 (2.3)、(2.4) 兩式, 我們可發現, $d = 0$ 時 ACF 遞減最為快速; $-0.5 < d < 0.5$ 但 $d \neq 0$ 時, ACF 遞減速度相對緩慢許多, 且 d 愈大時遞減速度愈慢。 $d = 0$ 時我們稱序列僅具短期記憶 (short memory); $-0.5 < d < 0$ 時稱序列具中等記憶 (intermediate memory), $0 < d < 0.5$ 時則稱序列具緩長記憶; 三者皆為定態且可逆的序列。值得一提的是, 當 $0 < d < 0.5$ 時,

$$\lim_{T \rightarrow \infty} \sum_{j=0}^T \rho_j = \lim_{T \rightarrow \infty} \sum_{j=0}^T |\rho_j| \rightarrow \infty。$$

這表示緩長記憶序列的自我相關係數即使落後期很長仍不夠小, 因此不能忽略不計。

表 2.1: 從移動平均係數看模型持續性

$j =$	1	2	100	200	300	400	1000
1 (單位根)	1	1	1	1	1	1	1
φ^j (緩長記憶)	0.49	0.36505	0.05275	0.03707	0.03015	0.02604	0.01632
0.99^j (短期記憶)	0.99	0.9801	0.36603	0.13398	0.04904	0.01795	0.00004

讓我們再透過移動平均表示法來看單位根序列、緩長記憶序列、及短期記憶序列的持續性差異：

$$y_{1,t} - \mu_1 = \sum_{j=0}^{\infty} \varepsilon_{1,t-j}, \quad (2.5)$$

$$y_{2,t} - \mu_3 = \sum_{j=0}^{\infty} \psi_j \varepsilon_{2,t-j} \psi_0 = 1, \quad \psi_j = \frac{j-1+0.49}{j} \psi_{j-1}, \quad (2.6)$$

$$y_{3,t} - \mu_2 = \sum_{j=0}^{\infty} (0.99)^j \varepsilon_{3,t-j}, \quad (2.7)$$

$y_{1,t}$, $y_{2,t}$, $y_{3,t}$ 分別為 I(1) 序列、I(0.49) 序列、及自迴歸參數為 0.99 之 AR(1) 序列。如表 2.1, 比較三序列不同落後期數之移動平均係數可發現, I(1) 序列中, 隨機衝擊的影響永遠不會減弱消失, 其餘兩者則皆會隨時間而逐漸減弱。自迴歸參數為 0.99 之 AR(1) 序列的短期持續性雖然強, 但由於遞減速度較快, 相距 1000 期後隨機衝擊的影響力遠小於 I(0.49) 序列; 這正是短期記憶與緩長記憶的主要差異所在。

但是金融報酬率的波動程度 (volatility) 為該金融商品風險程度 (risk) 的重要指標。文獻上最常被用來刻劃報酬率波動程度的模型莫過於 Engle (1982), 及 Bollerslev (1986) 所建立的「一般化自迴歸條件變異數非齊一」(generalized autoregressive conditional heteroskedasticity, 簡稱為 GARCH) 模型。de Lima, Briedt, and Crato (1994), Ding, Granger, and Engle (1993) 及 Harvey (1993) 等研究顯示, 許多總體或財務序列的波動程度都有高持續性的現象。但在 GARCH 模型中, ARMA 形式的條件變異數設定所捕捉的是波動程度的短期, 而非長期動態行為; 換言之, 用 GARCH 模型來刻劃屬於長期動態行為的長期持續性質並不恰當。為了同時捕捉波動程度的長短期動態行為, BBM/BM 在 GARCH 模型中加入部分整合設定而建立了「部分整合 GARCH」(fractionally integrated

GARCH, 簡稱 FIGARCH) 模型。在 GARCH 設定下, 波動程度的持續性係以幾何率遞減; 加入部分整合設定後的 FIGARCH 則以較緩慢的雙曲線率遞減。

假設 y_t 為我們所關心的序列, 如報酬率序列, 令 ε_t 為一以 σ^2 為變異數之白干擾序列, 且 ε_t 係序列 y_t 之干擾項 (innovations), 以 ARMA(l,m) 為例:

$$\phi(B)(y_t - \mu) = \theta(B)\varepsilon_t, \quad (2.8)$$

$\mu \equiv E(y_t)$ 是 y_t 的非條件平均數 (unconditional mean), $\phi(B) = 1 - \sum_{j=1}^l \psi_j B^j$ 為 l 階自迴歸多項式, 而 $\theta(B) = 1 + \sum_{j=1}^m \theta_j B^j$ 為 m 階移動平均多項式; 則所謂的 GARCH(p,q) 係指如下的設定:

$$h_t = \omega + \alpha(B)\varepsilon_t^2 + \beta(B)h_t, \quad (2.9)$$

其中, $h_t \equiv Var(y_t|Y^{t-1}) = Var(\varepsilon_t|Y^{t-1})$ 即所謂的條件變異數, $Y^{t-1} \equiv \{y_{t-1}, y_{t-2}, \dots, y_1\}$ 為各前期 y_t 所構成的訊息集合 (information set), $\omega = \sigma^2[1 - \beta(1) - \alpha(1)]$, 而 $\alpha(B) = \sum_{j=1}^q \alpha_j B^j$ 及 $\beta(B) = \sum_{j=1}^q \beta_j B^j$ 分別是落後運算子 B 的 q 階及 p 階多項式。

(2.9) 式可被改寫為 ε_t^2 的 ARMA 形式:

$$[1 - \alpha(B) - \beta(B)]\varepsilon_t^2 = \omega + [1 - \beta(B)]u_t, \quad (2.10)$$

其中, $u_t \equiv \varepsilon_t^2 - h_t$ 可被證明是平均數為 0、變異數為常數、非自我相關 (serially uncorrelated) 的序列, 因此可視為此 ARMA 形式中的干擾項。

相對於 GARCH 模型中 (2.10) 式的 ARMA 形式, BBM/BM 的 FIGARCH(p,d,q) 模型則可寫作 ε_t^2 的 ARFIMA 型式:

$$\psi(B)(1 - B)^d \varepsilon_t^2 = \omega + [1 - \beta(B)]u_t, \quad (2.11)$$

其中

$$\begin{aligned} \psi(B) &\equiv 1 - \sum_{j=1}^q \psi_j B^j, \\ (1 - B)^d &= \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(j+d)\Gamma(-d)} B^j \\ &= 1 - dB - \frac{d(1-d)}{2} B^2 - \frac{d(1-d)(2-d)}{6} B^3 - \dots \\ &= \pi(B) \end{aligned}$$

$$= \sum_{j=0}^{\infty} \pi_j B^j, \quad \pi_0 = 1, \quad \pi_j = \frac{j-1-d}{j} \pi_{j-1}.$$

$0 < d < 0.5$ 時表示 ε_t^2 具緩長記憶。相對於 GARCH 模型中的 (2.9) 式, (2.11) 式也可寫成下面的形式:

$$[1 - \beta(B)]h_t = \omega + [1 - \beta(B)]\varepsilon_t^2 - \psi(B)(1 - B)^d \varepsilon_t^2. \quad (2.12)$$

(2.11) 式與標準的 ARFIMA 形式最大的不同在於 ε_t^2 並未先與其期望值相減, 而代之以等號右邊的常數項值得注意的是, 這樣的形式不一致並非僅是「美感」上的問題, 它對模型估計本身也會造成不良影響。我們對 (2.11) 式作如下的修正:

$$\psi(B)(1 - B)^d (\varepsilon_t^2 - \sigma^2) = [1 - \beta(B)]y_t, \quad (2.13)$$

上式不但成為標準的 ARFIMA 形式, 更重要的是, 它也可以避免掉 (2.11) 式在估計步驟中所遭遇的若干樣本前數值 (pre-sample values) 選定上的問題, 並具有較理想的小樣本性質。同樣的, 我們可將 (2.13) 式改寫成

$$[1 - \beta(B)]h_t = -\beta(B)\varepsilon_t^2 + \varepsilon_t^2 - \psi(B)(1 - B)^d (\varepsilon_t^2 - \sigma^2). \quad (2.14)$$

雖然 (2.13) 式已成為標準的 ARFIMA 形式, (2.13) 式設定下的 FIGARCH, 與 ARFIMA 仍有一些結構上的差異。首先, 由上一節的討論我們知道, ARFIMA (p,0,q) 在定義上等同於 ARMA(p,q); 相對的, 即使 FIGARCH(p,0,q) 的確可寫成 ε_t^2 的 ARMA 形式, 由於 $\psi(B)$ 並不同於 $1 - \alpha(B) - \beta(B)$ FIGARCH(p,0,q) 雖可能仍保有 GARCH 形式, 卻不會全然等同於 GARCH(p,q)。另外, 在 ARFIMA 模型中, 我們並不須為其平均數或條件平均數加上任何限制條件, 然而在 FIGARCH 模型中, 我們卻有必要設下若干限制條件, 以保證非條件變異數及各期條件變異數皆非負數。

值得小心的是, GARCH 或 FIGARCH 所要描述的是 ε_t 之干擾項平方值, ε_t^2 的行為, 是一種「條件變異數方程式」設定至於 y_t 與干擾項 ε_t 之間的關係, 也就是「條件平均數方程式」, 則必須針對個案另做適當設定。在本節中我們以 ARMA(l,m) 描述 y_t 與其干擾項 ε_t 之間的關係, 而以 FIGARCH(p, d,q) 刻劃 ε_t^2 的行為, 我們可將這整套設定命名為 ARMA(l,m)-FIGARCH(p,d,q)。

如果要檢定緩長記憶假說，我們可先以緩長記憶模型估計部分差分係數 d ，再針對估計值進行假設檢定；如果我們的主要目的係在於估計本身，上述檢定程序無疑是最直接而符合直覺的，倘若我們只是單純地想分析資料的緩長記憶性質，耗費大量時間在模型的選擇和往往並非線性的估計上就顯得奢侈而多餘。尤其當我們需要重覆數百次、甚至數千次檢定時，上述先估計再檢定的程序將完全無法適用。基於上述理由，我們有必要介紹一套穩健 (robust) 而計算簡便的檢定統計量，幫助我們對資料具緩長記憶特性與否進行快速而確實的了解。常用的緩長記憶檢定法包括 Lo (1991) 修正 Hurst (1951) 之 R/S (rescaled range) 統計量而提出的「修正後 R/S 檢定」(modified R/S test)、Geweke and Porter-Hudak (1983) 的 頻譜迴歸 (spectral regression) 檢定 (簡稱為 GPH 檢定)、及 Lobato and Robinson (1997) 所提出的 LM 檢定。Cheung (1993b) 曾質疑 GPH 檢定及修正後 R/S 檢定的檢定力 (power); Robinson (1994) 認為修正後 R/S 檢定的效率性質 (efficiency properties) 值得懷疑; Lobato and Savin (1998) 指出，修正後 R/S 檢定統計量的漸近分配十分複雜，相對的，LM 檢定統計量的漸近分配則是常見的卡方分配。因此，以下我們將介紹 Lobato and Robinson (1997) 所提出的 LM 檢定。

令 x_t 為一弱定態緩長記憶 ARFIMA(p,d,q) 序列，則其頻譜密度函數 (spectral density function) $f(\lambda)$ 存在，且當 $d \neq 0$ 時滿足以下條件：

$$\lim_{\lambda \rightarrow 0^+} f(\lambda) \sim \lim_{\lambda \rightarrow 0^+} C_3 f(\lambda)^{-2d} \rightarrow \infty, \quad (2.15)$$

換言之，相對於表現在時間定義域 (time domain) 中 ACF 雙曲線率遞減或 $\sum_{j=0}^{\infty} \rho_j \rightarrow \infty$ 的特徵，緩長記憶序列表現在頻率定義域 (frequency domain) λ 的特徵則是：頻率很靠近 0 時，頻譜密度函數值有發散的傾向；Lobato and Robinson (1997) 的 LM 統計量事實上就是在檢查 $f(0^+)$ 是否有發散傾向。另外，由於 $f(0^+)$ 是否發散與序列的短、中期行為無關，以 $f(0^+)$ 是否發散作為標的的檢定準則會是一種不因短、中期設定差異而受影響的穩健準則 (robust procedure)。LM 統計量的計算方式是

$$LM = m \left(\frac{\sum_{j=1}^m v_j I(\lambda_j)}{\sum_{j=1}^m I(\lambda_j)} \right)^2, \quad (2.16)$$

其中

$$v_j = \log j - \frac{1}{m} \sum_{j=1}^m \log j,$$

$$I(\lambda_j) = \frac{1}{2\mu T} \left| \sum_{t=1}^T x_t e^{i\lambda_j t} \right|^2,$$

第 j 個頻率 $\lambda_j = 2\pi j/T$, T 是樣本數, m 是一個小於 T 的正整數且符合以下條件:

$$\lim_{T \rightarrow \infty} m \rightarrow \infty, \quad \lim_{T \rightarrow \infty} \frac{m}{T} \rightarrow 0.$$

(2.16) 式中唯一需要我們主觀給定的數字是 m , $[T/m]$ 代表的是檢定中每次納入考慮的最短期間 (shortest period)。統計量 LM 可看作 $f(0^+)$ 是的發散程度指標, 愈大的 LM 值反映著愈明顯的發散傾向。由於 d 係介於 0 至 0.5 間抑或等於 0 是 $f(0^+)$ 是否發散的唯一關鍵, 因此, 檢查 $f(0^+)$ 是否發散正相當於檢查序列係緩長或短期記憶。虛無假設及對立假設如下:

$$H_0 : d = 0 \text{ (短期記憶)} \quad H_1 : 0 < d < 0.5 \text{ (緩長記憶)}.$$

虛無假設成立時, LM 統計量的漸進分配為自由度 1 的卡方分配:

$$LM \xrightarrow{d} \chi_1^2.$$

最後必須強調的是, (2.25) 式所敘述的條件是 ARFIMA 序列之必要條件而非充分條件, 換言之, 若有一序列符合 (2.15) 式的要求, 即使它並非 ARFIMA 序列, 我們也承認它的緩長記憶性質; LM 統計量所檢定的是序列是否具緩長記憶而非序列是否為 ARFIMA 序列。

3 馬可夫變換模型

欲描述參數結構的變換, 一個可行的方式是以兩個以上的母體分別配適 (fit) 各個不同的狀態; 傳統的混合模型 (mixture model) 即可允許這種特殊配適。在簡單的混合模型中, 狀態的前後期變動是互相獨立的; 然而, 經濟狀態的變動往往並非毫無規律, 以景氣變動為例, 各期景氣勢必會持續一段期間, 絕非在好壞景氣間無規律地獨立變動。基於這項考量, 馬可夫變換模型修正狀態的變

動方式：假設狀態數目有限，視各期狀態為一馬可夫過程之各期實現值。所謂馬可夫過程係指一種特殊時間序列，對該序列未來實現值之預測只與序列的當期實現值有關。以我們所關心的狀態變換過程為例，令 s_t 代表無法觀察的隨機狀態變數，假設經濟在 k 個狀態間變換， s_t 的樣本空間 $\{s_t\} = \{1, 2, \dots, K\}$ ，不同的實現值各代表不同的狀態。若 s_t 係一馬可夫過程，則

$$P_r(s_{t+m}|Y^t, s_t, s_{t-1}, \dots) = P_r(s_{t+m}|s_t), \quad (3.1)$$

其中 m 為任意正整數， Y^t 與第 2 節之定義相同，為當期 (t 期) 的訊息集合。換言之，給定當期狀態，過去各期狀態及當期所有資訊都無法對未來狀態的預測有所貢獻。因此，我們可用下列狀態轉變矩陣 (the matrix of transition probability) 來描述狀態的變換過程：

$$P = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{k1} \\ p_{12} & p_{22} & \cdots & p_{k2} \\ \vdots & \vdots & \dots & \vdots \\ p_{1k} & p_{2k} & \cdots & p_{kk} \end{bmatrix},$$

其中 $p_{ij} = Pr(s_t = j|Y^t, s_{t-1} = i, s_{t-2} = \cdot, \dots) = Pr(s_{t-1} = j|s_{t-1} = i)$, $i, j, = 1, 2 \dots k_0$ 由於 $\sum_{j=1}^k p_{ij} = \sum_{j=1}^k Pr(s_t = j|s_{t-1} = i) \equiv 1$ 因此矩陣 P 每一直行的加總和都為 1。

Hamilton and Susmel (1994) 將上述馬可夫狀態變換的想法落實在股價指數報酬率的估計上，建立了 SWARCH-L(k,q) 模型：

$$\begin{aligned} y_t &= a_0 + a_1 y_{t-1} + \sqrt{g_{st}} \varepsilon_t, \\ \varepsilon_t &= \sqrt{h_t} v_t, \quad v_t \stackrel{iid}{\sim} (0, 1), \\ h_t &= b_0 + \sum_{j=1}^q b_j \varepsilon_{t-j}^2 + c D_{t-1} \varepsilon_{t-1}^2, \\ D_{t-1} &= \begin{cases} 1 & \varepsilon_{t-1} \leq 0, \\ 0 & \varepsilon_{t-1} > 0. \end{cases} \end{aligned} \quad (3.2)$$

這裡 s_t 的定義及其隨機行為與前述相同。實際報酬率資料往往有正負值比例相近的現象，因此假設 $Pr(D_{t-1} = 1) = Pr(D_{t-1} = 0) = 1/2$ ；換言之，

$E(cD_{t-1}) = c/2$ 為常數, $c > 0$ 時表示槓桿效果存在。 t 期狀態 $s_t = j$ 時, 條件變異數 $Var(y_t|Y^{t-1}) = g_j h_t$; $g_1 = 1 < g_2 < \dots < g_k$ 。換言之, 此模型的架構基本上是一個 AR(1)-ARCH(q) 模型: 以 AR(1) 描述報酬率 y_t 及其干擾項之間的關係, 而以 ARCH(q) 刻劃干擾項平方值的行為; 但在 ARCH 設定中加入了槓桿效果, 條件變異數的係數值則反映 k 個狀態的馬可夫變換而呈現比例性變動。若各期狀態已知, 我們可寫下近似對數概似函數

$$\begin{aligned} \ln L(\theta) &= \sum_{t=1} \ln l_t(\theta), \\ \ln l_t(\theta) &= -\frac{1}{2} \ln(2) - \frac{1}{2} \ln(g_{s_t} h_t) - \frac{1}{2} v_t^2, \end{aligned} \quad (3.3)$$

$\ln l_t(\theta)$ 是 t 期的對數條件密度函數, $\theta \equiv [a_0 \ a_1 \ g_2 \ \dots \ g_k \ b_0 \ \dots \ b_q \ c]'$ 為參數向量。由於只有 y_t 才是可觀察的變數, 我們必須將上式中的 v_t 及 h_t 都寫成 y_t 的函數:

$$v_t = \frac{y_t - a_0 - a_1 y_{t-1}}{\sqrt{g_{s_t} h_t}}, \quad (3.4)$$

$$\begin{aligned} h_t &= b_0 + \sum_{j=1}^q \left(\frac{y_{t-j} - a_0 - a_1 y_{t-j-1}}{\sqrt{g_{s_{t-j}}}} \right)^2 \\ &\quad + c D_{t-1} \left(\frac{y_{t-1} - a_0 - a_1 y_{t-2}}{\sqrt{g_{s_{t-1}}}} \right)^2. \end{aligned} \quad (3.5)$$

必須注意的是, 由於狀態變數 s_t 是觀察不到的隱藏變數, 我們無從得知 t 期的 g_{s_t} 究竟是 g_1, g_2, \dots, g_{k-1} 或是 g_k 。更精確地說, 由於 $\ln l_t(\theta)$ 是 s_t, v_t, h_t 的函數, v_t 是 s_t, h_t 的函數, 而 h_t 是 $s_{t-1}, s_{t-2}, \dots, s_{t-q}$ 的函數, 因此 $\ln l_t(\theta)$ 會是 $s_{t-1}, s_{t-2}, \dots, s_{t-q}$ 的函數; 考慮各種狀態組合, $\ln l_t(\theta)$ 共有 k^{q+1} 種可能性。

由上述討論我們知道 (3.3) 式是不可行的; 我們必須逐一計算各個時點下各個狀態組合的條件密度函數, 進而求得各個時點下的「期望條件密度函數」, 以建立一個可行的近似對數概似函數。為簡化分析起見, 我們定義一個新的狀態變數 S_t^* :

$$S_t^* = 1 + (s_t - 1)k^0 + (s_{t-1} - 1)k^1 + (s_{t-q} - 1)k^q, \quad (3.6)$$

S_t^* 的樣本空間 $\{S_t^*\} = \{1, 2, \dots, K^{q+1}\}$, 不同的實現值各代表上述不同的狀態組合。同樣的, 我們可定義 S_t^* 的狀態轉變矩陣:

$$P^* = [P_{ij}^*]_{k^{q+1} \times k^{q+1}} = [Pr(S_{t-1}^* = i)]_{k^{q+1} \times k^{q+1}}, \quad (3.7)$$

P_{ij}^* 是 P^* 中第 i 直行、第 j 橫列之元素。另定義兩個向量

$$\eta_t = [f(y_t | Y^{t-1}, S_t^* = j; \theta)]_{k^{q+1} \times 1}, \quad (3.8)$$

$$\Xi_{t|\tau} = [Pr(S_t^* = j | Y^\tau)]_{k^{q+1} \times 1}, \quad (3.9)$$

η_t , 是 t 期各狀態組合之條件密度函數所構成的向量, $\Xi_{t|\tau}$, 則是給定 τ 期訊息下, t 期各狀態組合之發生機率所構成的向量。我們必須自行給定 $\Xi_{1|0}$, 做為計算的開端, Hamilton and Susmel (1994) 將狀態轉變矩陣換算成所謂的遍態歷程機率 (ergodic probability), 作為 $\Xi_{1|0}$ 的給定值。給定 $\Xi_{1|0}$ 後, 我們就可以利用下面兩式遞迴求解:

$$\Xi_{t|t} = \frac{(\Xi_{t|t-1} \odot \eta_t)}{1'(\Xi_{t|t-1} \odot \eta_t)}, \quad (3.10)$$

$$\Xi_{t+1|t} = P^* \Xi_{t|t}, \quad (3.11)$$

其中, \odot 表示兩向量元素對元素相乘。 $1'(\Xi_{t|t-1} \odot \eta_t) = f(y_t | Y^{t-1}; \Theta)$ 即為 t 期的期望條件密度函數。 Θ 為所有參數組成的向量, 包括 P 中的 $(k-1)k$ 個參數及 θ 中的所有參數。將各期期望條件密度函數取對數相加即得到近似對數概似函數:

$$\ln L(\Theta) = \sum_{t=1}^T \ln f(y_t | Y^{t-1}; \Theta). \quad (3.12)$$

由上面的討論不難發現, 由於各期狀態組合數目為 k^{q+1} , 隨著狀態數目 k 或 ARCH 的階數 q 增多, 每期所需考慮的組合數目將會快速增加, 計算的複雜度亦會隨之大幅提高。一旦我們將 ARCH 設定改為 GARCH 甚至 FIGARCH, 由於 GARCH 與 FIGARCH 都可展開為 ARCH(∞), 實際運算時, t 期的組合數目都將會是 k^t 。當 $k = 3$ 時, 第 1000 期所必須考慮的不同組合有 3^{1000} 種, 第 2000 期則有 3^{2000} 種; 這種高繁複程度的計算並非現有技術所能處理。因此, 若要建立 GARCH 或 FIGARCH 設定下的馬可夫變換模型, 我們必須修正上述計算過程。

根據前述分析, 各期條件密度函數之所以會有太多種可能狀況, 最主要的關鍵在於函數中的 h_t 包含了各前期的 ε_t , 而各前期又 ε_t 都是該期狀態的函數, 從 h_t 而也會是各前期狀態的函數。因此, 若要大幅減少各期條件密度函數的數目, 在計算 h_t 時, 必須用一個並非「狀態的函數」的替代變數取代各前期的 ε_t 。本文建議使用的替代變數為

$$\tilde{\varepsilon}_t = \sum_{j=1}^k \left[Pr(s_t = j | Y^t) \left(\frac{y_t - a_0 - a_1 y_{t-1}}{\sqrt{g_j}} \right) \right], \quad (3.13)$$

換言之, 我們是將同一期不同狀態的 ε_t 以其所對應狀態的條件機率作為權數, 求算加權平均數, 作為 ε_t 的替代變數。Gray (1996) 及 Dueker (1997) 也都曾利用類似上述狀態條件機率加權平均的方式建立 GARCH 設定下的馬可夫變換模型。由於他們允許條件平均數方程式或條件變異數方程式中的係數都隨狀態而變動, 與我們僅捕捉條件變異數比例性變動的設定不同, 因此處理的細節略有差異, 但在方法論上三者是一致的。如此一來, 即使在 GARCH 或 FIGARCH 設定下, 計算各期條件密度函數時只須對應當期狀態考慮 k 種狀況即可。

4 台股指數報酬率與 SW-FIGARCH-L 模型

第1節曾提到, 1980年1月至1999年2月台股指數報酬率絕對值的 ACF 呈現十分緩慢的下降速度, 似乎有緩長記憶性質; 但將資料分成3個樣本數夠大的子資料後, 緩長記憶特徵卻變得較不明顯, 顯示樣本的緩長記憶特徵並不穩定。有鑑於股價指數報酬率係一重要經濟、財務變數, 我們有必要建立一個能夠同時估計其條件變異數狀態變換及長、短期行為的模型, 俾使我們對台股指數報酬率的動態性質獲致更深刻的了解。

4.1 緩長記憶檢定

在正式進入估計前, 我們應先確認台股指數報酬率資料是否真的呈現出條件變異數結構不穩定現象及緩長記憶現象。高櫻芬·呂仁廣·林建甫 (2001) 曾以平滑結構改變的 GARCH 模型研究 1990年1月4日至1998年1月24日的台股指數報酬率資料, 並進行各種 LM 檢定, 發現各檢定同樣都拒絕「條件變異數未發生結構改變」的虛無假設, 顯示台股指數報酬率資料的確呈現出條件

表 4.1: 台股指數報酬率絕對值及平方值資料的緩長記憶檢定

$m =$	絕對值資料				平方值資料			
	40	50	60	100	40	50	60	100
1987/3/20-1999/2/10	9.86	11.95	13.78	20.36	9.25	11.28	12.97	19.56
1987/3/20-1991/3/12	8.89	10.63	11.92	17.28	10.02	11.94	13.21	19.53
1991/3/13-1999/2/10	3.80	4.91	5.37	7.72	2.84	3.79	4.04	6.56

說明: 表中所列係 \sqrt{LM} ; 5% 顯著水準下之臨界值為 1.96。

變異數結構不穩定現象。本節進一步要確認的是台股指數報酬率絕對值及平方值資料是否呈現出緩長記憶現象。

根據圖 1, 資料期間由 1980/1/4-1999/2/10 縮短為 1987/3/20-1999/2/10 時, 緩長記憶特徵僅稍微減弱。但再進一步分割為 1987/3/20-1991/3/12 及 1991/3/13-1999/2/10 後, 兩段期間, 尤其是 1991/3/13-1999/2/10, 緩長記憶特徵明顯減弱。我們以 $m = 50, 60, 100$ 計算 1987/3/20-1999/2/10、1987/3/20-1991/3/12、1991/3/13-1999/2/10 段期間的緩長記憶 LM 統計值, 開平方根後列於表 4.1。各個期間、不同 m 值, 在 5% 顯著水準下都拒絕短期記憶虛無假設, 台股指數報酬率資料的確呈現出緩長記憶現象。然而, 與前述圖 1 觀察結果相同, 1991/3/13-1999/2/10 的緩長記憶特徵明顯減弱, 整個樣本的緩長記憶特徵並不穩定。

4.2 模型設定

SW(k)-FIGARCH-L(0,d,0) 與 SWARCH-L(k,q) 最大的不同在於條件變異數方程式的設定。以下我們將先定義 FIGARCH-L(0,d,0) 設定, 再進一步寫下整個模型。根據 (2.18) 式, 我們可寫下 FIGARCH(0,d,0) 模型:

$$(1 - B)^d (\varepsilon_t^2 - \sigma^2) = u_t, \quad (4.1)$$

$u_t \equiv \varepsilon_t^2 - h_t(1 - B)^d$ 的定義與 (2.4) 式相同。上式又可根據 (2.19) 式寫成

$$\begin{aligned} h_t &= \varepsilon_t^2 - (1 - B)^d (\varepsilon_t^2 - \sigma^2) \\ &= \sigma^2 - \sum_{j=1}^{\infty} \pi_j (\varepsilon_{jt}^2 - \sigma^2). \end{aligned} \quad (4.2)$$

在上式中加入槓桿效果即成爲我們所需要的 FIGARCH-L(0,d,0) 設定:

$$h_t = \sigma^2 - \sum_{j=1}^{\infty} \pi_j (\varepsilon_{t-j}^2 - \sigma^2) + (1 - bB)^{-1} \left(dD_{t-1}\varepsilon_{t-1}^2 - \frac{c}{2}\sigma^2 \right), \quad (4.3)$$

D_{t-1} 的定義與 (3.2) 式相同。我們將參數 b 命名爲「槓桿 GARCH 效果」。爲方便估計, 也可將等號兩邊同乘以 $(1 - bB)$ 而改寫爲:

$$\begin{aligned} h_t = (1 - b)\sigma^2 + bh_{t-1} - \sum_{j=1}^{\infty} \pi_j (\varepsilon_{t-j}^2 - \sigma^2) + b \sum_{j=1}^{\infty} \pi_j (\varepsilon_{t-j-1}^2 - \sigma^2) \\ + \left(dD_{t-1}\varepsilon_{t-1}^2 - \frac{c}{2}\sigma^2 \right), \end{aligned} \quad (4.4)$$

$d = 0$ 時, 上式可簡化爲

$$h_t = \left(1 - b - \frac{c}{2} \right) \sigma^2 + bh_{t-1} + cD_{t-1}\varepsilon_{t-1}^2, \quad (4.5)$$

這就是我們所謂的 GARCH-L(1,0)。

我們採 (3.2) 式的 SWARCH(k,q) 模型作爲基本架構, 以 (4.4) 式的 FIGARCH-L(0,d,0) 設定取代其中的 ARCH-L(q) 設定, 並將含常數項的 AR(1) 設定修改爲標準的 AR(1) 形式, 建立 SW(k)-FIGARCH-L(0,d,0) 模型如下:

$$\begin{aligned} y_t &= (1 - a)\mu + ay_{t-1} + \sqrt{g_{st}}\varepsilon_t, \\ \varepsilon_t &= \sqrt{h_t}v_t, \\ h_t &= (1 - b)\sigma^2 + bh_{t-1} - \sum_{j=1}^{\infty} \pi_j (\varepsilon_{t-j}^2 - \sigma^2) \\ &\quad + b \sum_{j=1}^{\infty} \pi_j (\varepsilon_{t-j-1}^2 - \sigma^2) + \left(cD_{t-1}\varepsilon_{t-1}^2 - \frac{c}{2}\sigma^2 \right), \end{aligned} \quad (4.6)$$

其中, μ 爲 y_t 之非條件平均數, s_t, v_t 的定義及其隨機行爲與 (3.2) 式相同。 $d = 0$ 時 FIGARCH-L(0,d,0) 會簡化爲 GARCH-L(1,0), 從而 $d = 0$ 時 SW(k)-FIGARCH-L(0,d,0) 也會簡化爲 SW(k)-GARCH-L(1,0); 此外, 當 $g_{st} \equiv 1$, 也就是無狀態變換時, SW(k)-FIGARCH-L(0,d,0) 會簡化爲 AR(1)-FIGARCH-L(0,d,0); $d = 0$ 且 $g_{st} \equiv 1$ 時則簡化爲 AR(1)-GARCH-L(1,0)。

我們所建立的 SW-FIGARCH-L 模型是 SW(k)-FIGARCH-L(0,d,0), 而非更一般化的 SW(k)-FIGARCH-L(p,d,q), 理由有二: 首先, 當 $d = 0$ 時, SW(k)-FIGARCH-L(0,d,0) 可簡化為 SW(k)-GARCH-L(1,0) 的形式; 換言之, SW(k)-FIGARCH-L(0,d,0) 中所捕捉的短期動態行為是各期的槓桿效果而非各期的第 1 階 ARCH 效果。在條件變異數方程式中, 槓桿效果及第 1 階 ARCH 效果都是 ε_{t-1}^2 的係數, 然而 Hamilton and Susmel (1994) 以 SWARCH-L (3,2)、SWARCH-L (4,2) 估計美國的股價指數報酬率資料, Lin (2000) 以 SWARCH-L (2,2)、SWARCH-L (3,2) 估計台股指數報酬率資料, 結果都顯示: 第 1 階 ARCH 效果十分不顯著; 相對的, 同是 ε_{t-1}^2 的係數, 槓桿效果卻非常顯著。我們也曾嘗試納入第 1 階 ARCH 係數, 估計 SW (3)-GARCH-L (1,1), 然而 ARCH 係數在估計過程中總是不停地向 0 靠近, 無法收斂。因此, 放棄第 1 階 ARCH 設定而僅以槓桿效果設定捕捉短期動態行為似乎是可以接受的。

此外, 一旦 p 或 q 不為 0, 部分差分係數數值的正負大小都將受到很大的限制。以 FIGARCH(1,d,1) 為例:

$$h_t = (1-\beta B)^{-1}(1-\psi B)(1-B)^d \sigma^2 + [1 - (1-\beta B)^{-1}(1-\psi B)(1-B)^d] \varepsilon_t^2。$$

為保證上式中的 h_t 不為負, 有必要限制 $0 < \psi < \beta \leq d < 1$ 。然而, 我們之所以要在馬可夫變換模型中加入部分整合設定, 主要是想藉由部分差分係數 d 捕捉 ε_t^2 「可能存在」的緩長記憶或中等記憶; 在估計中保留 $d \leq 0$ 的可能性是有必要的。 p 或 q 大於 0 的確可能有助於 ε_t^2 短期行為的捕捉, 但卻必須付出參數增加、模型複雜化、限制 $d > 0$ 等代價。相形之下, 保留部分差分係數數值正負的彈性, 而利用各期槓桿效果捕捉短期的行為, 似乎是一個較理想的選擇。

4.3 SW(k)-FIGARCH-L(0,d,0) 的估計方法

由於樣本前數值並不可得, 我們假設 $y_1 = \mu + \sqrt{g_{s1}}\varepsilon_1, h_1 = \sigma^2, \varepsilon_0 = 0$ 此外, 我們令 $\sum_{j=t}^{\infty} \pi_j (\varepsilon_{t-j}^2 - \sigma^2) = \sum_{j=t}^{\infty} \pi_j (\varepsilon_{t-j-1}^2 - \sigma^2) = 0$, 因此,

$$h_t = (1-b)\sigma^2 + bh_{t-1} - \sum_{j=1}^{t-1} \pi_j (\varepsilon_{t-j}^2 - \sigma^2) + b \sum_{j=1}^{t-1} \pi_j (\varepsilon_{t-j-1}^2 - \sigma^2) + \left(dD_{t-1}\varepsilon_{t-1}^2 - \frac{c}{2}\sigma^2 \right), \forall t \geq 2。$$

爲使計算各期條件密度函數時只須對應當期狀態考慮 k 種狀況, 在計算 h_t 時必須以加權平均替代變數 $\tilde{\varepsilon}_t$ 取代各前期的 ε_t :

$$h_t = (1-b)\sigma^2 + bh_{t-1} - \sum_{j=1}^{t-1} \pi_j (\tilde{\varepsilon}_{t-j}^2 - \sigma^2) + b \sum_{j=1}^{t-1} \pi_j (\tilde{\varepsilon}_{t-j-1}^2 - \sigma^2) + (dD_{t-1}\tilde{\varepsilon}_{t-1}^2 - c\sigma^2/2), \quad \forall t \geq 2. \quad (4.7)$$

定義 $\varepsilon_1 = y_1 - \mu$, $\varepsilon_t = y_t - (1-a)\mu - ay_{t-1} \forall t \geq 2$ 則在常態分配假設下, 給定訊息 Y^{t-1} , $s_t = j$ 之下, t 期的條件密度函數爲

$$f(y_t | Y^{t-1}, s_t = j; \theta) = \frac{1}{\sqrt{2\pi g_i h_t}} \exp\left(\frac{-\varepsilon_t^2}{2g_i h_t}\right), \quad j = 1, 2, \dots, k. \quad (4.8)$$

$\theta \equiv [\mu \ a \ g_2 \ \dots \ g_k \ \sigma^2 \ b \ c \ d]'$ 爲參數向量。向量 η_t 則定義爲

$$\eta_t = [f(y_t | Y^{t-1}, s_t = j; \theta)]_{k \times 1}, \quad (4.9)$$

令 ε_{tj} 代表 $s_t = j$ 時的 ε_t

$$\varepsilon_{tj} = \frac{\varepsilon_t}{\sqrt{g_i}}, \quad j = 1, 2, \dots, k, \quad (4.10)$$

而 $Y_t \equiv [\varepsilon_{tj}]_{k \times 1}$ 則爲 t 期各個 ε_{tj} 所組成的向量。另定義預測機率 (prediction probability) 向量爲

$$\Xi_{t|\tau} = [Pr(s_t = j | Y^\tau)]_{k \times 1}, \quad (4.11)$$

換言之, 預測機率向量即是給定 τ 期訊息下, t 期各狀態之發生機率所構成的向量。如此一來, 我們便可分別計算 t 期的期望條件密度函數及 $\tilde{\varepsilon}_t$ 如下:

$$f(y_t | Y^{t-1}; \Theta) = 1' (\Xi_{t|t-1} \odot \eta_t), \quad (4.12)$$

$$\tilde{\varepsilon}_t = 1' (\Xi_{t|t-1} \odot Y_t), \quad (4.13)$$

Θ 代表所有參數。同時亦可利用貝氏定理求得當期機率 (filtering probability) 向量:

$$\Xi_{t|t} = \frac{(\Xi_{t|t-1} \odot \eta_t)}{1' (\Xi_{t|t-1} \odot \eta_t)}. \quad (4.14)$$

利用當期機率向量及狀態轉變機率矩陣即可求得下一期的預測機率向量：

$$\Xi_{t+1|t} = P \Xi_{t|t} \quad (4.15)$$

以上各式構成一個遞迴系統：

- 將 $h_1 = \sigma^2$, $\varepsilon_1 = y_1 - \mu$, 代入 (4.8) 式求得 η_1 並利用 $\tilde{\varepsilon}_1$ 計算 Y_1 ;
- 給定第 1 期的預測機率向量 $\Xi_{1|0}$, 即可根據 (4.12)、(4.13)、(4.14)、(4.15) 式分別求得第 1 期的期望條件密度函數、 $\tilde{\varepsilon}_1$ 、第 1 期的當期機率向量 $\Xi_{1|1}$ 、以及第 2 期的預測機率向量 $\Xi_{2|1}$;
- (4.7) 式計算 h_t , 且計算 $\varepsilon_2 = y_2 - (1-a)\mu - ay_1$ 一起代入 (4.8) 式求得 η_2 並利 $\tilde{\varepsilon}_2$ 用計算 Y_2 ;
- 利用 (4.12)、(4.13)、(4.14)、(4.15) 式可分別求得第 2 期的期望條件密度函數、 $\tilde{\varepsilon}_2$ 、第 2 期的當期機率向量 $\Xi_{2|2}$ 、以及第 3 期的預測機率向量 $\Xi_{3|2}$;

換言之, 只要給定第 1 期的預測機率向量 $\Xi_{1|0}$, 即可透過上述反覆運算得到各期的期望條件密度函數 $f(y_t|Y^{t-1}; \Theta)$ 、當期機率向量 $\Xi_{t|t}$ 、以及預測機率向量。將各期期望條件密度函數取對數加總, 即可求得近似對數概似函數：

$$\ln L(\Theta) = \sum_{t=1}^T \ln f(y_t|Y^{t-1}; \Theta) \quad (4.16)$$

能使上述近似對數概似函數值達到最大的 Θ 值, $\hat{\Theta}$ 即是近似最大概似法的參數估計值。

當期機率顧名思義是從當期的角度看當期狀態變數的機率。除此之外, 我們也可以求算「事後機率」(smoothing probability), 從事後的角度看狀態變數的機率。根據 Hamilton (1994), 我們可利用各期當期機率向量及預測機率向量, 透過下式以遞迴方式逐次求得各期的事後機率向量 $\Xi_{t|T}$

$$\Xi_{t|T} = \Xi_{t|t} \odot \left\{ P' \left[\Xi_{t+1|T} (\div) \Xi_{t+1|t} \right] \right\}, \quad (4.17)$$

這裡 (\div) 表示兩向量元素對元素相除。

此外, 將上述估計步驟適當修改即可用以估計 SW(k)-GARCH-L(1,0) 以及 AR(1)-FIGARCH-L(0,d,0), 估計細節不再贅述。

表 4.2: 1991/3/13–1999/2/10 台股指數日報酬率的基本統計性質

平均數	標準差	偏態係數	峰態係數	最大值	最小值
0.0193	1.5228	-0.0814	5.1101	6.5420 (1991/8/22)	-6.9812 (1996/1/5)

4.4 實證結果

我們所選擇的樣本期間是 1991/3/13–1999/2/10, 基本統計特性見表 4.2。選擇理由有三: 首先, 根據圖 2.1 及第 1 節的 LM 檢定結果, 雖然 1991/3/13–1999/2/10 的台股指數日報酬率在平方或取絕對值後表現出緩長記憶特徵, 但相較於 1980/1/4–1999/2/10 或 1987/3/20–1999/2/10, 緩長記憶特徵已減弱許多。我們有必要探討 1991/3/13–1999/2/10 期間內是否還存在更多其他的結構改變點; 這段「子資料」的緩長記憶性質是否真的是狀態變換所造成的假性現象。其次, 馬可夫變換模型雖是一種描述多狀態結構改變的模型, 但由於我們實際能夠處理的狀態數目有限, 倘若樣本期間包含一段太過極端的資料, 資料自行分辨狀態的能力與彈性將大受影響。台股指數自 1990/2/10 的歷史高點 12495 急跌至 1990/10/1 的 2560 點後, 在 5 個月內又反彈上升約 2300 點。相對於往後數年, 這段期間報酬率高波動的持續時間相當長, 選擇樣本範圍時有必要避開這段較極端的期間。此外, 由於台灣股市有所謂的漲跌幅限制, 若樣本期間有多日出現「太多個股碰觸到漲跌幅限制」的現象, 這些特殊交易日彼此的相關性也可能會造成樣本表現出長時距依存特性; 1991/3/13 之後的交易日較無多日碰觸漲跌幅限制的現象, 以 1991/3/13–1999/2/10 為樣本期間進行估計較不會受到漲跌幅限制影響。

我們分別以 SW(3)-FIGARCH-L(0,d,0)、SW(3)-GARCH-L(1,0)、AR(1)-FIGARCH-L(0,d,0) 三種模型進行估計。參數估計結果列於表 4.3; 各種模型選擇及配適度的指標列於表 4.4; 殘差項、殘差項平方值自我相關的 Ljung-Box 檢定結果列於表 4.5、4.6; SW(3)-GARCH-L(1,0) 及 SW(3)-FIGARCH-L(0,d,0) 的事後機率分別繪於圖 2 及圖 3; 各模型的條件變異數則繪於圖 4。

根據估計結果, AR(1)-FIGARCH-L(0,d,0) 的部分差分係數估計值十分顯著地大於 0, 但 AR(1)-FIGARCH-L(0,d,0) 的部分差分係數估計值卻十分顯著地小於 0; 這表示, 同時考慮狀態變換後, 原本的緩長記憶特徵已不復存在, 序

表 4.3: 參數估計結果

	SW(3) -FIGARCH -L(0,d,0)		SW(3) -GARCH -L(1,0)		AR(1) -FIGARCH -L(0,d,0)	
	估計值	<i>t</i> 統計量	估計值	<i>t</i> 統計量	估計值	<i>t</i> 統計量
μ	0.0158	10.32	20.0160	0.51	-0.0167	-0.68
a	0.0444	2.22	0.0408	1.97	0.0508	1.95
g_2	3.0593	10.90	2.6671	12.62	—	—
g_2	12.3779	10.49	8.5715	10.87	—	—
σ^2	0.5585	10.85	0.5604	14.17	5.9652	8.66
b	0.7921	6.14	0.7311	12.38	0.7359	19.14
c	0.1766	22.45	0.1181	4.83	0.2223	6.15
d	-0.1469	-32.37	—	—	0.0952	6.71
p^{12}	0.0703	3.97	0.0490	1.55	—	—
p^{21}	0.0399	3.44	0.0100	1.10	—	—
p^{23}	0.0180	14.92	0.0151	2.59	—	—
p^{32}	0.0507	3.70	0.0341	2.56	—	—

說明: 1. 5% 顯著水準下臨界值為 ± 1.96 。但值得注意的是, 在 $d > 0$ 時的 FIGARCH 架構下, 的收斂速度較慢, 並非標準的。

2. 所報告的 t 統計量為原始計算的統計量, 學界有時使用 Bollerslev and Wooldridge (1992) 的頑強統計量, 我們曾嘗試, 其結果並無太大改變。因為此模型的變異數結構尚待更進一步研究, 因此報告原始的計算結果。
3. 轉態轉變矩陣的參數個數原有 $(3 - 1)3 = 6$ 個, 但實際估計時 p_{13} 及 p_{31} 總是不斷向 0 靠近而無法收斂。這可能是表示台股指數報酬率很少直接由高波動期變換為低波動期, 或直接由低波動期變換為高波動期。我們參考 Hamilton and Susmel (1994) 的作法, 將向 0 靠近無法收斂的狀態轉變機率直接設定為 0, 因此實際估計的狀態轉變矩陣參數只有 4 個。SW(3)-FIGARCH-L(0,d,0) 及 SW(3)-GARCH-L(1,0) 的狀態轉變機率矩陣分別為

$$\begin{pmatrix} 0.9297 & 0.0399 & 0 \\ 0.0703 & 0.9421 & 0.0507 \\ 0 & 0.0180 & 0.9493 \end{pmatrix} \text{ 及 } \begin{pmatrix} 0.9510 & 0.0100 & 0 \\ 0.0490 & 0.9749 & 0.0341 \\ 0 & 0.0151 & 0.9659 \end{pmatrix}。$$

列僅具中等記憶。

我們在 4.1 節以 LM 統計量檢定上述資料絕對值及平方值的緩長記憶現象, 檢定結果支持緩長記憶假說; 事實上我們也曾進一步以 ARFIMA(1,d,0) 及

表 4.4: 模型比較

	參數 總數	$\ln L$	AIC	Schwarz	$ LE $	$[LE]^2$	MAE	MSE
SW(3)-FIGARCH-L(0,d,0)	12	-3926.6	-3938.6	-3973.0	1.750	6.687	0.521	0.488
SW(3)-GARCH-L(1,0)	11	-3940.8	-3951.8	-3983.2	1.842	6.985	0.528	0.538
AR(1)-FIGARCH-L(0,d,0)	6	-4000.2	-4006.2	-4023.4	2.053	8.326	2.631	21.001

說明: 樣本數 $T = 2250$ 。 AIC 、Schwarz、 $|LE|$ 、 $[LE]^2$ 、 MAE 、及 MSE 的定義與 Hamilton and Susmel (1994) 完全相同: $AIC = \ln L -$ 參數總數; Schwarz = $\ln L -$ (參數總數/2) $\ln(T)$;

$$|LE| = T^{-1} \sum_{t=1}^T \left| \ln(\varepsilon_t^2) - \ln(h_t) \right|, \quad [LE]^2 = \sum_{t=1}^T \left| \ln(\varepsilon_t^2) - \ln(h_t) \right|^2,$$

$$MAE = T^{-1} \sum_{t=1}^T \left| \varepsilon_t^2 - \ln h_t \right|, \quad MSE = T^{-1} \sum_{t=1}^T \left| \varepsilon_t^2 - \ln h_t \right|^2.$$

因此, $\ln L$ 、 AIC 、Schwarz 愈大愈好, $|LE|$ 、 $[LE]^2$ 、 MAE 、 MSE 愈小愈好。

表 4.5: 殘差項的自我相關檢定

	$LB(2)$	$LB(5)$	$LB(10)$	$LB(20)$	$LB(40)$	$LB(80)$
SW(3)-FIGARCH-L(0,d,0)	2.62	7.85	15.56	37.10	63.34	96.83
SW(3)-GARCH-L(1,0)	2.89	8.54	14.79	35.10	60.50	93.23
AR(1)-FIGARCH-L(0,d,0)	2.29	6.72	14.93	37.17	69.04	103.38
5% 顯著水準臨界值	5.99	11.07	18.31	31.41	55.76	101.90

說明: Ljung-Box 檢定之虛無假設為序列不具自我相關; 統計量為

$$LB(m) = T(T+2) \sum_{j=1}^m \frac{1}{T-j} \rho_j^2 \sim \chi^2(m).$$

表 4.6: 殘差項平方值的自我相關檢定

	$LB(2)$	$LB(5)$	$LB(10)$	$LB(20)$	$LB(40)$	$LB(80)$
SW(3)-FIGARCH-L(0,d,0)	5.21	7.41	11.06	25.73	53.35	96.23
SW(3)-GARCH-L(1,0)	19.04	20.53	23.31	42.60	71.87	113.12
AR(1)-FIGARCH-L(0,d,0)	8.91	10.03	12.31	37.28	60.29	104.58
5% 顯著水準臨界值	5.99	11.07	18.31	31.41	55.76	101.90

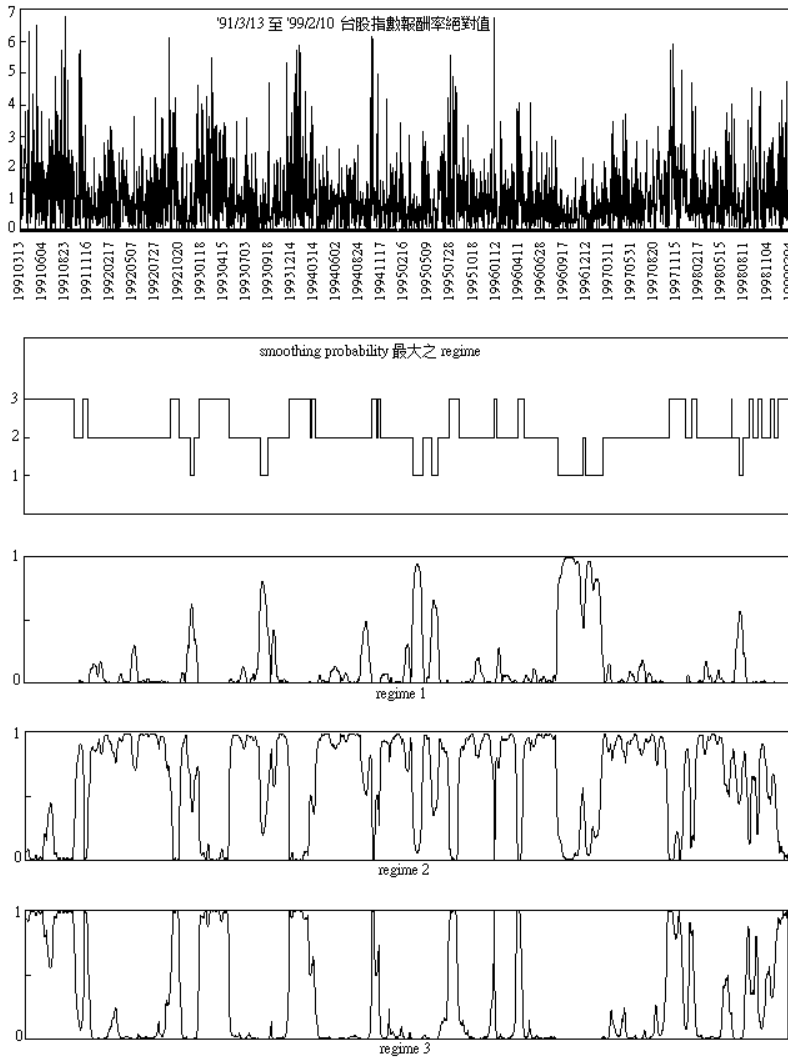


圖 2: SW(3)-GARCH-L(1,0) 之 smoothing probability

ARFIMA(0,d,0) 估計其絕對值、平方值資料，部分差分係數值介於 0.08 至 0.3 之間，且都顯著地大於 0；本節 AR(1)-FIGARCH-L(0,d,0) 的估計結果亦支持緩長記憶假說。種種估計、檢定結果都顯示，未考慮狀態變換之前，1991/3/13-

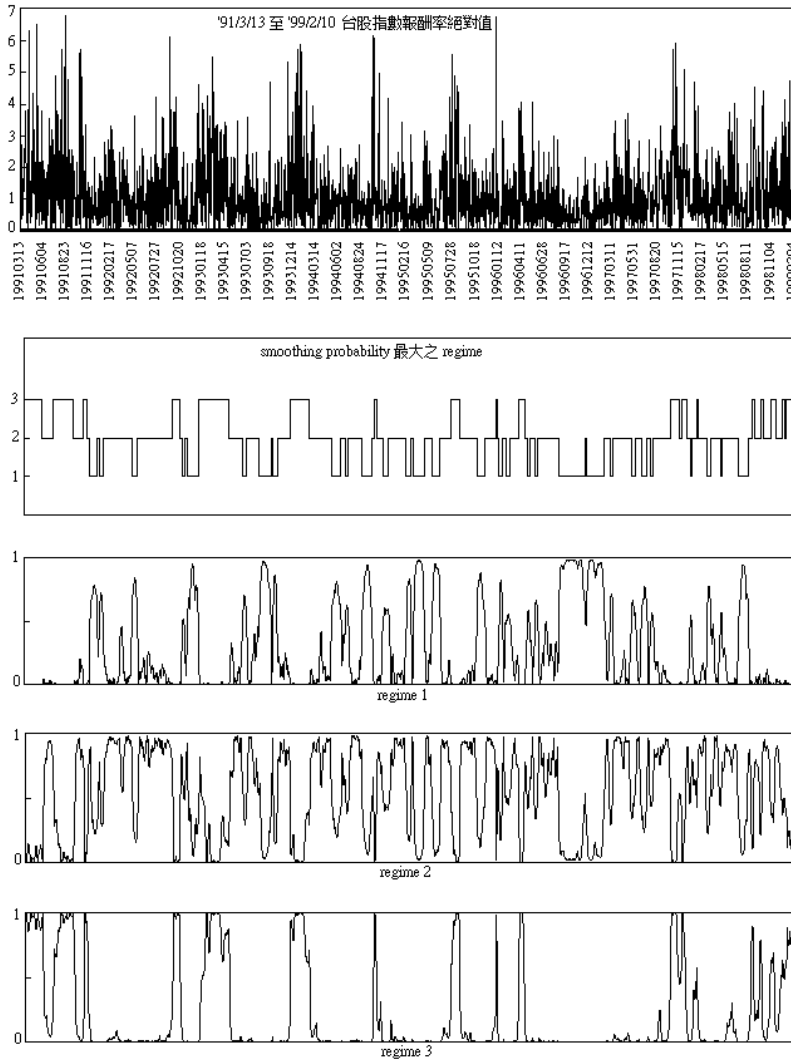


圖 3: SW(3)-FIGARCH-L(0,-0.1469,0) 之 smoothing probability

1999/2/10 期間台股指數報酬率的波動程度的確表現出緩長記憶特徵。然而，我們在之前的討論中曾經質疑台股指數報酬率的緩長記憶特徵並不穩定；當我們實際將狀態變換的因素納入，以本文所建立的 SW(3)-FIGARCH-L(0,d,0)

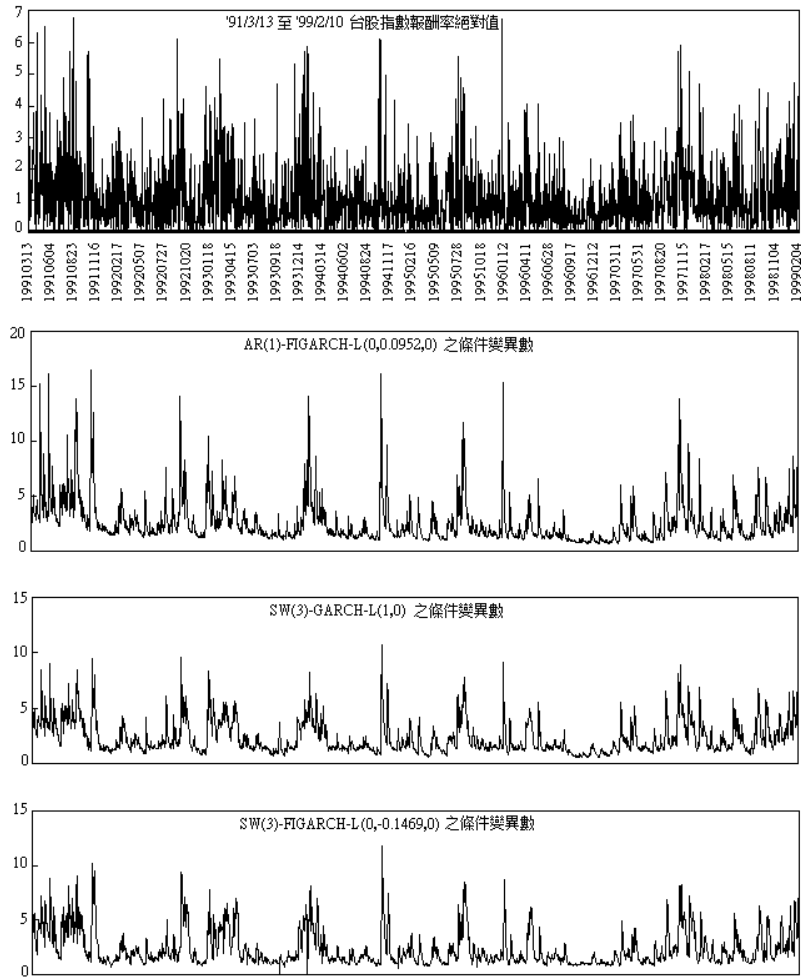


圖 4: 各模型條件變異數的比較

進行估計時，則發現緩長記憶特徵的確已不復存在。換言之，台股指數報酬率波動程度的緩長記憶性質很可能是狀態變換所造成的假性現象。我們將在下一章進一步證明狀態變換的確會造成假性緩長記憶，並指出台股指數報酬率絕對值 ACF 的型態及走勢都可從狀態變換的角度加以解釋。

比較表 4.3 中 SW(3)-FIGARCH-L(0,d,0) 及 SW(3)-GARCH-L(0,d,0) 的

狀態轉變機率矩陣可發現,前者 p_{11} 、 p_{22} 、 p_{33} 都較小,狀態持續性較低;再從圖 2、圖 3 的事後機率來看,SW(3)-FIGARCH-L(0,d,0) 中各狀態事後機率的持續性也相對較低。這可能是表示 SW(3)-GARCH-L(0,d,0) 所捕捉到的狀態變換特性中,有一部分在 SW(3)-FIGARCH-L(0,d,0) 中被改由部分差分係數來解釋,不再被認定為狀態變換。

根據表 4.4,所有指標一致顯示:可同時捕捉條件變異數狀態變換及長(中)短期行為的 SW(3)-FIGARCH-L(0,d,0) 對於 1980/1/4–1999/2/10 期間的台股指數報酬率具有最佳的配適能力及樣本內隔日預測 (one-period-ahead prediction) 能力;表現次之的是捕捉條件變異數狀態變換及短期行為的 SW(3)-GARCH-L(1,0);最後才是僅捕捉長(中)短期行為的 AR(1)-FIGARCH-L(0,d,0)。這似乎也告訴我們:關於台股指數報酬率的研究,中長期行為的捕捉的確有助於模型的配適,而狀態變換現象的刻劃則具有不可忽視的必要性。

此外我們要特別提出的是,一般以馬可夫變換模型估計總體或財務資料時,常會面臨一個窘境:該模型的主要解釋對象係序列干擾項平方值的動態性質,但殘差項的平方值,也就是模型未解釋部分的平方值,卻仍具高度的跨期相關;模型係為捕捉二階動差而設計,但估計後未解釋部分的二階動差卻仍十分不乾淨。我們認為這樣的窘境是未同時捕捉二階動差的中長期行為所造成的。首先,此外我們在第 1 節中曾特別指出,SWARCH (或 SWARCH-L) 模型及 SW-GARCH (或 SW-GARCH-L) 模型中 ARMA 形式的條件變異數設定所能捕捉的僅限於干擾項平方值的短期動態行為,倘若干擾項平方值具有顯著的「中長期性質」估計結果的殘差項平方值很可能會反映未被捕捉的中長期性質而呈現出高度跨期相關。我們可藉由概似比值檢定 (likelihood-ratio test) 佐證在狀態變換架構下捕捉中長期行為,也就是加入部分整合設定的必要性。SW(3)-FIGARCH-L(0,d,0) 及 SW(3)-GARCH-L(1,0) 的對數概似函數值分別為 -3926.6 及 -3940.8 ,概似比值統計值為 $2[-3926.6 - (-3940.4)] = 27.6$,大於 5% 顯著水準臨界值 3.841,換言之,在狀態變換架構下,部分差分係數十分顯著,加入部分整合設定以捕捉干擾項平方值的中長期行為的確有其必要性。因此,一旦未以部分整合設定捕捉干擾項平方值的中長期行為,就很可能會導致殘差項平方值仍具高度跨期相關。如表 4.6 所示,只捕捉短期行為的 SW(3)-GARCH-L(1,0) 具有高度殘差項平方值跨期相關,但同時捕捉中長期行為後,SW(3)-FIGARCH-L(0,d,0) 的殘差項平方值則完全通過了各種落後

表 5.1: 以非緩長序列資料進行 LM 檢定卻拒絕虛無假設的比例

顯著水準=	SW(3)-GARCH -L(1,0)			AR(1)-GARCH -L(1,0)		
	10%	5%	1%	10%	5%	1%
$ y_t $ 拒絕比例	0.8658	0.7938	0.6418	0.0602	0.0264	0.0056
y_t^2 之拒絕比例	0.7972	0.7116	0.5240	0.0558	0.0280	0.0068
同時拒絕比例	0.7838	0.6968	0.5092	0.0362	0.0162	0.0032

說明: 我們以上一節中實證估計得到的 SW(3)-GARCH-L(1,0) 模型及令 SW(3)-GARCH-L(1,0) 模型中的而得到的 AR(1)-GARCH-L(1,0) 模型作為 DGP, 各模擬出 5000 筆樣本數為 1,000 的資料, 取絕對值及平方值後進行 $m = 40$ 的 LM 檢定; 拒絕虛無假設表示有假性緩長記憶現象發生。

期數的 Ljung-Box 檢定。換言之, 我們所建立的 SW(3)-FIGARCH-L(0,d,0) 模型不但可避免以單純緩長記憶模型估計狀態變換母體所導致的假性緩長記憶問題, 更解決了未考慮顯著中長期性質之下馬可夫變換模型殘差項平方值高度跨期相關的窘境, 是一個相對優良的模型, 在財務資料的估計上十分具有發展潛力。

5 狀態變換與假性緩長記憶

為了證明條件變異數的狀態變換的確會造成假性緩長記憶現象, 我們以上一節中估計得到的 SW(3)-GARCH-L(1,0) 模型及令 SW(3)-GARCH-L(1,0) 模型中的而得到的 AR(1)-GARCH-L(1,0) 模型作為 DGP, 各模擬出 5,000 筆樣本數為 1,000 的資料, 取絕對值及平方值後進行 $m = 40$ 的 LM 檢定, 模擬結果列於表 5.1。根據模擬結果, 未考慮狀態變換時, AR(1)-GARCH-L(1,0) 資料的拒絕比例還在顯著水準之內, 但加入狀態變換後, SW(3)-GARCH-L(1,0) 資料的拒絕比例相當高, 顯示假性緩長記憶相當嚴重。換言之, 條件變異數的狀態變換的確會造成假性緩長記憶現象。

此外, 從表 4.1 中可發現, 1991/3/13-1999/2/10 期間, 報酬率絕對值的緩長記憶特徵比平方值顯著; 而表 5.1 的 SW(3)-GARCH-L(1,0) 資料雖是以「非緩長記憶模型」之估計結果為 DGP 而模擬出來的, 卻也同樣顯示出絕對值轉換之拒絕比例較平方值轉換更高的現象。換言之, 我們的「非緩長記憶模型」的確成功地捕捉到這段期間報酬率絕對值較平方值更具「緩長記憶」的現象。

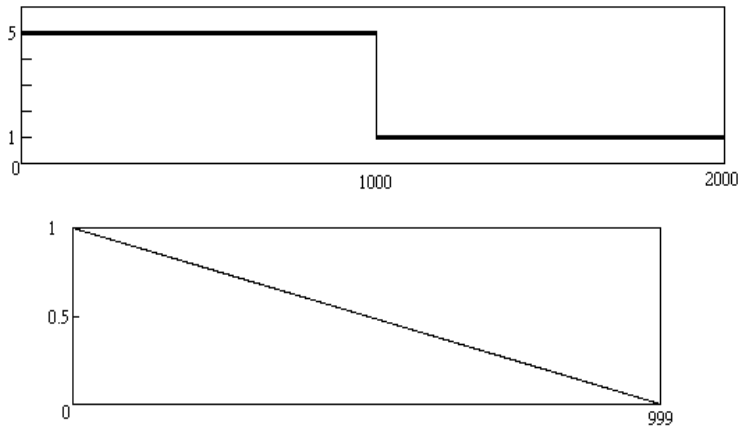
Ding, Granger, and Engle (1993) 的研究發現, 非但股價指數報酬率的絕對值資料具有緩長記憶性質, 將該絕對值資料進一步取各種大於 0 的整數或非整數次方後, 緩長記憶性質仍然存在。我們曾以 Ding, Granger, and Engle (1993) 所使用的 S&P 500 指數日報酬率以及本文所使用的台股指數日報酬率進行驗證, 結果顯示, 不但取各種大於 0 的整數或非整數次方後緩長記憶特徵仍都存在, 絕對值資料經自然對數 \ln 或指數 \exp 轉換後, 緩長記憶特徵亦仍存在。如果股價指數報酬率的緩長記憶特徵的確是狀態變換造成的結果, 我們必須從狀態變換的角度對上述現象提出適當的解釋。

倘若一序列在很長的一段期間之內一直維持在一定平均數水準附近, 而後平均數水準忽然改變, 之後序列也一直維持在新的平均數水準附近, 其 ACF 即會反映其平均數水準的狀態持續性而呈現緩慢下降的形態。為簡化說明, 如圖 5, 我們以 5 與 1 組成兩筆資料, 並分別計算其 ACF。根據計算結果, 資料數字一次永久性的改變會形成緩慢持續下降的 ACF, 資料在兩個數字間變換則會形成上下起伏的 ACF。在圖 6 中, 我們進一步模仿 1987/3/20–1991/3/12 及 1991/3/13–1999/2/10 兩段期間台股指數報酬率絕對值資料大致的平均數水準變動, 分別以 2 與 1、3 與 1 組成兩筆資料, 並分別計算其 ACF。結果顯示, 簡單數字變動資料的 ACF 與實際資料的 ACF 在形態及走勢上都有相當程度的一致性。換言之, 平均數水準狀態變換母體所產生的樣本 ACF 與緩長記憶母體的樣本 ACF 在形態及走勢上並不易分辨; 樣本 ACF 所表現出的長時距依存現象可能只是單純地反映著平均數水準的狀態持續性。

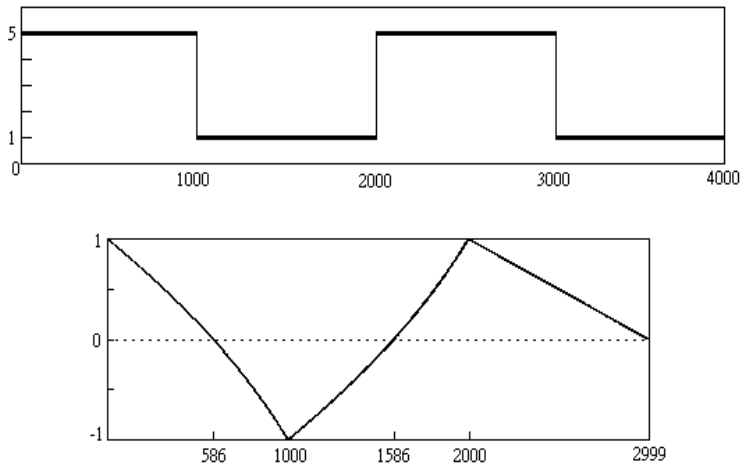
報酬率絕對值平均數的狀態變換現象事實上是報酬率波動程度的狀態變換所造成的結果: 令 $y_{1t} = g_1 \varepsilon_t$, $y_{2t} = g_2 \varepsilon_t$, $g_1 > g_2 > 0$ 其中 $\varepsilon_t = \sqrt{h_t} v_t$, v_t 為 $iid(0, 1)$ 序列, 則

$$E|y_{1t}| = g_1 E|\varepsilon_t| > g_2 E|\varepsilon_t| = E|y_{2t}|。$$

換言之, y_t 若在 y_{1t} 與 y_{2t} 間變換, 即條件變異數在 $g_1^2 h_t$ 與 $g_2^2 h_t$ 間變換, 則 $|y_t|$ 的平均數會在 $g_1 E|\varepsilon_t|$ 與 $g_2 E|\varepsilon_t|$ 間變換。事實上, 若條件變異數在 $g_1^2 h_t$ 與 $g_2^2 h_t$ 間變換, 將 $|y_t|$ 進一步作任意單調轉換 (monotonic transformation) 後, 轉換後序列的平均數也會隨而在不同水準間變換: 令 $M(|y_t|)$ 為 $|y_t|$ 的任意單調轉換, 則 $M(|y_{1t}|) > M(|y_{2t}|)$, 因此



以5與1組成一筆資料並計算其 ACF



以5與1組成一筆資料並計算其 ACF

圖 5: 報酬率資料波動程度之狀態變換反映在其絕對值資料上, 形成平均數水準的狀態變換。平均數水準一次永久性的改變會形成緩慢持續下降的 ACF。平均數水準在兩個狀態間變換則會形成緩慢持續上下起伏的 ACF。換言之, 報酬率絕對值資料的長時距依存現象, 可能只是單純地反映出其平均數水準的狀態持續性。

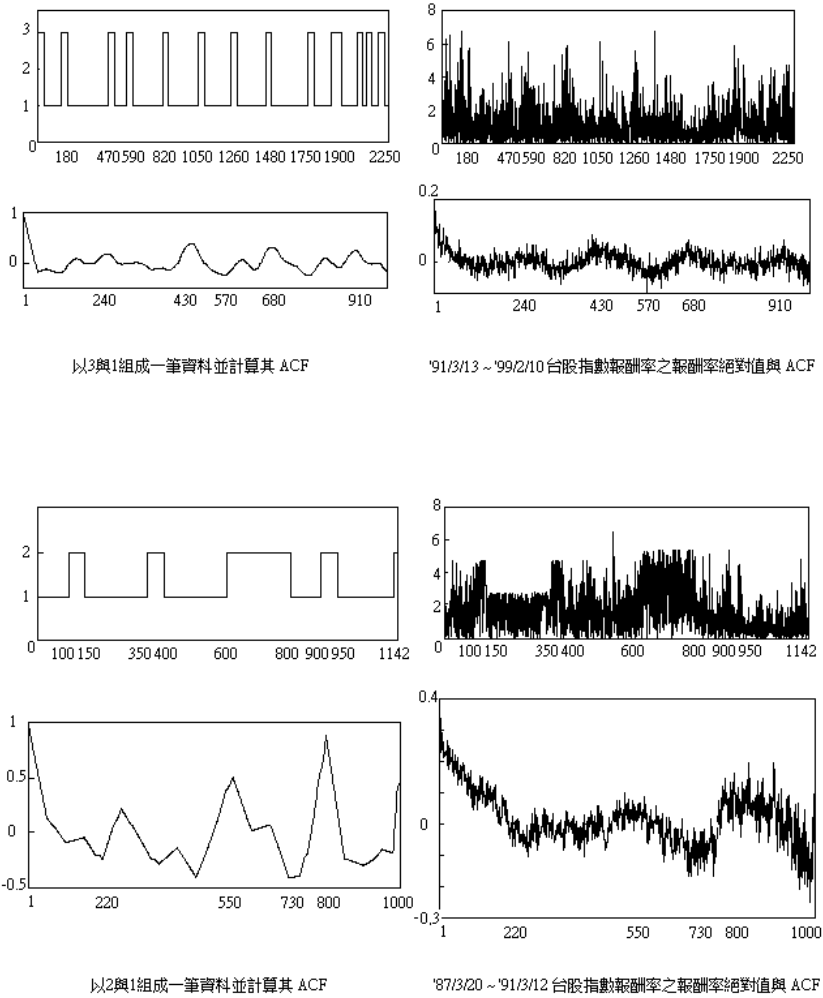


圖 6: 報酬率資料波動程度之狀態變換反映在其絕對值資料上, 形成平均數水準的狀態變換。1987/3/20 至 1991/3/12 與 1991/3/13 至 1999/2/10 兩段期間, 台股指數報酬率絕對值 ACF 的形態及走勢都能以平均數水準的狀態變換來說明。

$$\begin{aligned}
 E [M (|y_{1t}|)] &= \int_{-\infty}^{\infty} M (|y_{1t}|) f(\varepsilon_y) d\varepsilon_t > \int_{-\infty}^{\infty} M (|y_{2t}|) f(\varepsilon_y) d\varepsilon_t \\
 &= E [M (|y_{2t}|)]。
 \end{aligned}$$

綜上所述,倘若序列的條件變異數有狀態變換的現象,則該序列取絕對值或再進一步作次方、指數、對數等單調轉換後,轉換後的序列將會有平均數狀態變換的現象,因而形成疑似緩長記憶形態的 ACF。

為進一步證明台股指數報酬率絕對值資料的緩長記憶特徵的確可能來自於狀態變換,如圖 7,我們以兩種狀態變換模型作為 DGP 模擬 1991/3/13–1999/2/10 期間的台股指數報酬率絕對值資料,並計算 ACF,結果發現,模擬資料的 ACF 與真實資料的 ACF 十分神似,這顯示台股指數報酬率絕對值資料的緩長記憶特徵的確可用狀態變換現象來說明。結合第 4 節的估計、檢定結果,以及本節的模擬、檢定結果,我們懷疑台股指數報酬率絕對值資料表現出的緩長記憶特徵事實上是條件變異數的狀態變換所造成的假性緩長記憶現象;同時考慮狀態變換效果後,緩長記憶性質不復存在,僅具中等記憶。

6 結論

由於緩長記憶序列與狀態變換序列的樣本性質相當類似,以單純緩長記憶模型估計狀態變換母體很可能會得到假性的緩長記憶結果。真偽緩長記憶的辨別在經濟政策、財務決策、實證方法、及計量理論上都具有非常重要的意義,但辨別緩長記憶母體及狀態變換母體的計量方法卻付之闕如。一般以緩長記憶模型進行實證分析時,對此潛在問題,或未寄予相當重視,或僅主觀劃分結構改變點以為因應;一旦資料的確來自狀態變換母體,估計、檢定結果極可能仍會支持緩長記憶假說。有鑑於此,我們建立了 SW(k)-FIGARCH-L(0,d,0) 模型,同時估計條件變異數的狀態變換及長短期動態行為,使我們既可在狀態變換的前提下檢定緩長記憶假說,也能在考慮緩長記憶可能存在的前提下檢定是否有狀態變換發生;讓資料自行判定緩長記憶現象的真偽。

傳統條件變異數馬可夫變換模型所採用的是 ARMA 形式的條件變異數設定,僅能捕捉干擾項平方值的短期動態行為,倘若干擾項平方值具有顯著的中長期性質,估計結果的殘差項平方值很可能會反映未被捕捉的中長期性質而呈現出高度跨期相關;我們所建立的 SW(k)-FIGARCH-L(0,d,0) 模型修正此一缺失,引入部分整合設定,彌補了傳統模型在中長期行為刻劃上的不足。綜言之,SW(k)-FIGARCH-L(0,d,0) 模型不但可避免以單純緩長記憶模型估計狀態變換母體所導致的假性緩長記憶問題,更解決了未考慮顯著中長期性質之

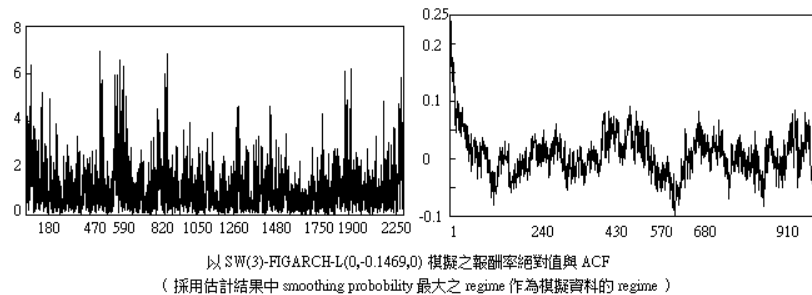
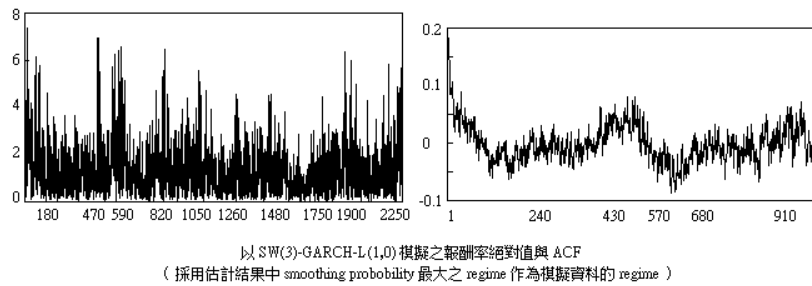
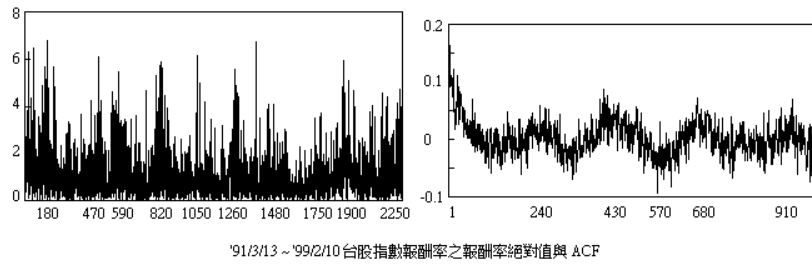


圖 7: 1991/3/13 至 1999/2/10 期間, 台股指數報酬率絕對值的 ACF 與狀態變換模型所模擬的 ACF 十分相似, 顯示台股指數報酬率絕對值的「緩長記憶現象」的確可能是條件變異數狀態變換所造成的假性緩長記憶現象。

下馬可夫變換模型未解釋部分平方值相當不乾淨的窘境, 是一個相對優良的模型, 在財務資料的估計上十分具有發展潛力。

根據實證結果, 未考慮狀態變換因素前, 1991/3/13–1999/2/10 期間, 台股指數日報酬率的波動程度的確顯示出緩長記憶特徵: 其絕對值及平方值資料一則完全拒絕 LM 檢定的非緩長記憶虛無假設, 另則皆在 ARFIMA(0,d,0) ARFIMA (1,d,0) 的初步估計中得到顯著大於 0 的部分差分係數; 直接以 AR(1)-FIGARCH-L(0,d,0) 估計該報酬率資料, 條件變異數方程式中的部分差分係數亦顯著大於 0。然而, 當我們改用 SW(3)-FIGARCH-L(0,d,0) 進行估計, 卻發現部分差分係數估計值十分顯著地小於 0。這似乎表示, 同時考慮狀態變換因素後, 資料波動程度的緩長記憶特徵不復存在, 僅具中等記憶。

我們以舉例論證方式指出, 平均數水準狀態變換母體所產生的樣本 ACF 與緩長記憶母體的樣本 ACF 在形態及走勢上並不易分辨, 樣本 ACF 所表現出的長時距依存現象可能只是單純地反映著平均數水準的狀態持續性。倘若序列的條件變異數有狀態變換的現象, 則該序列取絕對值或再進一步作次方、指數、對數等單調轉換後, 轉換後的序列將會有平均數狀態變換的現象, 因而形成疑似緩長記憶形態的 ACF。

為求明確揭示狀態變換假性緩長記憶問題的嚴重性, 我們以 SW(3)-GARCH-L(1,0) 作為 DGP 模擬出 5,000 筆樣本數為 1,000 的非緩長記憶資料, 並以其無狀態變換下的簡化模型 AR(1)-GARCH-L(1,0) 作為對照組, 逐一檢定緩長記憶特性。結果顯示, 未考慮狀態變換時, 資料的拒絕比例還在顯著水準之內, 但加入狀態變換後, 拒絕比例卻變得相當高; 狀態變換的確會造成相當嚴重的假性緩長記憶問題, 值得我們重視。

關於緩長記憶與狀態變換, 仍有部分值得深究的問題未及在本文探討。因為既然緩長記憶母體與狀態變換母體的樣本性質極為相近, 單獨捕捉其一勢必都會面臨「假性風險」; 以緩長記憶模型捕捉狀態變換母體會蒙受假性緩長記憶問題, 以狀態變換模型捕捉緩長記憶母體當然也同樣會蒙受假性狀態變換問題。若要進一步探討假性狀態變換問題, 設計一套穩健而計算簡便的狀態變換檢定方法便有其必要性。我們建立 SW(k)-FIGARCH-L(0,d,0) 模型的初衷不只是希望讓資料自行判定緩長記憶現象的真偽, 更希望能同時判定狀態變換現象的偽真。檢定 SW(k)-FIGARCH-L(0,d,0) 模型的部分差分係數相當於在狀態變換的架構下檢定緩長記憶的真偽; 檢定 SW(k)-FIGARCH-L(0,d,0) 模型的狀態轉變參數則相當於在緩長記憶的架構下檢定狀態變換的真偽。本文已在狀態變換的架構下檢定了緩長記憶的真偽, 然而在檢定狀態變換的真

偽方面,由於狀態轉變參數在單一狀態假設下無法被認定 (not identified),傳統的概似比值檢定法並不適用;目前只有 Hansen (1992) 及 Hamilton (1996) 兩種檢定法可以使用,卻都需要進行高度繁複耗時的電腦運算步驟。馬可夫模型狀態轉變參數顯著性檢定方法的改進,是我們的分析能否更加完備的主要關鍵。倘若我們真的解決了狀態轉變參數顯著性的檢定問題,則有兩件事情是我們必須儘速確認的:以 SW(k)-FIGARCH-L(0,d,0) 模型估計緩長記憶資料時,是否部分差分係數顯著大於0且同時狀態轉變參數不顯著;以 SW(k)-FIGARCH-L(0,d,0) 模型估計狀態變換資料時,是否狀態轉變參數顯著,且同時部分差分係數顯著小於0或不顯著。這值得以後深入的探討。

模型估計細節也仍有調整或改進的空間,例如:計算 h_t 時,不全面使用替代變數,讓前有限個落後項對應不同狀態組合而具數種組合可能性;改採 Hamilton (1994) 所建議的遍態歷境機率作為 $\Xi_{1|0}$ 的給定值;考慮報酬率分配的厚尾現象,我們可調整分配假設,改採 Student t 分配;放寬 $E(D_t) = 1/2$ 的假設,將 $E(D_t)$ 納入估計;或直接針對 ε_t^2 設計替代變數:

$$\tilde{\varepsilon}_t = \sum_{j=1}^k \left[p_r(s_t = j|Y^t) \left(\frac{y_t - (1-a)\mu - ay_{t-1}}{\sqrt{g_j}} \right) \right].$$

最後本文也可延伸出數個研究方向:如藉由條件變異數及事後機率,分析各財經或政治事件對報酬率波動程度的影響,並比較各事件影響力大小、影響力持續性、及影響形式的差異;比較各模型的樣本外預測 (forecast) 能力;探討條件變異數馬可夫變換架構下的財務避險策略及金融商品定價;同時考慮狀態變換因素,重新估計曾被論證或檢定具緩長記憶的總體或金融變數。

參考文獻

- 高櫻芬·呂仁廣·林建甫 (2001), “變異數結構的 SWARCH 模型估計: 台灣股票市場之模型研究”, 《證券市場發展季刊》, 49, 63-97.
- Anderson, M.K. (2000), “Do long-memory models have long memory?”, National Institute of Economic Research papers, number 25.
- Backus, D.K. and S.E. Zin (1993), “Long-memory inflation uncertainty, evidence from the term structure of interest rates”, *Journal of Money, Credit, and Banking*, 25, 681-700.

- Baillie, R.T. (1998), "Comment", *Journal of Business and Economic Statistics*, 16, 273–276.
- Baillie, R.T., T. Bollerslev and H.O. Mikkelsen (1996), "Fractionally integrated generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics*, 74, 3–30.
- Baillie, R.T., C.F. Chung and M.A. Tieslau (1996), "Analyzing inflation by the fractionally integrated ARFIMA-GARCH model", *Journal of Applied Econometrics*, 11, 23–40.
- Bollerslev, T. (1986), "Generalized autoregressive conditional heteroscedasticity", *Journal of Econometrics*, 31, 307–327.
- Bollerslev, T. and H.O. Mikkelsen (1996), "Modeling and pricing long memory in stock market volatility", *Journal of Econometrics*, 73, 151–184.
- Bollerslev, T. and J.M. Wooldridge (1992), "Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances", *Econometric Reviews*, 11, 143–172.
- Bos, C.S., P.H. Franses and M. Ooms (1998), "Long memory and level shifts: re-analyzing inflation rates", *Empirical Economics*, 24(3), 427–449.
- Cheung, Y.W. (1993a), "Long memory in foreign-exchange rates", *Journal of Business and Economic Statistics*, 11, 93–101.
- (1993b), "Tests for fractional integration: a Monte Carlo investigation", *Journal of Time Series Analysis*, 14, 331–345.
- de Lima, P., F.J. Breidt and N. Crato (1998), "The detection and estimation of long memory in stochastic volatility", *Journal of Econometrics*, 83(1–2), 325–348.
- Diebold, F.X. (1986), "Modelling the persistence of conditional variance: a comment", *Econometric Review*, 5, 51–56.
- Diebold, F.X. and G.D. Rudebusch (1989), "Long memory and persistence in aggregate output", *Journal of Monetary Economics*, 24, 189–209.
- Ding, Z., C.W.J. Granger and R.F. Engle (1993), "A long memory property of stock returns and a new model", *Journal of Empirical Finance*, 1, 83–106.
- Dueker, M.J. (1997), "Markov switching in GARCH processes and mean-reverting stock-market volatility", *Journal of Business and Economic Statistics*, 15, 26–34.
- Engle, R.F. (1982), "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation", *Econometrica*, 50, 987–1007.
- Engle, R.F. and C. Mustafa (1992), "Implied ARCH models from options prices",

- Journal of Econometrics*, 52, 280–311.
- Geweke, J.F. and S. Porter-Hudak (1983), “The estimation and application of long memory time series models”, *Journal of Time Series Analysis*, 1, 221–238.
- Granger, C.W.J. (1980), “Long memory relationships and the aggregation of dynamic models”, *Journal of Econometrics*, 14, 227–238.
- Granger, C.W.J. and R. Joyeux (1980), “An introduction to long-memory time series models and fractional differencing”, *Journal of Time Series Analysis*, 1, 15–29.
- Granger, C.W.J. and N. Hyung, (1999), “Occasional structural breaks and long memory”, UCSD Economics Discussion Papers with number, 99–14.
- Gray, S.F. (1996), “Modelling the conditional distribution of interest rates as a regime-switching process”, *Journal of Financial Economics*, 42, 27–62.
- Greene, M.T. and B.D. Fielitz (1977), “Long-term dependence in common stock returns”, *Journal of Financial Economics*, 4, 339–349.
- Hamilton, J.D. (1994), *Time Series Analysis*, New Jersey: Princeton University Press.
- (1996), “Specification testing in Markov-switching time-series models”, *Journal of Econometrics*, 70, 127–157.
- Hamilton, J.D. and R. Susmel (1994), “Autoregressive conditional heteroskedasticity and change in regime”, *Journal of Econometrics*, 64, 307–333.
- Harvey, A.C. (1993), “Long-memory in stochastic volatility”, manuscript, London School of Economics: London.
- Hansen, B.E. (1992), “The likelihood ratio test under nonstandard conditions: testing the Markov switching model of GNP”, *Journal of Applied Econometrics*, 7, 61–82.
- Hosking, J.R.M. (1981), “Fractional differencing”, *Biometrika*, 68, 165–176.
- (1984), “Modeling persistence in hydrological time series using fractional differencing”, *Water Resource Research*, 20, 1898–1908.
- Hurst, H.E. (1951), “Long term storage capacity of reservoirs”, *Transactions of American Society of Civil Engineering*, 116, 770–799.
- Lamoureux, C.G. and W.D. Lastrapes (1990), “Persistence in variance, structural change and the GARCH model”, *Journal of Business and Economic Statistics*, 8, 225–234.
- Lin, J.L. (2000), “Switching ARCH models of stock market volatility in Taiwan”, *Advances in Pacific Basin Business, Economics, and Finance*, 4, 1–21.

- Lo, A.W. (1991), "Long term memory in stock market prices", *Econometrica*, 59, 1279–1313.
- Lobato, I.N. and P.M. Robinson (1997), "A nonparametric test for $I(0)$, preprint", Department of Economics, University of Iowa.
- Lobato, I.N. and N.E. Savin (1998), "Real and spurious long-memory properties of stock-market data", *Journal of Business and Economic Statistics*, 16, 261–268.
- Lobato, I.N. and C. Velasco (2000), "Long memory in stock-market trading volume", *Journal of Business and Economic Statistics*, 18, 410–427.
- Nelson, C.R. and C.I. Plosser (1982), "Trends and random walks in macroeconomic time series: some evidence and implications", *Journal of Monetary Economics*, 10, 139–162.
- Perron, P. (1989), "The great crash, the oil price shock and the unit root hypothesis", *Econometrica*, 57, 1361–1401.
- Robinson, P.M. (1994), "Time series with strong dependence", in C.A. Sims ed., *Advances in Econometrics: Sixth World Congress*, 1, 47–95, Cambridge: Cambridge University Press.
- Sadique, S. and P. Silvapulle (2001), "Long-term memory in stock market returns: international evidence", *International Journal of Finance & Economics*, 6, 59–67.
- So, M. (2000), "Long-term memory in stock market volatility", *Applied Financial Economics*, 10, 519–524.
- Sowell, F. (1992), "Modeling long-run behavior with the fractional ARIMA model", *Journal of Monetary Economics*, 29, 277–302.
- Tsay, W.J. (2000), "The long memory story of the real interest rate", *Economics Letters*, 67, 325–330.
- Tsay, W.J. and C.F. Chung (2000), "The spurious regression of fractionally integrated processes", *Journal of Econometrics*, 96, 155–182.
- Zivot, E. and D.W.K. Andrews (1992), "Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis", *Journal of Business and Economic Statistics*, 10, 251–270.

投稿日期: 1999年11月4日, 接受日期: 2003年1月24日

Long Memory and Regime Switch

Jiun-Kuei Hsieh

Department and Graduate Institute of Finance, National Taiwan University

Chien-fu Jeff Lin

Department of Economics, National Taiwan University

The strong persistence in volatility of a variety of financial time series is well-known. To determine whether this persistence can be characterized as “long memory” is important in both financial and econometric modelling. In this paper, after summarizing Hamilton and Susmel’s (1994) regime switching SWARCH model and BBM’s (1996) long memory FIGARCH for volatility, we propose a general SW(k)-FIGARCH-L(0,d,0) model that allows the estimation of both regime switching parameters and the long memory parameter. In such a framework we are able to test whether the volatility still has long memory after regime switching has been considered and we show that regime switching in volatility can result in spurious long memory. Furthermore, we find the proposed model also solves one of the problems with the standard SWARCH model that squared residuals obtained from the SWARCH model estimation usually are highly correlated, which implies the simple regime switching mechanism is not able to characterize all the dynamics in volatility. Based on these encouraging results, we believe the proposed model is a promising tool in analyzing financial data. Our empirical analysis of the TAIEX data shows that the long memory in their volatility will reduce to intermediate memory after regime switching is considered, which represents an interesting example of spurious long memory in volatility that is caused by the regime switching.

Keywords: FIGARCH model, markov-switching model, SWARCH model,
SW(k)-FIGARCH-L model, long memory

JEL classification: C52