以無資料窺探偏誤的檢定評估共同基金績效

莊惠菁*　　　　管中閔
國立臺灣大學財務金融學系　　國立臺灣大學財務金融學系

摘要

本文採用新的計量檢定方法來評估臺灣共同基金的績效是否具有統計上的顯著性。爲避免在實證方法中常見的資料窺探 (data snooping) 所造成的推論錯誤, 我們採用 Hsu, Hsu, and Kuan (2008) 的逐步檢定法來認定具有顯著績效的基金。實證結果顯示, 臺灣股票型共同基金於 2002 到 2007 年間, 若以基金月報酬均數爲績效評估標準時, 沒有任何基金顯著的高於臺股加權指數等的大盤報酬, 而只有一至二支基金的夏普值顯著的超越大盤之夏普值。當比較標準爲三因子模型之報酬時, 只有不到百分之三的基金存在顯著的異常報酬。

關鍵詞: 共同基金、眞實性檢定、SPA 檢定、資料窺探、逐步檢定法

JEL **分類代號**: C12, C52, G11, G23

*聯繫作者: 莊惠菁, 國立臺灣大學財務金融學系, 106 臺北市羅斯福路四段一號。電話: (02) 3366-1072; E-mail: d94723007@ntu.edu.tw。

1 前言

共同基金的績效表現向來是財金實務界和學術界關注的焦點之一。在基金表現的評估方式上, 一般採用基金的月報酬率, 夏普值 (Sharpe ratio), 或是因子模型的異常報酬作爲績效指標, 然後根據指標值與比較基準的差距來判斷該基金是否值得投資。許多知名基金評比的網站會定期公佈這些績效指標值, 並據此將基金排名, 以供大衆做爲投資參考。[1] 然而這些績效指標是否眞的具有統計上的顯著性? 抑或只是基金在評估期間的運氣特別好, 而使得表現特別突出? 對於這些問題, 我們需要藉助適當的計量檢定方法來鑑別基金績效表現的眞實性。

由於財務資料難以重複實驗的侷限性, 個別的研究者通常利用相同的資料進行研究, 卻忽略所進行的個別檢定統計量之間的相關性, 而僅對部分的個別檢定結果做推論。文獻上將這種重複利用相同資料進行研究, 所產生對資料的過度推論或是推論錯誤, 稱之爲資料窺探 (data snooping) 問題。Lo and MacKinlay (1990) 證明, 若研究者事先依據以前研究推論所得的性質對資料加以排序, 所得到的統計量的性質將會改變, 進而影響研究者對參數顯著與否的判斷。Sullivan et al. (2001) 的研究顯示, 道瓊工業股價指數中著名的日期效應僅是對部分統計量做個別推論, 沒有考慮他們之間相關性之下的推論偏誤。Kadlec et al. (2006) 考慮各種不同的解釋變數, 被解釋變數和預測樣本內時間長度的排列組合之下, 發現資料窺探的問題存在於許多資產價格可預測性的研究文獻中。

當模型 (基金或是交易策略) 數目衆多時, 因資料窺探而產生的偏誤會更爲嚴重。舉例而言, 當虛無假設爲個別基金的報酬表現不如比較基準的報酬, 我們可以針對個別基金的績效, 分別執行顯著水準爲 5% 的檢定。若只有一支基金, 錯誤拒絕虛無假設的機率爲 5%; 但是當基金數爲 50, 且個別檢定統計量之間相互獨立時, 至少有一個錯誤拒絕虛無假設的機率增加爲 92.31% ($= 1 - (0.95)^{50}$)。換言之, 當基金數很多時, 我們幾乎總是可以找到一個比基準報酬表現優異的基金, 而實際上此基金的表現可能並不如比較基準。由此可知, 重複使用個別檢定很容易產生推論偏誤的問題。

在共同基金的分析中, 資料窺探偏誤的問題事實上更加複雜。不同基金可能投資於相

[1] 如臺灣共同基金績效評比: `http://140.112.111.12`; 基智網: `http://www.funddj.com`; 理柏: `http://www.lipperweb.com`; 晨星: `http://tw.morningstar.com` 等知名基金評比網站。

同的金融商品, 或者因爲從衆效果 (herding effect) 而投資於具有相同走勢之金融商品, 所以不同基金的表現並非相互獨立。Wermers (1999) 利用 1975 至 1994 年間美國的共同基金資料發現, 機構投資人 (或基金經理人) 對於小型股存在了顯著的從衆效果; 在國內基金市場亦有類似的討論, 如李春安與劉維琪 (2006) 與呂素蓮與李世英 (2008)。所以當我們利用基金的平均報酬 (或其他績效指標) 爲檢定統計量時, 忽略統計量之間的相關性亦造成統計上的推論偏誤。

由於資料窺探的問題不容忽視, White (2000) 提出「眞實性檢定法」(reality check, 以下簡稱 RC 檢定法)。RC 檢定法首先要盡可能將所有模型設定納入檢定問題中, 以避免只對部分模型結果做推論可能產生的問題。然後, RC 檢定法透過自我重抽樣 (bootstrap) 法建構虛無分布, 允許個別檢定之間存在相關性。這種檢定方法避免了資料窺探的問題, 使實證研究者得以正確判斷模型的眞實能力。然而, Hansen (2005) 指出 RC 檢定法的檢定力容易受到較差的模型影響, 因此提出了「卓越預測能力檢定法」(superior predictive ability test, 以下簡稱 SPA 檢定法) 作爲修正。SPA 檢定法採用標準化的統計量, 並利用更多樣本資訊重新調整虛無假設分布的位置, 從而提高了檢定力。

RC 和 SPA 檢定法的目的都在於衆多的模型之下, 驗證最佳的模型表現是否顯著優於比較基準模型。但在實務上, 除了表現最佳的基金外, 我們也常常希望知道是否還存在其他優於比較基準的模型。Romano and Wolf (2005) 根據 RC 檢定法提出了逐步多重檢定法 (以下簡稱爲 Step-RC 檢定法)。在控制錯誤拒絕發生的機率之下, Step-RC 檢定法可以透過多重階段, 逐步找出所有顯著優於比較基準的模型。然而, Step-RC 檢定法的基礎是 RC 檢定法, 所以其檢定力在各階段同樣容易受到較差模型影響。Hsu, Hsu, and Kuan (2008) 因此提出了 Step-SPA 檢定法, 將各階段的檢定以 SPA 檢定法的方式加以調整。他們的數學證明與模擬結果均顯示, Step-SPA 檢定法可以增加 Step-RC 的檢定力。

本文主要目的在於評估 2002 至 2007 年間台灣的股票型共同基金的績效。在國內文獻中, 林修葳與王佳眞 (2003), 王佳眞與徐辜元宏 (2004) 等利用績效指標的自我相關性來評估基金表現的持續性。我們則根據不同的績效指標與比較基準, 利用上述的 Step-SPA 檢定法來認定具有顯著績效的基金, 並分析基金於評估期間後表現的持續性。本文中我們

選擇臺股加權股價指數, 摩根史坦利資本國際 (MSCI) 的臺灣指數, 臺灣證券交易所的臺灣 50 (TW50) 指數當作比較基準的大盤報酬。實證結果顯示, 若以基金月報酬均數爲績效指標時, 沒有任何基金顯著的高於臺股加權指數等的大盤報酬, 而只有一至二支基金的夏普值顯著的超越大盤之夏普值。當比較標準爲三因子模型之報酬時, 只有不到百分之三的基金存在顯著的異常報酬。

本文結構如下。第二節介紹上述的檢定方法與其在財務實證上的應用; 第三節說明基金績效指標選擇; 第四節討論實證結果; 最後爲結論。

2 檢定方法與應用

本節中我們將詳細介紹 SPA 與 Step-SPA 檢定方法的應用及執行步驟。這些檢定不僅可以用來檢定基金績效的眞實性, 也可檢定衆多模型的預測能力, 或衆多交易策略的獲利能力。相關的實證研究可參見 Hansen and Lunde (2005), Hsu and Kuan (2005), 與 Qi and Wu (2006) 文中的討論。

令 $x_{t,j}$ 爲 t 時點 j 模型的資料, $x_{t,\phi}$ 爲比較基準模型的資料。我們想要檢定的未知母體參數爲 θ_j。當 θ_j 大於零時, 我們認定此模型表現超越比較基準模型; 而當 θ_j 小於零時, 我們認爲模型表現不如比較基準模型。舉例而言, 若我們想比較不同股票報酬模型的預測能力, 可定義 $x_{t,j} = L_{t,\phi} - L_{t,j}$, 其中 $L_{t,\phi} := L(\hat{r}_{t,\phi}, r_t)$ 爲一損失函數下, 比較基準模型之樣本外預測值 $\hat{r}_{t,\phi}$ 與實際樣本外報酬值 r_t 間的差距 (預測誤差), 而 $L_{t,j} := L(\hat{r}_{t,j}, r_t)$ 爲模型 j 之預測誤差。當 θ_j 爲 $x_{t,j}$ 的母體均數時, θ_j 大於零顯示, 平均而言, 比較基準之預測誤差大於模型 j 之預測誤差; θ_j 小於零表示 j 模型之預測能力不如比較基準模型。

除此之外, 在技術分析獲利能力的分析中, 可令 $x_{t,j} = \delta_{j,t-1} r_t$ 爲 j 策略下的實現報酬, 其中 $\delta_{j,t-1}$ 爲採用 j 策略時的交易訊號, 其值可能爲 1 (買進), 0 (賣出) 或 -1 (放空), 而 r_t 爲標的股票或基金等報酬。當 θ_j 爲 $x_{t,j}$ 之母體均數且大於零時, 則此交易策略平均而言會產生正的報酬; 反之, 此交易策略不會產生正的報酬。類似於技術分析的文獻, 我們亦可將基金的報酬視爲透過某一特定模型或交易策略之下實現的結果。本文中的 $x_{t,j}$ 定義爲基金 j 的報酬 (或是超額報酬), $x_{t,\phi}$ 爲臺股加權指數等的大盤指標之月報酬, θ_j 爲報酬之均數、夏普值或是異常報酬。

SPA 與 RC 檢定法的虛無假設爲, 沒有任何模型在以 θ_j 爲績效的衡量指標下優於基準模型:

$$H_0: \max_{j=1,\cdots,S} \theta_j \le 0;$$

拒絕虛無假設表示, S 種模型中至少有一個優於基準模型。令 $\hat{\theta}_j$ 與 $\hat{\sigma}_j^2$ 爲 θ_j 與 $\mathrm{Var}(\sqrt{T}\hat{\theta}_j)$ 的估計式。SPA 法的檢定統計量爲

$$T_{SPA} := \max\left(\max_{j=1,\cdots,S} \frac{\sqrt{T}\hat{\theta}_j}{\hat{\sigma}_j}, 0\right).$$

White (2000) 的 RC 檢定統計量則爲沒有標準化的統計量: $T_{RC} := \max_{j=1,\cdots,S} \sqrt{T}\hat{\theta}_j$。然而 Hansen (2005) 與 Romano and Wolf (2005) 均建議採用標準化的統計量, 使個別的 $\hat{\theta}_j/\hat{\sigma}_j$ 有相同的衡量尺度, 從而提昇檢定法在小樣本之下的檢定能力。我們在下一節中選擇績效指標作爲參數 θ_j 時, 將一併討論參數估計式和其標準差估計式的計算方式。

在虛無假設下, 檢定統計量的極限分布會受到未知參數的影響, 因此 RC 與 SPA 檢定法利用了自我重抽樣法來估計虛無分布。依循相同的做法, 我們採用了 Politis and Romano (1994) 的定態自我重抽樣 (stationary bootstrap) 方法來獲得原始資料的 B 個重製樣本 (replicated samples), $\{x_{t,j}^b\}_{b=1}^B$。定態自我重抽樣法與 Efron (1979) 的自我重抽樣法假設不同。Efron (1979) 假定資料型態是獨立且同態分布, 故可以利用抽出單一個樣本點後放回的方式重製樣本; 定態重抽樣法則以區塊 (block) 的抽出放回方式重製樣本, 以保持原始資料中的自我相關特性, 其中每一區塊包含的樣本點的數目服從參數爲 q 的幾何分布, 而 q 值代表了區塊期望長度的倒數。因此, 若 q 爲 1, 定態自我重抽樣法等同於 Efron (1979) 的重抽樣法。如果資料自我相關程度越強, 應使用越小的 q 值, 使得平均每一區塊中包含的樣本點數較多, 以捕捉這些相關性。爲使讀者清楚我們的作法, 我們將詳細的定態自我重抽樣法的執行方式列於附錄一。

給定重製樣本 $\{x_{t,j}^b\}_{b=1}^B$ 下, 我們可以計算每一重製樣本下的檢定統計量, 並且利用其經驗分布 (empirical distribution) 來估計虛無假設下檢定統計量的眞實分布。定義由第 b 個重製樣本計算所得的 SPA 檢定統計量爲

$$T_{SPA}^b := \max\left(\max_{j=1,\cdots,S} \frac{\sqrt{T}Z_j^b}{\hat{\sigma}_j^b}, 0\right),$$

其中 $Z_j^b := \hat{\theta}_j^b - \hat{\theta}_j \mathbf{1}_A$, $A := \left\{\sqrt{T}\hat{\theta}_j \geq -\sqrt{2\,\hat{\sigma}_j^2 \log\log T}\right\}$, $\hat{\theta}_j^b$ 與 $\hat{\sigma}_j^b$ 分別是利用第 b 個重製樣本計算而得的參數估計值與其標準差的估計值。$\mathbf{1}_A$ 爲指標函數; 若 A 事件成立, 則 $\mathbf{1}_A = 1$; 反之爲 0。如果 $Z_j^b = \hat{\theta}_j^b - \hat{\theta}_j$, 則上述統計量簡化爲 RC 檢定法的計算方式。重複上述的步驟於 B 個重製樣本, 我們可以計算其中 T_{SPA}^b 大於 T_{SPA} 的相對次數, 也就是利用自我重抽樣所得的 p 值:

$$p_{SPA} = \sum_{b=1}^{B} \frac{\mathbf{1}_{\{T_{SPA}^b > T_{SPA}\}}}{B}. \tag{1}$$

我們也可以將所有的 T_{SPA}^b 值排序後, 然後根據不同的尾端機率, 找到對應的自我重抽樣的臨界值, 並以此作爲拒絕虛無假設的判準。

比較 SPA 與 RC 檢定法, 除了是否採用標準化檢定統計量之外, SPA 檢定法的特色在於, 利用一個和樣本資訊有關的指標函數 $\mathbf{1}_A$ 來調整所估計之虛無分布的中心位置。因爲虛無假設爲一不等式之假設, 所以會受到未知母體參數的影響。透過這指標函數的調整, SPA 檢定法所估計的自我重抽樣下的虛無分布, 將較 RC 檢定法更接近虛無假設成立下的眞實分布。我們將在本段之最後討論 SPA 與 RC 檢定法之理論性質。

在實際的應用中, 研究者不僅想知道最佳之模型 (基金, 或是交易策略) 是否優於比較基準, 更希望能夠找出所有優於比較基準的模型。但前述之 SPA 或 RC 檢定法, 只能用來驗證表現“最佳”的模型是否眞實的優於比較基準模型。而 Step-RC 或 Step-SPA 檢定法則能透過不同階段, 逐步找出“所有”顯著優於比較基準的模型。作法上, Step-RC 或 Step-SPA 須先在檢定的第一階段決定一拒絕標準 (臨界值), 驗證是否有模型顯著的超越比較基準。然後, 刪除第一階段中所有顯著優於比較基準的模型資料, 再利用剩下的資料建構第二階段的新拒絕標準, 並以此標準驗證剩下的模型中是否仍存在優於比較基準的模型。其後的步驟可以類推, 直到無法拒絕任何個別的虛無假設爲止。

令個別假設爲 $H_{0,j} : \theta_j \leq 0$, $j = 1, \cdots, S$, 而 K_m 爲第 m 階段前未被拒絕的模型編號集合 (例如, $K_1 = \{1, \cdots, S\}$)。Hsu, Hsu, and Kuan (2008) 的 Step-SPA 檢定在第 m 階段的檢定步驟如下:

1. 將 $\hat{\theta}_i/\hat{\sigma}_i$, $i \in K_m$, 由大而小排列。

2. 根據 K_m 中的模型計算第 m 個階段的自我重抽樣臨界值: c_m (計算方法請見隨後的討論)。若模型 $i \in K_m$, 且 $\sqrt{T}\hat{\theta}_i/\hat{\sigma}_i$ 大於 c_m 時, 拒絕虛無假設 $H_{0,i}$。

3. 若有個別假設被拒絕, 這些模型將被刪除; 所有剩下的模型則構成新的 K_m。

我們重複執行以上步驟, 直到沒有任何個別假設被拒絕時, 檢定停止。

上述步驟中, 第 m 階段的臨界值 c_m 的計算方式如下。我們利用第 m 階段尚未被拒絕的原始資料 $x_{t,i}$, $i \in K_m$, 以定態自我重抽樣法重製 B 個樣本 $\{x_{t,i}^b\}_{b=1}^B$。舉例而言, 在第一階段時我們用所有資料製作 B 個重製樣本。若第一階段檢定結果拒絕了前五個模型, 則 $K_2 = \{6, \cdots, S\}$ 爲尚未被拒絕的模型編號集合。因此, 當第二階段時, 我們根據 $x_{t,i}$, $i \in K_2$, 的資料再抽出 B 個重製樣本, 以此類推。在第 m 階段, 定義檢定統計量 $T_{SPA}^{b,m}$ 爲

$$T_{SPA}^{b,m} := \max\left(\max_{i \in K_m} \frac{\sqrt{T} Z_i^b}{\hat{\sigma}_i^b}, 0\right),$$

$Z_i^b := \hat{\theta}_i^b - \hat{\theta}_i \mathbf{1}_A$, $A := \left\{\sqrt{T}\hat{\theta}_i \geq -\sqrt{2\,\hat{\sigma}_i^2 \log\log T}\right\}$。臨界值 c_m 爲 $\{T_{SPA}^{b,m}\}_{b=1}^B$ 的第 $1-\alpha$ 百分位數。若我們令 $Z_i^b := \theta_i^b - \hat{\theta}_i$, 則上述方法亦爲得到 Romano and Wolf (2005) 之 Step-RC 臨界值的方式。

最後, 我們簡短的介紹上述檢定方法的極限性質。假定估計式 $\hat{\boldsymbol{\theta}} := (\hat{\theta}_1, \cdots, \hat{\theta}_S)'$ 具有極限常態分布:

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}),$$

其中 $\boldsymbol{\theta} := (\theta_1, \cdots, \theta_S)$ 爲未知的母體參數。因爲虛無假設爲一不等式的假設, RC 檢定法 (或各階段的 Step-RC 檢定法) 依據 least favorable configuration (LFC) 的概念, 假定 $\boldsymbol{\theta} = \mathbf{0}$ 來建構檢定統計量的虛無分布。因此, RC 檢定在虛無假設下爲: $T_{RC} \xrightarrow{D} \max_{j=1,\cdots,S} W_j$, 其中 $(W_1, \cdots, W_S)$ 服從多元常態分布, 均數爲 $\mathbf{0}$ 向量, $\boldsymbol{\Omega}$ 爲變異數矩陣。我們將此性質表示爲 $T_{RC} \xrightarrow{D} \max(\mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}))$。然而 Hansen (2005) 證明, 在虛無假設下, 當 $\boldsymbol{\theta} \leq \mathbf{0}$, 但至少有一 θ_j 爲 0 的情況下, RC 檢定統計量的極限分布應爲 $\max(\mathcal{N}(\mathbf{0}, \boldsymbol{\Omega_0}))$, 其中 $\boldsymbol{\Omega_0}$ 爲 $\boldsymbol{\Omega}$ 矩陣的子矩陣, 且與 $\theta_j < 0$ 的模型無關。因此, 若檢定中包括了許多較差模型 ($\theta_j < 0$ 的模型), 然後根據 LFC ($\boldsymbol{\theta} = \mathbf{0}$) 來估計虛無分布, 則 RC 檢定的檢定力將因這些較差模型而降低。Hansen (2005) 甚至證明, 一旦有太多很差的模型, RC 檢定的檢定力甚至可能爲零。

爲了避免上述的問題, SPA 檢定法 (或各階段之 Step-SPA 檢定法) 透過一指標函數, 重新調整了所估計虛無分布的中心位置, 使其更接近虛無假設成立下的眞實分布。更精確的說, 我們將 SPA 法中重抽樣的的 Z_j^b 改寫爲

$$Z_j^b = \hat{\theta}_j^b - \hat{\theta}_j + \hat{\theta}_j \mathbf{1}_{A^c},$$

其中 A^c 爲 A 事件的補集, 而 $\hat{\theta}_j^b - \hat{\theta}_j$ 爲 RC 檢定中重抽樣的 Z_j^b。Hansen (2005) 證明 $\hat{\theta}_j \mathbf{1}_{A^c}$ 會強收斂 (converges almost surely) 於眞實的 θ_j。因此, 當虛無假設成立時, $\hat{\theta}_j \mathbf{1}_{A^c}$ 的值會小於或等於 0, 所以 SPA 方法下重抽樣的 Z_j^b 會小於 RC 方法下的 Z_j^b。由於這二種檢定方式都是利用 $\{Z_j^b\}_{b=1}^B$ 的特定分位數來建構拒絕標準, 故 SPA 檢定的拒絕標準會低於 RC 的拒絕標準, 而前者的檢定力也因此高於後者。

Hsu, Hsu, and Kuan (2008) 進一步證明 Step-SPA 檢定法具有以下的性質: 若眞實 $\theta_j > 0$, 樣本數趨近於無限大時, Step-SPA 拒絕個別檢定假設 $H_{0,j}$ 的機率趨近於 1。除此之外, 當控制至少有一錯誤拒絕的機率爲 α 時,[2] 則在任何一種 Romano and Wolf (2005) 所定義的檢定力之下, Step-SPA 檢定法的檢定力均高於 Step-RC 檢定法。關於上述 Step-SPA 檢定法更詳細的討論與數學證明, 請參見 Hsu, Hsu, and Kuan (2008)。

3 基金績效指標

我們考慮三種不同的基金績效衡量指標, 分別爲基金淨值的月報酬, 風險調整的夏普值, 與三因子模型下基金的異常報酬。令 $r_{t,j}$ 爲基金 j 的第 t 月月報酬, $r_{t,\phi}$ 爲比較基準之月報酬, 無風險利率爲 $r_{t,f}$。

當月報酬爲績效指標時, 上一節中檢定法的原始資料爲 $x_{t,j} = r_{t,j}$。在假定報酬資料爲定態序列之下, 報酬之母體均數爲 $\mu_j := \mathrm{E}(r_{t,j}), \forall t$, 則我們感興趣的檢定參數爲基金報酬平均數與比較標準報酬均數的差, 即 $\theta_j := \mu_j - \mu_\phi$。令 $\bar{x}_j$ 爲樣本平均數, 則 $\hat{\theta}_j = \bar{x}_j - \bar{x}_\phi$。

[2] 在多重檢定 (multiple test) 中,「至少有一錯誤拒絕的機率」(familywise error rate) 一般被定義爲 $\mathbb{P}(\text{reject at least one } H_{0,j} : j \in I_0)$, I_0 爲眞實假設的集合。當眞實假設的集合中只有一個假設時, 至少有一錯誤拒絕的機率等同於一般簡單檢定 (simple test) 中的型一誤差。

當我們採用定態自我重抽樣時, 依據 Hansen (2005) 的建議, 變異數估計式爲

$$\hat{\sigma}_j^2 = \frac{1}{T}\left[\hat{\gamma}_{0,j} + 2\sum_{i=1}^{T-1}\kappa(T,i)\,\hat{\gamma}_{i,j}\right], \tag{2}$$

式中 $\hat{\gamma}_{i,j} = \sum_{l=1}^{T-i}(d_{l,j}-\bar{d}_j)(d_{l+i,j}-\bar{d}_j)/T$, 而 $d_{t,j} := x_{t,j} - x_{t,\phi}$,

$$\kappa(T,i) = \frac{T-i}{T}(1-q)^i + \frac{i}{T}(1-q)^{T-i},$$

q 爲定態自我重抽樣法下幾何分布的參數。

以夏普値爲績效指標時, 檢定的原始資料爲基金之超額報酬: $x_{t,j} = r_{t,j}^e = r_{t,j} - r_{t,f}$。我們感興趣的檢定參數爲基金與大盤的夏普値差距:

$$\theta_j = \frac{\mu_j - \mu_f}{s_j} - \frac{\mu_\phi - \mu_f}{s_\phi},$$

其中 $s_j^2 = \mathrm{Var}(r_{t,j}), \forall t$。利用相對應的樣本平均數與樣本變異數來估計母體平均數和母體變異數, 可以得到 $\hat{\theta}_j$。爲了計算標準化的 SPA 檢定統計量, 必須先得到變異數估計式 $\hat{\sigma}_j^2$。我們採用類似於 Lo (2002) 與 Ledoit and Wolf (2008) 標準化資產報酬夏普値的方法來估計變異數; 細節請見附錄二的討論。

若是以三因子模型的異常報酬作爲基金績效評估標準時, 檢定的原始資料爲基金與大盤之超額報酬與因子之報酬。我們依循 Fama and French (1993) 的做法建構台灣市場之公司規模因子報酬 $r_{t,SMB}$、帳面/市値因子報酬 $r_{t,HML}$, 與市場投資組合超額報酬 $r_{t,\phi}^e := r_{t,\phi} - r_{t,f}$。詳細的原始資料與因子建構方式, 請見附錄三。對每一基金之超額報酬 $r_{t,j}^e$, 三因子模型爲

$$r_{t,j}^e = \alpha_j + \beta_{1j} r_{t,\phi}^e + \beta_{2j} r_{t,SMB} + \beta_{3j} r_{t,HML} + \varepsilon_{t,j}; \quad t = 1, \cdots, T.$$

我們所欲檢定的參數 θ_j 爲 α_j; 若 α_j 大於 0, 此基金報酬有高於三因子報酬的異常報酬存在。參數估計式 $\hat{\theta}_j$ 爲利用迴歸式估計的截距項 $\hat{\alpha}_j$。爲了計算標準化的 SPA 檢定統計量, 我們採用 HAC (heteroscedasticity and autocorrelation consistent) 的變異數估計式來估計 $\hat{\sigma}_j^2$, 以避免對誤差項的變異和自我相關性質做過多的假設。下一節實證研究中的 HAC 計算利用了 Bartlett 核函數 (kernel function), 且依照 Newey and West (1994) 文中建議的算法自動選擇頻寬 (bandwidth)。

4 實證結果

4.1 實證資料

本文的實證研究中, 基金淨值的月資料來源為臺灣經濟新報資料庫 (以下簡稱新報資料庫)。研究期間為 2002 年 11 月底至 2007 年 10 月底; 在研究期間結束前未清算或合併的國內股票與平衡型共同基金共 220 支, 每一支基金有 60 筆月資料。令 $r_{t,j}$ 為新報資料庫中股利調整後的基金淨值月報酬, $t = 1, \cdots, 60$, $j = 1, \cdots, 220$。我們選擇臺股加權指數, MSCI 臺灣指數, 與 TW50 指數的月報酬為基金績效的比較基準。本文中除了 TW50 指數資料來源為臺灣證券交易所網站外, 其餘資料皆來自於新報資料庫。比較基準指數之月報酬定義為 $r_{t,\phi}$。為計算超額報酬, 本文之無風險利率 $r_{t,f}$ 為可轉讓定期存單初級市場 1–90 天月息。[3]

表 1 為臺股加權, MSCI 臺灣, 與 TW50 指數月報酬的基本敍述統計量。我們發現臺股加權指數月報酬的平均報酬與夏普值, 都較 MSCI 臺灣指數與 TW50 指數報酬來得高。除此之外, 所有比較基準指數的月報酬均有輕微左偏的現象。計算 Phillips and Perron (1988) 的單根檢定發現, 比較基準指數的月報酬均在 1% 的顯著水準之下拒絕具有單根的假設。

表 2 為基金之超越大盤概況表。表 2 的上半部中,「總數」代表了基金的績效指標數值超越特定大盤數值的基金支數,「百分比」為總數佔所有基金支數的比例。舉例而言, 若以臺股加權月報酬為比較基準, 有超過 69% 的基金超越臺股加權月報酬; 若以夏普值為比較基準, 約有 61% 的基金超越臺股加權指數; 而三因子模型有 80% 的基金具有超越以臺股加權為市場因子的異常報酬。綜合而言, 若單純以績效指標的絕對數字來判斷國內共同基金中很大一部分在 2002–2007 年的績效都超越比較基準。

我們也對個別基金執行 t 檢定。檢定統計量為 $t_j := \sqrt{T}\hat{\theta}_j/\hat{\sigma}_j$。在顯著水準為 α 下, 利用常態分布的 $1-\alpha$ 百分位數 (即 $\Phi^{-1}(1-\alpha)$) 當作拒絕標準。若 $t_j > \Phi^{-1}(1-\alpha)$, 則認定 j 基金在顯著水準 α 下, 績效顯著優於比較基準。表 2 的下半部為執行個別 t 檢定的結果。當顯著水準為 5%,「顯著數」為 t 檢定統計量顯著的基金數目,「百分比」為顯

[3]我們也曾以平均銀行一個月定存利率作為無風險利率, 但實證的檢定結果與本文所報告的結果之間並無顯著的差異。

表 1: 指數月報酬的基本敍述統計量

	臺股加權	MSCI	TW50
平均數	1.366	1.208	1.278
標準差	4.971	5.085	4.926
夏普值	0.252	0.215	0.236
極大值	12.640	11.920	12.640
極小值	−11.620	−11.920	−11.620
偏態系數	−0.028	−0.053	−0.119
峰態系數	2.551	2.590	2.668
PP 統計量	−7.446	−7.854	−7.735
樣本數	60	60	60

註: PP 統計量爲 Phillips and Perron (1988) 單根檢定, 結果均在 1% 的顯著水準之下拒絕單根假設。

著的總數佔總基金數的比例。由表中我們可以看到, 若以 t 檢定作爲判準, 基金超越大盤的比例顯著的降低。例如, 以臺股加權指數爲比較基準, 只有 5% 的基金月報酬顯著優於臺股加權, 約 7% 的基金夏普值顯著超越比較基準的夏普值, 而約 21% 的基金具有異常報酬。

4.2 SPA 與 Step-SPA 檢定結果

我們再根據 SPA 與 Step-SPA 檢定來判斷基金績效的顯著性。這些檢定法的執行程式均利用 MATLAB 撰寫, 其中定態自我重抽樣的重製樣本數 $B = 1000$。爲了避免定態自我重抽樣法中參數 q 值選定的任意性, 我們利用 Politis and White (2004) 與 Patton et al. (2009) 的方式, 爲每一筆基金報酬資料分別估計最適的 q 值。這方法的假設是當參數 θ_j 爲樣本平均數時, 極小化由式 (2) 的 $\hat{\sigma}_j^2$ 估計眞實變異數 $\mathrm{Var}(\sqrt{T}\hat{\theta}_j)$ 所造成的均方誤差, 從而求得的最適 q 值。根據此方法, 我們發現大部分基金的最適 q 值都大於 0.285。[4] 因此我們以下會列出不同 q 值 $(0.33, 0.4, \cdots, 1$, 對應於期望區塊長度爲 $3, 2.5, \cdots, 1$ 個

[4]本文估計的 q 僅做爲參考之用。就作者所知, 目前文獻對定態自我重抽樣法下 q 值的選定, 只有在單變量資料下有一些分析討論 (如 Politis and White, 2004; Patton et al., 2009), 但在多變量資料中該如何選取 q, 文獻尚無相關討論。與本文類似的實證研究大多採用給定的 q 值, 如 0.5 或 0.9; 參見 Hansen and Lunde (2005), Hsu and Kuan (2005), 與 Qi and Wu (2006)。多變量資料下最適 q 值選定的問題是重要的研究議題, 但超出了本文的範圍。

表 2: 基金績效指標超越大盤概況

	臺股加權		MSCI		TW50	
	總數	百分比	總數	百分比	總數	百分比
平均數	153	69.545	166	75.454	159	72.272
夏普值	134	60.909	176	80.000	154	70.000
異常報酬	176	80.000	178	80.909	147	66.818
	顯著數	百分比	顯著數	百分比	顯著數	百分比
平均數	11	5.000	28	12.727	23	10.454
夏普值	15	6.818	24	10.909	16	7.272
異常報酬	46	20.909	47	21.363	29	13.181

註: 表中總數代表一績效指標下超越大盤的基金數目; 顯著數爲顯著水準爲 5% 之下, t 檢定顯著的基金個數。

月) 之下的檢定結果。

SPA 檢定結果列於表 3。與直接比較績效指標或是個別檢定所得到的結論不同的是, 若以平均報酬爲績效評估指標, 不論比較基準指數爲何 (也不論 q 值爲何), 我們發現即使在 10% 的顯著水準下, 並無任何基金顯著優於大盤指數月報酬。若是績效指標爲夏普值時, 不論比較基準指數爲何 (在大部分的 q 值之下), 基金編號 53 在顯著水準 10% 或是 5% 之下, 顯著超越大盤之夏普值。當利用三因子之異常報酬爲績效衡量標準時, 於 10% 的顯著水準下, 53 號基金在 MSCI 或 TW50 爲市場因子報酬時, 具有異常報酬。所以, SPA 檢定結果顯示, 當績效指標選定爲夏普值或異常報酬時, 的確有顯著優於大盤表現的基金存在。

我們更進一步利用 Step-SPA 檢定法來認定是否還有更多超越大盤績效的基金。由於 SPA 檢定就是 Step-SPA 第一步的檢定, 所以我們已經知道, 當績效指標選定爲月報酬均數時, 並無任何基金能在 10% 的顯著水準下顯著的超越大盤表現。若以夏普值爲基金績效指標, 在 5% 的顯著水準之下, Step-SPA 檢定方式還是只發現編號 53 號基金能超越大盤之夏普值。但在顯著水準爲 10% 時, 不論於任何 q 值下, 我們發現 53 號與 18 號基金都顯著優於 MSCI 指數之夏普值。除此之外, 10% 顯著水準下, Step-SPA 與 SPA 檢定的結果相同。

表 4 是以三因子異常報酬爲績效指標的 Step-SPA 檢定結果。表中以斜線分隔的基金

表 3: 五年基金績效的 SPA 檢定結果

	q	臺股加權 p_{SPA}	臺股加權 基金編號	MSCI p_{SPA}	MSCI 基金編號	TW50 p_{SPA}	TW50 基金編號
平均報酬	1.000	0.305	123	0.196	182	0.238	182
	0.666	0.250	182	0.156	182	0.160	182
	0.500	0.226	182	0.134	182	0.144	182
	0.400	0.251	182	0.127	182	0.150	182
	0.333	0.263	182	0.139	182	0.165	182
夏普值	1.000	0.082*	53	0.060*	53	0.042**	53
	0.666	0.143	53	0.048**	53	0.073*	53
	0.500	0.087*	53	0.060*	53	0.053*	53
	0.400	0.080*	53	0.056*	53	0.043**	53
	0.333	0.092*	53	0.036**	53	0.050*	53
異常報酬	1.000	0.088*	53	0.084*	53	0.126	53
	0.666	0.104	53	0.068*	53	0.128	53
	0.500	0.120	53	0.094*	53	0.090*	53
	0.400	0.079*	53	0.085*	53	0.077*	53
	0.333	0.066*	53	0.101	53	0.075*	53

註: 表中 q 為定態自我抽樣法的參數值; p_{SPA} 為 (1) 式的自我重抽樣 p 值; * 與 ** 分別表示在顯著水準 10% 與 5% 之下顯著。

表 4: 五年基金績效的 Step-SPA 的異常報酬檢定結果

q	臺股加權	MSCI	TW50
顯著水準 5%			
1.000	53, 182, 18	53, 182, 18	53, 182
0.666	53, 182, 18	53, 182/ 18	53, 182
0.500	53, 182, 18	53, 182/ 18	53, 182
0.400	53, 182, 18	53, 182/ 18	53, 182
0.333	53, 182, 18	53, 182/ 18	53, 182
顯著水準 10%			
1.000	53, 182, 18, 97, 26	53, 182, 18, 167, 97/ 158	53, 182, 97
0.666	53, 182, 18, 97, 26	53, 182, 18, 167, 97, 158	53, 182, 97
0.500	53, 182, 18, 97/ 26	53, 182, 18/ 167, 97, 158	53, 182/ 97
0.400	53, 182, 18, 97/ 26	53, 182, 18/ 167, 97, 158	53, 182, 97
0.333	53, 182, 18, 97, 26	53, 182, 18, 167, 97/ 158	53, 182, 97

註: q 爲定態自我抽樣法的參數值; 表中數字代表績效顯著的基金編號, 以斜線分隔的基金編號代表不同階段之下拒絕虛無假設的基金編號。

編號代表在不同階段之下拒絕虛無假設的基金編號。舉例而言, 在 10% 的顯著水準之下, $q = 0.5$, 市場報酬爲 TW50 指數時, 第一階段檢定認定 53 與 182 號基金具有顯著之異常報酬; 第二階段認定基金編號 97 具有顯著之異常報酬。由表 4 我們發現, 在顯著水準爲 5% 時, 53, 182 和 18 號三支基金在臺股加權與 MSCI 指數爲市場因子時, 均能顯著的拒絕虛無假設。而若以 TW50 爲市場因子時, 53 和 182 號基金具有異常報酬。若是我們將顯著水準變爲 10% 時, 當臺股加權爲市場因子時可以認定出 5 支基金具有異常報酬, 6 支基金在 MSCI 爲市場因子下具有異常報酬, 而只有 3 支基金在 TW50 爲市場因子下具有異常報酬。這些數字都遠低於個別檢定 (表 2) 所顯示的結果。

綜合而言, 當績效指標選定爲夏普值或是三因子模型的異常報酬時, Step-SPA 檢定法不論於第一階段或是後續的第二階段, 確實可以較 SPA 檢定法認定出更多優於比較基準的基金。而這些結果也顯示, 若僅直接比較績效指標, 或是藉由個別檢定來判斷基金績效, 我們可能會誤以爲許多基金都有超越大盤的表現。

表 5: 四年基金績效的 Step-SPA 檢定結果

	q	臺股加權	MSCI	TW50
夏普值	1.000	None	53	53
	0.666	53	53	53
	0.333	None	53	53
異常報酬	1.000	182, 53, 18	182, 53, 203	182, 53
	0.666	182, 53, 18	182, 53, 203	182, 53
	0.333	182, 53, 18	182, 53, 203	182, 53

註: 表中 q 爲定態自我抽樣法的參數值。表中數字代表 5% 之顯著水準下, 顯著優於大盤的基金編號。績效評估期間爲 2002 年 11 月至 2006 年 10 月。

4.3 基金績效的持續性

我們最後分析基金績效的持續性。除了前面所做的五年基金績效分析之外, 我們也利用 Step-SPA 檢定來評定基金四年績效表現 (2002 年 11 月至 2006 年 10 月)。我們也將對四年與五年績效中都有卓越表現的基金, 分別評估其持有一年的報酬或夏普值與大盤的差異。

以月報酬均數爲績效指標時, 我們發現沒有任何基金的四年績效顯著優於大盤表現; 此結果與五年績效的檢定結果相同。表 5 爲 5% 顯著水準下, 基金四年績效 (夏普值與異常報酬) 的 Step-SPA 檢定結果; 爲了節省篇幅, 我們只列出 $q = 0.33, 0.66, 1$ 的結果。由表中可以看出, 若以夏普值爲績效指標, 53 號基金不論在任何 q 之下均顯著優於 MSCI 與 TW50 之夏普值。若以三因子報酬爲比較標準, 我們發現當臺股加權爲市場因子, Step-SPA 檢定會指出 182, 53, 18 號基金具有顯著的異常報酬; 當 TW50 爲市場因子時, Step-SPA 只發現 182, 53 號基金有異常報酬。這些都與五年績效的檢定結果相同。當 MSCI 爲市場因子時, Step-SPA 檢定發現 182, 53, 203 號基金具有異常報酬。雖然五年績效的檢定認定的是 182, 53, 18 號基金, 但若將顯著水準改爲 10%, 則可發現 18 號基金的四年績效亦具有異常報酬。換言之, 根據基金四, 五年績效所做的檢定, 結果大體吻合。

我們於是對在四年與五年中都有顯著績效的 53, 182 與 18 號基金, 進一步評估其樣本外績效的持續性。我們選擇三種樣本外的分析期: 期間一爲四年後再持有基金一年 (2006

表 6: 基金的樣本外績效

績效指標	基金編號	臺股加權 期間一	臺股加權 期間二	臺股加權 期間三	MSCI 期間一	MSCI 期間二	MSCI 期間三	TW50 期間一	TW50 期間二	TW50 期間三
平均報酬	53	0.791	-0.726	-0.496	1.260	-0.479	-0.287	1.211	-0.862	-0.510
	182	0.692	0.601	0.624	1.160	0.829	0.833	1.112	0.463	0.600
	18	0.542	1.062	0.579	1.010	1.290	0.788	0.962	0.924	0.560
夏普值	53	0.353	-0.728	-0.498	0.471	-0.499	-0.289	0.447	-0.864	-0.517
	182	0.084	0.603	0.627	0.201	0.831	0.836	0.177	0.467	0.608
	18	0.153	1.059	0.576	0.271	1.288	0.785	0.247	0.923	0.558

註: 期間一爲 2006 年 11 月至 2007 年 10 月, 期間二爲 2007 年 11 月至 2008 年 10 月, 期間三爲 2007 年 11 月至 2009 年 2 月。

年 11 月至 2007 年 10 月), 期間二爲五年後再持有一年 (2007 年 11 月至 2008 年 10 月), 期間三爲五年後持有基金直到最近 (2007 年 11 月至 2009 年 2 月)。表 6 爲 53, 182 與 18 號基金在這三段樣本外期間的績效指標值。以 182 號基金的績效爲例, 在期間一內該基金與 MSCI 的月報酬之間, 12 個月的平均差距爲 1.16, 而夏普值的平均差距爲 0.20, 都是基金高於 MSCI。由表 6 可以看到, 182 與 18 號基金在這三個期間內, 不論在任何比較基準下都有超越大盤的表現。相對而言, 53 號基金雖然在期間一之內可以超越大盤, 但在其他兩個期間的表現卻不如大盤。[5] 這些結果顯示, Step-SPA 檢定在樣本內所認定的基金, 其樣本外的績效具有一定的持續性。

5 結論

如何正確評估共同基金的表現是來自運氣或是基金經理人的操作能力, 一直被認爲是件相當困難的工作。本文是國內文獻中, 首度應用考量過資料窺探問題的計量檢定方法的實證分析。我們採用了 Hsu, Hsu, and Kuan (2008) 的 Step-SPA 逐步檢定法, 透過不同的績效指標與比較基準, 來評估國內共同基金績效的顯著性。實證結果顯示, 國內股票型共同基金在月報酬方面並無任何基金顯著的優於大盤表現; 若以夏普值爲績效指標, 只有一

[5] 經查閱相關資料後發現, 53 號基金在樣本外期間曾經兩度更換經理人, 故該基金的樣本外績效可能與此有關。

至二支基金顯著超越大盤; 而不到百分之三的基金存在異常報酬。所以, 利用適當的計量檢定方法來驗證基金績效的顯著性, 我們得到與傳統作法非常不同的結論。

SPA 與 Step-SPA 檢定法雖可避免一般檢定法在資料窺探上的問題, 但本文的實證仍有一些侷限性。由於基金績效受交易成本等因素的影響, 所以更完整的績效評估應該設法考慮這些因素, 而這是本文中尚未做到的。在這個方向上未來的可能的研究議題還包括: 交易策略的獲利能力, 或是基金因子模型的預測能力等。這些將留待未來作更多與更深入的研究。

附錄一、定態自我重抽樣

爲簡化討論, 我們將原始資料排列爲矩陣 $X_{T\times J}$, 其中每一個元素爲 $x_{t,j}$。利用定態自我重抽樣法方法製作第 b 個重製樣本的方法如下: 從離散均勻分布, 值域爲 $\{1,2,...,T\}$ 的分布中抽出 T 個隨機樣本 $\{u_{b,t}\}_{t=1}^{T}$。從均勻連續值域爲 (0,1] 的分布中抽出 T 個隨機樣本 $\{v_{b,t}\}_{t=1}^{T}$。令 $\mathbf{x}_r := (x_{r,1},\cdots x_{r,J})$ 爲原始資料矩陣的第 r 個列向量; 定義第 b 個重製矩陣第 t 個列向量 $\mathbf{x}_t^b := \mathbf{x}_{\tau_{b,t}}$, $\tau_{b,1} = u_{b,1}$, 當 $t > 1$ 時,

$$\tau_{b,t} = \begin{cases} u_{b,t} & ,\text{if } \ v_{b,t} < q \\ \tau_{b,t-1}\cdot \mathbf{1}_{\{\tau_{b,t-1}<T\}} + 1 & ,\text{if } \ v_{b,t} \geq q. \end{cases}$$

因此, 重製樣本爲 $X_{T\times J}^b := [\mathbf{x}_1^b,\ldots,\mathbf{x}_T^b]'$, 個別元素爲 $x_{t,j}^b$。上面的抽樣方法等同於將許多不同長度的區塊相連結, 而每一個區塊長度爲服從獨立且同分布, 成功機率爲 q 的幾何分布; 每一區塊長度的期望值爲 $1/q$。

其他形式的自我重抽樣法與其統計性質, 尤其是考量過資料形態具有自我相關性質的區塊重抽樣方式, 如移動區塊自我重抽樣法 (moving blocks bootstrap), 循環式自我重抽樣法 (circular bootstrap) 等, 可參見 Lahiri (2003) 專書中的討論。

附錄二、夏普值變異數估計

定義 $\mathrm{E}(r_{t,j}^2) = \gamma_j$, $j = 1,2,\cdots,S,\phi$ 爲基金月報酬與比較基準月報酬的二階動差。令 $Y_t = (r_{t,j}, r_{t,\phi}, r_{t,j}^2, r_{t,\phi}^2, r_{t,f})'$; $\boldsymbol{\beta} = (\mu_j,\mu_\phi,\gamma_j,\gamma_\phi,\mu_f)'$。令 $\boldsymbol{\varphi}_t(Y_t,\boldsymbol{\beta}) = Y_t - \boldsymbol{\beta}$, 由定義, 我們可得母體動差條件 $\mathrm{E}[\boldsymbol{\varphi}_t(Y_t,\boldsymbol{\beta}_0)] = \mathbf{0}$, 其中 $\boldsymbol{\beta}_0$ 爲眞實母體參數。令 $\hat{\boldsymbol{\beta}}_T$ 爲滿足相對應樣本動差條件 $\sum_{t=1}^T \boldsymbol{\varphi}_t(Y_t,\boldsymbol{\beta})/T = \mathbf{0}$ 的解。在一些正規的假設條件之下, 我們可以得到

$$\sqrt{T}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0) \xrightarrow{D} \mathcal{N}(\mathbf{0},\mathbf{\Sigma_0}),$$

其中

$$\mathbf{\Sigma_0} := \lim_{T\to\infty} E\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{\varphi}_t(Y_t,\boldsymbol{\beta}_0)\boldsymbol{\varphi}_s(Y_s,\boldsymbol{\beta}_0)'\right].$$

夏普値的差為 $\boldsymbol{\beta}$ 的連續函數, 可表示為

$$f(\boldsymbol{\beta}) := \frac{\mu_j - \mu_f}{\sqrt{\gamma_j - \mu_j^2}} - \frac{\mu_\phi - \mu_f}{\sqrt{\gamma_\phi - \mu_\phi^2}}.$$

利用一階泰勒展開逼近,

$$\sqrt{T}(f(\hat{\boldsymbol{\beta}}_T) - f(\boldsymbol{\beta}_0)) \xrightarrow{D} \mathcal{N}(0, \nabla_{\boldsymbol{\beta}} f(\boldsymbol{\beta}_0)' \, \boldsymbol{\Sigma}_{\mathbf{0}} \, \nabla_{\boldsymbol{\beta}} f(\boldsymbol{\beta}_0)),$$

$\nabla_{\boldsymbol{\beta}} f(\boldsymbol{\beta}) := (\partial f(\boldsymbol{\beta})/\partial\mu_j, \partial f(\boldsymbol{\beta})/\partial\mu_\phi, \partial f(\boldsymbol{\beta})/\partial\gamma_j, \partial f(\boldsymbol{\beta})/\partial\gamma_\phi, \partial f(\boldsymbol{\beta})/\partial\mu_f)'$。令 $\hat{\boldsymbol{\Sigma}}_T$ 為 $\boldsymbol{\Sigma}_{\mathbf{0}}$ 的估計式, 則變異數估計式 $\hat{\sigma}_j^2$ 為 $\nabla_{\boldsymbol{\beta}} f(\hat{\boldsymbol{\beta}}_T)' \, \hat{\boldsymbol{\Sigma}}_T \, \nabla_{\boldsymbol{\beta}} f(\hat{\boldsymbol{\beta}}_T)$。

為避免對變異數與自我相關結構做過多的假設, 我們利用 HAC 變異數估計方式計算 $\hat{\boldsymbol{\Sigma}}_T$,

$$\hat{\boldsymbol{\Sigma}}_T = \frac{T}{T-5} \sum_{j=-T+1}^{T-1} k\left(\frac{j}{M}\right) \hat{\Gamma}_T(j),$$

$$\hat{\Gamma}_T(j) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} \boldsymbol{\varphi}_t(\hat{\boldsymbol{\beta}}_T)\boldsymbol{\varphi}_{t-j}(\hat{\boldsymbol{\beta}}_T)' & \text{for } j \geq 0 \\ \frac{1}{T}\sum_{t=-j+1}^{T} \boldsymbol{\varphi}_{t+j}(\hat{\boldsymbol{\beta}}_T)\boldsymbol{\varphi}_t(\hat{\boldsymbol{\beta}}_T)' & \text{for } j < 0, \end{cases}$$

其中 $k(\cdot)$ 為核函數, M 為頻寬。本文 HAC 的計算方式為利用 Bartlett 核函數並依照 Newey and West (1994) 文中方法自動選擇頻寬。詳細的 HAC 變異數估計式算法、核函數的選擇、與頻寬之選擇的討論, 可參考 Andrews and Monahan (1992) 與 Newey and West (1994)。

附錄三、因子報酬建構

我們利用新報資料庫中之臺灣財經資料庫的「TEJ Profile」、「TEJ 上市櫃 + 曾經上市櫃公司資料庫」、「TEJ Equity: 上市 (櫃) 調整股價 (月) 除權息調整」、「TEJ 下市、管理股票、全額交割股」等子資料庫之資料建構因子之月報酬。因子報酬的計算中包含上市與上櫃的公司。不同於 Fama and French (1993) 於每年度 6 月底更新組成公司成分, 我們於每年度 12 月底更新一次因子組成之公司, 刪除於資產負債表相關資訊有缺失值, 或月報酬與市值資訊有缺失值, 或是已經下市、下櫃的公司。

我們定義帳面價值爲資產負債表上股東權益總額加上遞延所得稅, 扣除特別股股本之值。市值爲「TEJ Equity」資料庫中各公司之市值資料; 則一公司的「帳面市值比」即爲帳面價值與市值二者相除。我們可利用所有當年度因子組成成分之公司的市值中位數與帳面市值比的 30 與 70 分位數, 將成分公司分爲六類資產組合。分別爲市值大與小 (Big/Small) 的公司, 與帳面市值比屬於高中低 (Value/Neutral/Growth) 的公司。如歸類爲 Small Value (SV) 的公司爲市值小於中位數, 帳面市值比高於 70 分位數的公司; Big Neutral (BN) 組合內公司的市值大於中位數, 且帳面市值比落於 30 與 70 分位數之內。

在每月底, 我們定義 $\bar{r}_{t,\cdot}$ 爲各類資產組合以市值加權之平均報酬。舉例而言, $\bar{r}_{t,BN}$ 爲 BN 資產組合於 t 月以市值加權之平均報酬; 如此一來, 我們可計算公司規模因子與帳面/市值因子報酬。公司規模因子報酬 $r_{t,SMB}$ 定義爲三種市值小之公司組合之報酬平均數減去三種大市值資產組合報酬平均數。而帳面/市值因子報酬 $r_{t,HML}$ 定義爲高帳面/市值比組合與低帳面市值比組合報酬平均數之差異。即

$$\begin{aligned} r_{t,SMB} &:= \frac{1}{3}(\bar{r}_{t,SV} + \bar{r}_{t,SN} + \bar{r}_{t,SG}) - \frac{1}{3}(\bar{r}_{t,BV} + \bar{r}_{t,BN} + \bar{r}_{t,BG}); \\ r_{t,HML} &:= \frac{1}{2}(\bar{r}_{t,SV} + \bar{r}_{t,BV}) - \frac{1}{2}(\bar{r}_{t,SG} + \bar{r}_{t,BG}). \end{aligned}$$

在財務實證研究中, 因子模型的研究文獻衆多。有興趣的讀者可參考 Kenneth R. French 教授個人網頁上的相關資訊與介紹。

參考文獻

王佳眞、徐辜元宏 (2004),「風險值的應用與臺灣共同基金績效指標之持續性」, 臺大管理論叢, 第 14 卷第 2 期, 23-48.

李春安、劉維琪 (2006),「基金從衆、基金聲譽價値與基金折價關係之研究」, 證券市場發展季刊, 第 18 卷第 1 期, 107-142.

呂素蓮、李世英 (2008),「三項分配下共同基金經理人之從衆行爲-以臺灣股票型基金爲例」, 交大管理學報, 第 28 卷第 2 期, 41-72.

林修葳、王佳眞 (2003),「臺灣共同基金績效持續性之研究」, 管理學報, 第 20 卷第 4 期, 655-688.

Andrews, D. W. K. and J. C. Monahan (1992). "An improved heteroscedasticity and autocorrelation consistent covariance matrix estimator," *Econometrica*, **60**, 953–966.

Efron, B. (1979). "Bootstrap methods: Another look at the Jackknife," *The Annals of Statistics*, **7**, 1–26.

Fama, E. F. and K. R. French (1993). "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, **33**, 3–56.

Hansen, P. R. (2005). "A test for superior predictive ability," *Journal of Business and Economic Statistics*, **23**, 365–380.

Hansen, P. R. and A. Lunde (2005). "A forecast comparison of volatility models: Does anything beat a GARCH (1,1)?," *Journal of Applied Econometrics*, **20**, 873–889.

Hsu, P.-H. and C.-M. Kuan (2005). "Reexamining the profitability of technical analysis with data snooping checks," *Journal of Financial Econometrics*, **3**, 606–628.

Hsu, P.-H., Y.-C. Hsu, and C.-M. Kuan (2008). "Testing the predictive ability of technical analysis using a new stepwise test without data snooping bias," *Working Paper*.

Kadlec, R. K., J. McConnell, R. Rau, and V. Singal (2006). "Is time-series-based predictability evident in real time?," *Journal of Business*, **79**, 1263–1292.

Lahiri, S. N. (2003). "Resampling Methods for Dependent Data," *Springer*.

Ledoit, O. and M. Wolf (2008). "Robust performance hypothesis testing with the Sharpe ratio," *Journal of Empirical Finance*, **15**, 850–859.

Lo, A. W. and A. C. MacKinlay (1990). "Data-snooping biases in tests of financial asset pricing models," *Review of Financial Studies*, **3**, 431–467.

Lo, A. W. (2002). "The statistics of Sharpe ratios," *Financial Analysts Journal*, **58**, 36–52.

Newey, W. K. and K. D. West (1994). "Automatic lag selection in covariance matrix estimation," *Review of Economic Studies*, **61**, 631–653.

Patton, A., Politis, D.N., and H. White (2009). "Correction: Automatic block-length selection for the dependent bootstrap," to appear in *Econometric Reviews*, Vol. 28, Issue 4.

Phillips, P. C. B. and P. Perron (1988). "Testing for a unit root in time series regression," *Biometrika*, **75**, 335–346.

Politis, D. N. and J. P. Romano (1994). "The stationary bootstrap," *Journal of the American Statistical Association*, **89**, 1303–1313.

Politis, D. N. and H. White (2004). "Automatic block-length selection for the dependent bootstrap," *Econometric Reviews*, **23**, 53–70.

Qi, M. and Y. Wu (2006). "Technical trading-rule profitability, data snooping, and reality check: Evidence from the foreign exchange market," *Journal of Money, Credit, and Banking*, **38**, 2135–2158.

Romano, J. P. and M. Wolf (2005). "Stepwise multiple testing as formalized data snooping," *Econometrica*, **73**, 1237–1282.

Sullivan, R., A. Timmermann, and H. White (2001). "Dangers of data mining: the case of calendar effects in stock returns," *Journal of Econometrics*, **105**, 249–286.

Wermers, R. (1999). "Mutual fund herding and the impact on stock prices," *Journal of Finance*, **54**, 581–622.

White, H. (2000). "A reality check for data snooping," *Econometrica*, **68**, 1097–1126.