

台灣股市波動結構改變之分析：平滑轉換迴歸的 LM 檢定與 Markov Switching GARCH 模型的應用

呂仁廣*

摘 要

本文以 AR(1)-GARCH(1,1) 模型，來檢定和估計台灣股市的波動情況，所使用的資料為西元 2000 年 1 月 4 日至 2011 年 10 月 25 日台灣加權股價指數。實證方法分成兩個步驟，第一步驟是先檢定台灣股市波動之結構是否發生改變，我們採取的是林建甫與張焯然(1997)以及高櫻芬、呂仁廣與林建甫(2001)的方法，也就是使用 Lin and Terasvirta (1994)的平滑轉換迴歸架構中的 LM 檢定，透過 Logistic 分配函數，來偵測 AR(1)-GARCH(1,1) 模型中的條件變異數方程式是否發生結構改變的情況，結果發現常態分配下的 GARCH 模型在統計檢定上確實發生結構改變，不論是單調、周而復始、非單調的結構改變均有可能。第二步驟是針對台灣股市波動結構改變時點的估計，我們使用的 Markov Switching GARCH (1,1)模型確實對台股資料配適得比較好；最後，我們利用平滑機率值找出台灣股市高波動狀態所發生的時段，並得知這些時段也代表當時的台灣股市處於高風險的時期。

關鍵詞：結構改變檢定、Markov Switching、GARCH

*華夏技術學院 資管系助理教授（通訊作者）

電子郵件：LueRenGuang@gmail.com

收稿日期：2011.11.15

修改日期：2012.04.20

接受日期：2012.06.19

Structural Change Analysis of Taiwan's Stock Market Volatility with LM Test of Smooth Transition Regression and Markov Switching GARCH Model

Ren-Guang Lue*

Abstract

In this paper, we use AR(1)-GARCH(1,1) model to test and estimate Taiwan's Stock Market Volatility with the TSE sample period from January 4, 2000 to October 25, 2011. Empirical method is divided into two steps. First, we apply the method of Lin and Zhang (1997) and Gau, Lue, and Lin (2001), which is LM test of STR in Lin and Terasvirta (1994) through Logistic function, to test whether structural change happens in Taiwan's Stock Market Volatility. From our empirical results, we find GARCH model do happen to change structurally, in all possible forms. Second, we try to detect the change-point of volatility structural change with the Markov Switching GARCH (1,1) model, and find that it do fit better than GARCH(1,1). Finally, we use the smoothing probabilities to get the periods of high-volatility of Taiwan's stock market, during which periods are known to represent high-risk period of Taiwan's stock market.

Key words: structural change test 、 Markov Switching 、 GARCH

* Assistant Professor, Department of Information Management, Hwa Hsia Institute of Technology. (correspondence author)

壹、前言

財務理論的開宗明義說到，若市場具有資訊效率，則一旦市場達到均衡狀態時，就無套利空間。此時市場均衡價格就會產生，這也是資產定價理論的基礎。然而如果市場在未達到均衡狀態之前，市場將會存在低風險但可以套利的機會，這也就是財務市場上常常發生投資人群聚(Herding)或一窩巢追漲、追跌的行為現象，以致形成財務資料獨有的報酬波動聚集(volatility clustering)之出象。計量學界也由原本只關注時間序列模型的殘差項是否為白噪音，殘差項只有常數下的非條件變異數情況，隨著人們愈趨於關注風險間的跨時相關，計量學界也演變成由 Engle (1982) 和 Bollerslev (1986) 領軍，發現殘差項也可以被允許有隨時間變動的條件變異數的設定，讓我們得以運用他們所發展出來的 GARCH 模型於金融市場上波動度(volatility)的估計與預測，如果將此代表波動度的標準差數值應用在衍生性金融商品的評價上，例如選擇權，通貨交換，換匯換利等，就很可能讓投資人得以有進一步的套利的空間。然而，Lamoureux and Lastrapes (1990) 在用 GARCH 模型分析美國股票報酬率資料時，發現其實很多情況下，條件變異數會發生高持續性的現象，即所謂的自積 GARCH 過程(Integrated-GARCH process)，而且學界發現這個自積 GARCH 過程的產生很有可能是因為條件變異數方程式的參數發生結構性改變。如果文章中的實證分析沒有考慮到 GARCH 模型結構可能發生改變的現象，則對於波動度的估計和預測就會發生錯誤。Chu (1995) 進一步利用記分的弱收斂性導出其文章中 GARCH 結構改變檢定統計量 LM 的漸近分配。林建甫與張焯然(1997) (以下簡稱林-張(1997))則是使用 Lin and Terasvirta (1994)的平滑轉換迴歸(smooth transition regression)架構中的 LM 檢定來確認條件變異數方程式是否發生結構改變。

但是計量上若是只有關心 GARCH 模型的結構是否改變還不夠，還必須知道這些改變發生的時點(change-point)和發生機率，才能進一步推算條件標準差的估計值，所以 Kokoszka and Leipus (1999)在結構改變檢定工作做完後，又緊接著在他們的另外一篇文章 Kokoszka and Leipus (2000)當中估計 GARCH 模型結構改變發生的時點。高櫻芬、呂仁廣與林建甫(2001) (以下簡稱高-呂-林)則是將上述兩種工作放在同一篇文章一起做，而他們所用來估計 GARCH 結構改變發生時點的模型則是 Hamilton and Susmel (1994)發展出來的 SWARCH 模型，此方法是將馬可夫鏈(Markov Chain)與 ARCH 模型做結合的估計方式。Cai (1994)也利用類似 SWARCH 的模型來估計美國政府公債殖利率的波動程度。最後完成將馬可夫鏈與 GARCH 模型做結合的則是 Gray (1996)以及 Haas, Mittnik, and Paoletta (2004)等人。¹

本文將跟隨高櫻芬、呂仁廣與林建甫(2001)的做法，先進行台灣股市 GARCH 模型結構改變的檢定，再利用 Markov Switching GARCH 模型做 GARCH 模型結構改變時點的估計，以及它的發生機率。因為只有結合了這兩種步驟，實證分析才不致於因為一下子跳入非線性程度複雜的 Markov Switching GARCH 模型中，缺乏先驗統計推論上對於 GARCH 結構改變的檢定，引起資料探勘(data mining)的問題。

我們文章的安排如下，第貳節將說明如何利用平滑轉換迴歸架構中的 LM 檢定，來偵測條件變異數方程式是否發生連續型的結構改變；第參節將闡述 Markov Switching GARCH 模型的估計原理，以及那些運用馬可夫鏈工具的相關估計方法之間的不同；第肆節是把上述兩種工作，應用到台灣股市的實證分析，包括

¹ 對於上述文章所用方法在細節上的比較和說明，我們會放在第參節中。

波動是否發生結構改變，及結構改變時點的估計，第伍節是本文的結論。

貳、股市波動之結構是否發生改變的檢定

股市波動結構是否發生改變，我們採取的是林-張(1997)以及高-呂-林(2001)的方法，也就是使用 Lin and Terasvirta (1994)的平滑轉換迴歸架構中的 LM 檢定，利用 Logistic 分配函數，來偵測條件變異數方程式是否發生連續型的結構改變。²令描述台股報酬率的條件平均數方程式為

$$R_t = \mu_t + \varepsilon_t, \quad (1)$$

R_t 是台股報酬率， μ_t 是模型描述 R_t 平均數的部分，³ ε_t 為模型的干擾項，我們假設干擾項 ε_t 為服從平均數是零、變異數是 h_t 的常態分配，而 h_t 可以用一般化自我迴歸條件異質(general autoregressive heteroscedasticity, GARCH)模型來描述。由於本文變數符號眾多，我們將運用 Bollerslev (1986)和林-張(1997)的符號定義，簡潔說明這個檢定方法的原理。以下是我們針對台股報酬率的條件方程式設定，即給定(1)式干擾項的條件分配為

$$\varepsilon_t | \Psi_{t-1} \sim N(0, h_t), \text{ or } \varepsilon_t | \Psi_{t-1} \sim t(m),$$

即 h_t 為 ε_t 的條件變異數，並且它的隨機過程可以寫成

² 由於 Logistic 分配函數為連續函數，所以此結構改變檢定可偵測到更細微的連續型結構改變，當然它也包括任何離散改變的情況，是故它是一個適合本文研究目的的檢定方法。

³ 一般而言， $\mu_t \sim ARMA(p, q)$ 過程，為簡化分析，在第肆節的實證部分，我們依照大部分的文獻對於報酬率的處理假設 $\mu_t \sim AR(1)$ 。

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} = z_t' \omega. \quad (2)$$

其中 $\omega = (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$ 為參數向量，而 $z_t = (1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2, h_{t-1}, \dots, h_{t-p})'$ 為變數向量， Ψ_{t-1} 為 $t-1$ 的期訊息集合， m 為 t 分配的自由度。如果(2)式的 GARCH 過程發生結構改變，我們可以看出參數向量 ω 發生位置上的結構改變。根據平滑轉換迴歸原理，我們可以將條件變異數 h_t 的方程式寫成更為一般化的

$$h_t = z_t' \omega_1 + z_t' \omega_2 F(s_t), \quad (3)$$

這時，結構改變是否發生的關鍵就在於 ω_2 是否為零。第(3)式中的 s_t 為平滑轉換迴歸運用在變數第二階動差下的轉換變數(transition variable)，在這裡可以視其為控制迴歸轉換機制下的狀態變數(state variable)。函數 F 在本文中被設定成 s_t 的 logistic 分配函數⁴

$$F(s_t) = \frac{1}{1 + \exp[-s_t]}, \quad (4)$$

而狀態變數 s_t 雖然是不可觀察，但在 Lin and Terasvirta (1994)中可以被設計成時間變數 t 的高階多項式

$$s_t = s(t) = r(t^k + c_1 t^{k-1} + \dots + c_{k-1} t + c_k) = r \mathbf{K}_t' \mathbf{c}, \quad (5)$$

其中向量 $\mathbf{K}_t = (1, t, \dots, t^k)'$ 為時間變數 t 的高階多項式變數的向量， $\mathbf{c} = (c_k, c_{k-1}, \dots, 1)'$ 為其係數向量。第(5)式中的高階多項式前有一個參數

⁴ 狀態變數 s_t 是由平滑轉換迴歸中的轉換參數應用在結構改變檢定議題而來，在本文中的第貳節中的 $s_t \in [0, 1]$ 與第參節中的 $s_t \in \{1, 2\}$ ，乃因為計量的設計而有不同的值域。

r ，它決定的是結構改變的速度，即若 $r \rightarrow \infty$ 時， ω 的結構改變為間斷式的跳躍；若 $r \rightarrow 0$ 時， ω 結構改變沒有發生；而當 $r \in (0, \infty)$ 時， ω 的結構改變為平滑轉換式的結構改變。

然而在推導檢定統計量的過程中，第(3)式中的 ω_2 將無法認定(not identified)，為了解決這個擾嚷參數(nuisance parameter)的問題，Lin and Terasvirta (1994)、林-張(1997)以及高-呂-林(2001)的方法是將 $F(s_t)$ 往下平移1/2，定義成新的函數 $\tilde{F}(s_t)$ ，並再用一階泰勒展開逼近 $\tilde{F}(s_t)$ ，將相同項目的係數重新組合後，條件變異數方程式可以得到以下含有 \otimes (Kronecker 乘數)的參數化形式

$$\begin{aligned} h_t &= z_t' \varphi_1 + (K_t \otimes z_t)' \varphi_2, \\ &= (1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2, h_{t-1}, \dots, h_{t-p})' \varphi_1 \\ &\quad + (z_t', t z_t', t^2 z_t', \dots, t^k z_t')' \varphi_2 \end{aligned} \quad (6)$$

此時模型的參數向量則變成了 $\varphi = (\varphi_1', \varphi_2')'$ ，也就是我們將 GARCH 過程中參數是否有發生結構改變的虛無假設從 $H_0: \omega_2 = 0$ 改變成

$$H_0: \varphi_2 = 0. \quad (7)$$

至於結構改變的形式則是由時間變數 t 的高階多項式所決定。當階數 $k=1$ 時，結構改變是屬於單調轉換；當階數 $k=2$ 時，結構改變是週而復始的結構改變；當 $k=3$ 時，則結構改變可以涵蓋所有非單調的轉換。若用矩陣符號來表示此虛無假設，即令 $\mathbf{R} = (\mathbf{I}, \mathbf{0})$ ，則虛無假設用向量形式可以表示成

$$H_0: \mathbf{R}\varphi = (\mathbf{0}, \mathbf{I}) \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \mathbf{0}. \quad (8)$$

在虛無假設 $H_0: \varphi_2 = 0$ 之下，如果估計這個 GARCH 模型的條件對數概似函數 $L_T(\varphi)$ 所導出的 Hessian 矩陣的期望值為 \mathbf{H} ，而導出的訊息矩陣(information matrix)為 \mathbf{B} ，則根據 Lin and Terasvirta (1994)和林-張(1997)的推導，檢定 GARCH 模型結構改變虛無假設 $H_0: \varphi_2 = 0$ 的拉氏乘數(Lagrange Multiplier, LM)檢定統計量為自由度 $k \times (p+q+1)$ 的卡方分配，即⁵

$$\begin{aligned} LM &= T [\nabla L_T(\varphi)]' \tilde{\mathbf{H}}_T^{-1} (\mathbf{R} \tilde{\mathbf{C}}_T \mathbf{R}') \mathbf{R} \tilde{\mathbf{H}}_T^{-1} [\nabla L_T(\varphi)] \\ &\sim \chi_{k \times (p+q+1)}^2 \end{aligned} \quad (9)$$

其中 $\tilde{\mathbf{H}}_T$ 是 $\mathbf{H}_T(\varphi) \equiv E[\nabla^2 L_T(\varphi)]$ 的一致性估計式， $\tilde{\mathbf{C}}_T$ 是參數 φ 的變異數-共變異數矩陣的一致性估計式，意即 $\mathbf{C}_T(\varphi)^{-1/2} \sqrt{T}(\tilde{\varphi} - \varphi) \xrightarrow{D} N(0, \mathbf{I})$ ，而且這個 $\mathbf{C}_T(\varphi)$ 的組成可以分成三個成份，即 $\mathbf{C}_T(\varphi) \equiv \mathbf{H}_T(\varphi)^{-1} \mathbf{B}_T(\varphi) \mathbf{H}_T(\varphi)^{-1}$ ，中間的成分在計量理論上被定義成訊息矩陣，或被稱之為 $\nabla L_T(\varphi)$ 的變異數-共變異數矩陣，即 $\mathbf{B}_T(\varphi) \equiv \text{var}[\nabla L_T(\varphi)]$ ，而它的一致性估計式則為 $\tilde{\mathbf{B}}_T$ ，屬於 $\tilde{\mathbf{C}}_T$ 的中間部分。⁶

參、股市波動之結構改變的時點估計

如果股市波動發生結構改變，對於結構改變的時點估計，本文採取的是 Markov Switching GARCH 的方法，利用狀態變換的機率來做偵測。我們假設波動結構改變成兩種狀態，即波動高的狀態和波動低的狀態。令 s_t^* 為

⁵ 如前所述， p, q 為 ARMA 過程的階數， k 為時間 t 的高階多項式的最高階次方數。

⁶ 在這裡，我們用 $L_T(\varphi)$ 表示樣本數為 T ，未知參數為 φ 的條件對數概似函數；以 $\nabla L_T(\varphi)$ 表示它的梯度向量 (gradient vector)；以 $\nabla^2 L_T(\varphi)$ 表示它對 φ 的二階導數矩陣。

第 t 期的狀態變數，當波動程度為高時，其值為 1，反之為 2。在上一節中的檢定方法，我們假設狀態變數可以用時間變數的高階多項式來逼近它，而在本節中所用來估計結構改變時點的狀態變數是無法觀察到的，需要用最大概似法在估計資料時自動決定。假設 s_t^* 可以被一馬可夫鏈所刻劃，即

$$\begin{aligned} \Pr(s_t^* = j | s_{t-1}^* = i, s_{t-2}^* = k, \dots) \\ = \Pr(s_t^* = j | s_{t-1}^* = i) \end{aligned} \quad (10)$$

上式的意義為這一期的狀態變數為何決定於前一期，而非前幾期。為了簡化符號，本文一律將第(10)式稱為由狀態 i 至狀態 j 的移轉機率 p_{ij} 。然而，模型為了完整描述由高波動至高波動、高波動至低波動、低波動至低波動、低波動至高波動四種情況，上式的移轉方式將被下面新的狀態變數 s_t 予以更完整的敘述：

$$\begin{cases} s_t = 1, & \text{if } s_t^* = 1 \text{ and } s_{t-1}^* = 1, \\ s_t = 2, & \text{if } s_t^* = 2 \text{ and } s_{t-1}^* = 1, \\ s_t = 3, & \text{if } s_t^* = 1 \text{ and } s_{t-1}^* = 2, \\ s_t = 4, & \text{if } s_t^* = 2 \text{ and } s_{t-1}^* = 2. \end{cases} \quad (11)$$

此時，新的狀態變數 s_t 可以有四種情況，由前一期的狀態和這一期的狀態共同決定。如此一來馬可夫鏈的移轉矩陣(transition matrix)，以下用就 \mathbf{P} 代表，可以進一步寫成

$$\begin{aligned} \mathbf{P} &= \begin{bmatrix} p_{11} & p_{21} & p_{31} & p_{41} \\ p_{12} & p_{22} & p_{32} & p_{42} \\ p_{13} & p_{23} & p_{33} & p_{43} \\ p_{14} & p_{24} & p_{34} & p_{44} \end{bmatrix} \\ &= \begin{bmatrix} p_{11}^* & 0 & p_{11}^* & 0 \\ p_{12}^* & 0 & p_{12}^* & 0 \\ 0 & p_{21}^* & 0 & p_{21}^* \\ 0 & p_{22}^* & 0 & p_{22}^* \end{bmatrix}. \end{aligned} \quad (12)$$

上式每一行的加總值為 1。這個移轉矩陣 \mathbf{P} 將是計算模型的條件對數概似函數、以及狀態發生機率的基礎。⁷

因為我們使用的是最大概似法來估計 Markov Switching GARCH 模型，所以首先我們必須先建構 Markov Switching GARCH 模型的條件概似函數。如前所述，股市報酬率會遵循第(1)式的過程，而針對第(1)式的誤差項 ε_t ，如果要將其條件分配設定成會隨著不同狀態而變動的情況，必須要將此誤差項中會隨著狀態改變的部分分解出來。文獻上對此處理的方式各有不同，我們在此節用比較好理解的方法，即 Hamilton and Susmel (1994)一文所使用的，來加以闡述。Hamilton and Susmel (1994)將誤差項 ε_t 拆解成兩部分，目的是為了分開波動率大小為不同狀態的係數 $\sqrt{g_{s_t}}$ 和純粹處理股市衝擊的部分 $\tilde{\varepsilon}_t$ ⁸

$$\varepsilon_t = \sqrt{g_{s_t}} \times \tilde{\varepsilon}_t \quad (13)$$

而且就純粹處理股市衝擊的部分 $\tilde{\varepsilon}_t$ 來描述條件概似函數的分配。

$$\tilde{\varepsilon}_t = \sqrt{h_t} \cdot v_t, \text{ with } v_t \stackrel{i.i.d.}{\sim} N(0,1) \text{ or } v_t \stackrel{i.i.d.}{\sim} t(m). \quad (14)$$

此時，條件變異數方程式就可以進一步寫成

$$\begin{aligned} h_t &= \alpha_{0,s_t} + \sum_{i=1}^q \alpha_{i,s_t} \tilde{\varepsilon}_{t-i}^2 + \sum_{i=1}^p \beta_{i,s_t} h_{t-i} \\ &= z_t' \omega_{s_t}, \quad s_t = 1, 2. \end{aligned} \quad (15)$$

⁷ 由(12)式可知，四種情況又回到了兩狀態的移轉樣貌，是故以下的推導， s_t 仍維持只有 1 和 2 兩種狀態數值。

⁸ 狀態 1 的 g_1 被標準化為 1，狀態 2 的 g_2 則由估計資料時自動決定。詳細原理請參考 Hamilton and Susmel (1994)。

第(15)式中參數向量爲

$$\omega_{s_t} = (\alpha_{0,s_t}, \alpha_{1,s_t}, \dots, \alpha_{q,s_t}, \beta_{1,s_t}, \dots, \beta_{p,s_t})',$$

即它的值隨著 s_t 而有所不同。由於本文假設狀態空間裡有兩個狀態，所以其值爲 1 和 2，分別代表高波動和低波動的世界。由於狀態變數 s_t 是無法觀察到的，但可以用馬可夫鏈的原理加上資料特性來決定，亦即 s_t 的估計將是透過以下的最大概似函數的求算過程來達成。而因爲模型牽涉到馬可夫鏈和 GARCH 過程，所以 Hamilton (1994) 處理這種最大概似函數的方法是又把它分解成兩個重要的向量來計算，即代表參數部分的 η_t 和代表狀態部分的 $\xi_{t|t}$ ，而後者更是計算狀態變數的平滑機率基礎。

由於 Markov Switching GARCH 模型的估計過程非常複雜，文獻上使用的演算法 (algorithm) 在細微處雖然各有不同，如 Cai (1994)、Hamilton and Susmel (1994)、Gray (1996)、Haas, Mittnik, and Paoletta (2004) 等人，但是它們主要的精神是相同的。用 Hamilton (1994) 所揭櫫的原則，第(1)式下的條件殘差機率分配可以寫成

$$\varepsilon_t | \Psi_{t-1} \sim \begin{cases} f(\varepsilon_t | s_t = 1, \Psi_{t-1}) & \text{with probability } \pi_{1t} \\ f(\varepsilon_t | s_t = 2, \Psi_{t-1}) & \text{with probability } \pi_{2t} \end{cases}$$

其中 $f(\cdot)$ 爲常態或 t 分配的機率密度函數， π_{1t} 、 π_{2t} 爲狀態 1、2 在第 t 期發生機率，即 $\pi_{1t} \equiv \Pr(s_t = 1 | \Psi_{t-1})$ 、 $\pi_{2t} \equiv \Pr(s_t = 2 | \Psi_{t-1})$ 。再定義 η_t 爲不同狀態下殘差的機率分配所組成的向量

$$\eta_t = \begin{bmatrix} f(\varepsilon_t | s_t = 1; \theta) \\ f(\varepsilon_t | s_t = 2; \theta) \end{bmatrix}, \quad (16)$$

也就是 η_t 是由變數 ε_t 在不同狀態下的條件分配爲元素所組成，所以模型中的主要參數 $\theta = (\gamma', \omega_{s_t}', \sigma^2)'$ 就可以內含其中。至於馬可夫鏈若要進入概似函數，還必須讓它可以在時間序列上做更新，所以還必須再定義一個具有馬可夫鏈型式的向量 ξ_t ，即

$$\xi_t = \begin{cases} (1, 0)', & s_t = 1; \\ (0, 1)', & s_t = 2. \end{cases} \quad (17)$$

當我們對 ξ_{t+1} 取給定 $s_t = i$, $i = 1, 2$ 下的條件期望值時

$$E(\xi_{t+1} | s_t = i) = \begin{bmatrix} p_{i1} \\ p_{i2} \end{bmatrix}, \quad i = 1, 2. \quad (18)$$

也就是 $E(\xi_{t+1} | s_t = i)$ 爲移轉矩陣 \mathbf{P} 的第 i 行。因此我們可以得出

$$E(\xi_{t+1} | \xi_t) = \mathbf{P}\xi_t \quad (19)$$

的結果，根據馬可夫鏈的定義，上式可寫成

$$\xi_{t+1|t} \equiv E(\xi_{t+1} | \xi_t, \xi_{t-1}, \dots) = \mathbf{P}\xi_t. \quad (20)$$

而第(20)式也意味著它是一個向量自我迴歸過程

$$\xi_{t+1} = \mathbf{P}\xi_t + \mathbf{v}_{t+1}, \quad \mathbf{v}_{t+1} \equiv E(\xi_{t+1} | \xi_t) - \mathbf{P}\xi_t. \quad (21)$$

將模型中可以觀察到的資料，根據第(16)式，代入 η_t 中，再結合 $\xi_{t|t-1}$ ，根據 Hamilton(1994)的推導，可以得到

$$\hat{\xi}_{t|t} = \frac{(\hat{\xi}_{t|t-1} \odot \eta_t)}{\mathbf{1}'(\hat{\xi}_{t|t-1} \odot \eta_t)} \quad (22)$$

其中 \odot 表示向量元素對向量元素的乘法(element-by-element multiplication)。我們可看出 $\hat{\xi}_{t|t-1}$ 是 $\xi_{t|t-1}$ 的估計值，表示概似函數在運算前， $\hat{\xi}_{t|t-1}$ 必須優先於 η_t 估計。透過移轉矩陣 \mathbf{P} ，可以對 $\hat{\xi}_{t|t}$ 加以進一步更新，亦即

$$\hat{\xi}_{t+1|t} = \mathbf{P} \cdot \hat{\xi}_{t|t} \quad (23)$$

上式即計算狀態變數的過濾機率(filtering probability) $\Pr(s_t = i | \mathbf{Y}^t; \theta)$, $i = 1, 2$ 的基礎，其中 \mathbf{Y}^t 表示前面 t 期可觀察到的樣本資料集合。此時，對數概似函數 $L(\theta)$ 可以寫成(以常態分配為例)

$$\begin{aligned} L(\theta) &= \sum_{t=1}^T \log f(\varepsilon_t; \theta) \\ &= \sum_{t=1}^T \log \left[\mathbf{1}'(\hat{\xi}_{t|t-1} \odot \eta_t) \right] \\ &= \sum_{t=1}^T \log \left[\pi_{1t} \frac{1}{\sqrt{2\pi h_{1t}}} \exp\left(\frac{-\varepsilon_t^2}{2h_{1t}}\right) \right. \\ &\quad \left. + \pi_{2t} \frac{1}{\sqrt{2\pi h_{2t}}} \exp\left(\frac{-\varepsilon_t^2}{2h_{2t}}\right) \right], \quad (24) \end{aligned}$$

其中 $f(\varepsilon_t; \theta)$ 可以用矩陣方式的 $\mathbf{1}'(\hat{\xi}_{t|t-1} \odot \eta_t)$ 來計算，也牽涉到常態分配或 t 分配的函數形式；而 h_{1t} 、 h_{2t} 分別為狀態 1、2 下的條件變異數，由於 GARCH 模型具有遞迴的(recursive)的架構，造成其與馬可夫鍊結合時，統計上會

出現估計上的困難，而這個困難在 Hamilton and Susmel (1994)以及 Cai (1994)都無法解決，因為他們的估計方法是路徑相依(path dependent)，是故本文採取的是 Gray (1996)的做法，將在本節後面說明。至此，狀態變數的平滑機率(smoothing probability)估計值的計算方式可以寫成⁹

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} \odot \left\{ \mathbf{P}' \cdot \left[\hat{\xi}_{t+1|T} (\div) \hat{\xi}_{t+1|t} \right] \right\} \quad (25)$$

其中 (\div) 表示向量元素對向量元素的除法(element-by-element division)，算出來的向量元素即狀態變數的平滑機率估計值 $\Pr(s_t = i | \mathbf{Y}^T; \theta)$, $i = 1, 2$ ，其中 \mathbf{Y}^T 表示全部 T 期可觀察到的樣本資料集合。然而上述在計算概似函數值和平滑機率值的過程中，如果要把 Hamilton and Susmel (1994)的 SWARCH 模型要擴展成 Markov Switching GARCH 模型，如果移轉矩陣的維度(dimension)不斷增加，將造成計算過程中電腦的記憶體無法負荷，以 GARCH(1,1)模型為例

$$h_t = \alpha_{0,s_t} + \alpha_{1,s_t} \varepsilon_{t-1}^2 + \beta_{1,s_t} h_{t-1} \quad (26)$$

將(26)式反覆代入，則第 t 期的條件標準差 h_t 會決定於一歷史序列，即

$$\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_0, s_t, s_{t-1}, \dots, s_1\} \quad (27)$$

致使估計程序太過龐大而無法計算。解決此問題的方法，文獻上有兩種，第一種是由 Gray

⁹ 請注意過濾機率、平滑機率以及 π_{1t} 、 π_{2t} 均為狀態變數的條件機率，但是所用訊息集合不同。

(1996)所提出來的，他的方法是將第(26)式改寫成

$$h_t = \alpha_{0,s_t} + \alpha_{1,s_t} \varepsilon_{t-1}^2 + \beta_{1,s_t} \tilde{h}_{t-1} \quad (28)$$

其中 \tilde{h}_{t-1} 被設計成以 $\hat{\xi}_{t|t}$ 來建構，即

$$\tilde{h}_{t-1} = \sum_{i=1}^2 \hat{\xi}_{i,t-1|t-2} (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \tilde{h}_{t-1}) \quad (29)$$

也就是因為 \tilde{h}_{t-1} 是訊息集合 Ψ_{t-1} 資料中的函數，所以 h_t 決定於狀態 s_t 即可，而不被前面 $\{s_{t-1}, s_{t-2}, \dots, s_1\}$ 影響。另一種方法是 Haas, Mittnik, and Paolella (2004)所提出來的，他們將(26)式改回原來 GARCH(1,1)的遞迴方式

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (30)$$

再放入 $\varepsilon_t = \sqrt{h_{s_t}} v_t$, $s_t = 1, 2$, $v_t \sim N(0,1)$ 的限制式，也就是想辦法做到 h_t 只是訊息集合 Ψ_{t-1} 資料中變數的函數，而不受到前面狀態序列 $\{s_{t-1}, s_{t-2}, \dots, s_1\}$ 的影響。

有鑑於文獻上對於解決 GARCH 遞迴結構結合馬可夫鍊在估計上的困難大多採取 Gray (1996)的做法，本文也將從善如流。如此一來，用來建構殘差項條件分配的訊息集合 Ψ_{t-1} 中，將不必再包括狀態變數序列元素 $\{s_{t-1}, s_{t-2}, \dots, s_1\}$ ，也就是前一期的 sigma-代數並不會再由狀態變數序列來產生，而在模型估計過程中也可以不需要第(13)、(14)式拆解的動作。此時，透過 Gray (1996)的邏輯，第(1)、(2)式可以合併寫成(以常態分配為例)

$$R_t = \mu_t + \sqrt{h_{s_t}} v_t, \text{ with } v_t \stackrel{i.i.d.}{\sim} N(0,1) \quad (31)$$

則根據條件變異數的定義，

$$\begin{aligned} h_t &= E[R_t - E(R_t) | \Psi_{t-1}]^2 \\ &= E[R_t^2 | \Psi_{t-1}] - E[R_t | \Psi_{t-1}]^2 \\ &= \pi_{1t} (\mu_t^2 + h_{1t}) + \pi_{2t} (\mu_t^2 + h_{2t}) - (\pi_{1t} \mu_t + \pi_{2t} \mu_t)^2 \end{aligned} \quad (32)$$

此時，條件變異數 h_t 不再是路徑相依，而且如果進一步建立下一期兩個狀態的條件變異數 h_{t+1} 、 h_{2t+1} 可以只使用第 t 期的 h_t 即可。這個意思即為，如果關注的是第 t 期的 h_{it} , $i=1, 2$ ，則它可以用

$$h_{it} = \alpha_{0,i} + \alpha_{1,i} \varepsilon_{t-1}^2 + \beta_{1,i} h_{i,t-1}, \quad i = s_t = 1, 2 \quad (33)$$

計算而得。而第(33)式就是第(24)式條件對數概似函數中 h_{1t} 、 h_{2t} 的代入方式。第(33)式中等號右邊第三項 $h_{i,t-1}$ 可以按照第(32)式的定義如下計算

$$\begin{aligned} h_{i,t-1} &= \pi_{1,t-1} (\mu_{t-1}^2 + h_{1,t-1}) + \pi_{2,t-1} (\mu_{t-1}^2 + h_{2,t-1}) \\ &\quad - (\pi_{1,t-1} \mu_{t-1} + \pi_{2,t-1} \mu_{t-1})^2 \end{aligned} \quad (34)$$

肆、實證分析

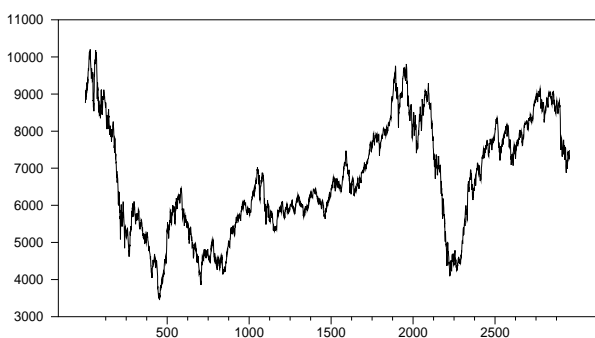
我們以文獻中處理股市報酬率最常用模型 AR(1)-GARCH(1,1)，來檢定和估計台灣股市，¹⁰即第貳節 R_t 的條件平均數方程式為 $\mu_t = \gamma_0 + \gamma_1 R_{t-1}$ 。本文所使用的資料為台灣經濟新報資料庫中，變數代號為 Y9999 的台灣加權股價指數，樣本期間為西元 2000 年 1 月 4 日

¹⁰ 文章中使用 AR(1)-GARCH(1,1) 模型者包括 Gray(1996)、林-張(1997)、高-呂-林(2001)等人，這樣做的好處是可以達成模型精減原則 (parsimonious principle)。

至 2011 年 10 月 25 日，共 2954 個交易日，也就是有 2954 筆樣本數。至於的每日交易時點，我們取的是當天的收盤價。令 P_t 是第 t 期的台灣加權股價指數，則第 t 期的台股報酬率可以寫成

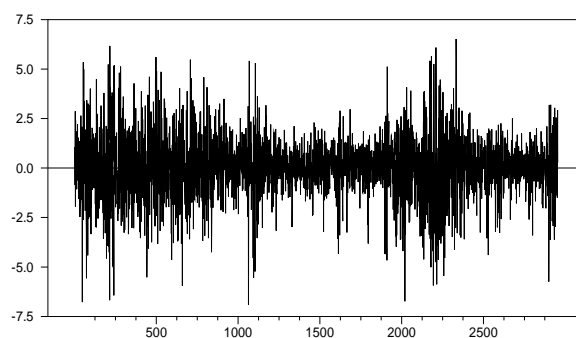
$$R_t = (\ln P_t - \ln P_{t-1}) \times 100。 \quad (35)$$

對於 P_t 和 R_t 的圖形，我們分別用圖一和圖二表示。



圖一、台灣加權股價指數走勢圖

圖一為台灣加權股價指數的走勢，因為股市交易日並非每天都有，所以圖形中的橫軸表示第 t 期的交易日，也就是第一筆樣本 ($t=1$) 表示 2000 年 1 月 4 日，最後一筆樣本 ($t=2954$) 表示 2011 年 10 月 25 日。從圖中可看出在這 11 年當中台股上下振盪，尤如隨機漫步，圖形中的最高點，發生在 2000 年 2 月 17 日 ($t=30$)；最低點發生在 2001 年 10 月 4 日 ($t=453$)。經過第(35)式的轉換後，台股報酬率 R_t 的圖形放在圖二，由圖形可看出台股報酬率大的波動跟隨著大的波動，小的波動跟隨著小的波動。



圖二、台股報酬率

首先，我們先用 ARCH-LM 檢定，來看如果股價指數報酬率模型是 AR(1) 時，模型殘差是否會發生 ARCH 效果。AR(1) 的估計結果如下：

表 1 台股報酬率的模型為 AR(1) 時的結果

	AR(1)
γ_0	-0.005 (0.028)
γ_1	0.063 (0.018)***

註：弧號內數字表示參數估計式的標準差。一個星號 * 表示 10% 的顯著水準，兩個星號 ** 表示 5% 的顯著水準，以及三個星號 *** 表示 1% 的顯著水準

利用利用表一報酬率 AR(1) 模型的殘差，跑 ARCH-LM 檢定，若落後(lag)項取 4 期，得到結果如下：

表 2 針對台股報酬率的模型為 AR(1) 時的殘差平方自我相關檢定

	ARCH-LM 檢定統計量
Lag 1 to Lag 4	57.491***

註：星號的定義如表一

表二得知，殘差項平方後自我相關的檢定，都嚴格拒絕係數等於 0 的虛無假設，表示報酬率 AR(1)模型的殘差出現波動聚集的現象。所以我們進一步估計 AR(1)-GARCH(1,1)模型，看看結果如何，

表 3 台股報酬率的模型為 AR(1) -GARCH(1,1) 時的估計結果

	常態	t 分配
γ_0	0.050** (0.022)	0.064*** (0.020)
γ_1	0.069*** (0.019)	0.051*** (0.016)
α_0	0.020*** (0.006)	0.012** (0.005)
α_1	0.076*** (0.010)	0.005*** (0.009)
β_1	0.917*** (0.011)	0.935*** (0.010)
d.f. (m)		7.449*** (0.994)
$L_T(\omega)$	-5161.392	-5119.759

註：星號的定義如表一

雖然 AR(1)-GARCH(1,1)模型似乎較 AR(1)模型對資料更為配適，但是從表中看出無論條件分配是常態還是 t 分配，條件方程式的後兩項的係數加總， $\alpha_1 + \beta_1$ ，均大於 0.9，接近於 1，表示條件變異數的持續性很高，資料有可能出現 IGARCH 的現象，所以我們進一步進行 GARCH 結構改變的檢定。以 $\varepsilon_t | \Psi_{t-1} \sim N(0, h_t)$ 為例，台股報酬率條件於前一期訊息集合的條件機率分配可以寫成

$$f(R_t | \Psi_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} \exp \left[-\frac{(R_t - \mu_t)^2}{2h_t} \right] \quad (36)$$

當我們用最大概似法估計時，須將上式改寫成條件對數概似函數，即上式對每一個樣本做對數加總，可得

$$L_T(\theta) = \frac{1}{T} \left\{ \sum_{t=1}^T \log f(R_t | \Psi_{t-1}) \right\} \\ = \frac{1}{T} \left[-\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t) - \frac{1}{2} \sum_{t=1}^T \frac{(R_t - \mu_t)^2}{h_t} \right]$$

其中 $\theta = (\gamma', \phi')'$ 包括條件平均數方程式參數向量 γ' 和經過平滑轉換後的條件變異數方程式(第(6)式)之參數向量 ϕ' 。以上式分別對條件平均數方程式及條件變異數的參數分別求取一階和二階微分，可以分別得到參數向量的點估計值及其 Hessian 矩陣下的假設檢定的結果。由於 H_T 和 B_T 在此情況下均具有區塊對角性(block diagonality)，而本文重點在於條件變異數的結構改變，以下我們將針對經過平滑轉換後的條件變異數方程式(第(6)式)之參數向量 ϕ' 做說明。將其一階微分整理如下，由於上式第一項為常數，並不會影響一階微分結果，是故我們可以定義第 t 期的條件對數概似函數為

$$l_t(\phi) = -(1/2) \log(h_t) - (1/2)(\varepsilon_t)^2 / h_t \quad (37)$$

全期的條件對數概似函數可寫成較簡潔的

$$L_T(\phi) = \frac{1}{T} \sum_{t=1}^T l_t(\phi) \quad .$$

第(37)式若對平滑轉換後的條件變異數方程式(第(6)式)的參數向量做一階微分，則為

$$\begin{aligned}\nabla l_t(\varphi) &= -\frac{1}{2} \frac{1}{h_t} \frac{\partial h_t}{\partial \varphi} + \frac{1}{2} \frac{\varepsilon_t^2}{h_t} \frac{1}{h_t^2} \frac{\partial h_t}{\partial \varphi} \\ &= \frac{1}{2} \frac{1}{h_t} \frac{\partial h_t}{\partial \varphi} \left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \\ &= \frac{1}{2} \frac{1}{h_t} \left(t^j z_t + \sum_{i=1}^p \beta_i \frac{\partial h_{t-i}}{\partial \varphi} \right) \left(\frac{\varepsilon_t^2}{h_t} - 1 \right), \\ &\quad j = 0, 1, \dots, k.\end{aligned}\quad (38)$$

我們使用林-張(1997)和高-呂-林(2001)的 LM1 和 LM2 兩種檢定統計量，¹¹來進行條件變異數方程式是否發生結構改變的檢定。LM1 是針對殘差的條件概似函數是常態分配情況下所做的檢定，此時，因為模型並沒有出現動態誤設(dynamic misspecification)的情況，條件變異數方程式之參數訊息矩陣等式(information matrix equation)得以成立，亦即 $\mathbf{H}_T(\varphi) + \mathbf{B}_T(\varphi) = \mathbf{0}$ ，此時第(9)式的 LM 檢定統計量可以化減成

$$LM_T = 2T \nabla L_T(\varphi)' \mathbf{H}_T^{-1} \nabla L_T(\varphi) \quad (39)$$

$\mathbf{H}_T(\varphi) = -\mathbf{B}_T(\varphi) = -1/T \sum_{t=1}^T E \left[\nabla l_t(\varphi) \nabla l_t(\varphi)' \right]$ ，如果我們使用 Bollerslev (1986)以及林-張(1997)的符號，可以將上面第(39)式改寫成本文的第一種 LM 檢定統計量，LM1，

$$LM1 \equiv \frac{1}{2} f_0' Z_0 (Z_0' Z_0)^{-1} Z_0' f_0 \sim \chi_{k \times (p+q+1)}^2 \quad (40)$$

此處 f_0 是由 $f_t \equiv \varepsilon_t^2/h_t - 1$ 為元素所組成的行向量， Z_0 是由 $Z_t \equiv (1/h_t)(\partial h_t/\partial \varphi)$ 為元素所組成

的行向量，亦即

$$f_0 \equiv \left(\frac{\varepsilon_1^2}{h_1} - 1, \dots, \frac{\varepsilon_T^2}{h_T} - 1 \right)', \quad Z_0 \equiv \left(\frac{1}{h_1} \frac{\partial h_1}{\partial \varphi}, \dots, \frac{1}{h_T} \frac{\partial h_T}{\partial \varphi} \right)'$$

上述檢定結構改變的 LM1 檢定統計量與傳統 Bollerslev (1986)所用的 LM 檢定統計量不同的地方在於第(38)式，即第 t 期概似函數之一階微分式，中的第三個等式之第一項具有時間多項式因子的 $t^j z_t$ ，而非 Bollerslev (1986)的 z_t 。我們在進行程式計算 LM1 時，將利用本文 AR(1)-GARCH(1,1)模型所估出樣本內配適值 $\{h_t\}_{t=1}^T$ 、 $\{\varepsilon_t\}_{t=1}^T$ ，進而得出 f_0 ，並再算出 $\{\partial h_t/\partial \varphi\}_{t=1}^T$ 序列，並進而得出 Z_0 ，最後計算第(40)式即可得 LM1 數值。根據第貳節的說明，LM1 將同樣會服從 $k \times (p+q+1)$ 自由度的卡方分配，我們將檢定結果放在表四。由表四得知，除了 $k=3$ 下的 99%顯著水準沒有結構改變外，其餘情況皆告訴我們 GARCH 模型在統計檢定上發生結構改變，不論是單調、周而復始、非單調的結構改變均有可能。

表 4 $H_0: \varphi_2 = 0$ 的第一種檢定

臨界值	10%	5%	1%	LM1
$k = 1$	6.251	7.814	11.344	21.02
$k = 2$	10.644	12.591	16.811	
$k = 3$	14.683	16.918	21.665	

上述 LM1 檢定統計量奠基於常態分配的假設下，但財務資料的分配有時不是常態分配，如果我們使用非常態分配的條件對數概似函數來求解 GARCH 模型時，訊息矩陣等式此時就不會成立，但是如同 Chu (1995)說的，結構改變的 LM 檢定統計量弱收斂於平分貝索過

¹¹ 本文的 LM1 是林-張(1997)和高-呂-林(2001)文中的 LM1a。

程(squared Bessel process)的關鍵變數向量 V_t ，仍是一個平賭差分數列(Martingale Difference Sequence)，即

$$\begin{aligned} E(V_t | \Psi_t) &= E(f_t Z_t | \Psi_t) \\ &= E\left[\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\left(\frac{1}{h_t} \frac{\partial h_t}{\partial \varphi}\right) \middle| \Psi_t\right] \\ &= \mathbf{0}。 \end{aligned}$$

若 V_t 的共變數矩陣為 \mathbf{V}_T 即

$$\mathbf{B}_T \equiv \frac{1}{T} \left(\sum_{t=1}^T V_t \right) \left(\sum_{t=1}^T V_t \right)'$$

若它在最大概似法下的一致性估計式為 $\tilde{\mathbf{V}}_T$ ，則利用參數的變異-共變數矩陣原理，將之代入，可得本文的第二種檢定統計量

LM2 =

$$\begin{aligned} &\frac{1}{T} \left[\nabla L_T(\tilde{\varphi}) \right]' \tilde{\mathbf{H}}_T^{-1} \mathbf{R}' \left(\mathbf{R} \tilde{\mathbf{H}}_T^{-1} \tilde{\mathbf{B}}_T \tilde{\mathbf{H}}_T^{-1} \mathbf{R}' \right) \\ &\mathbf{R} \tilde{\mathbf{H}}_T^{-1} \left[\nabla L_T(\tilde{\varphi}) \right] \sim \chi_{k \times (p+q+1)}^2 \end{aligned} \quad (41)$$

第(41)式的 LM2 則是針對殘差的條件概似函數是非常態分配情況下所做的檢定統計量，根據第貳節的說明，LM2 仍然會服從 $k \times (p+q+1)$ 自由度的卡方分配，我們將 LM2 的檢定結果放在表五。

表 5 $H_0: \varphi_2 = 0$ 的第二種檢定

臨界值	10%	5%	1%	LM2
$k = 1$	6.251	7.814	11.344	6.54
$k = 2$	10.644	12.591	16.811	
$k = 3$	14.683	16.918	21.665	

由表五得知，若條件分配為非常態，GARCH 只有在 $k=1$ 下的 10% 顯著水準發生結構改變，而且這種結構改變只能是單調型的結構改變。

因為由表四和表五得知條件分配為常態時，發生條件變異數結構改變的可能性較高，因此我們接下來的 Markov Switching GARCH 模型，來估計條件變異數結構改變的時點時，所用的殘差條件分配將使用常態的設定，才不會發生資料探勘的問題。

我們對於台股波動結構改變的時點估計，所使用的 Markov Switching GARCH 模型如下式：

$$\begin{aligned} R_t &= \mu_t + \varepsilon_t \\ &= \gamma_0 + \gamma_1 R_{t-1} + \varepsilon_t \\ \varepsilon_t &= \sqrt{h_t} \cdot v_t, \quad v_t \sim N(0,1) \\ h_t &= \alpha_{0,s_t} + \alpha_{1,s_t} \varepsilon_{t-1}^2 + \beta_{1,s_t} h_{t-1}, \quad s_t = 1, 2. \end{aligned} \quad (42)$$

由於我們使用 Gray (1996) 的方法來估計 Markov Switching GARCH 模型，所以第(42)式的參數並沒有包括 g_{s_t} ，而且也沒有牽涉到 $\tilde{\varepsilon}_t$ 的分解。¹²此外，從第(42)式中我們可看出，條件方程式的係數下標都有狀態變數 s_t ，也就是這些參數的估計值將會隨著狀態的不同而有不同。因為本文分析的狀態有兩個，所以在 GARCH(1,1) 的設定下，條件變異數方程式的未知參數將有六個。我們將估計結果放在表六中。

¹² 那是屬於 Hamilton and Susmel (1994) 一文才有的參數。

表 6 台股報酬率的模型為 Markov Switching GARCH(1,1)時的估計結果

	AR(1)-SW(2)-GARCH(1,1)
γ_0	0.06198*** (0.022)
γ_1	0.06190*** (0.019)
$\alpha_{0,1}$	0.573*** (0.026)
$\alpha_{0,2}$	1.721*** (0.089)
$\alpha_{1,1}$	0.051*** (0.0006)
$\alpha_{1,2}$	0.041*** (0.013)
$\beta_{1,1}$	0.252*** (0.024)
$\beta_{1,2}$	0.628*** (0.022)
P_{11}	0.986*** (0.002)
P_{12}	0.016*** (0.003)
$L_T(\omega_s)$	-5117.283

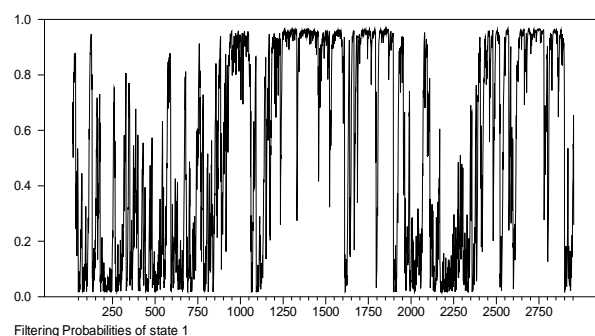
註：星號的定義如表一

由表六中可看出大部分的參數估計值均為顯著，且 Markov Switching GARCH(1,1)模型的對數概似函數值高於表三的無論是條件分配常態還是 t 分配的 AR(1)-GARCH(1,1)，顯示出 Markov Switching GARCH(1,1)模型確實對台股資料配適得比較好。此外， P_{11} 的機率估計值很高，表示 R_t 一旦進入高波動狀態，則持續較長的時間，所以下面的分析會以高波動狀態為主。

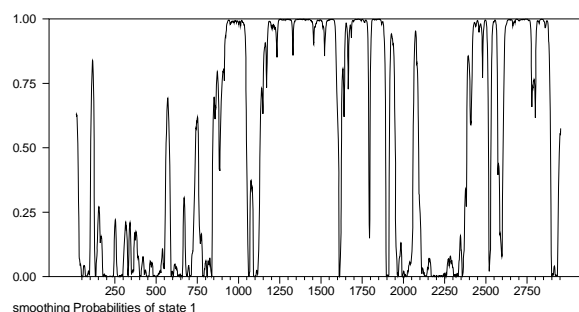
我們將台灣股市高波動狀態的過濾機率

$\Pr(s_t = 1 | \mathbf{Y}^T; \boldsymbol{\theta})$ 的圖形放在圖三，以及高波動狀態的平滑機率 $\Pr(s_t = 1 | \mathbf{Y}^T; \boldsymbol{\theta})$ 的圖形放在圖四。從圖形中可看出，就高波動狀態而言，過濾機率值的變動幅度比平滑機率值的變動幅度來得大，由於過濾機率 $\Pr(s_t = 2 | \mathbf{Y}^T; \boldsymbol{\theta})$ 和平滑機率 $\Pr(s_t = 2 | \mathbf{Y}^T; \boldsymbol{\theta})$ 之低波動狀態的機率剛好是 1 分別減去上述 $s_t = 1$ 兩者，所以就低波動狀態而言，過濾機率值的變動幅度也是會比平滑機率來得大。這是因為平滑機率的計算會以全部樣本資料做為訊息，較為客觀，圖形看起來也較「平滑」，故我們所使用的分析也會以平滑機率為主。

如果我們將高狀態的平滑機率值大於 50% 情況，即 $\Pr(s_t = 1 | \mathbf{Y}^T; \boldsymbol{\theta}) \geq 0.5$ ，的交易日當做是台灣股市高波動發生的時間，而把 $\Pr(s_t = 2 | \mathbf{Y}^T; \boldsymbol{\theta}) > 0.5$ 情況的交易日當做是低波動發生的時間，則我們可以把高波動狀態所發生的時段找出來。以下我們將高波動的發生時點標示成 $t|_{s_t=1}$ ，及將高波動發生的時段標示成 $[t|_{s_t=1, \text{start}}, t|_{s_t=1, \text{end}}]$ ，並得到表七的結果。



圖三 高波動狀態過濾機率圖



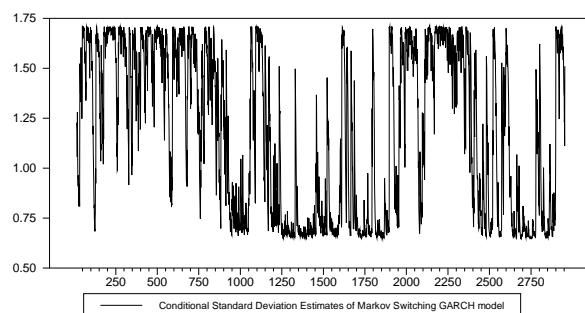
圖四 高波動狀態平滑機率圖

表 7 利用平滑機率值找出台灣股市高波動狀態所發生的時段

時段	$[t _{s_t=1, \text{start}}, t _{s_t=1, \text{end}}]$	交易日
1	[1,26]	2000 年 1 月 4 日至 2000 年 2 月 11 日
2	[104,124]	2000 年 5 月 25 日至 2000 年 6 月 21 日
3	[738,757]	2002 年 11 月 26 日至 2002 年 12 月 23 日
4	[844,1048]	2003 年 5 月 6 日至 2004 年 3 月 1 日
5	[1131,1606]	2004 年 6 月 25 日至 2006 年 5 月 29 日
6	[1625,1790]	2006 年 6 月 26 日至 2007 年 2 月 26 日
7	[1798,1892]	2007 年 3 月 8 日至 2007 年 7 月 23 日
8	[1916,1950]	2007 年 8 月 24 日至 2007 年 10 月 16 日
9	[2060,2092]	2008 年 3 月 28 日至 2008 年 5 月 15 日
10	[2385,2516]	2009 年 7 月 16 日至 2010 年 1 月 19 日
11	[2531,2572]	2010 年 2 月 8 日至 2010 年 4 月 16 日
12	[2604,2893]	2010 年 6 月 1 日至 2011 年 7 月 28 日
13	[2946,2954]	2011 年 10 月 13 日至 2011 年 10 月 25 日

表七是利用平滑機率值找出台灣股市高波動狀態所發生的時段，而這些時段也代表當時的台灣股市處於高風險的時期。例如時段 4 剛好位於台灣總統大選前，當時藍綠政黨競爭激烈，也造成股市的動盪不安。

如同前言提及的，波動率的預測和掌握對財務理論和投資實務是很重要的，所以我們亦將 Markov Switching GARCH 模型對條件標準差估計值放在圖五中。



圖五 台股 Markov Switching GARCH(1,1) 模型所估計的條件標準差

從圖五中，可明顯看出當條件標準差估計值高，或台股波動度高時，台灣股市未必有較大的機率值位於高波動狀態，這是因為在 Gray (1996) 的方法下，條件標準差估計值是用條件變異數方程式的參數估計式，乘上高波動狀態的過濾機率(或平滑機率)值，以及乘上低波動狀態的過濾機率(或平滑機率)值，兩項加總而得，也就是因為條件標準差是隨機變數，計算它估計值的方法就是去算它的期望值。所以投資人如果要用 Markov Switching GARCH 模型來進行市場套利分析時，除了條件標準差估計值外，也必須將平滑機率和過濾機率考慮進去。

最後，我們進行本文中三個模型的診斷 (Diagnostic Checking)，使用的是 Ljung-Box Q 統計量，分別針對條件分配是常態、t 分配、以及 Markov Switching AR(1)-GARCH(1,1) 模型的 $\{\varepsilon_t/\sqrt{h_t}\}$ 和 $\{\varepsilon_t^2/h_t\}$ 兩序列，看三模型的此兩序列之無自我相關的虛無假設是否能不被拒絕。我們使用自我相關的延遲(lag)項是依照 Gray (1996) 的做法，為 1、2、3、5、10 和 15。

我們將三模型的 $\{\varepsilon_t/\sqrt{h_t}\}$ 序列之 LB 檢定結果放在表八，從表八可以看出在本文 AR(1)-GARCH 的設定下，無論條件分配是常態、t 分配、還是 Markov Switching GARCH 模型，均在延遲項為 15 時， $\{\varepsilon_t/\sqrt{h_t}\}$ 序列發生

自我相關，顯示三者條件平均數的實證表現上相差無幾。

表 8 $\{\varepsilon_t/\sqrt{h_t}\}$ 序列之 LB 檢定結果

延遲項	條件分配為常態		條件分配為 Student t		Markov Switching GARCH	
	LB	P 值	LB	P 值	LB	P 值
1	0.33	[0.56]	2.26	[0.13]	1.86	[0.17]
2	0.33	[0.84]	2.27	[0.32]	3.32	[0.18]
3	3.10	[0.37]	5.18	[0.15]	5.32	[0.14]
5	3.74	[0.58]	5.89	[0.31]	5.96	[0.30]
10	7.00	[0.72]	9.23	[0.51]	13.14	[0.21]
15	25.2	[0.04]	27.4	[0.02]	45.5	[0.00]

至於三個模型的 $\{\varepsilon_t^2/h_t\}$ 序列之 LB 檢定結果放在表九，在 AR(1)-GARCH 的設定下，當條件分配為常態時，LB 延遲項為 1、3、5 的 $\{\varepsilon_t^2/h_t\}$ 序列均發生自我相關；而當條件分配為 t 分配時，LB 延遲項在 3、5、10 下的 $\{\varepsilon_t^2/h_t\}$ 序列均有自我相關；最後我們看 Markov Switching GARCH 模型，其 $\{\varepsilon_t^2/h_t\}$ 序列在這些延遲項下都沒有自我相關。如此可知，Markov Switching GARCH 模型在條件變異數方程式表現得最好，顯示條件變異數方程式有了結構改變的考量，確實讓 AR(1)-GARCH(1,1) 設定將台股報酬率資料解釋得更好。

表 9 $\{\varepsilon_t^2/h_t\}$ 序列之 LB 檢定結果

延遲項	條件分配為常態		條件分配為 Student t		Markov Switching GARCH	
	LB	P 值	LB	P 值	LB	P 值
1	2.70	[0.09]	0.87	[0.34]	2.42	[0.11]
2	3.11	[0.21]	2.58	[0.27]	4.37	[0.11]
3	9.52	[0.02]	12.0	[0.00]	5.14	[0.16]
5	11.9	[0.03]	14.8	[0.01]	5.90	[0.31]
10	14.0	[0.16]	17.5	[0.06]	8.90	[0.54]
15	16.5	[0.34]	20.8	[0.14]	9.67	[0.83]

伍、結論

本文以 AR(1)-GARCH(1,1) 模型，來檢定和估計台灣股市的波動情況，研究所使用的資料為西元 2000 年 1 月 4 日至 2011 年 10 月 25 日台灣加權股價指數。實證方法分成兩個步驟，第一步驟是先檢定台灣股市波動之結構是否發生改變，我們採取的是林建甫與張焯然 (1997) 以及高櫻芬、呂仁廣與林建甫 (2001) 的方法，也就是使用 Lin and Terasvirta (1994) 的平滑轉換迴歸架構中的 LM 檢定，透過 Logistic 分配函數，來偵測台股報酬率之 AR(1)-GARCH(1,1) 設定中的條件變異數方程式是否發生結構改變的情況，我們使用的檢定統計量有 LM1 和 LM2 兩種，結果發現 GARCH 模型在統計檢定上發生結構改變，不論是單調、周而復始、非單調的結構改變均有可能，而且殘差條件分配用常態分配的設定將會有結構改變的情況。第二步驟是針對台灣股市波動結構改變時點的估計，我們使用的是條件常態分配的 Markov Switching GARCH (1,1) 模型，實證結果顯示 Markov Switching GARCH(1,1) 模型確實對台股資料配適得比較好；此外，移轉機率 P_{11} 的機率估計值很高，表示 R_t 一旦進入高波動狀態，則持續較長的時間；最後，我們利用平滑機率值找出台灣股市高波動狀態所發生的時段，並得知這些時段也代表當時的台灣股市處於高風險的時期。最後，我們進行本文中三個模型的診斷 (Diagnostic Checking)，使用的是 Ljung-Box Q 統計量，分別針對條件分配是常態、t 分配、以及 Markov Switching AR(1)-GARCH(1,1) 模型的標準化殘差和標準化殘差平方兩序列，結果 Markov Switching GARCH 模型在條件變異數方程式表現得最好，顯示條件變異數方程式有了結構改變的考量，確實讓 AR(1)-GARCH(1,1) 設定將台股報酬率資料解釋得最好。

參考文獻

中文部份

- 林建甫、張焯然 (1997)。GARCH模型條件變異數結構變動的檢定。《經濟論文》，25：2，201-225。
- 高櫻芬、呂仁廣、林建甫 (2001)。變異數結構改變的SWARCH模型估計：台灣股價報酬之實證研究。《證券市場發展季刊》，13：1，63-98。

英文部分

- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 301-327.
- Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics*, 12, 309-316.
- Chu, C. J. (1995). Detecting Parameter Shift in Generalized Autoregressive Conditional Heteroskedasticity Models. *Econometric Review*, 14:2, 241-266.
- Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica*, 50, 987-1008.
- Gray, S. F. (1996). Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process. *Journal of Financial Economics*, 42, 27-62.
- Haas, M., Mitnik, S., & Paoletta, M. S. (2004). A New Approach to Markov-Switching GARCH Models. *Journal of Financial Econometrics*, 2, 493-530.
- Hamilton, J. D. (1994). Time Series Analysis, Princeton. NJ: Princeton University Press.
- Hamilton, J. D., & Susmel, R. (1994). Autoregressive Conditional Heteroskedasticity and Changes in Regime. *Journal of Econometrics*, 64, 307-333.
- Kokoszka, P., & Leipus, R. (1999). Testing for parameter changes in ARCH models. *Lithuanian Mathematical Journal*, 39, 231-247.
- Kokoszka, P., & Leipus, R. (2000). Change-point estimation in ARCH models. *Bernoulli*, 6, 1-28.
- Lamoureux, C., & Lastrapes, W. (1990). Persistence in Variance, Structural Change and the GARCH Model. *Journal of Business and Economic Statistics*, 8, 255-234.
- Lin, C. F., & Terasvirta, T. (1994). Testing the Constancy of Regression Parameters against Continuous Structural Change. *Journal of Econometrics*, 62, 211-228.

