

出國報告（出國類別：其他）

參加瑞士央行基金會舉辦之
「總體經濟預測」課程研習報告書

服務機關：中央銀行

姓名職稱：朱浩榜/經濟研究處四等專員

派赴國家：瑞士

出國期間：106年6月17日至7月3日

報告日期：106年9月

目 錄

壹、前言.....	1
貳、常見的總體經濟預測模型.....	2
一、參數隨時間變動之模型.....	2
二、混合頻率模型.....	11
參、預測績效評估.....	18
一、相對預測績效.....	18
二、絕對預測績效.....	21
肆、預測模型之應用：即時預報台灣的經濟成長率.....	23
一、簡介.....	23
二、模型與資料說明.....	24
三、即時預報結果.....	25
四、結果探討.....	32
伍、結論與建議.....	34
參考資料.....	35

參加瑞士央行基金會舉辦之 「總體經濟預測」課程研習報告書

壹、前言

瑞士央行基金會（Foundation of Swiss National Bank）格岑塞訓練中心（Study Center Gerzensee）於 2017 年 6 月舉辦「總體經濟預測」（Macroeconomic Forecasting）研習課程。課程講師為蘇黎世聯邦理工學院（ETH Zürich）之 Daniel Kaufmann 博士、義大利 Bocconi 大學之 Massimiliano Marcellino 教授，以及西班牙 Pompeu Fabra 大學之 Barbara Rossi 教授，課程內容涵蓋時間序列分析、常見的總體經濟預測模型，以及預測績效之評估等議題；另亦邀請瑞典央行的經濟學家介紹該行經濟預測之分析架構，並進行經驗分享。

本次研習參加學員共 22 人，包括本行，以及來自奧地利、亞塞拜然、保加利亞、加拿大、丹麥、愛沙尼亞、喬治亞、德國、冰島、印尼、哈薩克、吉爾吉斯、拉脫維亞、馬爾他、巴基斯坦、菲律賓、塞爾維亞、西班牙、瑞士、突尼西亞、烏干達等國中央銀行之代表。

本報告共分為 5 個部分。除前言外，第 2 部分為常見的總體經濟預測模型，包括參數隨時間變動（time-varying）之模型，以及混合頻率資料（mixed frequency data）模型等。第 3 部分簡介預測績效之評估，包括模型間之預測表現，以及預測值本身之性質等評估。充分利用頻率較高且較為即時的經濟指標，將有助掌握當前的經濟情勢，故第 4 部分嘗試運用混合頻率模型，就台灣經濟成長率之即時預報（nowcast）進行實證分析。最後，第 5 部分為結論與建議。

貳、常見的總體經濟預測模型

一、參數隨時間變動之模型

在基本的線性迴歸模型中，通常會假設模型的參數是穩定不變的。然而，政策改變、全球化、新的生產方法等，或是原物料價格飆漲、金融市場崩盤等事件，皆可能造成模型參數變動，而無法以線性模型進行估計及預測，此時便可能須放寬參數穩定不變之限制。以下將依序介紹門檻（threshold）及平滑移轉（smooth transition）自我迴歸模型、馬可夫轉換（Markov-switching）模型、狀態空間（state space）模型等參數可隨時間變動之模型¹。

（一）門檻及平滑移轉自我迴歸模型

首先，假設模型如下：

$$y_t = \begin{cases} \phi_{0,1} + \phi_{1,1}y_{t-1} + \varepsilon_t, & q_t \leq c \\ \phi_{0,2} + \phi_{1,2}y_{t-1} + \varepsilon_t, & q_t > c \end{cases} \quad (2.1.1)$$

此為門檻自我迴歸（Threshold Autoregressive, TAR）模型，其模型參數取決於變數 q_t 是否高於門檻值 c 。其經濟意涵在於，當 q_t 高於某個門檻值 c 時， y_t 的動態調整過程將改變。例如，當通貨膨脹率高於某一水準，則可能發生惡性通膨（hyperinflation），通膨動態將改變；或是當財政赤字高於某一門檻值，則消費者與廠商將預期政府會增稅或縮減公共支出而改變其行為，進而影響經濟成長。

若 $q_t = y_{t-1}$ ，則（2.1.1）式將為自激門檻自我迴歸（Self-Exciting Threshold Autoregressive, SETAR）模型。可再改寫為

$$y_t = (\phi_{0,1} + \phi_{1,1}y_{t-1})(1 - I(y_{t-1} > c)) + (\phi_{0,2} + \phi_{1,2}y_{t-1})I(y_{t-1} > c) + \varepsilon_t. \quad (2.1.2)$$

¹ 本節內容主要參考 Ghysels and Marcellino（2017）與 Study Center Gerzensee 研習課程講義。

其中， I 為指標函數（indicator function）。改寫成（2.1.2）式將較易於估計。由於在 TAR 或 SETAR 模型中，不同參數值之間的轉換並不連續，故亦可考慮較為平滑的轉換形式。例如，可將 I 換成 Logistic 函數

$$G(q_t; \gamma, c) = \frac{1}{1 + \exp(-\gamma[q_t - c])},$$

此即為平滑移轉自我迴歸（Smoothed Transition Autoregressive, STAR）模型。參數 γ 可決定模型的平滑程度，當 γ 越大，則與 TAR 模型越相近；當 $\gamma = 0$ ，則將轉回線性模型。值得注意的是：（1）在 TAR 或 STAR 模型中，自我相關係數（autocorrelation）或偏自我相關係數（partial autocorrelation）將隨時間改變；（2）落後期數仍可使用資訊評選準則（information criteria）決定；（3）模型的狀態（regime）可多於 2 種。

在 TAR 的模型估計上，由於 $I(y_{t-1} > c)$ 僅為虛擬變數，故可逕以最小平方法（ordinary least squares, OLS）估計模型的參數 $\phi = (\phi_{0,1}, \phi_{0,2}, \phi_{1,1}, \phi_{1,2})$ 。由於代入不同的 c 值可得到相應的估計值 $\hat{\phi}(c)$ 、誤差變異數 $\hat{\sigma}^2(c)$ ，故若在某一範圍內，不斷代入不同的 c 值並計算其 $\hat{\sigma}^2(c)$ ，便可找出使 $\hat{\sigma}^2(c)$ 最小的 c 值 \hat{c} ，以及相應的 $\hat{\phi}(\hat{c})$ 與 $\hat{\sigma}^2(\hat{c})$ 。而 STAR 模型因無解析解（analytical solution），故須以數值分析（numerical analysis）求解。至於起始值，則可以在某一範圍內以「網格搜尋」（grid search）的方式給定不同的 γ 、 c ，此時由於 $G(y_{t-1}; \gamma, c)$ 已知，故可以 OLS 估計式估計參數 $\hat{\phi}(\gamma, c)$ ，並利用殘差計算 $\hat{\sigma}^2(\gamma, c)$ 。找出可使 $\hat{\sigma}^2(\gamma, c)$ 最小的 γ 、 c ，便以其相應的 $\hat{\phi}(\gamma, c)$ 做為起始值。

最後，應注意在非線性模型中，不可以迭代（iterative）的方式進行向前 2 期以上的預測。若 y_t 為 y_{t-1} 的非線性函數 $F(y_{t-1}; \theta)$ ，則

$$\hat{y}_{t+2} = \mathbb{E}[F(y_{t-1}; \theta)] \neq F(\mathbb{E}[y_{t-1}]; \theta) = F(\hat{y}_{t+1}; \theta).$$

因此，應以直接預測（direct forecast）的方式進行預測²：

$$y_{t+h} = F(y_t; \theta) + u_t.$$

（二）馬可夫轉換模型

在 TAR 或 STAR 模型中，模型參數取決於可觀察的變數。然而，若參數取決於未知且不可觀察的變數 S_t ，則需就其性質做一些假設，如 S_t 為連續（continuous）或間斷（discrete）變數等。當 S_t 為間斷變數，則可假設 S_t 為馬可夫鏈（Markov chain），此即馬可夫轉換模型³。該模型的經濟意涵在於， S_t 可視為無法直接觀察的經濟狀態，如景氣擴張或衰退，或是金融部門面臨高風險、中風險或低風險等。

首先，假設 S_t 為隨機整數 $\{1, 2, \dots, N\}$ ， S_t 出現狀態 j 的機率為

$$\mathbb{P}(S_t = j | S_{t-1} = i, S_{t-2} = k, \dots) = \mathbb{P}(S_t = j | S_{t-1} = i) = p_{ij},$$

亦即，僅有前一期的狀態會影響對本期 S_t 狀態的機率。此時， S_t 即為 N 種狀態的馬可夫鏈； $\{p_{ij}\}_{i,j=1,\dots,N}$ 為轉換機率（transition probability），表示狀態 i 之下期出現狀態 j 之機率，且 $\sum_{j=1}^N p_{ij} = 1$ 。接著，可進一步將轉換機率寫成轉換矩陣（transition matrix）

$$P = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{N1} \\ p_{12} & p_{22} & \cdots & p_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1N} & p_{2N} & \cdots & p_{NN} \end{bmatrix}.$$

其次，假設一個維度 $N \times 1$ 的向量

$$\xi_t = \begin{cases} (1, 0, \dots, 0)', & \text{if } S_t = 1 \\ (0, 1, \dots, 0)', & \text{if } S_t = 2 \\ \vdots & \\ (0, 0, \dots, 1)', & \text{if } S_t = N \end{cases} \quad (2.2.1)$$

² 關於 TAR 或 STAR 模型之預測應用，可參考 Ferrara et al. (2015)。

³ 關於馬可夫轉換模型，可進一步參閱 Hamilton (1994) 第 22 章。

則其非條件機率

$$\mathbb{E}(\xi_t) = \begin{bmatrix} \mathbb{P}(S_t = 1) \\ \mathbb{P}(S_t = 2) \\ \vdots \\ \mathbb{P}(S_t = N) \end{bmatrix} = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_N \end{bmatrix},$$

而其條件機率

$$\mathbb{E}(\xi_{t+1}|S_t = i) = \begin{bmatrix} p_{i1} \\ p_{i2} \\ \vdots \\ p_{iN} \end{bmatrix},$$

即為轉換矩陣 P 的第 i 欄。因此，

$$\mathbb{E}(\xi_{t+1}|\xi_t) = P\xi_t.$$

根據上述性質，可得

$$\begin{aligned} \xi_{t+1} &= P\xi_t + u_{t+1}, \\ u_{t+1} &= \xi_{t+1} - \mathbb{E}(\xi_{t+1}|\xi_t, \xi_{t-1}, \dots). \end{aligned} \quad (2.2.2)$$

亦即， ξ_{t+1} 可寫成向量自我迴歸 (Vector Autoregression, VAR) 的形式，而 u_{t+1} 則為向前一期的預測誤差。再依此 VAR 形式，可進一步得到

$$\mathbb{E}(\xi_{t+m}|\xi_t, \xi_{t-1}, \dots) = P^m \xi_t.$$

由於在 $S_{t+m} = j$ 時， ξ_{t+m} 的第 j 個元素為 1、其餘元素為 0，故 $\mathbb{E}(\xi_{t+m}|\xi_t, \xi_{t-1}, \dots)$ 的第 j 個元素的意涵為：給定第 t 期的狀態為 i ，則系統在第 $t + m$ 期狀態為 j 的機率。因此，向前 m 期的轉換矩陣即為 P^m 。

假設應變數 y_t 係取決於本身的落後項 $\{y_{t-1}, y_{t-2}, \dots\}$ 、外生變數 $\{x_t, x_{t-1}, x_{t-2}, \dots\}$ 、不同的狀態 S_t ，且模型的誤差項為獨立且分配一致 (independent and identically distributed, i.i.d.)，則當第 t 期且系統出現在狀態 j 時， y_t 的條件密度 (conditional density) 函數為

$$f(y_t|S_t = j, x_t, Y_{t-1}; \alpha).$$

其中， $Y_{t-1} = \{y_{t-1}, y_{t-2}, \dots, x'_{t-1}, x'_{t-2}, \dots\}$ 、 α 為包括係數及誤差變異數在內的模型參數。至於在 t 期且系統出現在狀態 j 的機率，則為

$$\mathbb{P}(S_t = j|Y_t; \theta) = \frac{\mathbb{P}(S_t = j|Y_{t-1}; \theta)f(y_t|S_t = j, x_t, Y_{t-1}; \alpha)}{\sum_{j=1}^N \mathbb{P}(S_t = j|Y_{t-1}; \theta)f(y_t|S_t = j, x_t, Y_{t-1}; \alpha)}, \quad (2.2.3)$$

其中， θ 包括 S_t 的轉換機率 p_{ij} 與模型參數 α 。接著，可仿（2.2.1）式將不同狀態的條件機率堆疊成維度 $N \times 1$ 的向量

$$\xi_{t|t} = \begin{bmatrix} \mathbb{P}(S_t = 1|Y_t; \theta) \\ \mathbb{P}(S_t = 2|Y_t; \theta) \\ \vdots \\ \mathbb{P}(S_t = N|Y_t; \theta) \end{bmatrix}.$$

若給定 Y_t ，並對（2.2.2）式取條件期望值，則可得

$$\mathbb{E}(\xi_{t+1}|Y_t) = P\mathbb{E}(\xi_t|Y_t) + \mathbb{E}(v_{t+1}|Y_t),$$

或是

$$\xi_{t+1|t} = P\xi_{t|t}. \quad (2.2.4)$$

（2.2.4）式即為未來各個狀態之機率，可用於未來狀態之預測。在給定（2.2.3）式及（2.2.4）式的條件下，只要有起始值 $\xi_{1|0}$ 即可計算 $\xi_{t|t}$ （及 $\xi_{t+1|t}$ ）；而前述的非條件機率 π_j ，則為起始值的可能選項。

在模型估計上，可藉最大概似估計法（maximum likelihood estimation, MLE）估計參數 θ 。模型的對數概似（log likelihood）函數為

$$\mathcal{L}(\theta) = \sum_{t=1}^T \log f(y_t|x_t, Y_t; \theta),$$

模型的估計值

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \mathcal{L}(\theta) \quad \text{s.t.} \quad \sum_{j=1}^N \pi_j = 1, \quad \pi_j \geq 0,$$

可利用最大期望演算法 (Expectation Maximization algorithm, EM 演算法) 以迭代的方式, 或是以數值分析求得。而在模型的預測上, 若 Y_t 、 x_{t+1} 已知, 但未來的狀態 S_{t+1} 未知, 則 y_t 向前一期的預測值將為

$$\mathbb{E}(y_{t+1} | x_{t+1}, Y_t; \alpha) = \sum_{j=1}^N \mathbb{P}(S_{t+1} = j | Y_t; \theta) \mathbb{E}(y_{t+1} | S_{t+1} = j, x_{t+1}, Y_t; \alpha).$$

若將未來不同狀態的預測值 $\mathbb{E}(y_{t+1} | S_{t+1} = j, x_{t+1}, Y_t; \alpha)$ 堆疊成維度 $N \times 1$ 的向量 h_t , 則 $\mathbb{E}(y_{t+1} | x_{t+1}, Y_t; \alpha) = h_t' \xi_{t+1|t}$ 。類似地, 可利用同樣的作法得到向前 m 期的預測值。

最後, 只要樣本夠多且參數變動夠大, 則可用赤池氏資訊評選準則 (Akaike Information Criterion, AIC)、貝氏資訊評選準則 (Bayesian Information Criterion, BIC) 或 Hannan–Quinn 等評選準則決定模型的狀態數目 N 與落後期數。

(三) 狀態空間模型

經濟理論往往含有無法直接觀察到的變數, 如潛在產出、自然失業率等。若模型含有無法觀察到的連續變數, 則可利用狀態空間模型加以估計⁴。一般而言, 狀態空間模型可寫成以下形式:

$$\alpha_t = T\alpha_{t-1} + R\eta_t, \quad (2.3.1)$$

$$y_t = Z\alpha_t + S\xi_t, \quad (2.3.2)$$

$$\begin{bmatrix} \eta_t \\ \xi_t \end{bmatrix} \sim \text{i.i.d. } \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & H \end{bmatrix} \right). \quad (2.3.3)$$

其中, α_t 為無法直接觀察的狀態變數 (state variable) 向量、 y_t 為可

⁴ 關於狀態空間模型, 可進一步參閱 Hamilton (1994) 第 14 章。

觀察到的變數向量；而 (2.3.1) 式稱為狀態方程式 (state equation)、(2.3.2) 式稱為轉換 (transition) 或測量方程式 (measurement equation)，描繪可觀察到的變數與狀態變數之間的線性關係；誤差項 η_t 、 ξ_t 則為平均數 0、常態分配、彼此無關且符合 i.i.d. 的隨機變數。狀態空間模型的應用極為廣泛，許多模型皆可表示成狀態空間模型的形式。

狀態空間模型可利用 Kalman Filter 進行預測及求解。首先，可將 (2.3.1) 式至 (2.3.3) 式之狀態及測量方程式表示成常態密度函數的形式：

$$\begin{aligned} f(\alpha_t | \alpha_{t-1}, Y_{t-1}) &= \mathcal{N}(T\alpha_{t-1}, RQR'), \\ f(y_t | \alpha_t, Y_{t-1}) &= \mathcal{N}(Z\alpha_t, SHS'), \end{aligned}$$

其中， $Y_{t-1} = (y_{t-1}, y_{t-2}, \dots)$ 。

其次，可在給定過去觀察資訊的條件下，將狀態變數表示成

$$f(\alpha_t | Y_{t-1}) = \mathcal{N}(\alpha_{t|t-1}, P_{t|t-1}). \quad (2.3.4)$$

其中，

$$\begin{aligned} \alpha_{t|t-1} &= \mathbb{E}(\alpha_t | Y_{t-1}) = T\alpha_{t-1|t-1}, \\ P_{t|t-1} &= \text{var}(\alpha_t - \alpha_{t|t-1}) = TP_{t-1|t-1}T' + RQR'. \end{aligned}$$

由於可利用 (2.3.4) 式建構向前一期的預測模型，故 (2.3.4) 式亦稱為「預測方程式」(prediction equation)。在觀察到當期的資訊 y_t 之後，可得到

$$f(\alpha_t | Y_t) = \mathcal{N}(\alpha_{t|t}, P_{t|t}), \quad (2.3.5)$$

其中，

$$\begin{aligned} \alpha_{t|t} &= \alpha_{t|t-1} + P_{t|t-1}Z'F_t^{-1}v_t, \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1}Z'F_t^{-1}ZP_{t|t-1}, \end{aligned}$$

且 $v_t = y_t - y_{t|t-1}$ 、 $F_t = \mathbb{E}(v_t v_t')$ 。(2.3.5) 式說明如何運用 y_t 的訊息，更新在 $t-1$ 期時對 α_t 所做的預測，故亦稱為「更新方程式」(updating equation)。由上可知，預測方程式將取決於更新方程式內的 $\alpha_{t-1|t-1}$ 及 $P_{t-1|t-1}$ ，而更新方程式則取決於前一期的 $\alpha_{t|t-1}$ 、 $P_{t|t-1}$ ，以及 v_t 與 F_t 。

接著，在給定過去觀察資訊的條件下， y_t 的「概似方程式」(likelihood equation) 為

$$f(y_t|Y_{t-1}) = \mathcal{N}(y_t|y_{t|t-1}, F_t), \quad (2.3.6)$$

其中，

$$\begin{aligned} y_{t|t-1} &= \mathbb{E}(Z\alpha_t + S\xi_t|Y_{t-1}) = Z\alpha_{t|t-1}, \\ F_t &= \mathbb{E}(v_t v_t') = ZP_{t|t-1}Z' + SHS'. \end{aligned}$$

同樣地，可利用 (2.3.6) 式預測向前一期的 y_t ，且該式取決於 $\alpha_{t|t-1}$ 、 $P_{t|t-1}$ 。此外，聯合概似函數 (joint likelihood) 可表為

$$f(y_1, \dots, y_T) = \prod_{t=1}^T f(y_t|Y_{t-1}), \quad (2.3.7)$$

只要給定 $f(y_1)$ ，即可得到聯合概似函數。

最後，關於 α_t 的初始條件，一個常見的方式是選取 α_t 的非條件平均數及變異數。根據 (2.3.1) 式及 (2.3.3) 式，可知

$$\alpha_{1|0} = \mathbb{E}(\alpha_t) = 0. \quad (2.3.8)$$

若定義 $P_{1|0} = \mathbb{E}(\alpha_t' \alpha_t)$ ，則可得

$$\text{vec}(P_{1|0}) = [\mathbb{I} - (T \otimes T)]^{-1} \text{vec}(RQR'), \quad (2.3.9)$$

其中， $\text{vec}(P_{1|0})$ 為 $P_{1|0}$ 垂直堆疊而成的向量。

綜上所述，可藉由迭代的方式，逐步得到預測值及概似函數：

1. 利用 (2.3.8) 式、(2.3.9) 式之初始條件得到 $\alpha_{2|1}$ 、 $P_{2|1}$ 。

2. 利用 $\alpha_{2|1}$ 、 $P_{2|1}$ 及 (2.3.6) 式之概似方程式，得到 $y_{2|1}$ 、 F_2 。
3. 利用 $\alpha_{2|1}$ 、 $P_{2|1}$ 、 $y_{2|1}$ 、 F_2 及 (2.3.5) 式之更新方程式，得到 $\alpha_{2|2}$ 、 $P_{2|2}$ 。
4. 利用 $\alpha_{2|2}$ 、 $P_{2|2}$ 及 (2.3.4) 式之預測方程式，得到 $\alpha_{3|2}$ 、 $P_{3|2}$ 。
5. 重複上述步驟直到最後一個樣本點 $t = T$ ，即可獲得所有的預測值及概似函數，並利用 (2.3.7) 式得到：

$$\log \mathcal{L}(\theta) = -\frac{NT}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log |F_t| - \frac{1}{2} \sum_{t=1}^T v_t' F_t v_t.$$

同樣地，除了預測向前一期的 y 、 α ，亦可進行向前 h 期的預測。根據 (2.3.1) 式， $\alpha_{t+h} = T^h \alpha_t + R\eta_{t+h} + TR\eta_{t+h-1} + \dots + T^{h-1}R\eta_{t+1}$ ，故 α 向前 h 期的預測為

$$\mathbb{E}(\alpha_{t+h} | Y_t) = \alpha_{t+h|t} = T^h \alpha_{t|t},$$

預測誤差 $u_{t+h} = \alpha_{t+h} - \alpha_{t+h|t}$ ，其變異數為

$$\text{var}(u_{t+h}) = T^h P_{t|t} T^{h'} + RQR' + \dots + T^{h-1}RQR'T^{h-1'}.$$

其次，據 (2.3.2) 式， $y_{t+h} = Z\alpha_{t+h} + S\xi_{t+h}$ ，故 y 向前 h 期的預測為

$$\mathbb{E}(y_{t+h} | Y_t) = y_{t+h|t} = Z\alpha_{t+h|t},$$

預測誤差 $v_{t+h} = y_{t+h} - y_{t+h|t}$ ，其變異數為

$$\text{var}(v_{t+h}) = Z \text{var}(u_{t+h}) Z' + SHS'.$$

值得注意的是，從上述 Kalman Filter 的計算過程中，可知所有的方程式均給定至前一期的資訊 Y_{t-1} ，再以當期的資訊 y_t 進行更新及下一期之預測。但事實上，除了預測之外，亦可使用全期的樣本資訊 Y_T 進行狀態變數的事後 (ex post) 估計，此即「Kalman smoother」⁵。

⁵ 詳見 Harvey (1993) 第 4 章。

二、混合頻率模型

經濟成長率等資料通常係按季發布，因此，若要評估當前的經濟成長率，則僅能借助具有同時或領先性質，且頻率較高的經濟指標。當樣本的資料頻率不同時，逕將高頻率資料加總成低頻率資料的作法雖然簡單，但卻可能損失部分資訊，甚至變更資料的生成機制（generating mechanism）。因此，如何建構模型，利用較高頻率之資訊對較低頻率之資料進行預測或即時預報，便為重要的課題。本節將介紹橋樑方程式（bridge equations）、MIDAS（Mixed Data Sampling）模型，以及混合頻率 VAR（Mixed frequency VAR, MF-VAR）等 3 種常見的混合頻率模型⁶。

在此將符號定義如下： t 為低頻資料的時間單位，且 $t = 1, \dots, T$ ； m 為時間單位內，高頻資料出現的次數； y_t^L 為低頻變數； $x_{t-j/m}^H$ 為高頻變數，而 $(t - j/m)$ 表過去第 j 個出現的變數，且 $j = 0, 1, 2 \dots$ 。因此，若資料型態為季與月頻率之混合，則在第 t 季時，最後出現、第二個出現及第一個出現的月資料將分別為 $x_{t/3}^H$ 、 $x_{t-1/3}^H$ 及 $x_{t-2/3}^H$ 。

（一）橋樑方程式

橋樑方程式係聯結高頻資料與低頻資料的線性迴歸模型，用於建構兩者之間的統計關係，以便在高頻率的資料發布時（如工業生產指數），即時估計低頻資料（如 GDP）。

由於在進行即時預報時，高頻率的資料往往僅部分可得，故橋樑方程式之估計往往包含兩步驟。以季與月之混頻資料為例，首先，預測當季尚未發布的月資料（可用單變量或 VAR 等時間序列模型進

⁶ 本節內容主要參考 Ghysels and Marcellino（2017）、Study Center Gerzensee 研習課程講義，以及 Foroni and Marcellino（2013）。

行預測)，再加總成季資料 $x_t^L = \sum_{i=0}^{m-1} a_i x_{t-i/m}^H$ ；其次，再以橋樑方程式建構兩者之間的關係：

$$y_t^L = a + bx_t^L + u_t^L. \quad (2.4.1)$$

其中， u_t^L 為符合 i.i.d. 之誤差項。雖然 (2.4.1) 式為靜態的形式，但在實際運用時，可再進一步擴充，如納入 y_t^L 或 x_t^L 的落後項等。

至於高頻變數之選擇，則可利用樣本內 (in-sample) 或樣本外 (out-of-sample) 的方式，以資訊評選準則或均方根差 (root mean square error, RMSE) 等方法加以篩選。

(二) MIDAS 模型

在估計橋樑方程式之前，尚需先進行高頻變數之預測及加總；相對地，MIDAS 模型則藉由較為精簡 (parsimonious) 的落後分配多項式 (distributed lag polynomial) 的形式，直接建構高頻與低頻變數之間的關係 (Ghysels et al., 2004)。

假設要預測向前 h 期的低頻變數，則 MIDAS 模型可表示為

$$y_{t+h}^L = a_h + b_h \mathbb{C}(L^{1/m}; \theta) x_t^H + \varepsilon_{t+h}^L, \quad (2.5.1)$$

其中， $\mathbb{C}(L^{1/m}; \theta) = \sum_{k=0}^K c(k; \theta) L^{k/m}$ 且 $\mathbb{C}(1; \theta) = \sum_{k=0}^K c(k; \theta) = 1$ ； $L^{k/m} x_t^H = x_{t-k/m}^H$ 。而 $c(k; \theta)$ 可有許多不同的選擇，其中，「exponential Almon Lag」的形式頗為常見：

$$c(k; \theta) = \frac{\exp(\theta_1 k + \dots + \theta_Q k^Q)}{\sum_{k=0}^K \exp(\theta_1 k + \dots + \theta_Q k^Q)}$$

其優點在於極具彈性，僅需少數參數即可調整函數型態 (通常設定 $Q = 2$)，包括遞增、遞減或峰型 (hump-shaped) 等。

除了 exponential Almon Lag，其他常見的形式包括⁷：

1. Beta Lag：基於 Beta 函數，

$$c(k; \theta_1, \theta_2) = \frac{f\left(\frac{k}{K}, \theta_1; \theta_2\right)}{\sum_{k=0}^K f\left(\frac{k}{K}, \theta_1; \theta_2\right)},$$

$$f(x, a, b) = \frac{x^{a-1}(1-x)^{b-1}\Gamma(a+b)}{\Gamma(a)\Gamma(b)}.$$

其中， $\Gamma(a) = \int_0^{\infty} e^{-x} x^{a-1} dx$ （Gamma 函數）。可設定 $\theta_1 = 1$ 且 $\theta_2 > 1$ ，則僅需估計 θ_2 即可。

2. 線性（linear）： $c(k; \theta) = \frac{1}{K}$ ，此形式並無包含其他參數。
3. 雙曲（hyperbolic）：

$$c(k; \theta) = \frac{g\left(\frac{k}{K}, \theta\right)}{\sum_{k=0}^K g\left(\frac{k}{K}, \theta\right)},$$

$$g(k, \theta) = \frac{\Gamma(k + \theta)}{\Gamma(k + 1)\Gamma(\theta)}.$$

由於 Gamma 函數僅有一個參數需估計，故設定上較不若 Beta Lag 有彈性。

4. 幾何（geometric）：

$$c(k; \theta) = \frac{\theta^k}{\sum_{k=0}^K \theta^k}$$

且 $|\theta| \leq 1$ 。

在 MIDAS 模型的估計上，(2.5.1) 式可用非線性最小平方法（nonlinear least square, NLS）估計（Ghysels et al., 2004; Andreou et al., 2010）。值得注意的是，在 MIDAS 模型中，增加多項式的落後期數

⁷ 詳見 Ghysels et al. (2006)、Ghysels et al. (2009)。

並不會使參數的數目增加，故無法直接使用 AIC、BIC 或 HQ 等資訊評選準則決定模型的落後期數。

至於 MIDAS 模型的預測，由於在給定不同的預測期間(horizon)下，均須利用不同的資訊集合進行估計，並得到不同的參數估計值，故參數估計值僅適用該特定的預測期間；亦即，對不同的預測期間而言，參數估計值均須重新估計。因此，MIDAS 模型係以 (2.5.1) 式的形式進行直接預測。

在 (2.5.1) 式中，可知解釋變數同時含有係數 b_h 與各落後期的權重 $\mathbb{C}(L^{1/m}; \theta)$ (且權重相加須等於 1)。事實上，若建構 MIDAS 模型的目的係用於預測，則毋須將 b_h 與 θ 分開估計。亦即，改為估計

$$y_{t+h}^L = a_h + \mathbb{C}(L^{1/m}; \tilde{\theta}_h) x_t^H + \varepsilon_{t+h}^L, \quad (2.5.2)$$

並略為調整各落後期的權重。如 exponential Almon Lag，可改為：

$$\mathbb{C}(L^{1/m}; \tilde{\theta}_h) = \sum_{k=1}^K \exp(\tilde{\theta}_{h1}k + \dots + \tilde{\theta}_{hQ}k^Q) L^{(k-1)/m}.$$

然而，若欲測試高頻變數是否顯著，則須估計 b_h 並檢定其是否為 0。具體而言，可依據 Davies(1987)之建議，給定一組 $\tilde{\theta}_h^i = \{\tilde{\theta}_h^1, \tilde{\theta}_h^2, \dots, \tilde{\theta}_h^G\}$ 並計算 $b_h(\tilde{\theta}_h^i)$ 及其 t -檢定統計量 $t_{b_h(\tilde{\theta}_h^i)}$ ，再得到

$$t_{max} \equiv \sup_i t_b(\tilde{\theta}^i).$$

而 t_{max} 在虛無假設(null hypothesis)下的分配則可以模擬的方式得到，詳見 Hansen (1996)。

若在低頻資料的單位時間內，高頻資料出現的次數 m 不大(如季與月之混頻資料)，則可不必予以(2.5.1)式中 $b_h \mathbb{C}(L^{1/m}; \theta)$ 或(2.5.2)式中 $\mathbb{C}(L^{1/m}; \tilde{\theta}_h)$ 的限制，此即未受限的 MIDAS (unrestricted MIDAS,

U-MIDAS) 模型 (Forni et al., 2015)。以季與月之混頻資料為例，模型可表示為

$$y_{t+h}^L = \alpha_h + \sum_{i=1}^k \sum_{j=0}^{l_i} c_{i,j} x_{i,t-j/m}^H + \varepsilon_{t+h}^L. \quad (2.5.3)$$

若模型的誤差項 ε_{t+h}^L 沒有序列相關 (serial correlation)，則 (2.5.3) 式可逕以 OLS 估計。

最後，(2.5.1) 式或 (2.5.2) 式僅為 MIDAS 模型的基本型式，實際用於預測時，可在 MIDAS 模型加入低頻變數 y_{t+h}^L 的落後項，使其成為 AR-MIDAS 模型。此外尚有許多不同型式的延伸，如可結合馬可夫轉換模型、平滑移轉函數或因子模型 (factor model) 等，詳見 Forni and Marcellino (2013) 之介紹。

(三) 混合頻率 VAR 模型

混合頻率的資料亦可利用多變量模型加以處理，如 MF-VAR。一般而言，MF-VAR 處理混頻資料的做法可分為「參數導向」(parameter-driven) 與「樣本導向」(observation-driven) 等兩種。參數導向的作法係運用狀態空間模型，將低頻資料視為高頻資料，並將部分觀察值視為資料上的遺漏 (missing data)，再萃取無法直接觀察到的狀態資訊；而樣本導向則係依據樣本觀察值建構模型，作法類似一般的 VAR 模型。以下將依序簡介參數導向及樣本導向之 MF-VAR。

首先，假設 y_t^L 為每季的 GDP 成長率，在 $t = 3, 6, 9, \dots, T$ 時可得到觀察值，而每月的 GDP 成長率 y_t^* 則無法直接觀察。其次，令每季的 GDP (取對數值) 為 Y_t^L ，每月無法直接觀察的 GDP (取對數值) 為 Y_t^* 、 $Y_{t-1/3}^*$ 及 $Y_{t-2/3}^*$ ，且 Y_t^L 為 Y_t^* 、 $Y_{t-1/3}^*$ 及 $Y_{t-2/3}^*$ 的幾何平均。因此，

若將 Y_t^L 對其落後3期取差分為 y_t^L (即 $y_t^L = \Delta_3 Y_t^L$)、將 Y_t^* 取差分為 y_t^* (即 $y_t^* = \Delta Y_t^*$)，則

$$\begin{aligned} y_t^L &= \frac{1}{3}(y_{t/3}^* + y_{t-1/3}^* + y_{t-2/3}^*) + \frac{1}{3}(y_{t-1/3}^* + y_{t-2/3}^* + y_{t-3/3}^*) \\ &\quad + \frac{1}{3}(y_{t-2/3}^* + y_{t-3/3}^* + y_{t-4/3}^*) \\ &= \frac{1}{3}y_{t/3}^* + \frac{2}{3}y_{t-1/3}^* + y_{t-2/3}^* + \frac{2}{3}y_{t-3/3}^* + \frac{1}{3}y_{t-4/3}^*. \end{aligned} \quad (2.6.1)$$

其次，建構雙變數的VAR，連結無法直接觀察到的月GDP成長率 y_t^* ，以及相應的經濟指標 x_t^H ：

$$\phi(L) \begin{bmatrix} y_t^* - \mu_y^* \\ x_t^H - \mu_x^H \end{bmatrix} = u_t^H, \quad (2.6.2)$$

其中， $u_t^H \sim \mathcal{N}(0, \Sigma)$ 。接著，令VAR的落後期數為4期⁸，且

$$z_t = \begin{bmatrix} y_t^* - \mu_y^* \\ x_t^H - \mu_x^H \end{bmatrix}, \quad s_t = \begin{bmatrix} z_t \\ \vdots \\ z_{t-4} \end{bmatrix},$$

則可將(2.6.1)式與(2.6.2)式以狀態空間模型的形式表示：

$$s_t = F s_{t-1/3} + G v_t, \quad (2.6.3)$$

$$\begin{bmatrix} y_t^L - \mu_t^L \\ x_t^H - \mu_x^H \end{bmatrix} = H s_t, \quad (2.6.4)$$

其中， $\mu_t = 3\mu_y^*$ 、 $v_t \sim \mathcal{N}(0, \mathbb{I}_2)$ ，且

$$F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}; \quad F_1 = [\phi_1 \quad \phi_2 \quad \phi_3 \quad \phi_4 \quad 0_{2 \times 2}]; \quad F_2 = [\mathbb{I}_8 \quad 0_{8 \times 2}];$$

$$G = \begin{bmatrix} \Sigma^{\frac{1}{2}} \\ 0_{8 \times 2} \end{bmatrix}; \quad H = [H_0 \quad \dots \quad H_4],$$

而

$$H(L) = \begin{bmatrix} 1/3 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2/3 & 0 \\ 0 & 0 \end{bmatrix} L + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} L^2 + \begin{bmatrix} 2/3 & 0 \\ 0 & 0 \end{bmatrix} L^3 + \begin{bmatrix} 1/3 & 0 \\ 0 & 0 \end{bmatrix} L^4.$$

⁸ 若落後期數大於4期，詳參Mariano and Murasawa (2010)。

在無法直接觀察到 y_t^L 的時間點($t \neq 3, 6, 9, \dots$)，即將其視為遺漏值，且為符合 i.i.d. 與常態分配之隨機變數。(2.6.3) 式與 (2.6.4) 式之狀態空間模型可以 Kalman Filter 估計；若參數太多導致 MLE 不易收斂，則可考慮以 EM 演算法估計⁹。

樣本導向的 MF-VAR 則直接將不同頻率的資料堆疊成 VAR 的形式。同樣地，若以季與月之混頻資料為例，則模型可表示為

$$\begin{bmatrix} x_{t+1-2/3}^L \\ x_{t+1-1/3}^L \\ x_{t+1/3}^L \\ y_{t+1}^L \end{bmatrix} = \Phi \begin{bmatrix} x_{t-2/3}^L \\ x_{t-1/3}^L \\ x_{t/3}^L \\ y_t^L \end{bmatrix} + \varepsilon_{t+1}. \quad (2.6.5)$$

其中， Φ 為 4×4 的係數矩陣，且 ε_{t+1} 為維度 4×1 的向量，代表各變數的衝擊。藉由 (2.6.5) 式，可知雙變數之混頻資料將在堆疊後轉成 4 個變數之 VAR；且模型的最後一式即為 U-MIDAS 的模型迴歸式，而前三式則可衡量低頻變數對高頻變數的影響。

若欲將樣本導向的 MF-VAR 應用於即時預報，則須將 (2.6.5) 式轉換為下三角矩陣的結構式 VAR (structural VAR)：

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ a_{21} & 1 & 0 & 0 \\ a_{31} & a_{32} & 1 & 0 \\ a_{41} & a_{42} & a_{43} & 1 \end{bmatrix} \begin{bmatrix} x_{t+1-2/3}^L \\ x_{t+1-1/3}^L \\ x_{t+1/3}^L \\ y_{t+1}^L \end{bmatrix} = \Phi_{t-1} \begin{bmatrix} x_{t-2/3}^L \\ x_{t-1/3}^L \\ x_{t/3}^L \\ y_t^L \end{bmatrix} + \tilde{\varepsilon}_{t+1}.$$

最後一式即可用於即時預報，而前三式則可用於高頻資料之預測。

在模型的估計上，即與一般的 VAR 模型相同；而若 VAR 模型的維度較大，則可利用貝氏 (Bayesian) 方法估計，如 McCracken et al. (2015)。

⁹ 詳參 Dempster et al. (1977) 與 Stock and Watson (2002)。

參、預測績效評估

一、相對預測績效

若欲比較兩個預測模型的表現，通常可就模型的預測誤差（預測值與實際值之差距）設定某一損失函數（loss function），再比較該損失函數之大小，以判斷何者的預測績效較佳；而損失函數之間的差異是否顯著，則有賴統計檢定之輔助。以下將簡介常用於評估相對預測績效之統計檢定¹⁰。

（一）非相容模型

首先，令 e_{it} 為模型 i 之預測誤差， $g(\cdot)$ 為損失函數，可為均方差（mean square error, MSE）、平均絕對誤差（mean absolute error, MAE）等，而 $\mathbb{E}g_{it}$ 則用於衡量模型 i 的預測績效。當有兩個非相容（nonnested）的預測模型，且其預測誤差的損失函數之差為¹¹

$$f_t = g(e_{1t}) - g(e_{2t}),$$

此時，若虛無假設

$$\mathbb{E}(f_t) = 0$$

成立，則兩個模型的預測能力沒有顯著差異。若 f_t 符合共變異定態（covariance stationarity），則可得到

$$\sqrt{P}(\bar{f} - \mathbb{E}(f_t)) \sim_A \mathcal{N}(0, V). \quad (3.1.1)$$

其中， P 為預測誤差的樣本數、 \bar{f} 為 f_t 的平均數（ $P^{-1} \sum f_t$ ）、 $V = \sum_{j=-\infty}^{\infty} \mathbb{E}(f_t - \mathbb{E}(f_t))(f_{t-j} - \mathbb{E}(f_t))$ 。當 f_t 不具序列相關時， $V = \mathbb{E}(f_t - \mathbb{E}(f_t))^2$ ；而當 f_t 具有序列相關，則可改以Newey–West HAC

¹⁰ 本節內容主要參考 Study Center Gerzensee 研習課程講義與 West (2006)。

¹¹ 隨 g 的設定不同， f_t 可為其他不同的形式，如預測誤差具序列相關（ $e_{1t+1}e_{1t}$ ）等。

估計式估計。若 \widehat{V} 為 V 的一致性估計式，則檢定統計量為

$$DM = \frac{\bar{f}}{\sqrt{\widehat{V}}} \sim_A \mathcal{N}(0,1),$$

此即 Diebold–Mariano 檢定。實務上，若損失函數為 MSE 或 MAE，則可逕將 f_t 對常數項進行迴歸，並使用 HAC 標準誤計算其 t -統計量。

值得注意的是，模型參數的估計誤差亦可能影響 e_t ，進而影響 f_t 及 V ；而僅有在符合下列三種情形之一時，模型參數的估計誤差不會造成影響：

1. f_t 的期望值對參數的偏微分為 0。以線性模型 ($y_t = X'_{it}\beta_i^* + e_{it}$) 為例，若其損失函數為 MSE 的形式，則

$$f_t = e_{1t}^2 - e_{2t}^2 = (y_t - X'_{1t}\beta_1^*)^2 - (y_t - X'_{2t}\beta_2^*)^2. \quad (3.1.2)$$

接著，對 (3.1.2) 式取期望值後，再對 β 偏微分，可得

$$\frac{\partial \mathbb{E}f_t(\beta^*)}{\partial \beta} = [-2\mathbb{E}(X'_{1t}e_{1t}), 2\mathbb{E}(X'_{2t}e_{2t})]'. \quad (3.1.3)$$

此時，只要解釋變數 X_{it} 與預測誤差 e_{it} 無關，(3.1.3) 式即為 0。

2. 預測模型的樣本內觀察值 R 夠大，使

$$\frac{P}{R} \rightarrow 0$$

成立。因為相對較大的 R 將使 β 的產生不確定性較小。文獻研究顯示¹²，只要 $P/R < 0.1$ ，即可忽略該影響。

3. 某些特定情況，包括參數估計值係以遞迴 (recursive) 方式估得及預測誤差符合條件變異數齊一 (conditionally homoskedastic)；而損失函數則須為平均預測誤差 ($\mathbb{E}e_t$)、向前一期預測誤差的序列

¹² 可參考 West (1996, 2001)，或 Clark and McCracken (2001)、McCracken (2007) 等。

相關 ($\mathbb{E}e_{t+1}e_t$)，或是模型的預測誤差與另一個模型的預測值不相關等 ($\mathbb{E}e_t X'_{2t} \beta_2^* = 0$)。

若不符合上述三種情形，則需考慮調整 (3.1.1) 式中之 V 及檢定統計量，詳見 West (1996)。

(二) 相容模型

若欲比較預測績效的兩個模型為相容 (nested) 時，則 Diebold–Mariano 檢定將無法適用。假設有兩個模型：

$$\text{模型 1: } y_t = \beta_{10} + \beta_{11}y_{t-1} + e_{1t} = X'_{1t}\beta_1^* + e_{1t};$$

$$\text{模型 2: } y_t = \beta_{20} + \beta_{21}y_{t-1} + \beta_{22}x_{t-1} + e_{2t} = X'_{2t}\beta_2^* + e_{2t}.$$

在 x 對 y 沒有「Granger 因果關係」(Granger causality) 的虛無假設下， β_{22} 將為 0，使得 $\beta_2^{*'} = (\beta_1^{*'}, 0)$ 、 $X'_{1t}\beta_1^* = X'_{2t}\beta_2^*$ ，並使模型 1 與模型 2 的誤差項相同，導致 $f_t = 0$ (例如，若損失函數為 MSE，則 $e_{2t}^2 - e_{1t}^2 = 0$ ；若為 MAE，則 $|e_{1t}| - |e_{2t}| = 0$)。因此，Diebold–Mariano 檢定無法適用於兩個相容模型。

為瞭解於模型 2 加入的變數 x 是否提供有助預測的額外資訊，Clark and McCracken (2001) 從預測涵蓋性 (forecast encompassing) 的角度，建議比較 \hat{e}_{1t} 與 $(\hat{e}_{1t} - \hat{e}_{2t})$ 的共變異數，並提出檢定統計量

$$\text{ENC-NEW} = \bar{f} = \frac{P^{-1} \sum_{t=R}^T \hat{e}_{1t+1} (\hat{e}_{1t+1} - \hat{e}_{2t+1})}{\hat{\sigma}^2}.$$

其中， $\hat{\sigma}^2 = P^{-1} \sum_{t=R}^T \hat{e}_{2t+1}^2$ ， P 為預測誤差的樣本數， R 為模型的樣本內觀察值。若拒絕虛無假設，則表示加入額外變數 x 的確有助提升預測績效。然而，ENC-NEW 檢定僅適用於線性模型且預測期間為向前一期的情況，且其檢定統計量不為常態分配，使用限制較多。

另一方面，Clark and West (2007) 則建議，當損失函數為 MSE 時，可調整含有額外變數的模型 2 之 MSE，以考量模型 2 因納入無關的變數而產生的干擾 (noise)。亦即，檢定

$$\hat{\sigma}_1^2 - \hat{\sigma}_{2-\text{adj}}^2 = P^{-1} \sum_{t=R}^T \hat{e}_{1t+h}^2 - \left[P^{-1} \sum_{t=R}^T \hat{e}_{2t+h}^2 - P^{-1} \sum_{t=R}^T (\hat{y}_{1t+h} - \hat{y}_{2t+h})^2 \right],$$

其中， \hat{y}_{t+h} 為向前 h 期的預測值。實務上，可將 $(\hat{e}_{1t+h}^2 - [\hat{e}_{2t+h}^2 - (\hat{y}_{1t+h} - \hat{y}_{2t+h})^2])$ 對常數項進行迴歸，再計算其 t -統計量即可。

二、絕對預測績效

除了模型間的相對預測績效之外，模型預測值本身的性質往往亦為關注的重點；而其性質包括不偏性 (unbiasedness)、理性 (rationality)、效率性 (efficiency) 及預測誤差不具序列相關等，皆可用於績效之衡量。說明如下¹³：

1. 不偏性：若預測誤差的期望值為 0，則表示預測值沒有偏誤。實務上，可將預測誤差對常數項進行迴歸，並使用 HAC 標準誤計算其 t -統計量¹⁴。
2. 理性：考慮迴歸式

$$y_t = \delta_0 + \delta_1 \hat{y}_{t|t-h} + u_{t,h}.$$

當常數項 δ_0 為 0 時，預測值具有不偏性；當 $\delta_1 = 1$ ，預測值為最適；而當 $\delta_0 = 0$ 且 $\delta_1 = 1$ 成立時，預測值具有理性性質。因此，可對 $[\delta_0, \delta_1] = [0, 1]$ 進行聯合檢定，若檢定結果不拒絕虛無假設，則其理性性質成立。

¹³ 本節內容主要參考 Study Center Gerzensee 研習課程講義與 West (2006)。

¹⁴ 如前所述，若預測值係用模型以參數及解釋變數推得，則當 $P/R \geq 0.1$ 時，或須考慮調整其 HAC 估計標準誤。有關效率性之討論亦同。

3. 效率性：若解釋變數已無額外有助預測的資訊，則符合效率性。具體作法上，可將預測誤差對預測模型的解釋變數進行迴歸，並使用 HAC 標準誤計算其 t -統計量。
4. 誤差不具序列相關：若預測值係利用遞迴方式產生，則逕將預測誤差對當期誤差進行迴歸，並計算其 t -統計量即可；若預測值係利用滾動 (rolling) 樣本的方式產生，則須將預測誤差對當期誤差及解釋變數進行迴歸，再計算其 t -統計量¹⁵。

最後，值得說明的是，預測績效檢定的種類眾多，各有其適用的情形與條件。例如，當預測誤差本身呈現不穩定 (instability) 的狀態時，前述關於絕對預測績效之檢定或樣本內「Granger 因果關係」檢定等即無法適用。關於預測績效評估近來的發展及應用，可參見 Clark and McCracken (2013)；而關於預測不穩定時的績效評估，則可參見 Rossi (2013)。

¹⁵ 此作法適用於預測誤差符合條件變異數齊一的情況。若預測誤差為條件變異數不齊一 (conditionally heteroscedastic)，或須另行調整其標準誤。

肆、預測模型之應用：即時預報台灣的經濟成長率

在許多國家，GDP 的資料頻率為季度且發布時間較為落後，故可利用頻率較高、較為即時，且具有領先或同時性質的經濟指標或相關資料，掌握經濟活動的狀態，此即「即時預報」，實務上可利用混合頻率模型加以處理。因此，本節擬應用 MIDAS 與橋樑方程式等常見的混頻模型，就台灣經濟成長率的即時預報進行實證分析。

一、簡介

我國 GDP 的統計資料係由主計總處編製、發布，惟其發布時間較為落後。以 2017 年第 2 季 GDP 為例，主計總處於 7 月 28 日發布其概估統計值，而初步統計值則須至 8 月 18 日方才發布。此時，已可獲得第 2 季許多與 GDP 高度相關且高頻經濟指標（如出口總額、工業生產指數等月資料），故可充分利用此資訊以評估當下的 GDP。

近來，國內應用月資料預測當季 GDP 的文獻包括張志揚(2013)與林依伶(2013)。張志揚(2013)以零售及餐飲業營業額、資本設備進口額，以及海關進出口總額等月資料，分別作為民間消費、資本形成，以及輸出入等季資料的高頻指標，再藉由橋樑方程式建構兩者之間的關係，並將 GDP 各組成分累加以預測經濟成長率。另一方面，林依伶(2013)考量農曆春節效果等季節因素，建構高頻指標的月模型並據以進行預測，再將高頻指標的實際值或預測值透過橋樑方程式代入總體季模型，以預測 GDP 及其組成分。

前述文獻均利用橋樑方程式建構指標與目標變數（即經濟成長率）之間的關係，故仍須先就高頻指標進行預測，再將其轉換為低頻。由於 MIDAS 模型可直接建構當期及落後期的高頻指標與未來經濟成長率之間的關係，故本節擬嘗試利用 MIDAS 模型，並仿照 Kuzin et al.

(2011) 的作法，以單一高頻指標進行經濟成長率之即時預報，並比較不同模型的預測績效。但另一方面，橋樑方程式的作法仍廣為各國央行使用（如 Bundesbank, 2013），故除 MIDAS 模型外，亦使用橋樑方程式進行即時預報，以利比較。

二、模型與資料說明

首先考慮使用的是基本形式的 MIDAS 模型，其設定如 (2.5.1) 式。而在時間序列模型中，落後期的被解釋變數往往具有顯著的解釋能力，故亦考慮納入 AR 項之 AR-MIDAS 模型 (Clements and Galvão, 2008) 與 ADL-MIDAS 模型 (Andreou et al., 2013)。其間差異在於，AR-MIDAS 模型的 AR 項為受限 (restricted)，但 ADL-MIDAS 模型的 AR 項則未受限 (unrestricted)。

其次，經濟成長率為季資料，而欲使用之高頻指標為月資料，在頻率差距不會太大的情況下，亦可考慮使用 U-MIDAS 模型，如 (2.5.3) 式；以及將 U-MIDAS 納入 AR 項的 AR-U-MIDAS 模型。至於橋樑方程式的模型設定，本節將依 (2.4.1) 式建構基本的橋樑方程式模型 (BE)，並另建構納入 AR 項的橋樑方程式模型 (AR-BE)。

由於高頻指標係用於即時預報經濟成長率，故須與經濟活動密切相關，方有助於預報。在此考慮 3 項經濟指標，說明如下：

1. 工業生產指數年增率：經濟部統計處按月發布上月之工業生產指數，資料僅落後一個月。其與經濟成長率走勢一致，可用於衡量整體經濟活動，常作為經濟成長率的代理變數。
2. 出口數量指數年增率：財政部統計處按月發布上月之出口數量指數，資料僅落後一個月。由於我國相當依賴國際貿易，出口表現往往影響經濟成長，且經濟成長率為「實質」的概念，故選擇出

口數量指數之年增率。

3. 實質零售及餐飲業營業額年增率：經濟部統計處按月發布上月之零售及餐飲業營業額，資料僅落後一個月。由於民間消費占 GDP 比率超過五成，且零售及餐飲業營業額為衡量民間消費的重要指標，故取其年增率作為高頻指標。然而，經濟部統計處係發布零售及餐飲業營業額之名目值，故此處另以 CPI 平減為實質值。

資料的樣本期間為 2001 年第 1 季至 2017 年第 2 季，主因零售及餐飲業營業額僅能溯自 1999 年 1 月，且為取年增率並保留部分樣本作為落後期之故。另外，此處使用的資料均為修正後資料，而非即時資料 (real-time data)¹⁶。

三、即時預報結果

(一) 單一指標

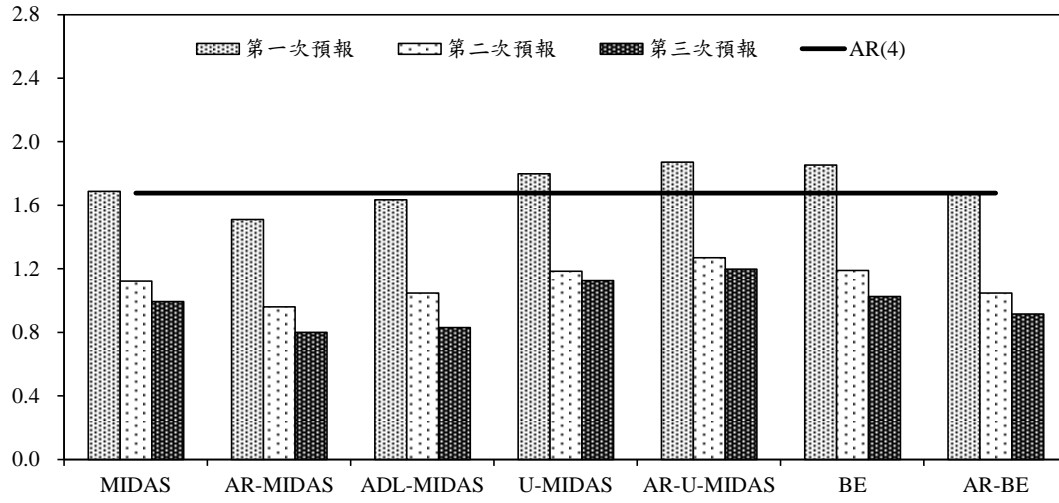
為評估各混頻模型即時預報的結果，首先以 2001 年第 1 季至 2012 年第 4 季的資料進行樣本內估計，再以遞迴的方式，逐月對該季進行樣本外即時預報，故每季均有 3 次的即時預報值。另外，此處亦以 AR(4) 模型逐季進行樣本外預測，作為各混頻模型即時預報結果的比較基準 (benchmark)。至於各模型中月資料的落後期，為納入其過去一年內的資訊，均選擇落後 12 期。

圖 1 為各模型分別使用工業生產指數年增率、出口數量指數年增率及實質零售及餐飲業營業額年增率等 3 種高頻指標，得到的經濟成長率即時預報值之 RMSE。圖中亦繪示 AR 模型向前一季的經濟成長率預測值之 RMSE，以利比較。

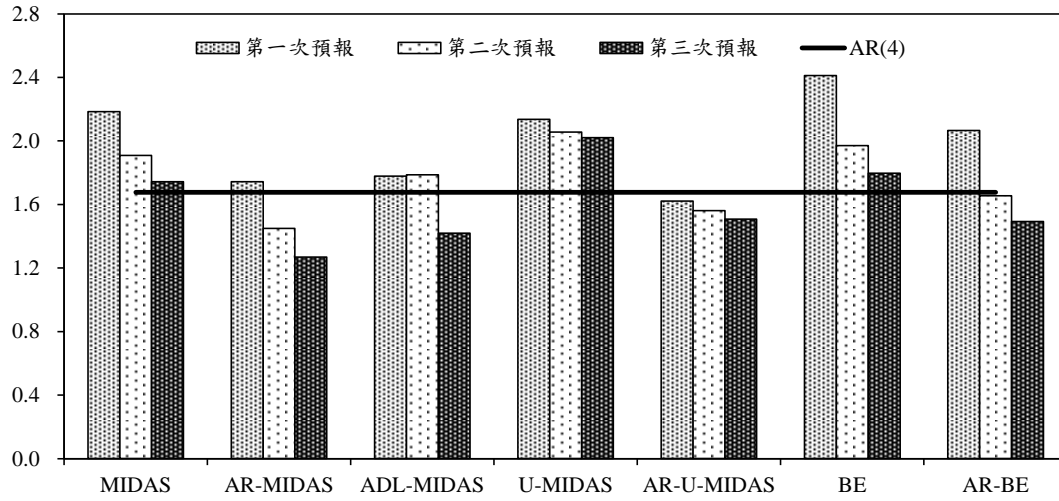
¹⁶ 僅 2017 年第 2 季之經濟成長率為即時資料 (主計總處 8 月 18 日公布之初步統計數)。

圖 1：經濟成長率即時預報值之 RMSE (單一指標)

(a) 工業生產指數年增率



(b) 出口數量指數年增率



(c) 實質零售及餐飲業營業額年增率

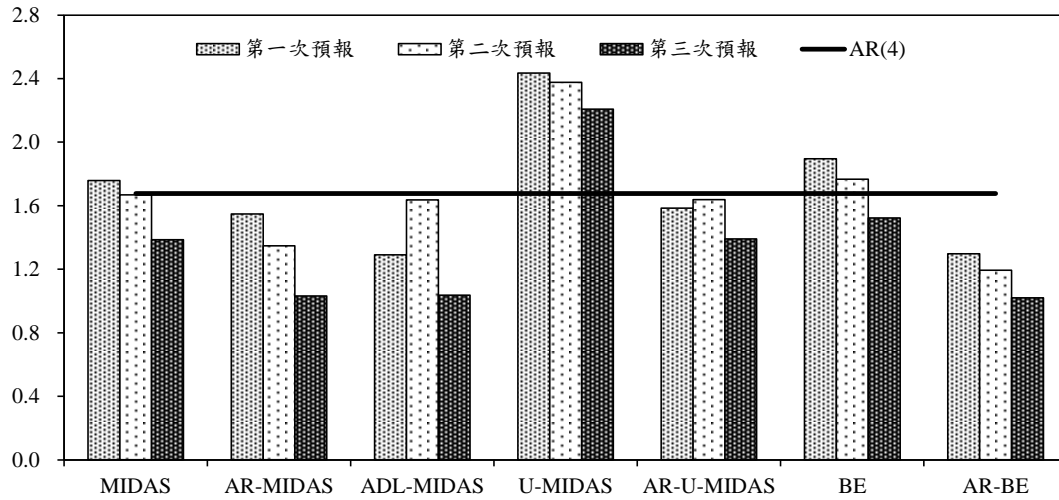


表 1 則為各模型使用 3 種不同高頻指標的即時預報值，與 AR 模型預測值的相對 RMSE。而為了解各模型的即時預報績效是否與 AR 模型的預測值顯著不同，表中另以 Diebold–Mariano 檢定檢驗兩者的 MSE 是否顯著異於 0。

根據即時預報的結果，可知無論使用的模型與高頻指標為何，3 次即時預報的 RMSE 多呈現下降的趨勢，顯示隨當季的月資料越多、資訊越豐富，即時預報的績效多可獲得改善。

在 3 種不同高頻指標中，以工業生產指數年增率的預報績效最佳。表 1 顯示，工業生產指數年增率的第一次即時預報值多與 AR 預測值無顯著差異，但在月資料增加後，第二次即時預報值的 RMSE 即已大幅下降，且績效顯著優於 AR 預測值。到第三次預報時，其 RMSE 仍持續下降、顯著程度亦續升。預報績效次佳者為實質零售及餐飲業營業額年增率，其第一次預報值亦多與 AR 預測值無顯著差異；而在第二、三次預報時，則有部分模型的績效轉為顯著。出口數量指數年增率的預報績效則較差，其第一次預報值多較 AR 預測值為差；第二、三次預報則轉為不顯著，僅 AR-MIDAS 模型在第三次預報時，在 10% 的顯著水準下顯著優於 AR 預測值。

而在各模型中，普遍以 AR-MIDAS 模型的 RMSE 最小且多呈顯著，ADL-MIDAS 模型次之，顯示納入 AR 項的 MIDAS 模型的表現最佳。AR-BE 模型亦表現頗佳，顯示橋樑方程式仍具有相當的預測能力。最後，U-MIDAS 模型的 RMSE 偏大，且其預報值大多無法顯著優於 AR 預測值，即時預報的表現較其他模型為差。

表 1：經濟成長率即時預報值之相對 RMSE（單一指標）

(a) 工業生產指數年增率

	第一次預報	第二次預報	第三次預報
MIDAS	1.007	0.670*	0.593**
AR-MIDAS	0.901*	0.573**	0.477***
ADL-MIDAS	0.974	0.625**	0.495***
U-MIDAS	1.073	0.707*	0.672**
AR-U-MIDAS	1.117*	0.757**	0.715**
BE	1.105	0.709*	0.612**
AR-BE	0.995	0.625***	0.547***

(b) 出口數量指數年增率

	第一次預報	第二次預報	第三次預報
MIDAS	1.302**	1.138	1.040
AR-MIDAS	1.040	0.864	0.757*
ADL-MIDAS	1.060	1.065	0.846
U-MIDAS	1.274*	1.226	1.205
AR-U-MIDAS	0.967	0.931	0.899
BE	1.438**	1.175	1.071
AR-BE	1.231*	0.986	0.890

(c) 實質零售及餐飲業營業額年增率

	第一次預報	第二次預報	第三次預報
MIDAS	1.048	0.995	0.826
AR-MIDAS	0.923	0.803**	0.615***
ADL-MIDAS	0.769	0.975	0.617***
U-MIDAS	1.452	1.417	1.316
AR-U-MIDAS	0.945	0.977	0.829
BE	1.130	1.054	0.908
AR-BE	0.773*	0.712***	0.608***

說明：1.表中數值為各模型經濟成長率即時預報值之 RMSE 除以 AR 模型向前一期經濟成長率預測值之 RMSE。

2.***、**與*分別表示 Diebold–Mariano 檢定統計量在 1%、5%與 10%的水準下顯著，損失函數設定為 MSE 的形式。

(二) 混合不同指標

由於結合不同的預測值可能有助改善預測結果，故本節亦仿照 Kuzin et al. (2011) 的作法，混合 (pool) 不同指標的即時預報結果。混合的方法包括簡單平均與中位數；此外，亦可賦予過去配適較為良好的指標較高的權重，故另以前 4 季 MSE 的倒數作為權重，計算不同預測值的加權平均¹⁷。

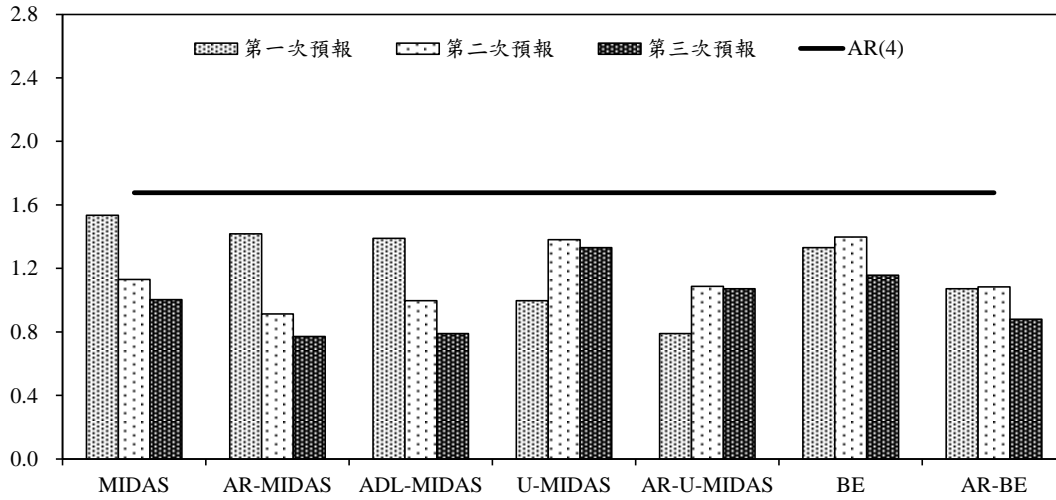
圖 2 繪示各模型混合前述 3 項不同指標的即時預報值之 RMSE，而其預報值與 AR 預測值的相對 RMSE 則列示於表 2。同樣地，表 2 提供 Diebold–Mariano 檢定統計量的顯著性，以瞭解各模型的即時預報績效是否與 AR 預測值顯著不同。

從圖 2 及表 2 可知，各模型採簡單平均或中位數的方法混合不同指標後，RMSE 多呈下降、顯著程度多呈上升，預報績效多獲得改善，表示混合不同的預測值的確有所助益，且其中又以簡單平均的整體表現較佳。其次，不同模型仍以納入 AR 項的 MIDAS 模型表現最佳（包括 AR-與 ADL-MIDAS），AR-BE 次之。另一方面，加權平均的表現較簡單平均及中位數略差。但因在此僅選取 3 個變數，故結果可能隨所採變數之增減而改變。

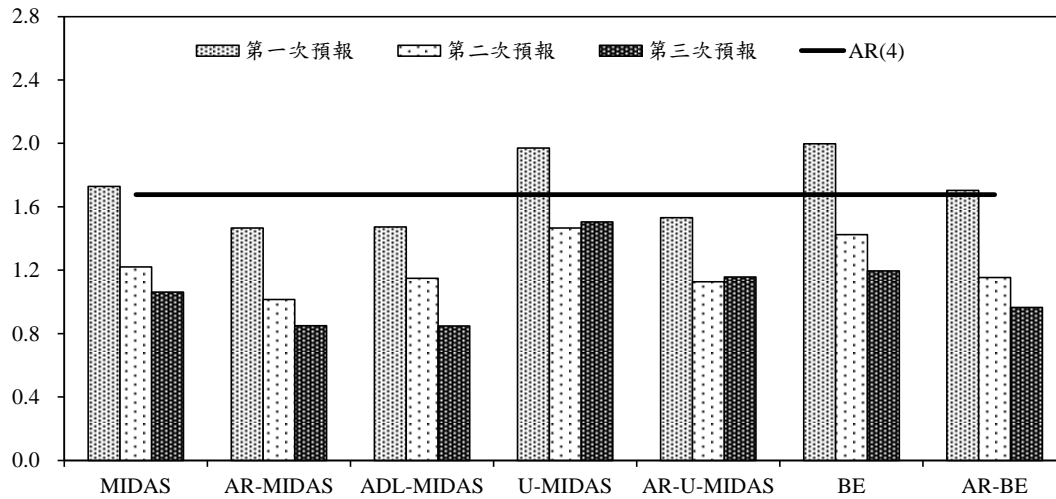
¹⁷ 如 Kuzin et al. (2011)、Schumacher (2016) 等。

圖 2：經濟成長率即時預報值之 RMSE (混合不同指標)

(a) 簡單平均



(b) 加權平均



(c) 中位數

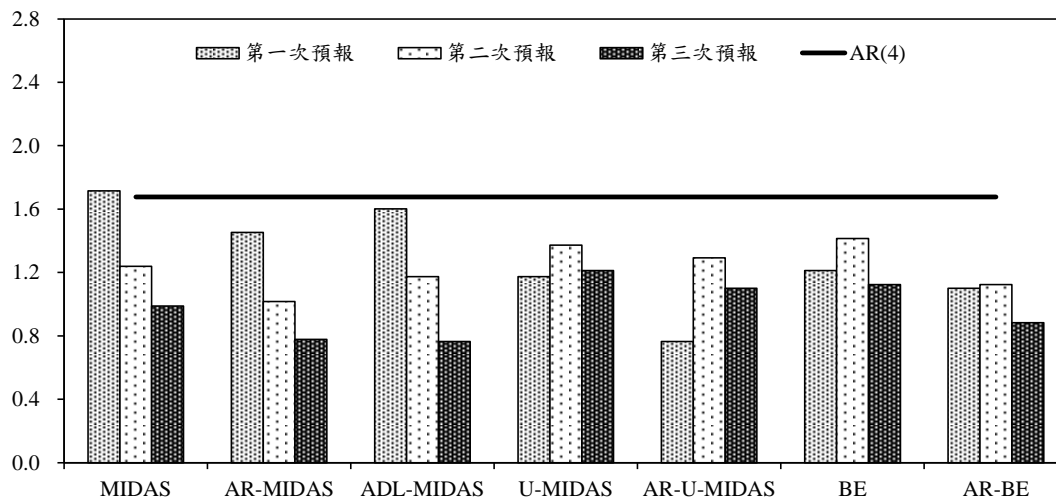


表 2：經濟成長率即時預報值之相對 RMSE（混合不同指標）

(a) 簡單平均

	第一次預報	第二次預報	第三次預報
MIDAS	0.916	0.674**	0.599***
AR-MIDAS	0.847***	0.545***	0.460***
ADL-MIDAS	0.829***	0.595***	0.471***
U-MIDAS	0.595***	0.824	0.795*
AR-U-MIDAS	0.471***	0.649***	0.640***
BE	0.795*	0.834	0.691**
AR-BE	0.640***	0.647***	0.525***

(b) 加權平均

	第一次預報	第二次預報	第三次預報
MIDAS	1.031	0.729*	0.633**
AR-MIDAS	0.875***	0.606***	0.507***
ADL-MIDAS	0.878***	0.686**	0.506***
U-MIDAS	1.176	0.875	0.898
AR-U-MIDAS	0.914	0.672***	0.691**
BE	1.192*	0.850	0.713**
AR-BE	1.016	0.689***	0.576***

(c) 中位數

	第一次預報	第二次預報	第三次預報
MIDAS	1.024	0.739*	0.590***
AR-MIDAS	0.867***	0.607***	0.465***
ADL-MIDAS	0.956	0.700**	0.457***
U-MIDAS	0.700**	0.819	0.724**
AR-U-MIDAS	0.457***	0.771**	0.657**
BE	0.724**	0.844	0.671**
AR-BE	0.657**	0.671***	0.527***

說明：1.表中數值為各模型經濟成長率即時預報值之 RMSE 除以 AR 模型向前一期經濟成長率預測值之 RMSE。

2. ***、**與*分別表示 Diebold–Mariano 檢定統計量在 1%、5%與 10%的水準下顯著，損失函數設定為 MSE 的形式。

四、結果探討

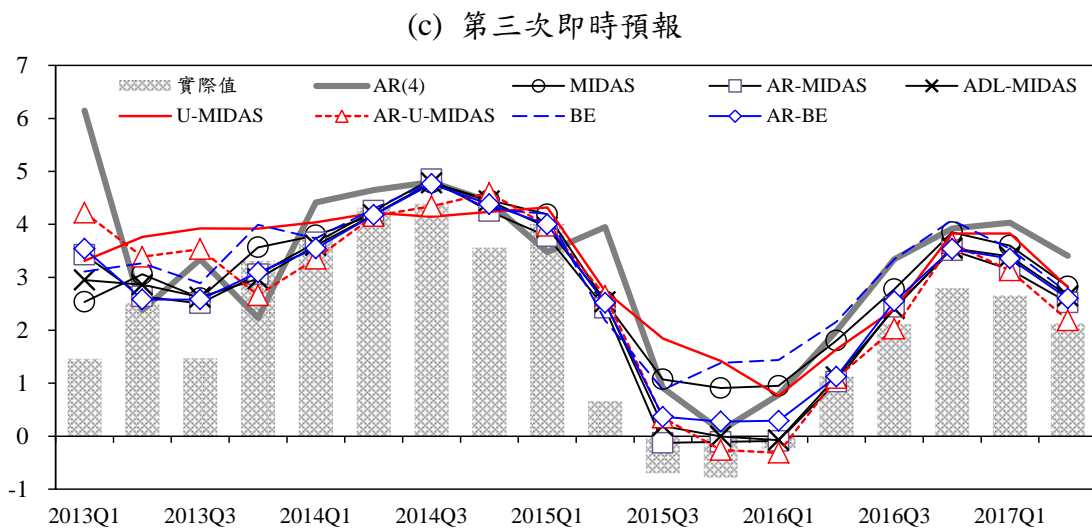
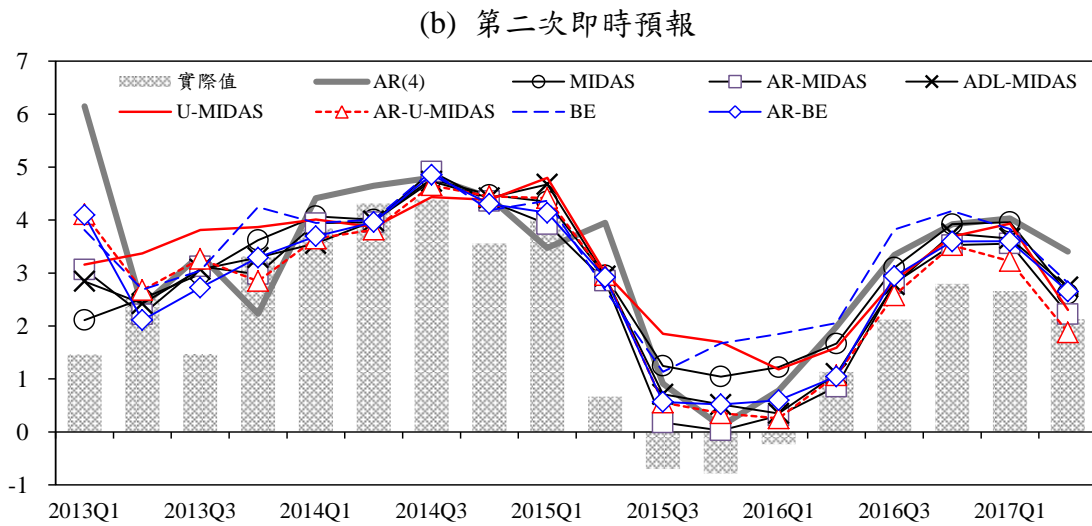
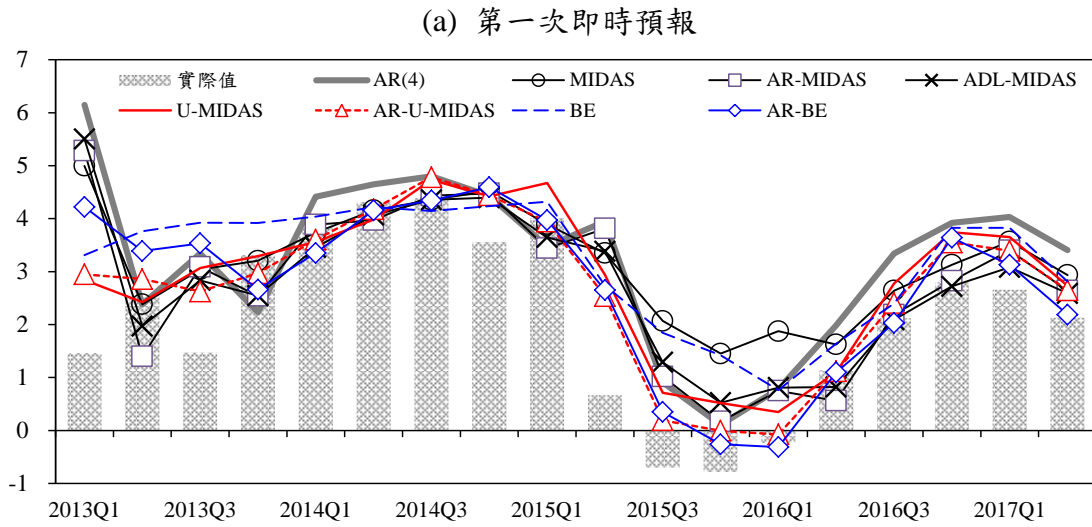
為比較經濟成長率即時預報值與實際值之差異，茲將兩者走勢繪示於圖 3。囿於篇幅限制，且整體而言，各模型以簡單平均的方式混合不同指標的預報績效最好，故圖中之即時預報值係取不同指標之簡單平均。

觀察圖 3，可發現隨月資料增加，各模型即時預報的結果將漸趨一致，其中以第一次至第二次預報的趨同程度較大，第二次至第三次預報的趨同程度則較為有限。其次，混頻模型的確有助掌握當前經濟情勢。例如，2015 年第 2 季的經濟成長率大幅下降，各預報值亦趨降，而 AR 預測值則無法即時預警。接著，前述 RMSE 結果顯示納入 AR 項模型（如 AR-與 ADL-MIDAS，以及 AR-BE）的預報績效較為良好，可能係因在經濟成長率變動幅度較大的期間（如 2015 年第 2 季至 2016 年第 1 季，經濟成長率大幅下降，其後復明顯回升），其預報值較能貼近實際值。此外，雖各模型的預報值多優於 AR 預測值，惟近來較易高估，此情形可能與模型設定或變數挑選有關¹⁸。

雖然本節的實證結果顯示，即時預報模型的確能顯著提升當季的預測績效，惟結果仍可更臻完善。例如，此處僅使用 3 種高頻指標，且選擇方式較為任意（arbitrary），故可選取更多高頻指標，或以較嚴謹的方式挑選其他指標；而除了取平均值或中位數外，亦可考慮以其他方式混合不同的預測值。另外，此處橋樑方程式的高頻指標僅以單變量模型推估其預測值，若改用 VAR 等多變量模型，亦可能有助提升橋樑方程式的預測能力。

¹⁸ 此處僅使用 3 種高頻指標，且若僅混合利用工業生產指數年增率、實質零售及餐飲業營業額年增率計算之即時預報值，則近年高估的情形將大幅改善。

圖 3：經濟成長率實際值與即時預報值（以簡單平均混合不同指標）



伍、結論與建議

本次研習內容包括經濟預測的基本概念，主要涵蓋參數隨時間變動之模型、混合頻率模型等常見的總體經濟預測模型，以及常用的預測績效評估方法及統計檢定等議題。參與本次研習課程，尚可藉此與其他學員進行經驗交流，了解其他國家的預測實務。在交流的過程中，可發現對許多國家的央行而言，專業判斷對經濟預測仍相當重要，而良好的預測模型亦有助專業判斷，兩者可相輔相成。

充分利用頻率較高且較為即時的經濟指標，有助掌握當前的經濟情勢；因此，本報告應用 MIDAS 模型與橋樑方程式等混頻模型，建構高頻指標與低頻變數之間的關係，對台灣經濟成長率進行即時預報。結果發現，即時利用高頻指標，或是混合不同高頻指標的資訊，的確有助提升預測績效；而在各模型中，又以 AR-與 ADL-MIDAS 模型表現最佳。未來仍可嘗試不同的模型設定，使結果更臻完善，或發展不同的即時預報模型等，以提升對經濟成長的短期預測能力。

另外，除了評估不同預測模型之間的相對預測績效，如前所述，亦可評估模型（或機構）預測值本身的性質。例如，預測值是否符合不偏性、理性、效率性，或是預測誤差是否具有序列相關等。近年國內此方面的文獻仍為數不多，故檢討預測值本身的絕對預測績效，亦為未來可考慮的研究方向之一。

本次課程講師均為相關領域之專家，學識及經驗均極為豐富，授課內容充實且深入；且主辦單位亦另邀請瑞典央行的經濟學家介紹該行經濟預測之分析架構，並進行經驗分享，對預測實務之精進頗有助益。因此，參與本次研習課程，當有助提升總體經濟預測方面之研究及分析能力。

參考資料

- 林依伶 (2013), 「台灣總體經濟即期季模型—考量月指標之農曆春節效果」, 中央銀行季刊, 35 (3), 61–92。
- 張志揚 (2013), 「台灣總體經濟即期季模型之建立—運用月資料改善國民所得預測」, 中央銀行季刊, 35 (3), 37–60。
- Andreou, E., E. Ghysels, and A. Kourtellos (2010), “Regression Models with Mixed Sampling Frequencies,” *Journal of Econometrics*, 158(2), 246–61.
- Andreou, E., E. Ghysels, and A. Kourtellos (2013), “Should Macroeconomic Forecasters Use Daily Financial Data and How?” *Journal of Business and Economic Statistics*, 31(2), 240–251.
- Bundesbank (2013), “Forecasting Models in Short-Term Business Cycle Analysis—A Workshop Report,” *Deutsche Bundesbank Monthly Report*, September 2013, 69–83.
- Clark, T. E., and Mccracken, M. W. (2001), “Tests of Equal Forecast Accuracy and Encompassing for Nested Models,” *Journal of Econometrics*, 105(1), 85–110.
- Clark, T. E., and Mccracken, M. W. (2013), Advances in Forecast Evaluation, in: G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Volume 2B, pp.1107–1201 (Elsevier).
- Clark, T. E., and West, K. D. (2007), “Approximately Normal Tests for Equal Predictive Accuracy in Nested Models,” *Journal of Econometrics*, 138(1), 291–311.
- Clements, M. P., and A. B. Galvão (2008), “Macroeconomic Forecasting with Mixed Frequency Data: Forecasting US Output Growth,” *Journal*

- of Business and Economic Statistics*, 26(4), 546–554.
- Davies, R. B. (1987), “Hypothesis Testing when a Nuisance Parameter is Present only under the Alternative,” *Biometrika*, 74(1), 33–43.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), “Maximum Likelihood from Incomplete Data via the EM Algorithm,” *Journal of the Royal Statistical Society: Series B*, 39(1), 1–38.
- Ferrara, L., M. G. Marcellino, and M. Mogliani (2015), “Macroeconomic Forecasting during the Great Recession: The Return of Non-Linearity?” *International Journal of Forecasting*, 31(3), 664–679.
- Froni, C., and M. G. Marcellino (2013), “A Survey of Econometric Methods for Mixed Frequency Data,” Norges Bank Working Paper.
- Froni, C., M. G. Marcellino, and C. Schumacher (2015), “Unrestricted Mixed Data Sampling (MIDAS): MIDAS Regressions with Unrestricted Lag Polynomials,” *Journal of the Royal Statistical Society: Series A*, 178(1), 57–82.
- Ghysels, E., and M. G. Marcellino (2017), *Applied Economic Forecasting Using Time Series Methods*. Oxford University Press (forthcoming).
- Ghysels, E., A. Rubia, and R. Valkanov (2009), “Multi-Period Forecasts of Volatility: Direct, Iterated, and Mixed-Data Approaches,” Discussion Paper UNC/UCSD.
- Ghysels, E., P. Santa-Clara and R. Valkanov (2004), “The MIDAS Touch: Mixed Data Sampling Regressions,” Discussion Paper UNC/ UCLA.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2006), “Predicting Volatility: Getting the Most Out of Return Data Sampled at Different Frequencies,” *Journal of Econometrics*, 131(1), 59–95.

- Hamilton, J. D. (1994), *Time Series Analysis*. Princeton University Press, Princeton.
- Hansen, B. E. (1996), “Inference when a Nuisance Parameter is not Identified under the Null Hypothesis,” *Econometrica* 64(2), 413–430.
- Harvey, A. C. (1993), *Time Series Models*. Harvester, New York.
- Kuzin, V., M. Marcellino, and Schumacher, C. (2011), “MIDAS vs. Mixed-Frequency VAR: Nowcasting GDP in the Euro Area,” *International Journal of Forecasting*, 27(2), 529–542.
- Mariano, R., and Y. Murasawa (2010), “A Coincident Index, Common Factors, and Monthly Real GDP,” *Oxford Bulletin of Economics and Statistics*, 72(1), 27–46.
- McCracken, M. W. (2007), “Asymptotics for Out-of-Sample Tests of Granger Causality,” *Journal of Econometrics*, 140(2), 719–52.
- McCracken, M. W., M. Owyang, and T. Sekhposyan (2015), “Real-Time Forecasting with a Large, Mixed Frequency, Bayesian VAR,” Federal Reserve Bank of St. Louis Working Paper 2015-030A.
- Rossi, B. (2013), *Advances in Forecasting Under Instability*, in: G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Volume 2B. Elsevier, Amsterdam, pp.1201–1321.
- Schumacher, C. (2016), “A Comparison of MIDAS and Bridge Equations,” *International Journal of Forecasting*, 32(2), 257–270.
- Stock, J. H., and M. W. Watson (2002), “Macroeconomic Forecasting Using Diffusion Indexes,” *Journal of Business and Economic Statistics*, 20(2), 147–162.
- Study Center Gerzensee (2017), 「Macroeconomic Forecasting」 研習課

程講義。

West, K. D. (1996), “Asymptotic Inference about Predictive Ability,”
Econometrica, 64(5), 1067–1084.

West, K. D. (2001), “Tests of Forecast Encompassing when Forecasts Depend on Estimated Regression Parameters,” *Journal of Business and Economic Statistics*, 19(1), 29–33.

West, K. D. (2006), Forecast Evaluation, in: G. Elliott, C. Granger, and A. Timmermann (ed.), *Handbook of Economic Forecasting*, Volume 1. Elsevier, Amsterdam, pp.99–134.