

# 運用數位分析工具掌握臺灣景氣走勢\*

廖 銘 傳\*\*

壹、前言

貳、文獻回顧

參、研究方法

肆、實證分析

伍、結論與建議

## 摘 要

本研究運用大量經濟變數及其落後期，預測未來 1 個月景氣收縮機率，採用每月新增資料方式，進行模擬即時分析。將訓練資料以 LASSO 迴歸挑選變數後，分別使用 Probit 模型、彈性網路 (elastic net)、隨機森林 (random forest) 建立模型，再以驗證資料預測未來 1 個月景氣收縮機率，使用 ROC 曲線下面積 (area under the curve, AUC) 評估預測表現。

實證結果顯示，隨機森林、彈性網路預測表現較好，1982 年資料集之最佳模型為隨機森林，AUC 可達 92.9%；2001 年資料集之最佳模型為彈性網路，AUC 可達 91.6%，且兩種最佳模型預測之景氣收縮機率，皆能對應我國景氣峰谷認定結果。

觀察臺灣第 14 次景氣循環後之景氣走勢，2018 年 3 月起景氣收縮機率走揚，反應美中貿易紛爭爆發；2020 年 1 月起景氣收縮機率上升，則顯示武漢肺炎 (COVID-19) 疫情對我國經濟之衝擊。

\* 本文參加國發會 2020 年研究發展作品評選，榮獲數位經濟與創新科技類佳作獎。

\*\* 作者為經濟發展處科員。本文係筆者個人觀點，不代表國發會意見，若有疏漏之處當屬筆者之責。

## Use Artificial Intelligence Methods to Grasp Taiwan's Business Cycle

Ming-Chuan Liao

*Officer*

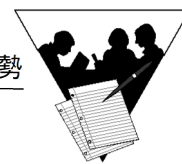
*Economic Development Department, NDC*

### Abstract

This study uses a large number of economic variables and their lags to predict the contraction probability in the next month, and uses monthly new data to conduct simulation and real-time analysis. We use the LASSO regression to select important variables in training data, and then apply the Probit, elastic net, and random forest model to predict the contraction probability in the next month in the validation data. We use the area under the ROC curve (AUC) to evaluate forecast performance.

The empirical results show that random forest and elastic net models perform better. The best model in the data set of 1982 is random forest, with AUC reaching 92.9%; the best model in the data set of 2001 is elastic net, with AUC reaching 91.6%. The contraction probabilities of the two best models can correspond to the peaks and troughs of Taiwan's business cycle.

Observing the trend after the 14th Taiwan's business cycle, the contraction probability has increased since March 2018, reflecting the beginning of the U.S.-China trade war; the contraction probability has soared since January 2020, indicating the influence of the COVID-19.



## 壹、前言

### 一、研究緣起

Burns and Mitchell (1946) 描述景氣循環 (business cycle) 為國家總體經濟活動之波動，許多經濟活動大約同時發生復甦、擴張，隨後發生收縮、衰退現象，一個景氣循環包含由谷底 (trough) 至高峰 (peak) 之擴張期 (expansion) 及高峰至谷底之收縮期 (contraction)。

國家發展委員會 (以下稱國發會) 主要業務之一為掌握景氣動向，並依據當前經濟情勢，適時規劃經濟發展對策。故國發會按月發布景氣對策信號及景氣指標，藉由景氣對策信號之更迭，以及領先、同時、落後指標之趨勢變化，掌握當前景氣狀態。

隨著數位時代來臨，數據呈現爆炸性成長，且資料分析方法亦有長足進步，如人工智慧 (artificial intelligence, AI)、機器學習 (machine learning, ML)、深度學習 (deep learning, DL)、文字探勘 (text mining) 方法，皆在各領域取得重要成果。

由於主要國家及國際機構開始運用新方法進行經濟預測，本研究嘗試結合大量經濟變數及數位分析工具，預測未來 1 個月景氣收縮機率，輔助判讀景氣走勢。

### 二、研究目的

臺灣屬於小型開放經濟體，經濟景氣與全球連動性高，加以臺灣在各個經濟發展階段，景氣循環態樣有所不同，如何掌握景氣動態成為一項挑戰。本研究目的之一為建立預測機制，以有效掌握景氣動態或提前預警景氣轉折出現。

另由於資料分析方法推陳出新，OECD 表示機器學習方法為未來具有潛力之分析工具，且國內經濟預測使用非傳統計量方法之成果尚少，本研究嘗試利用機器學習方法中的彈性網路（elastic net）及隨機森林（random forest）進行預測，並與傳統計量方法比較。

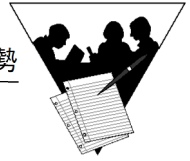
除此之外，本研究考慮大量變數攜帶之豐富資訊，於分析過程挑選具有代表性之變數進行預測，並觀察變數與景氣狀態之領先、同時、落後特性，挑選之重要變數亦可列入景氣指標構成項目檢討參考。

## 貳、文獻回顧

本研究目的為即時掌握景氣走勢，相關文獻可分為景氣狀態預測、景氣轉折點認定兩種，前者使用具有領先性的經濟變數進行預測，後者則觀察經濟體系各部門變數之趨勢變化，認定景氣高峰或谷底，以下分別探討相關文獻。

景氣狀態預測方面，Rudebusch and Williams（2009）發現僅使用殖利率曲線之預測模型，其預測衰退期的表現，比總體經濟專家更好，總體經濟專家可能低估殖利率曲線攜帶之資訊。Berge（2015）評估幾種經濟指標之景氣轉折點預測能力，發現最佳模型考慮變數非線性行為，且包含殖利率曲線及具有領先性之變數。

Chauvet and Potter（2005）考慮誤差項具序列相關，且各循環週期具有不同變異數之延伸 Probit 模型，發現配適效果較標準模型好。Kauppi and Saikkonen（2008）以動態 Probit 模型預測景氣衰退期，結果顯示動態模型之預測表現較靜態模型佳，且發現



利差為重要變數，若模型加入景氣狀態落後期，可提升預測表現。

Fossati (2016) 分別使用 4 個變數估計之動態因子、102 個變數萃取之共同因子，運用標準 Probit 模型、自迴歸 Probit 模型、馬可夫轉換模型 (Markov switching models) 預測景氣衰退機率。實證結果顯示，使用較多變數萃取之共同因子進行分析，其馬可夫轉換模型的樣本外預測結果，較接近 NBER 公布的衰退期。

Ng (2013) 使用機器學習之 boosting 預測景氣衰退，自 132 個時間序列及其落後期，總計 1,500 個變數進行篩選，發現重要變數少於 10 個，其中 Aaa 利差為預測未來 3 個月或 6 個月衰退之重要變數，風險債券、5 年期利差則為預測未來 12 個月衰退之重要變數；除此之外，當利差不能提供資訊時，某些就業變數對預測最近 2 次景氣衰退有所幫助。

國內文獻方面，黃裕烈 (2019) 依據芝加哥聯邦準備銀行做法，篩選 273 個財經變數，建構金融情勢指數 (Financial Conditions Index, FCI)，再進一步運用 FCI、領先指標、同時指標及工業生產指數，以 ProbVAR 模型預測景氣衰退機率，研究發現加入 FCI 能改善模型預測效果。

徐士勛等人 (2020) 運用 Stock and Watson (2002) 提出之 EM 演算法，萃取經濟趨勢調查之共同因子，並運用動態 Probit 模型，以領先指標或同時指標結合共同因子，預測未來 1、3、6 個月之景氣衰退機率。研究發現趨勢調查共同因子可提升領先指標之預測能力，故能提供觀測景氣動向之額外資訊。

景氣轉折點認定方面，Bry and Boschan ( 1971 )、Harding and Pagan ( 2002 ) 均提出事後認定方法，觀察一個綜合各項經濟活動之變數，以認定景氣循環轉折點。

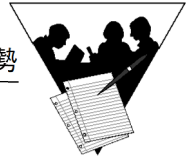
Bry and Boschan ( 1971 ) 以 12 個月移動平均、Spencer 曲線平滑化數列，再計算 3 至 6 個月短期移動平均，並搭配限制條件認定轉折點。Harding and Pagan ( 2002 ) 修正 Bry and Boschan ( 1971 ) 方法，刪除 Spencer 曲線平滑化過程，先決定可能的轉折點集合，確保高峰與谷底交替出現，再考量循環持續時間、深度及完整性後，去除不必要的轉折點。

許多文獻運用馬可夫轉換模型認定景氣轉折點，如 Chauvet ( 1998 ) 萃取總體經濟變數之共同因子，允許共同因子考慮擴張期、收縮期之轉換結構，以馬可夫轉換過程結合動態因子模型認定景氣轉折點。該方法認定之轉折點與 NBER 結果相近，且發現共同因子與美國商務部公布之同時指標相似。

Chauvet and Hamilton ( 2006 ) 使用狀態空間馬可夫轉換模型認定景氣轉折點，並發展以 GDP 為基礎之「每季即時衰退機率指標」、以多指標為基礎之「每月即時衰退機率指標」，實證結果顯示有助即時認定景氣狀態。

Chauvet and Piger ( 2008 ) 使用 Harding and Pagan ( 2002 ) 無母數方法及馬可夫轉換動態因子模型，即時認定美國景氣轉折點。實證結果發現，馬可夫轉換動態因子模型認定谷底速度較 NBER 快，且兩種方法皆可認定 NBER 公布之轉折點，不會有偽陽性問題。

Stock and Watson ( 2014 ) 觀察美國工業生產、非農業部門就業、實質製造業及批發零售業銷售、實質個人所得淨額等 4 類別



計 270 個總體經濟變數，建立大量經濟變數之轉折點抽樣分配後，以分層隨機抽樣、類別領先落後特性進行調整。實證結果顯示，以調整後之轉折點抽樣分配的眾數，做為整體景氣轉折點，最適合描述美國景氣循環狀態。

周大森（2019）延伸 Stock and Watson（2014）的做法，觀察生產、銷售、勞動市場、所得等 4 類別計 234 個經濟變數，以及加上對外貿易等 5 類別計 338 個經濟變數，建立轉折點之抽樣分配後，以分層隨機抽樣、類別領先落後特性、訊噪比（signal-noise ratio, SNR）進行調整，發現以訊噪比調整之轉折點抽樣分配的中位數，做為臺灣景氣循環轉折點，具有較小的認定誤差。

Giusto and Piger（2017）將擴張期及收縮期視為兩種狀態，將認定景氣轉折點轉換為分類問題，利用美國非農就業人數、工業生產指數、排除移轉性收入的實質個人所得、實質製造業銷售額做為分類依據，以機器學習方法中的學習向量量化法（learning vector quantization, LVQ），即時認定景氣轉折點。實證結果顯示，學習向量量化法能有效且即時認定景氣收縮期。

周大森（2019）參考 Giusto and Piger（2017）做法，利用基準循環之構成項目，包括實質 GDP、工業生產指數、非農業部門就業人數、製造業銷售量指數、批發零售及餐飲營業額、實質海關出口值，即時認定景氣轉折點。實務上落後轉折時點 6 個月，即可認定該次轉折點，故能有效提升即時性。

機器學習方法應用文獻方面，Jung, Patnam and Ter-Martirosyan（2018）使用彈性網路、超級學習（super learner）、循環神經網路（recurring neural network, RNN），預測包含先進及新興經濟體等 7 個國家之經濟成長率。實證結果顯示，3 種機器學習方法的預測表現，均優於 IMF 世界經濟展望的預測值，雖然

機器學習方法不能完全取代專家預測，但在經濟預測領域可提供額外有價值的資訊。

Woloszko (2020) 發展適應樹 (adaptive trees)，處理經濟預測遭遇之非線性及未知結構改變問題，透過梯度提升樹 (gradient boosted trees) 加權近期觀測值，藉以反應結構變化。實證研究對 6 個國家未來 3 至 12 個月之經濟成長率，進行模擬即時分析，結果適應樹預測表現與 OECD 模型相近，且較 AR (1)、隨機森林、梯度提升樹來得好。

經由文獻回顧探討，多種機器學習方法如彈性網路、隨機森林已用於經濟預測，且殖利率曲線、利差等變數對景氣狀態預測或轉折點認定有所幫助，亦建議篩選變數或萃取共同因子，以納入大量變數攜帶資訊，本研究將參考相關文獻做法，進行後續分析。

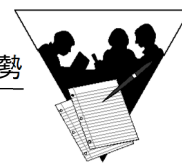
## 參、研究方法

### 一、研究流程

首先定義景氣狀態為  $Y_t$ ， $Y_t = 1$  代表收縮期， $Y_t = 0$  代表擴張期，共有  $T + 1$  期資料，每期有  $M$  個經濟變數。由於經濟變數常受到季節、交易日、春節效應影響而產生季節性變化，增加研究分析的困難，故本研究對經濟變數進行季節調整，使其更能反應景氣循環之變動狀況。

本研究為掌握景氣動態，關注各變數組成之循環成分，參考國發會編製景氣指標方法，以二階段 HP-filter 將經濟變數取出循環成分，另參考周大森 (2019) 學習向量量化法之變數處理方法，計算循環成分之月增率，將經濟變數循環成分月增率  $x_t$  用於





後續分析。

另因不同經濟變數對景氣具有相異的領先、同時、落後特性，故將落後期納入考量，定義變數集合  $\omega_t = \{x_t, x_{t-1}, \dots, x_{t-L}\}$ ，包含當期至落後  $L$  期變數，計有  $P = M \times (L + 1)$  個變數。

由於採取當期至落後  $L$  期變數，預測未來 1 期景氣狀態，將景氣狀態  $Y_{t+1}$  與變數集合  $\omega_t$  之關係以數學式表示，其中  $f(\cdot)$  代表使用的模型， $\beta$  代表模型參數， $\varepsilon_{t+1}$  則代表誤差項：

$$Y_{t+1} = f(x_t, x_{t-1}, \dots, x_{t-L}; \beta) + \varepsilon_{t+1}, t = L + 1, L + 2, \dots, T$$

為即時掌握景氣走勢，本研究進行模擬即時分析，每月加入 1 期資料後，重新以二階段 HP-filter 取出循環成分，再計算循環成分月增率後，納入落後期得到新的變數集合  $\omega_t$ 。

將已認定景氣狀態期間做為訓練資料，經 LASSO 迴歸挑選變數後，以 Probit 模型、彈性網路 (elastic net)、隨機森林 (random forest) 建立預測模型，再將未認定景氣狀態期間做為驗證資料，以驗證資料預測未來 1 個月景氣收縮機率，並使用 AUC 評估預測表現。(詳圖 1)

## 二、資料處理

### (一) 季節調整

許多經濟變數受到生產週期、風俗習慣或假期影響，具有季節性變化，將會影響景氣狀態判斷，季節調整即是以統計方法進行調整，使得經濟變數較能反應景氣變動狀況。本研究使用 R 軟體之 seasonal 套件，考慮春節效果、交易日效果及不同 ARIMA 模型，實作 X-13-ARIMA 季節調整程序。

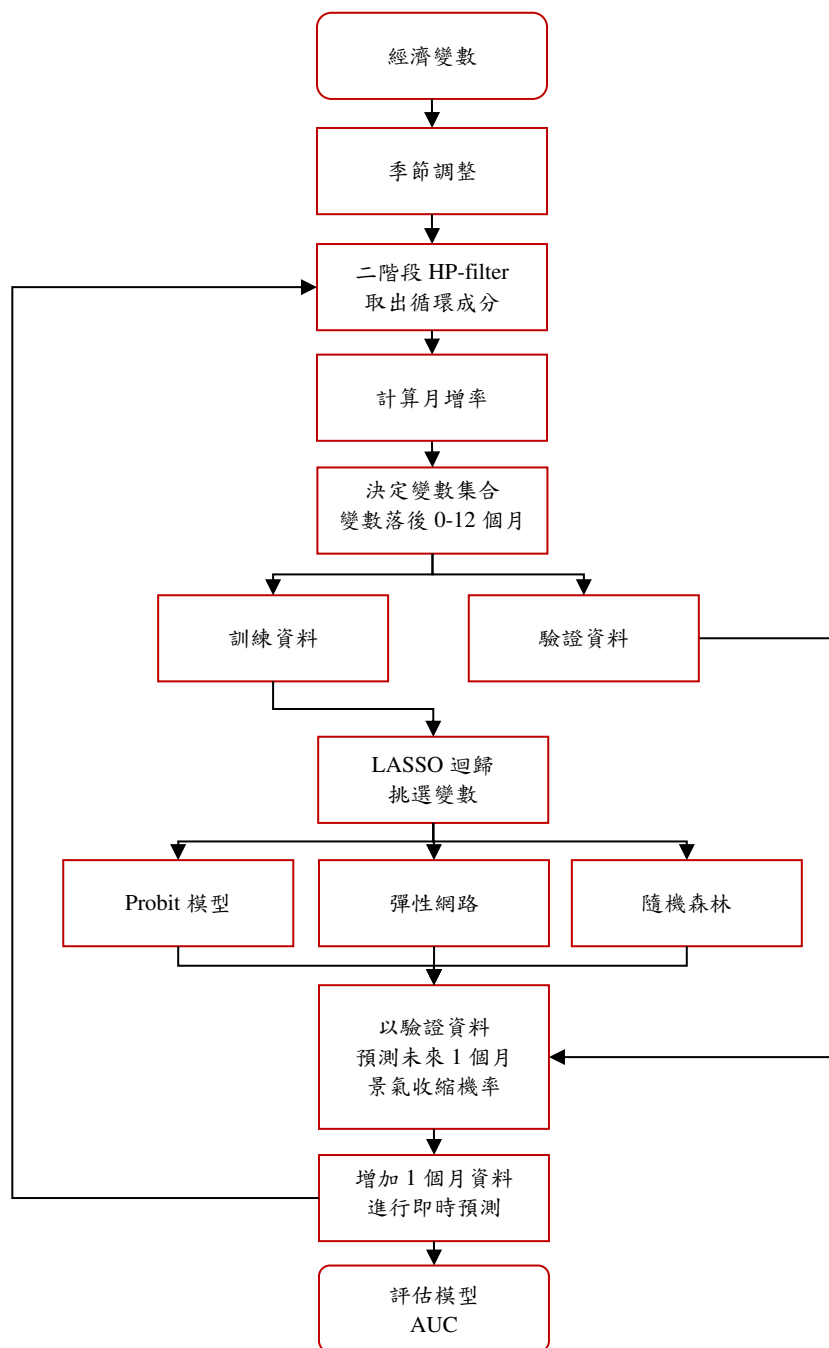
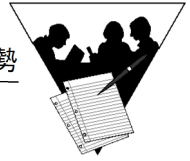


圖 1 研究流程



## (二) 二階段 HP-filter

由於我國景氣循環採取 OECD 成長循環觀點，且經濟變數具有長期趨勢及短期波動，故本研究以 HP-filter (Hodrick-Prescott, 1997) 去除變數之長期趨勢後，再將去除長期趨勢之數列，以 HP-filter 去除短期波動，得到經濟變數之循環成分。

## (三) 計算月增率

由於景氣狀態與循環成分變動密切相關，本研究參考周大森 (2019) 學習向量量化法之變數處理方法，計算經濟變數之循環成分月增率，進行後續分析。

## 三、LASSO 迴歸

本研究之目的為預測未來 1 個月景氣狀態，故將景氣狀態  $Y_{t+1}$  做為反應變數，其中  $Y_{t+1} = 1$  表示收縮期， $Y_{t+1} = 0$  表示擴張期，以大量變數及其落後期之變數集合  $\omega_t = \{x_t, x_{t-1}, \dots, x_{t-L}\}$  做為解釋變數。

由於大量變數之間經常有共線性問題，參數估計將會不穩定，文獻主要有兩種處理方式，分別為萃取共同因子及篩選變數，如 Chauvet (1998)、Chauvet and Hamilton (2006)、Fossati (2016)、徐士勛等人 (2020) 萃取變數之共同因子，納入大量變數攜帶資訊；Ng (2013) 則使用 boosting 方法直接篩選重要變數。

除此之外，Tibshirani (1996) 提出 LASSO 迴歸，將線性迴歸之誤差平方和 (residual sum of squares, RSS) 加入限制式，係數絕對值之和必須小於某個參數  $\lambda$  值。透過此種設定，某些參數估計值會收縮 (shrinkage) 至 0，達成變數選取程序，此方法亦

能推廣至廣義線性模型 ( generalized linear model, GLM ) 及決策樹 ( decision tree )。

本研究考慮臺灣在各個經濟發展階段，景氣循環態樣有所不同，加以 Smeekes and Wijler ( 2018 ) 發現，相較於動態因子模型，LASSO 迴歸在大量數據具有較好的穩健性，故本研究之模擬即時分析，使用 LASSO 迴歸選取變數，以因應不同階段可能有不同重要變數之現象。

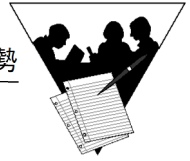
假設訓練資料由第  $L+1$  期至第  $T'$  期，將景氣狀態  $Y_{t+1}$  做為反應變數，變數集合  $\omega_t$  寫成向量形式  $\omega_t = [x_t, x_{t-1}, \dots, x_{t-L}]^T$  做為解釋變數，建立羅吉斯迴歸 ( logistic regression )：

$$P(Y_{t+1} = 1) = p_{t+1} = \frac{\exp(\omega_t^T \beta)}{1 + \exp(\omega_t^T \beta)}$$

羅吉斯迴歸對應之對數概似比函數 ( log likelihood function )  $\ell(\beta)$  為以下形式：

$$\begin{aligned} \ell(\beta) &= \sum_{t=L+1}^{T'} [Y_{t+1} \log(p_{t+1}) + (1 - Y_{t+1}) \log(1 - p_{t+1})] \\ &= \sum_{t=L+1}^{T'} \left[ Y_{t+1} \log \left( \frac{p_{t+1}}{1 - p_{t+1}} \right) + \log(1 - p_{t+1}) \right] \\ &= \sum_{t=L+1}^{T'} [Y_{t+1} \omega_t^T \beta - \log(1 + \exp(\omega_t^T \beta))] \end{aligned}$$

若將  $\ell(\beta)$  加入係數絕對值之和小於  $\lambda$  值的限制，重新定義對數概似比函數  $\ell_{LASSO}(\beta)$ ，則  $\ell_{LASSO}(\beta)$  將有 LASSO 迴歸結構，再以最大概似法得到參數估計值  $\hat{\beta}_j$ ：



$$\ell_{LASSO}(\beta) = \sum_{t=L+1}^{T'} [Y_{t+1} \omega_t^T \beta - \log(1 + \exp(\omega_t^T \beta))] - \underbrace{\lambda \sum_{j=1}^{M(L+1)} |\beta_j|}_{\text{LASSO 迴歸懲罰項}}$$

LASSO 迴歸參數 $\lambda$ 值必須事先決定， $\lambda$ 值愈大代表懲罰效果愈大，代表有愈多參數估計值收縮至 0。本研究運用 R 軟體之 `cv.glmnet` 函數，以交叉驗證 (cross validation) 方法決定適當 $\lambda$ 值，使 $\ell_{LASSO}(\beta)$ 為最大。

得到適當 $\lambda$ 值之後，重新運用訓練資料，以 LASSO 迴歸選取參數估計值不為 0 之變數，變數集合由 $\omega_t = \{x_t, x_{t-1}, \dots, x_{t-L}\}$ 縮小至 $\omega'_t = \{x'_{1,t}, x'_{2,t}, \dots, x'_{p,t}\}$ ，變數個數由 $M \times (L+1)$ 個減少至 $p$ 個。

#### 四、Probit 模型

由於景氣狀態分為擴張期及收縮期，故預測未來 1 個月景氣狀態可轉換為分類問題，將景氣狀態 $Y_{t+1}$ 做為反應變數，變數集合 $\omega'_t = \{x'_{1,t}, x'_{2,t}, \dots, x'_{p,t}\}$ 為解釋變數，以 Probit 模型進行預測。

由於景氣僅有擴張期 ( $Y_{t+1} = 0$ ) 及收縮期 ( $Y_{t+1} = 1$ ) 兩種狀態，與無法觀測到的潛在變數 (latent variable)  $Y_{t+1}^*$ 有關，可運用指標函數 (indicator function) 表示：

$$Y_{t+1} = \begin{cases} 1, & \text{若 } Y_{t+1}^* > 0 \\ 0, & \text{若 } Y_{t+1}^* \leq 0 \end{cases}$$

潛在變數 $Y_{t+1}^*$ 與變數集合 $\omega'_t$ 具有線性關係：

$$Y_{t+1}^* = \beta_0 + \beta_1 x'_{1,t} + \dots + \beta_p x'_{p,t} + \varepsilon_{t+1}, \quad t = L+1, L+2, \dots, T'$$

$$\varepsilon_{t+1} \sim iid N(0, 1)$$

Probit 模型之機率分配，可運用景氣狀態 $Y_{t+1}$ 與潛在變數 $Y_{t+1}^*$ 的關係進行推導，其中 $\Phi$ 為標準常態累積分配函數：

$$\begin{aligned} P(Y_{t+1} = 1) &= P(Y_{t+1}^* > 0) \\ &= P(\beta_0 + \beta_1 x'_{1,t} + \cdots + \beta_p x'_{p,t} + \varepsilon_{t+1} > 0) \\ &= P[\varepsilon_{t+1} > -(\beta_0 + \beta_1 x'_{1,t} + \cdots + \beta_p x'_{p,t})] \\ &= \Phi(\beta_0 + \beta_1 x'_{1,t} + \cdots + \beta_p x'_{p,t}), \quad t = L + 1, L + 2, \dots, T' \end{aligned}$$

在參數估計方面，由於 Probit 模型為非線性模型，必須以最大概似估計法（maximal likelihood estimation, MLE）或非線性最小平方方法（nonlinear least squares, NLS）進行估計。

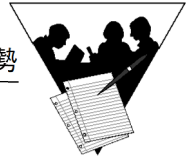
本研究以 R 軟體之 glm 函數進行估計，該函數透過 IRLS（iteratively reweighted least squares）演算法，得到參數最大概似估計值 $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ ，再預測未來 1 個月景氣收縮機率 $\hat{p}_{t+1}$ ：

$$\hat{p}_{t+1} = \Phi(\hat{\beta}_0 + \hat{\beta}_1 x'_{1,t} + \cdots + \hat{\beta}_p x'_{p,t})$$

## 五、彈性網路

Zou and Hastie（2005）提出彈性網路，主要概念為同時考慮脊迴歸（ridge regression）及 LASSO 迴歸懲罰項，可處理大量變數之共線性問題，亦能做為變數選取方法。實證研究及模擬結果顯示，相對於 LASSO 迴歸，彈性網路的預測效果通常較好，且當變數個數遠大於觀測值時，彈性網路將特別有用。

假設訓練資料由第 $L + 1$ 期至第 $T'$ 期，景氣狀態 $Y_{t+1}$ 為反應變數，變數集合 $\omega'_t$ 寫成向量形式 $\omega'_t = [x'_{1,t}, x'_{2,t}, \dots, x'_{p,t}]^T$ 為解釋變數，對數概似比函數 $\ell_{ElasticNet}(\beta)$ 與 LASSO 迴歸相似，再加入



脊迴歸懲罰項：

$$\ell_{ElasticNet}(\boldsymbol{\beta}) = \sum_{t=L+1}^{T'} \left[ Y_{t+1} \boldsymbol{\omega}'_t \boldsymbol{\beta} - \log \left( 1 + \exp(\boldsymbol{\omega}'_t \boldsymbol{\beta}) \right) \right] -$$

$$\lambda \sum_{j=1}^p \left[ \underbrace{\frac{(1-\alpha)}{2} (\beta_j)^2}_{\text{脊迴歸懲罰項}} + \underbrace{\alpha |\beta_j|}_{\text{LASSO 迴歸懲罰項}} \right]$$

彈性網路之參數 $\lambda$ 值及 $\alpha$ 值必須事先決定，其中 $\lambda$ 值效果與 LASSO 迴歸相同，數值愈大代表有愈多參數估計值收縮至 0，此部分亦以交叉驗證方法決定適當 $\lambda$ 值，使 $\ell_{ElasticNet}(\boldsymbol{\beta})$ 為最大。

$\alpha$ 值介於 0 至 1，為脊迴歸與 LASSO 迴歸懲罰項之權重，當 $\alpha = 0$ 時，會成為脊迴歸之特例，而當 $\alpha = 1$ 時，會成為 LASSO 迴歸之特例。為探討不同 $\alpha$ 值之預測效果，本研究在後續章節將進行模擬， $\alpha$ 值由 0 開始，每隔 0.1 設定為新的 $\alpha$ 值，直到 $\alpha = 1$ 為止，分別建立彈性網路進行分析。

設定適當 $\lambda$ 值及 $\alpha$ 值之後，以最大概似法得到彈性網路參數估計值 $\hat{\beta}_j$ ，將 $p$ 個參數估計值寫成向量形式 $\hat{\boldsymbol{\beta}}$ ，預測未來 1 個月景氣收縮機率 $\hat{p}_{t+1}$ ：

$$\hat{p}_{t+1} = \frac{\exp(\boldsymbol{\omega}'_t \hat{\boldsymbol{\beta}})}{1 + \exp(\boldsymbol{\omega}'_t \hat{\boldsymbol{\beta}})}$$

## 六、隨機森林

Breiman (2001) 提出隨機森林，為一種決策樹之組合預測方法，主要概念為以抽出放回之隨機抽樣，抽出多組隨機樣本後，分別建立 CART (classification and regression tree) 決策樹，

再以決策樹投票結果進行分類或預測，透過投票平均可改善準確性及控制過度配適（over-fitting）問題。

由於進行抽出放回之隨機抽樣，隨機森林的決策樹之間彼此獨立，且具有相同的分配。隨機森林之誤差決定於個別樹的強度，以及決策樹之間的相關性，隨著決策樹個數增加，隨機森林之誤差會收斂到極限。

假設訓練資料由第 $L+1$ 期至第 $T'$ 期，景氣狀態 $Y_{t+1}$ 為反應變數，變數集合 $\omega'_t = \{x'_{1,t}, x'_{2,t}, \dots, x'_{p,t}\}$ 計有 $p$ 個解釋變數。隨機森林每次挑選之變數比例為 $p'/p$ ，故挑出 $p'$ 個變數建立決策樹，共進行 $m$ 次隨機抽樣，總計產生 $m$ 個決策樹。（詳圖2）

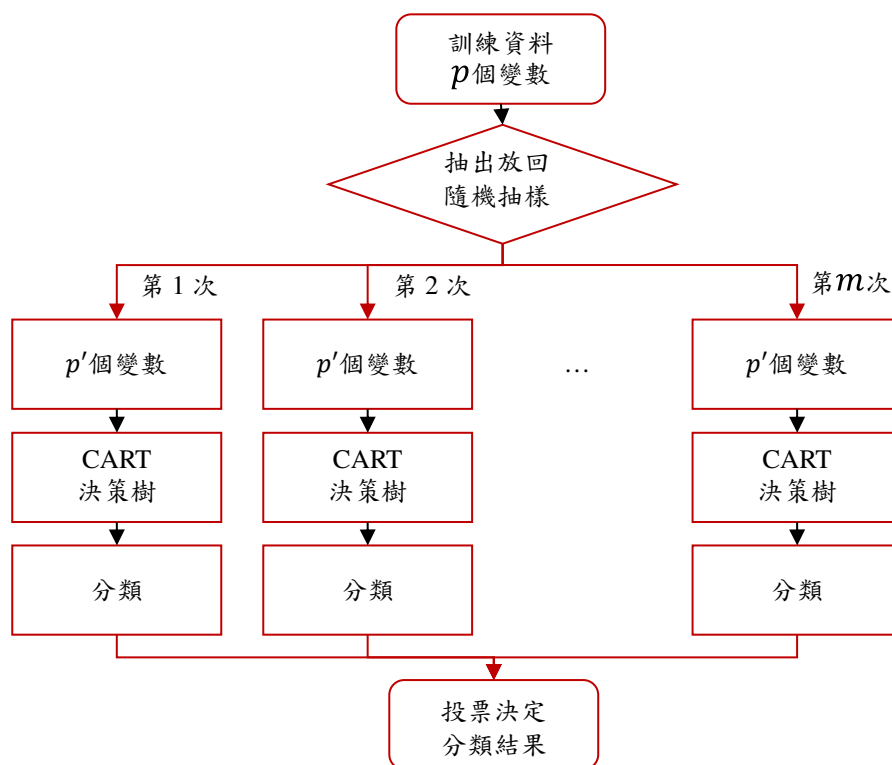
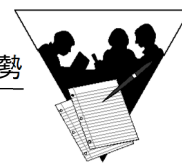


圖2 隨機森林示意圖





隨機森林須事先決定 2 個參數，包含挑選變數比例  $p'/p$ 、進行  $m$  次隨機抽樣，本研究以網格搜尋 ( grid search ) 決定參數組合，設定挑選變數比例  $p'/p=1/5$ 、 $1/4$ 、 $1/3$ 、 $1/2$ 、 $1$ ，進行  $m=500$ 、 $1000$ 、 $1500$ 、 $2000$  次隨機抽樣，以不同參數組合之隨機森林預測未來 1 個月景氣收縮機率。

## 七、預測能力評估

將訓練資料以 LASSO 迴歸選取重要變數後，使用 Probit 模型、彈性網路、隨機森林建立預測模型，再以驗證資料預測未來 1 個月景氣收縮機率。得到景氣收縮機率後，需要設定門檻值判斷景氣是否收縮，若景氣收縮機率大於或等於門檻值，則預測為收縮期，反之則為擴張期。

將景氣狀態之預測及實際結果進行交叉分析，並以混淆矩陣 ( confusion matrix ) 呈現 ( 詳表 1 )：

表 1 混淆矩陣狀態

|    |     | 實際                           |                              |
|----|-----|------------------------------|------------------------------|
|    |     | 收縮期                          | 擴張期                          |
| 預測 | 收縮期 | 真陽<br>( true positive, TP )  | 偽陽<br>( false positive, FP ) |
|    | 擴張期 | 偽陰<br>( false negative, FN ) | 真陰<br>( true negative, TN )  |

進一步定義真陽率 ( true positive rate, TPR )、偽陽率 ( false positive rate, FPR )、真陰率 ( true negative rate, TNR )、偽陰率 ( false negative rate, FNR ) 如下：

- ◆ 真陽率：實際為收縮期，且正確預測為收縮期的比率。

$$\text{真陽率} = \frac{TP}{TP + FN}$$

- ◆ 偽陽率：實際為擴張期，但錯誤預測為收縮期的比率。

$$\text{偽陽率} = \frac{FP}{FP + TN}$$

- ◆ 真陰率：實際為擴張期，且正確預測為擴張期的比率。

$$\text{真陰率} = \frac{TN}{FP + TN}$$

- ◆ 偽陰率：實際為收縮期，錯誤預測為擴張期的比率。

$$\text{偽陰率} = \frac{FN}{TP + FN}$$

由於設定不同門檻值，可得到不同混淆矩陣，若以真陽率為縱軸，偽陽率為橫軸，可得到 ROC 曲線 ( receiver operating characteristic curve )。由於 ROC 曲線為不同門檻值之真陽率及偽陽率組合，進一步定義 ROC 曲線下面積 ( area under the curve, AUC )，做為評估預測表現之標準。

AUC 的範圍介於 0% 至 100%，若模型分類完全正確，真陽率為 100% 且偽陽率為 0%，AUC 為 100%；若模型為完全隨機猜測，則真陽率等於偽陽率，AUC 為 50%。一般預測模型的 AUC 介於 50% 至 100%，數值愈大愈好，Greiner 等人 ( 2000 ) 指出，AUC 介於 50.0% 至 70.0%、70.0% 至 90.0%、90.0% 至 100.0% 分別具有低度、中度、高度準確性。( 詳圖 3 )

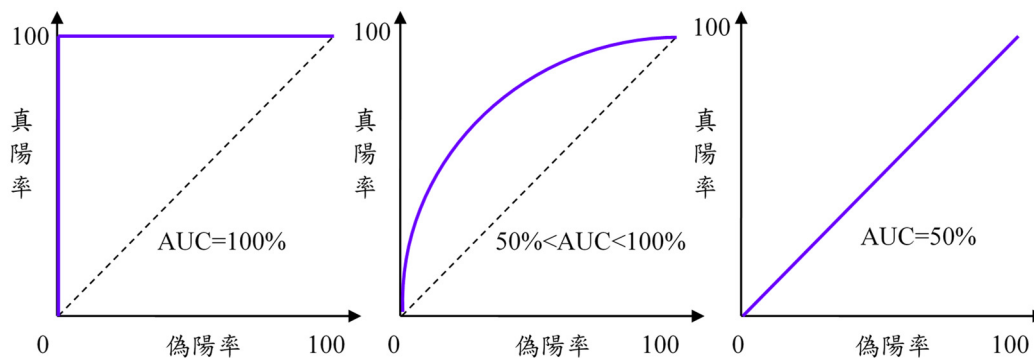
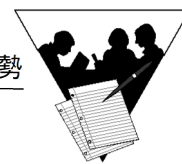


圖 3 ROC 曲線下面積示意圖

## 肆、實證分析

為充分運用大量經濟數據豐富之資訊，本研究蒐集建築、銷售、物價、貿易、生產、就業、交通、稅賦、信心、工商業及金融等 11 類別計 89 個變數，檢視各經濟變數是否具有季節性，若有則進行李節調整。

由於各經濟變數之起始月份不同，本研究設定 1982 年 1 月至 2020 年 7 月之 44 個變數為 1982 年資料集、2001 年 1 月至 2020 年 7 月之 89 個變數為 2001 年資料集，分別進行分析。(詳表 2、表 3)

表 2 1982 年資料集之變數集合

| 編號 | 類別 | 變數名稱         | 編號 | 類別 | 變數名稱                |
|----|----|--------------|----|----|---------------------|
| 1  | 建築 | 建築物開工樓地板面積   | 23 | 就業 | 勞動力參與率              |
| 2  | 銷售 | 製造業銷售價值      | 24 | 就業 | 失業率                 |
| 3  | 銷售 | 製造業內銷價值      | 25 | 就業 | 工業及服務業受僱員工人數        |
| 4  | 銷售 | 製造業直接外銷價值    | 26 | 就業 | 工業及服務業總工時           |
| 5  | 銷售 | 製造業銷售量指數     | 27 | 就業 | 工業及服務業加班工時          |
| 6  | 物價 | 消費者物價指數      | 28 | 就業 | 工業及服務業經常性薪資         |
| 7  | 物價 | 核心消費者物價指數    | 29 | 就業 | 工業及服務業淨進入率          |
| 8  | 物價 | 躉售物價指數       | 30 | 就業 | 製造業單位產出勞動成本指數       |
| 9  | 物價 | 進口物價指數 (新台幣) | 31 | 就業 | 工業勞動生產力指數           |
| 10 | 物價 | 出口物價指數 (新台幣) | 32 | 交通 | 鐵路貨運噸數              |
| 11 | 貿易 | 海關出口值 (美元)   | 33 | 稅賦 | 關稅                  |
| 12 | 貿易 | 海關進口值 (美元)   | 34 | 稅賦 | 貨物稅                 |
| 13 | 貿易 | 出口數量指數       | 35 | 稅賦 | 證券交易稅               |
| 14 | 貿易 | 進口數量指數       | 36 | 稅賦 | 娛樂稅                 |
| 15 | 貿易 | 純貿易條件        | 37 | 信心 | 製造業營業氣候測驗點          |
| 16 | 貿易 | 所得貿易條件       | 38 | 金融 | M1B 日平均             |
| 17 | 生產 | 工業生產指數       | 39 | 金融 | 準貨幣日平均              |
| 18 | 生產 | 製造業生產價值指數    | 40 | 金融 | 存款貨幣日平均             |
| 19 | 生產 | 製造業存貨價值      | 41 | 金融 | 金融業隔夜拆款利率           |
| 20 | 生產 | 電力 (企業) 總用電量 | 42 | 金融 | 商業本票 31-90 天利率 (次級) |
| 21 | 就業 | 勞動力          | 43 | 金融 | 短期票券市場 - 期末發行餘額     |
| 22 | 就業 | 非農就業人數       | 44 | 金融 | 股價指數                |

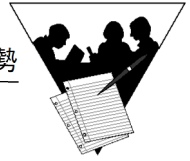


表 3 2001 年資料集之變數集合

| 編號 | 類別 | 變數名稱            | 編號 | 類別  | 變數名稱                              |
|----|----|-----------------|----|-----|-----------------------------------|
| 1  | 建築 | 建築物開工樓地板面積      | 46 | 就業  | 失業給付初次認定核付件數                      |
| 2  | 建築 | 核發建築物使用執照總樓地板面積 | 47 | 就業  | 失業給付核付金額                          |
| 3  | 銷售 | 製造業銷售價值         | 48 | 交通  | 鐵路貨運噸數                            |
| 4  | 銷售 | 製造業內銷價值         | 49 | 交通  | 公路貨運噸數                            |
| 5  | 銷售 | 製造業直接外銷價值       | 50 | 交通  | 汽車登記數                             |
| 6  | 銷售 | 製造業銷售量指數        | 51 | 交通  | 機車登記數                             |
| 7  | 銷售 | 批發業營業額          | 52 | 稅賦  | 關稅                                |
| 8  | 銷售 | 零售業營業額          | 53 | 稅賦  | 貨物稅                               |
| 9  | 銷售 | 餐飲業營業額          | 54 | 稅賦  | 證券交易稅                             |
| 10 | 銷售 | 綜合商品零售業營業額      | 55 | 稅賦  | 娛樂稅                               |
| 11 | 物價 | 消費者物價指數         | 56 | 信心  | 消費者信心指數                           |
| 12 | 物價 | 核心消費者物價指數       | 57 | 信心  | 消費者信心指數-未來半年國內物價水準                |
| 13 | 物價 | 躉售物價指數          | 58 | 信心  | 消費者信心指數-未來半年家庭經濟狀況                |
| 14 | 物價 | 進口物價指數 (台幣)     | 59 | 信心  | 消費者信心指數-未來半年國內經濟景氣                |
| 15 | 物價 | 出口物價指數 (台幣)     | 60 | 信心  | 消費者信心指數-未來半年國內就業機會                |
| 16 | 貿易 | 外銷訂單金額 (美元)     | 61 | 信心  | 消費者信心指數-未來半年投資股票時機                |
| 17 | 貿易 | 外銷訂單動向指數        | 62 | 信心  | 消費者信心指數-未來半年購買耐久性財貨時機             |
| 18 | 貿易 | 海關出口值 (美元)      | 63 | 信心  | 製造業營業氣候測驗點                        |
| 19 | 貿易 | 海關進口值 (美元)      | 64 | 信心  | 服務業營業氣候測驗點                        |
| 20 | 貿易 | 機械及電機設備進口值 (美元) | 65 | 信心  | 營建業營業氣候測驗點                        |
| 21 | 貿易 | 半導體設備進口值 (美元)   | 66 | 工商業 | 公司動態-現有登記家數                       |
| 22 | 貿易 | 出口數量指數          | 67 | 工商業 | 公司動態-新設立家數                        |
| 23 | 貿易 | 進口數量指數          | 68 | 工商業 | 公司動態-新設立資本額                       |
| 24 | 貿易 | 純貿易條件           | 69 | 工商業 | 商業動態-現有登記家數                       |
| 25 | 貿易 | 所得貿易條件          | 70 | 金融  | M1B 日平均                           |
| 26 | 生產 | 工業生產指數          | 71 | 金融  | M2 日平均                            |
| 27 | 生產 | 製造業生產價值指數       | 72 | 金融  | 準貨幣日平均                            |
| 28 | 生產 | 製造業存貨價值         | 73 | 金融  | 存款貨幣日平均                           |
| 29 | 生產 | 電力 (企業) 總用電量    | 74 | 金融  | 消費者貸款                             |
| 30 | 就業 | 勞動力             | 75 | 金融  | 全體金融機構放款與投資                       |
| 31 | 就業 | 非農就業人數          | 76 | 金融  | 五大銀行新承做放款金額                       |
| 32 | 就業 | 勞動力參與率          | 77 | 金融  | 五大銀行新承做放款利率                       |
| 33 | 就業 | 失業率             | 78 | 金融  | 金融業隔夜拆款利率                         |
| 34 | 就業 | 工業及服務業受僱員工人數    | 79 | 金融  | 商業本票 31-90 天利率 (次級)               |
| 35 | 就業 | 工業及服務業總工時       | 80 | 金融  | 10 年期中央政府公債次級市場利率                 |
| 36 | 就業 | 工業及服務業加班工時      | 81 | 金融  | 長短期利差 (10 年期公債利率-金融業隔夜拆款利率)       |
| 37 | 就業 | 工業及服務業加班費       | 82 | 金融  | 長短期利差 (10 年期公債利率-31 到 90 天商業本票利率) |
| 38 | 就業 | 工業及服務業經常性薪資     | 83 | 金融  | 短期票券市場 - 月底發行餘額                   |
| 39 | 就業 | 工業及服務業淨進入率      | 84 | 金融  | 短期票券市場 - 當期交易金額 (買入+賣出)           |
| 40 | 就業 | 製造業單位產出勞動成本指數   | 85 | 金融  | 股價指數                              |
| 41 | 就業 | 工業勞動生產力指數       | 86 | 金融  | 股票市場成交總值                          |
| 42 | 就業 | 平均失業週數          | 87 | 金融  | 信用卡預借現金金額                         |
| 43 | 就業 | 新登記求僱倍數         | 88 | 金融  | 信用卡簽帳金額                           |
| 44 | 就業 | 求職就業率           | 89 | 金融  | 金融機構自動櫃員機交易金額                     |
| 45 | 就業 | 求才利用率           |    |     |                                   |

本研究採用模擬即時分析，將已認定景氣狀態期間做為訓練資料，未認定景氣狀態期間做為驗證資料，而為觀測景氣循環變化，必須使用足夠長度資料做為訓練資料。

以 1982 年資料集為例，首先以 1982 年 1 月至 1998 年 11 月資料做為訓練資料建立模型，預測未來 1 個月景氣收縮機率，如以 1982 年 1 月資料，預測 2 月景氣收縮機率；再以訓練資料建立之模型，運用至驗證資料，如以 1998 年 12 月資料預測 1999 年 1 月景氣收縮機率。(詳表 4、圖 4)

表 4 我國自第 9 次景氣循環起之峰谷及認定月份

| 循環次序    | 高峰      | 認定月份    | 谷底      | 認定月份    |
|---------|---------|---------|---------|---------|
| 第 5 循環  | 1980.01 | -       | 1983.02 | -       |
| 第 6 循環  | 1984.05 | -       | 1985.08 | -       |
| 第 7 循環  | 1989.05 | -       | 1990.08 | -       |
| 第 8 循環  | 1995.02 | -       | 1996.03 | -       |
| 第 9 循環  | 1997.12 | 2000.06 | 1998.12 | 2000.06 |
| 第 10 循環 | 2000.09 | 2004.12 | 2001.09 | 2004.12 |
| 第 11 循環 | 2004.03 | 2008.04 | 2005.02 | 2009.03 |
| 第 12 循環 | 2008.03 | 2009.03 | 2009.02 | 2011.09 |
| 第 13 循環 | 2011.02 | 2014.11 | 2012.01 | 2014.11 |
| 第 14 循環 | 2014.10 | 2016.01 | 2016.02 | 2018.02 |

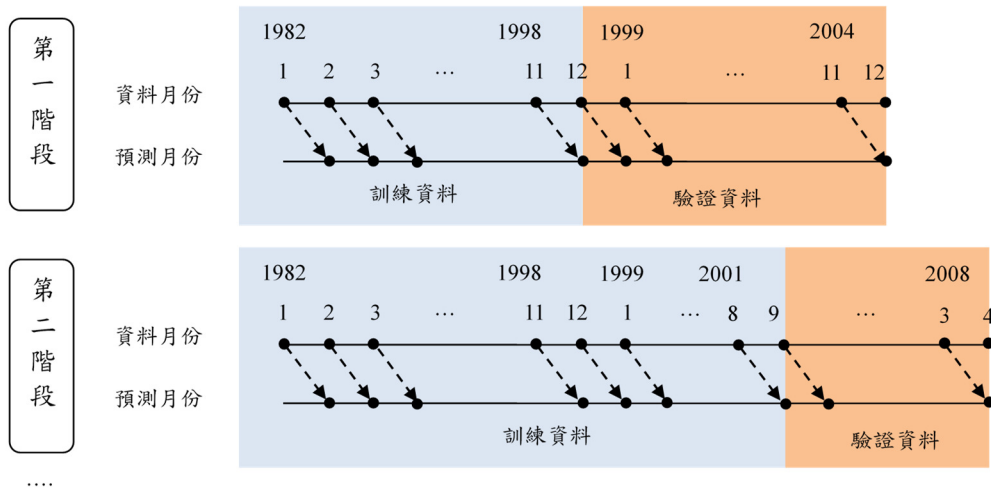
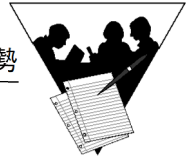


圖 4 模擬即時分析示意圖

由於下次景氣峰谷認定月份為 2004 年 12 月，故自 1998 年 12 月起至 2004 年 11 月止，每月新增 1 期資料，重新以二階段 HP-filter 取出循環成分，再計算循環成分月增率，納入落後期得到變數集合。以 1982 年 1 月至 1998 年 11 月資料做為訓練資料建立模型後，再運用驗證資料預測最新 1 個月景氣收縮機率。

國發會於 2004 年 12 月認定 2000 年 9 月為高峰、2001 年 9 月為谷底，將已知景氣狀態加入訓練資料，自 2001 年 9 月至 2008 年 3 月止，每月新增 1 期資料，重複上述過程得到變數集合後，以 1982 年 1 月至 2001 年 8 月資料做為訓練資料建立模型後，再運用驗證資料預測最新 1 個月景氣收縮機率；其餘資料期間在景氣峰谷認定後，均以此種方法加入訓練資料，進行後續建模及預測。

由於不同階段之模擬即時分析，可能預測同月份之景氣狀態，如上述兩階段均預測 2001 年 10 月至 2004 年 12 月景氣收縮機率，本研究紀錄最新預測結果，並與實際景氣狀態比較。

2001 年資料集亦有類似做法，首先以 2001 年 1 月至 2009 年 1 月資料做為訓練資料建立模型，再運用至驗證資料，如以 2009 年 2 月資料預測 2009 年 3 月景氣收縮機率，其餘處理方法與 1982 年資料集相同。

### 一、1982 年資料集

針對 1982 年資料集進行模擬即時分析，將訓練資料以 LASSO 迴歸選取變數後，分別以 Probit 模型、彈性網路、隨機森林建立模型，再以驗證資料預測未來 1 個月景氣收縮機率，並與實際景氣狀態比較。由於模型參數會影響預測表現，故本研究就各模型設定不同參數，重新進行模擬即時分析，再以 AUC 評估預測表現。

#### (一) Probit 模型

變數集合考慮落後 0 至 12 個月變數，研究結果顯示，若變數集合包含愈多落後期，不一定能增加預測表現，其中以變數集合包含落後 5 個月變數，AUC 達到 79.3% 為最高，惟所有變數集合之 AUC 皆未達八成。(詳表 5)

表 5 Probit 模型預測能力-AUC

單位：%

|     | 變數落後期個數(月) |      |      |      |      |      |      |      |      |      |      |      |      |
|-----|------------|------|------|------|------|------|------|------|------|------|------|------|------|
|     | 0          | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   |
| AUC | 79.1       | 76.2 | 76.0 | 79.1 | 76.9 | 79.3 | 76.1 | 73.4 | 74.1 | 73.7 | 77.8 | 75.6 | 71.1 |





## (二) 彈性網路

變數集合考慮落後 0 至 12 個月變數，彈性網路 $\alpha$ 值考慮由 0 至 1，研究結果顯示，若變數集合不包含落後期，彈性網路之預測表現最佳，AUC 皆在 88.0% 以上。

彈性網路 $\alpha$ 值愈接近 1，預測表現愈差，主因可能為本研究先以 LASSO 迴歸選取變數，再使用彈性網路進行預測。彈性網路 $\alpha$ 值愈接近 1，代表愈接近 LASSO 迴歸，故脊迴歸懲罰項的效果愈小，較無法提供模型更多幫助。

整體來說，變數集合包含落後 3 個月變數，彈性網路 $\alpha$ 值設定為 0 之預測表現最好，AUC 可達 90.0%，且不同參數組合之彈性網路 AUC 均超過八成。(詳表 6)

表 6 彈性網路預測能力-AUC

單位：%

| $\alpha$ | 變數落後期個數(月) |      |      |      |      |      |      |      |      |      |      |      |      |
|----------|------------|------|------|------|------|------|------|------|------|------|------|------|------|
|          | 0          | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   |
| 0.0      | 89.7       | 88.4 | 89.4 | 90.0 | 89.0 | 88.8 | 87.4 | 87.3 | 86.7 | 87.2 | 86.6 | 86.9 | 84.1 |
| 0.1      | 88.4       | 85.5 | 86.7 | 87.4 | 88.3 | 87.5 | 86.6 | 86.1 | 85.8 | 86.6 | 86.5 | 86.0 | 83.4 |
| 0.2      | 88.5       | 85.4 | 86.7 | 87.4 | 88.1 | 87.5 | 86.6 | 86.0 | 85.6 | 86.6 | 86.3 | 86.0 | 83.2 |
| 0.3      | 88.3       | 85.4 | 86.5 | 87.4 | 88.0 | 87.3 | 86.5 | 86.0 | 85.5 | 86.6 | 86.2 | 85.9 | 83.2 |
| 0.4      | 88.2       | 85.3 | 86.3 | 87.2 | 87.8 | 87.1 | 86.6 | 86.1 | 85.3 | 86.7 | 86.2 | 85.8 | 83.2 |
| 0.5      | 88.2       | 85.3 | 86.2 | 87.2 | 87.7 | 87.0 | 86.7 | 86.2 | 85.2 | 86.7 | 85.9 | 85.6 | 83.3 |
| 0.6      | 88.2       | 85.2 | 86.0 | 87.1 | 87.7 | 86.9 | 86.6 | 86.1 | 85.2 | 86.9 | 85.9 | 85.8 | 83.4 |
| 0.7      | 88.2       | 85.2 | 86.3 | 87.1 | 87.5 | 87.0 | 86.8 | 86.2 | 85.1 | 86.9 | 85.8 | 86.1 | 83.6 |
| 0.8      | 88.2       | 85.4 | 86.5 | 86.9 | 87.6 | 87.2 | 86.8 | 86.3 | 85.2 | 86.9 | 85.7 | 86.5 | 83.7 |
| 0.9      | 88.2       | 85.8 | 86.3 | 86.4 | 87.6 | 87.0 | 86.7 | 86.1 | 84.8 | 86.7 | 85.3 | 87.2 | 84.1 |
| 1.0      | 88.4       | 86.1 | 85.7 | 86.4 | 87.1 | 87.0 | 86.0 | 85.5 | 84.4 | 85.7 | 84.3 | 86.8 | 83.8 |

### (三) 隨機森林

變數集合亦考慮落後 0 至 12 個月變數，每次隨機抽樣挑選之變數比例 $p'/p$ ，設定為 1/5、1/4、1/3、1/2、1，進行 $m$ 次隨機抽樣，產生之 $m$ 個決策樹設定為 500、1000、1500、2000 個。

研究結果顯示，若變數集合包含落後 6 個月變數，隨機森林之預測表現最佳，AUC 皆在 90.0%以上。每次隨機抽樣挑選之變數比率，對預測表現較有影響，若挑選的變數比率愈低，代表最後 $m$ 個決策樹彼此差異較大，預測表現通常較好。

若分別進行 500、1000、1500、2000 次隨機抽樣生成隨機森林，其預測表現大致相同，代表進行 500 次隨機抽樣之隨機森林，已具有可信的預測結果。

整體來說，變數集合包含落後 6 個月變數，挑選的變數比率為 1/4，進行 1500 次隨機抽樣之隨機森林預測表現最佳，AUC 可達到 92.9%，且不同參數組合之隨機森林 AUC 皆超過八成五。(詳表 7)

### (四) 綜合比較

比較最佳參數設定之各模型預測表現，以隨機森林表現最佳，AUC 可達 92.9%，彈性網路次之，AUC 為 90.0%，兩者均具有高度準確性；Probit 模型表現最差，AUC 為 79.3%，僅具有中度準確性。(詳表 8)

再以最佳參數設定之各模型，預測 1999 年 1 月至 2016 年 2 月驗證資料期間之景氣收縮機率，與我國第 10 至 14 次景氣循環峰谷認定結果比較，發現隨機森林對應性最佳，彈性網路次之，Probit 模型最差，故使用隨機森林進行預測，較能掌握景氣走勢變化。(詳圖 5、6、7)

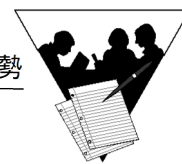


表 7 隨機森林預測能力-AUC

單位：%

| 決策樹<br>個數 | 挑選變數<br>比率 | 變數落後期個數(月) |      |      |      |      |      |       |      |      |      |      |      |      |
|-----------|------------|------------|------|------|------|------|------|-------|------|------|------|------|------|------|
|           |            | 0          | 1    | 2    | 3    | 4    | 5    | 6     | 7    | 8    | 9    | 10   | 11   | 12   |
| 500       | 1/5        | 91.8       | 91.7 | 91.6 | 90.3 | 89.0 | 91.9 | 92.3  | 91.3 | 91.0 | 90.3 | 89.7 | 90.8 | 88.4 |
|           | 1/4        | 92.2       | 91.4 | 91.3 | 89.8 | 89.6 | 91.7 | 92.0  | 91.1 | 91.0 | 89.6 | 90.0 | 90.5 | 88.4 |
|           | 1/3        | 91.7       | 91.1 | 91.4 | 90.1 | 90.0 | 91.1 | 92.5  | 90.8 | 91.1 | 89.5 | 89.8 | 90.7 | 88.0 |
|           | 1/2        | 91.3       | 91.0 | 91.2 | 89.2 | 89.3 | 91.1 | 91.4  | 90.6 | 90.1 | 88.8 | 89.1 | 89.6 | 87.2 |
|           | 1          | 90.5       | 89.4 | 89.8 | 87.4 | 88.9 | 90.7 | 90.9  | 89.5 | 88.7 | 87.6 | 88.4 | 88.5 | 85.8 |
| 1000      | 1/5        | 91.7       | 91.9 | 91.6 | 90.5 | 89.3 | 92.4 | 92.87 | 91.5 | 90.8 | 90.1 | 89.9 | 90.7 | 88.2 |
|           | 1/4        | 91.9       | 91.8 | 91.7 | 90.5 | 89.8 | 91.7 | 92.3  | 91.1 | 91.1 | 89.8 | 89.7 | 90.8 | 88.2 |
|           | 1/3        | 91.9       | 91.5 | 91.2 | 90.0 | 89.3 | 91.8 | 91.7  | 91.2 | 90.2 | 89.6 | 89.8 | 89.9 | 88.0 |
|           | 1/2        | 91.6       | 91.0 | 90.5 | 89.3 | 89.3 | 90.9 | 91.6  | 90.5 | 90.2 | 88.7 | 89.6 | 89.5 | 87.4 |
|           | 1          | 90.7       | 89.3 | 90.0 | 87.4 | 89.4 | 90.2 | 91.3  | 89.3 | 88.8 | 87.6 | 88.6 | 88.7 | 85.3 |
| 1500      | 1/5        | 92.1       | 91.5 | 91.6 | 90.6 | 89.8 | 92.0 | 92.5  | 91.6 | 90.9 | 90.1 | 89.7 | 90.4 | 88.3 |
|           | 1/4        | 92.1       | 91.3 | 91.5 | 90.1 | 89.4 | 91.5 | 92.90 | 91.4 | 91.0 | 89.8 | 89.9 | 89.8 | 88.2 |
|           | 1/3        | 91.7       | 91.2 | 91.1 | 89.6 | 89.5 | 92.0 | 92.3  | 91.2 | 90.5 | 89.7 | 89.5 | 89.9 | 87.9 |
|           | 1/2        | 91.7       | 90.9 | 90.7 | 89.2 | 89.3 | 91.2 | 91.7  | 90.1 | 90.1 | 88.9 | 89.1 | 90.3 | 87.5 |
|           | 1          | 90.4       | 89.5 | 90.3 | 87.8 | 88.5 | 90.9 | 90.8  | 89.2 | 88.9 | 87.8 | 88.4 | 89.0 | 85.4 |
| 2000      | 1/5        | 91.7       | 91.9 | 91.4 | 90.5 | 89.8 | 92.0 | 92.3  | 91.5 | 91.2 | 90.0 | 90.1 | 90.8 | 88.3 |
|           | 1/4        | 92.0       | 91.7 | 91.4 | 90.2 | 89.7 | 91.9 | 92.5  | 91.3 | 90.8 | 89.9 | 90.0 | 90.6 | 88.0 |
|           | 1/3        | 91.9       | 91.3 | 91.0 | 89.8 | 89.4 | 91.3 | 92.2  | 91.3 | 90.2 | 89.4 | 89.9 | 90.0 | 87.9 |
|           | 1/2        | 91.6       | 90.9 | 90.7 | 88.7 | 89.1 | 91.8 | 91.6  | 90.3 | 90.3 | 88.8 | 89.1 | 88.7 | 87.3 |
|           | 1          | 90.3       | 89.4 | 90.0 | 87.1 | 88.9 | 91.0 | 90.7  | 89.1 | 89.0 | 87.5 | 88.7 | 88.1 | 85.6 |

表 8 各模型最佳參數設定預測表現

單位：%

| 編號 | 研究方法      | 參數設定                                     | AUC  |
|----|-----------|------------------------------------------|------|
| 1  | Probit 模型 | 變數落後期 5 個月                               | 79.3 |
| 2  | 彈性網路      | 變數落後期 3 個月<br>$\alpha = 0$               | 90.0 |
| 3  | 隨機森林      | 變數落後期 6 個月<br>決策樹個數 1500 個<br>挑選變數比率 1/4 | 92.9 |

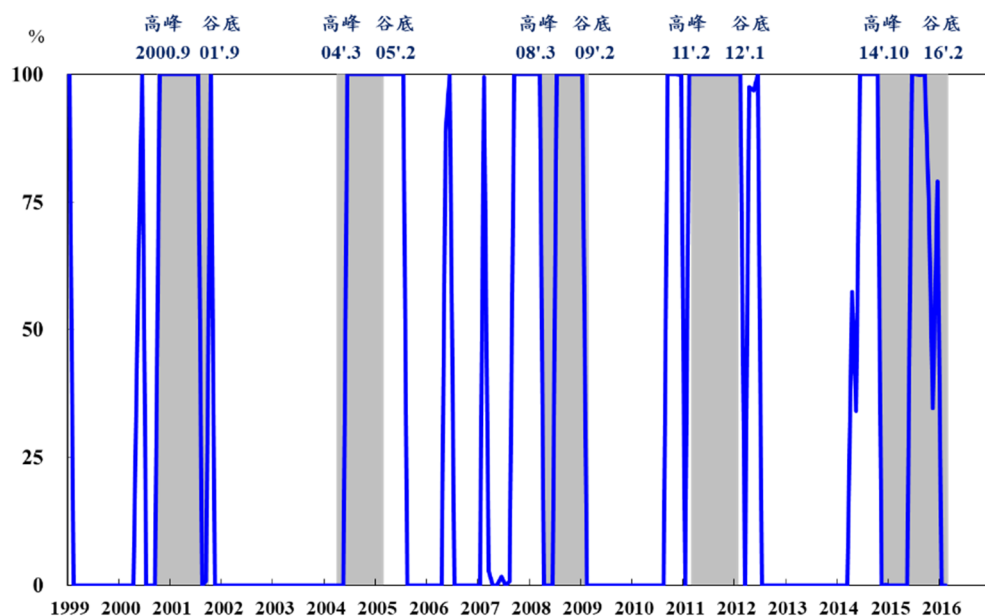


圖 5 Probit 模型預測未來 1 個月景氣收縮機率

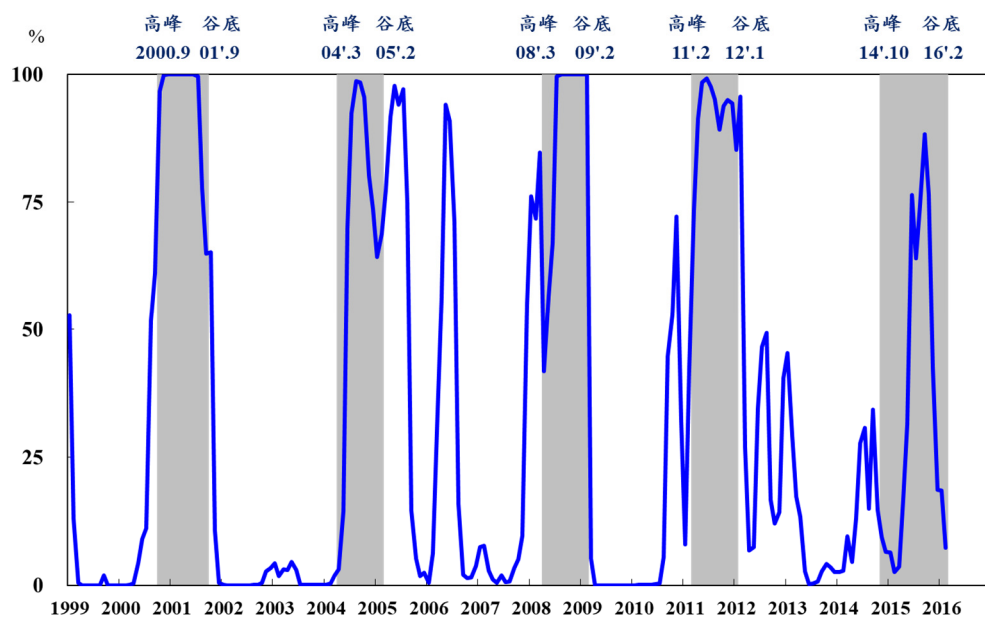


圖 6 彈性網路預測未來 1 個月景氣收縮機率

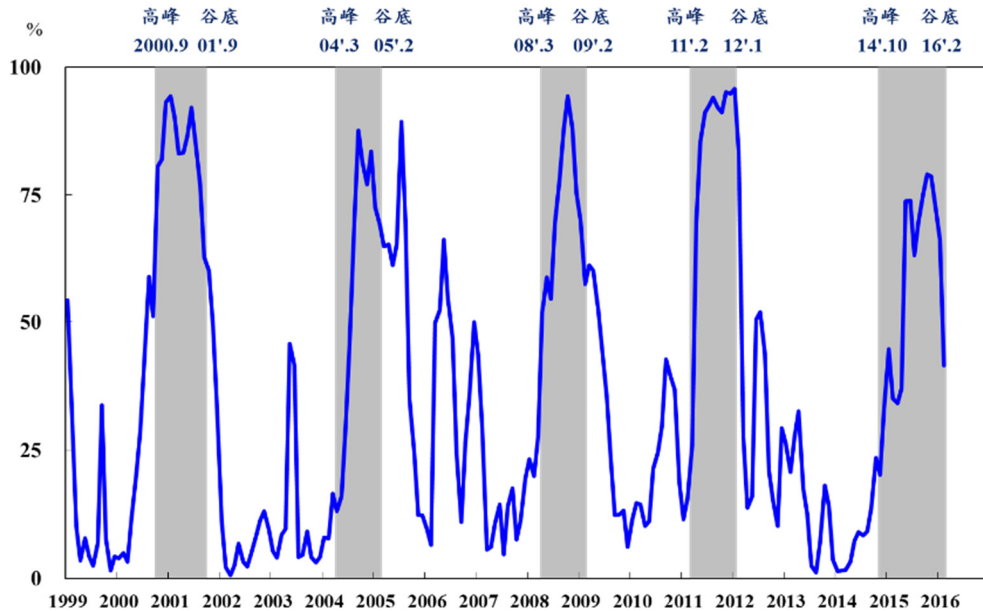
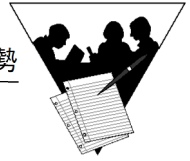


圖 7 隨機森林預測未來 1 個月景氣收縮機率

#### (五) 近期景氣走勢

以 1982 年資料集之最佳模型，亦即變數集合包含落後 6 個月變數、決策樹個數為 1500 個、挑選變數比率為 1/4 之隨機森林，預測第 14 次景氣循環後之景氣走勢。以下呈現兩種結果，第一種為本研究採取之即時預測方法，每月重新進行二階段 HP-filter 取出循環成分，並紀錄最新 1 個月預測結果；第二種為直接預測方法，以 1982 年 1 月至 2016 年 2 月資料建立模型，預測 2016 年 3 月至 2020 年 7 月之景氣收縮機率。

即時預測與直接預測結果相近，惟即時預測每月重新使用二階段 HP-filter 取出循環成分，其結果易受末端值影響，故最新 1 期預測結果變化較大，波動較為明顯；直接預測採用截至 2020

年 7 月最新資料進行二階段 HP-filter，再以 1982 年 1 月至 2016 年 2 月之景氣狀態資料建立模型，一次性預測 2016 年 3 月至 2020 年 7 月景氣收縮機率，波動較為平滑。實務上可採取目前已知最新景氣狀態資料，針對尚未認定景氣狀態區間進行預測。（詳圖 8、圖 9）

觀察第 14 次景氣循環後之景氣走勢，2018 年 3 月起景氣收縮機率走揚，反應美中貿易紛爭爆發，美國對中國大陸進口商品額外加徵關稅，中國大陸亦提出對應措施<sup>1</sup>；2018 年 12 月景氣收縮機率達到最高後下降，反應美中領導人於 G20 舉行會談，雙方暫停新的貿易措施，設定 90 天談判期限；2019 年 4 月起景氣收縮機率維持低點，則反應廠商回流效應提高，在台產能比重增加，加以轉單效應、供應鏈重組態勢明顯所致。

2020 年 1 月起景氣收縮機率逐漸上升，反應中國大陸爆發武漢肺炎疫情，隨後蔓延至歐美消費市場，經濟活動陷入停滯或衰退，進而衝擊我國經濟成長動能；2020 年 4 月景氣收縮機率達到最高後下降，則反應我國防疫措施得宜，經濟活動開始復甦所致，惟截至 2020 年 7 月，景氣收縮機率仍偏高。

<sup>1</sup> 2018 年 3 月 22 日美國宣布為回應中國大陸對智慧財產權的侵犯，將依據 1974 年貿易法第 301 條，對進口商品加徵關稅，7 月 6 日起對 340 億美元商品加徵 25% 關稅，8 月 23 日起對 160 億美元商品加徵 25% 關稅，中國大陸亦提出對應加稅措施。9 月 24 日起美國對中國大陸進口之 2000 億美元商品加徵 10% 關稅，中國大陸亦對美國進口之 600 億美元商品加徵 5% 或 10% 關稅。2018 年 12 月 1 日美中領導人於 G20 舉行會談，雙方宣布暫停新的貿易措施，並設定 90 天談判期限；美國貿易代表署於 2019 年 3 月 2 日宣布不調高加徵 10% 關稅稅率，以因應美中貿易持續談判進程，惟 5 月 11 日美國將中國大陸進口之 2,000 億美元商品加徵稅率，由 10% 調升至 25%，6 月 1 日中國大陸亦將美國進口之 600 億美元商品加徵稅率，提高為 5% 至 25%。

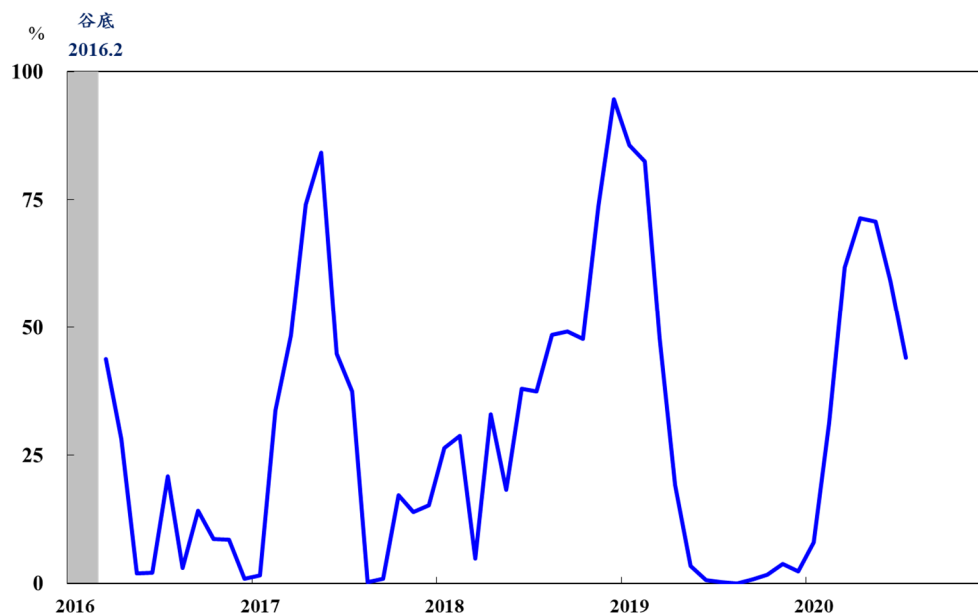
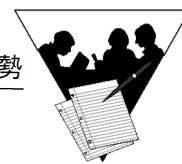


圖 8 近期景氣走勢（即時預測）

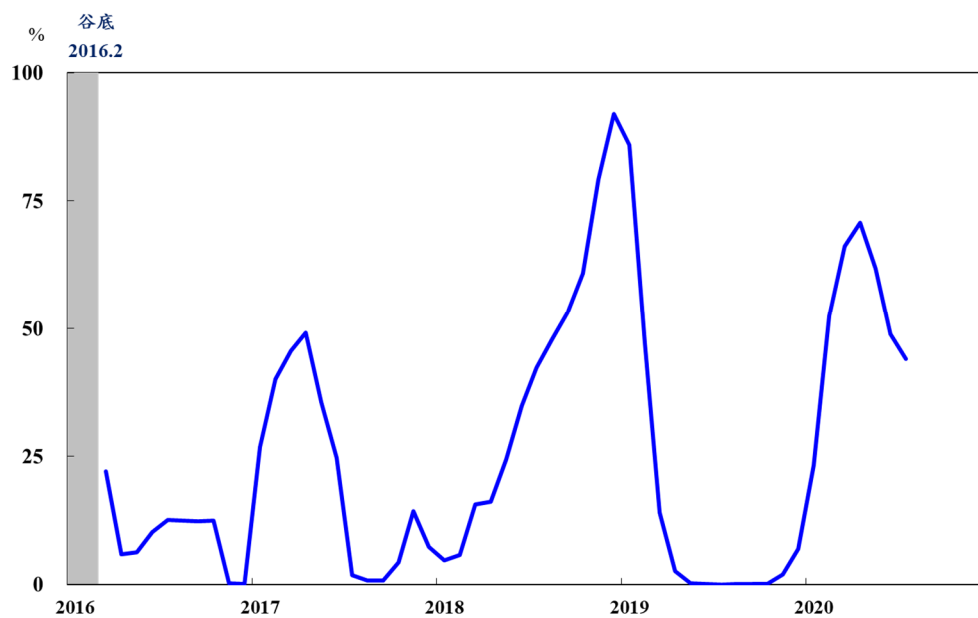


圖 9 近期景氣走勢（直接預測）

## 二、2001 年資料集

本研究亦針對 2001 年資料集進行模擬即時分析，相關方法與 1982 年資料集相同，惟使用 2001 年 1 月至 2020 年 7 月資料及 89 個變數進行分析，特色為時間數列較短且變數較多，以下就各模型設定不同參數，重新進行模擬即時分析，以 AUC 評估預測表現。

### (一) Probit 模型

變數集合考慮落後 0 至 12 個月變數，研究結果顯示，若變數集合包含愈多落後期，不一定能增加預測表現，其中以變數集合包含落後 11 個月變數，AUC 達到 78.1% 為最高，惟所有變數集合之 AUC 皆未達八成。(詳表 9)

表 9 Probit 模型預測能力-AUC

單位：%

|     | 變數落後期個數(月) |      |      |      |      |      |      |      |      |      |      |      |      |
|-----|------------|------|------|------|------|------|------|------|------|------|------|------|------|
|     | 0          | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   |
| AUC | 58.2       | 65.4 | 67.6 | 71.4 | 52.1 | 69.7 | 62.5 | 71.9 | 66.2 | 66.7 | 73.3 | 78.1 | 73.2 |

### (二) 彈性網路

變數集合考慮落後 0 至 12 個月變數，彈性網路 $\alpha$ 值考慮由 0 至 1。研究結果顯示，若變數集合包含落後 2 個月變數，彈性網路之預測表現最佳，AUC 皆在 85.0% 以上；而彈性網路 $\alpha$ 值愈接近 1，其預測表現愈差，與 1982 年資料集相同。

整體來說，變數集合包含落後 2 個月變數，彈性網路 $\alpha$ 值設定為 0 之預測表現最好，AUC 可達 91.6%，且不同參數組合之彈性網路 AUC 皆為七成至九成。(詳表 10)





表 10 彈性網路預測能力-AUC

單位：%

| $\alpha$ | 變數落後期個數(月) |      |      |      |      |      |      |      |      |      |      |      |      |
|----------|------------|------|------|------|------|------|------|------|------|------|------|------|------|
|          | 0          | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   |
| 0.0      | 84.8       | 87.5 | 91.6 | 89.6 | 90.4 | 89.6 | 82.0 | 79.3 | 81.1 | 75.7 | 72.5 | 81.0 | 79.5 |
| 0.1      | 83.2       | 86.9 | 89.4 | 87.1 | 84.9 | 87.9 | 82.7 | 80.7 | 80.4 | 74.7 | 71.4 | 78.8 | 79.6 |
| 0.2      | 82.7       | 86.5 | 88.7 | 85.9 | 84.9 | 87.6 | 82.5 | 80.5 | 79.8 | 74.1 | 71.2 | 78.6 | 79.7 |
| 0.3      | 82.2       | 86.2 | 88.3 | 85.5 | 84.7 | 87.3 | 82.3 | 80.6 | 79.0 | 73.3 | 71.2 | 78.7 | 79.6 |
| 0.4      | 82.2       | 86.0 | 87.7 | 85.2 | 84.5 | 87.3 | 82.1 | 80.5 | 79.0 | 73.0 | 71.5 | 78.4 | 79.7 |
| 0.5      | 82.0       | 85.8 | 87.3 | 85.2 | 84.0 | 87.7 | 82.0 | 80.4 | 78.9 | 72.4 | 71.8 | 78.0 | 79.5 |
| 0.6      | 81.9       | 85.3 | 87.2 | 84.9 | 83.6 | 87.7 | 82.1 | 80.0 | 78.3 | 73.6 | 71.9 | 78.0 | 79.5 |
| 0.7      | 81.5       | 85.3 | 87.2 | 85.1 | 83.1 | 87.7 | 82.2 | 79.5 | 78.0 | 73.4 | 72.1 | 77.8 | 79.3 |
| 0.8      | 80.9       | 84.7 | 87.1 | 84.9 | 82.7 | 87.3 | 81.1 | 78.8 | 76.9 | 73.2 | 72.1 | 76.9 | 78.8 |
| 0.9      | 81.1       | 83.9 | 87.0 | 85.4 | 81.8 | 86.5 | 79.9 | 79.0 | 76.5 | 73.0 | 73.0 | 76.0 | 77.5 |
| 1.0      | 79.3       | 82.7 | 85.9 | 84.7 | 78.8 | 84.3 | 78.5 | 80.3 | 75.0 | 75.6 | 74.4 | 74.9 | 76.9 |

### (三) 隨機森林

變數集合考慮落後 0 至 12 個月變數，若變數集合包含落後 6 個月變數，隨機森林之預測表現最佳，AUC 皆在 88.0% 以上；若分別進行 500、1000、1500、2000 次隨機抽樣生成隨機森林，其預測表現大致相同。

隨機森林設定不同挑選變數比率，在不同變數集合之預測表現有所差異。變數集合包含落後 0 至 6 個月變數，不同挑選變數比率之預測表現較無規律；變數集合包含落後 7 個月以上變數，若挑選變數比率愈高，則預測表現愈好。

本研究推測 2001 年資料集之變數較多，若變數集合包含過多落後期，在 LASSO 迴歸選取變數階段，可能無法蒐集足夠資訊。由於隨機森林階段會隨機挑選變數建立決策樹，若挑選變數

比率較少，則變數代表性將會不足，故在變數集合包含較多落後期的情況下，增加挑選變數比率，能提升隨機森林預測表現。

整體來說，變數集合包含落後 6 個月變數，挑選之變數比率為 1/3，且進行 2000 次隨機抽樣之隨機森林預測表現最好，AUC 可達到 90.5%，且不同參數組合之隨機森林 AUC 皆為七成至九成。(詳表 11)

表 11 隨機森林預測能力-AUC

單位：%

| 決策樹<br>個數 | 挑選變數<br>比率 | 變數落後期個數(月) |      |      |      |      |      |      |      |      |      |      |      |      |
|-----------|------------|------------|------|------|------|------|------|------|------|------|------|------|------|------|
|           |            | 0          | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   |
| 500       | 1/5        | 86.6       | 87.7 | 88.5 | 86.3 | 87.0 | 88.6 | 89.8 | 85.3 | 84.7 | 78.3 | 74.7 | 80.6 | 79.9 |
|           | 1/4        | 86.6       | 87.9 | 87.9 | 87.8 | 88.6 | 88.5 | 90.4 | 85.5 | 87.1 | 79.1 | 72.2 | 80.8 | 81.0 |
|           | 1/3        | 86.3       | 87.9 | 88.4 | 86.9 | 88.2 | 89.6 | 89.6 | 86.8 | 87.9 | 82.5 | 73.9 | 80.9 | 81.6 |
|           | 1/2        | 88.0       | 87.6 | 88.4 | 87.8 | 88.3 | 89.0 | 90.4 | 88.0 | 87.5 | 86.7 | 76.4 | 82.6 | 82.7 |
|           | 1          | 83.6       | 87.1 | 88.4 | 87.5 | 89.2 | 89.1 | 90.3 | 87.3 | 87.3 | 87.2 | 79.6 | 83.5 | 84.3 |
| 1000      | 1/5        | 85.4       | 87.8 | 88.1 | 85.9 | 87.4 | 89.3 | 88.9 | 84.8 | 84.8 | 79.3 | 73.9 | 79.9 | 78.3 |
|           | 1/4        | 87.5       | 88.3 | 88.5 | 87.3 | 87.9 | 89.2 | 89.5 | 86.1 | 86.5 | 80.1 | 73.6 | 80.8 | 80.6 |
|           | 1/3        | 87.1       | 87.4 | 88.4 | 87.7 | 87.9 | 89.9 | 88.9 | 86.4 | 87.1 | 82.3 | 74.8 | 80.9 | 82.0 |
|           | 1/2        | 85.5       | 87.6 | 88.9 | 88.3 | 88.5 | 89.6 | 90.1 | 87.5 | 88.3 | 86.2 | 76.8 | 82.9 | 83.0 |
|           | 1          | 84.2       | 86.2 | 88.3 | 88.0 | 89.1 | 88.7 | 89.9 | 87.4 | 87.5 | 89.7 | 78.8 | 83.8 | 84.6 |
| 1500      | 1/5        | 87.3       | 87.6 | 87.9 | 85.8 | 87.1 | 89.0 | 89.3 | 84.7 | 85.2 | 79.5 | 73.0 | 80.0 | 79.6 |
|           | 1/4        | 87.0       | 88.1 | 88.3 | 86.8 | 87.5 | 89.5 | 88.9 | 86.0 | 86.3 | 79.3 | 72.8 | 81.4 | 80.3 |
|           | 1/3        | 86.6       | 87.7 | 87.9 | 87.0 | 88.3 | 89.5 | 89.5 | 87.1 | 87.0 | 81.9 | 74.5 | 81.8 | 81.8 |
|           | 1/2        | 87.4       | 87.5 | 88.2 | 88.2 | 89.1 | 89.9 | 90.5 | 87.5 | 88.4 | 84.0 | 77.3 | 82.5 | 82.0 |
|           | 1          | 83.8       | 85.9 | 88.9 | 88.0 | 88.9 | 89.5 | 90.1 | 86.5 | 87.4 | 87.5 | 78.4 | 83.6 | 84.7 |
| 2000      | 1/5        | 85.8       | 88.0 | 88.0 | 86.0 | 87.5 | 88.5 | 89.5 | 84.7 | 85.2 | 78.2 | 74.1 | 80.9 | 79.6 |
|           | 1/4        | 87.0       | 88.0 | 88.5 | 86.9 | 87.5 | 88.9 | 90.2 | 85.7 | 86.7 | 79.3 | 73.1 | 81.0 | 81.1 |
|           | 1/3        | 87.1       | 87.7 | 88.6 | 86.9 | 87.6 | 89.4 | 90.5 | 87.2 | 87.5 | 83.6 | 74.9 | 81.0 | 81.7 |
|           | 1/2        | 85.6       | 87.5 | 88.6 | 87.7 | 88.6 | 89.9 | 90.1 | 88.1 | 88.0 | 84.5 | 77.1 | 82.6 | 82.0 |
|           | 1          | 84.0       | 86.3 | 87.9 | 87.9 | 87.8 | 89.3 | 90.1 | 87.4 | 88.6 | 89.6 | 78.2 | 83.4 | 85.1 |



#### (四) 綜合比較

比較最佳參數設定之各模型預測表現，以彈性網路表現最佳，AUC 可達 91.6%，隨機森林次之，AUC 為 90.5%，兩者均具有高度準確性；Probit 模型表現最差，AUC 為 78.1%，僅具有中度準確性。(詳表 12)

再以最佳參數設定之各模型，預測 2009 年 3 月至 2016 年 2 月驗證資料期間之景氣收縮機率，與我國第 12 至 14 次景氣循環峰谷認定結果比較，發現彈性網路對應性最佳，隨機森林次之，Probit 模型最差，故使用彈性網路進行預測，較能即時掌握景氣動態。(詳圖 10、11、12)

表 12 各模型最佳參數設定預測表現

| 編號 | 研究方法      | 參數設定                                     | 單位：% |
|----|-----------|------------------------------------------|------|
|    |           |                                          | AUC  |
| 1  | Probit 模型 | 變數落後期 11 個月                              | 78.1 |
| 2  | 彈性網路      | 變數落後期 2 個月<br>$\alpha = 0$               | 91.6 |
| 3  | 隨機森林      | 變數落後期 6 個月<br>決策樹個數 2000 個<br>挑選變數比率 1/3 | 90.5 |

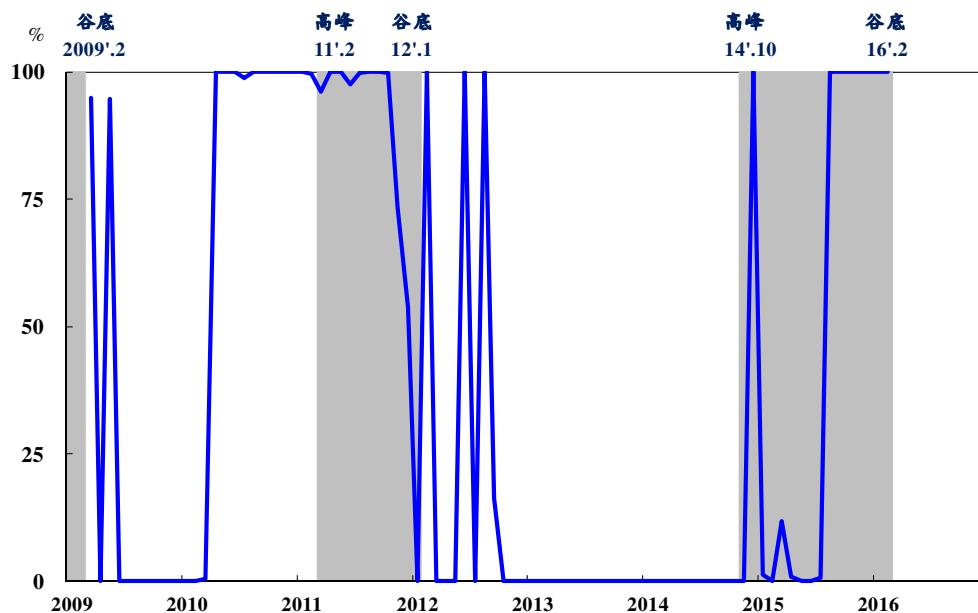


圖 10 Probit 模型預測 1 個月後景氣收縮機率

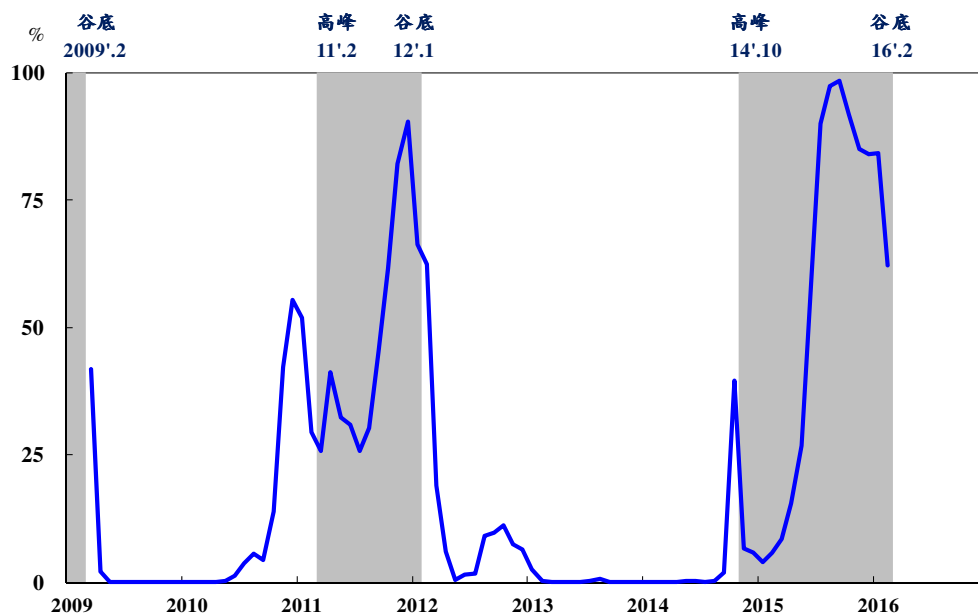


圖 11 彈性網路模型預測 1 個月後景氣收縮機率

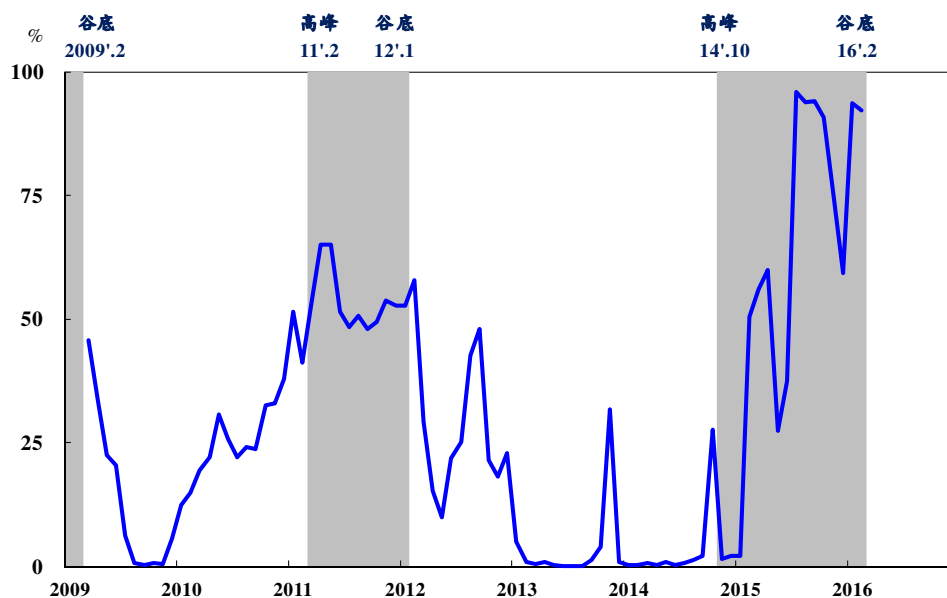
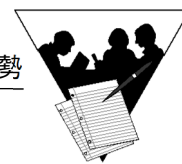


圖 12 隨機森林模型預測 1 個月後景氣收縮機率

### (五) 近期景氣走勢

以 2001 年資料集之最佳模型，亦即變數集合包含落後 2 個月變數、 $\alpha$  值為 0 之彈性網路，預測第 14 次景氣循環後之景氣走勢，比照 1982 年資料集運用即時預測、直接預測方法，預測 2016 年 3 月至 2020 年 7 月之景氣收縮機率。研究結果顯示，即時預測與直接預測結果相近，且即時預測波動較為明顯，直接預測結果則較為平滑，與 1982 年資料集結果相同。(詳圖 13、14)

觀察第 14 次景氣循環後之景氣走勢，2018 年 2 月起景氣收縮機率走揚，反應美中貿易紛爭進程，惟與 1982 年資料集結果不同之處，在於 2018 年 5 月至 8 月景氣收縮機率較低；2019 年 1 月景氣收縮機率達到最高後下降，反應美中領導人於 G20 舉行會談，雙方暫停新的貿易措施之成果；2019 年 5 月起景氣收縮機率維持低點，則反應廠商回流效應及供應鏈重組態勢。

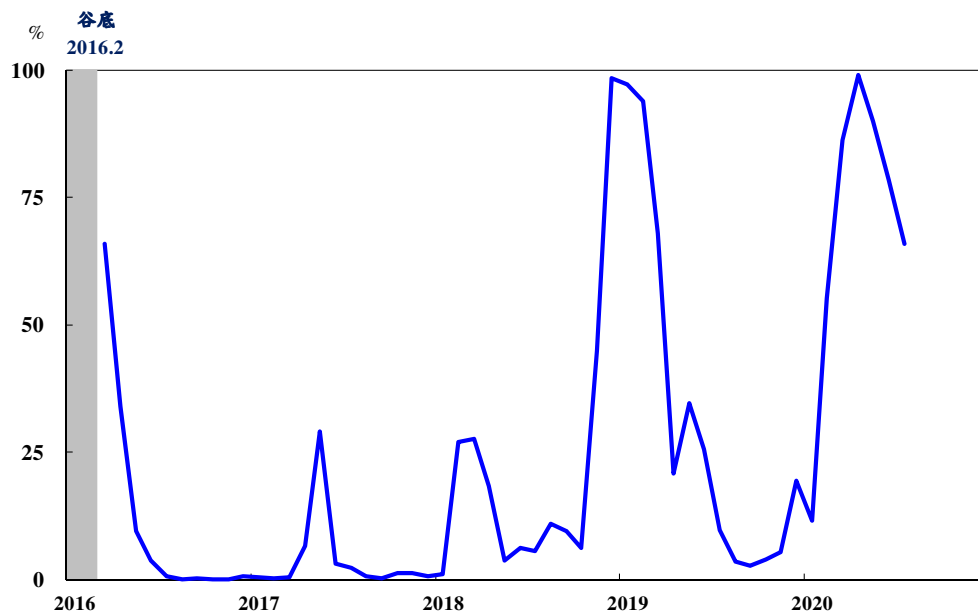


圖 13 近期景氣走勢（即時預測）

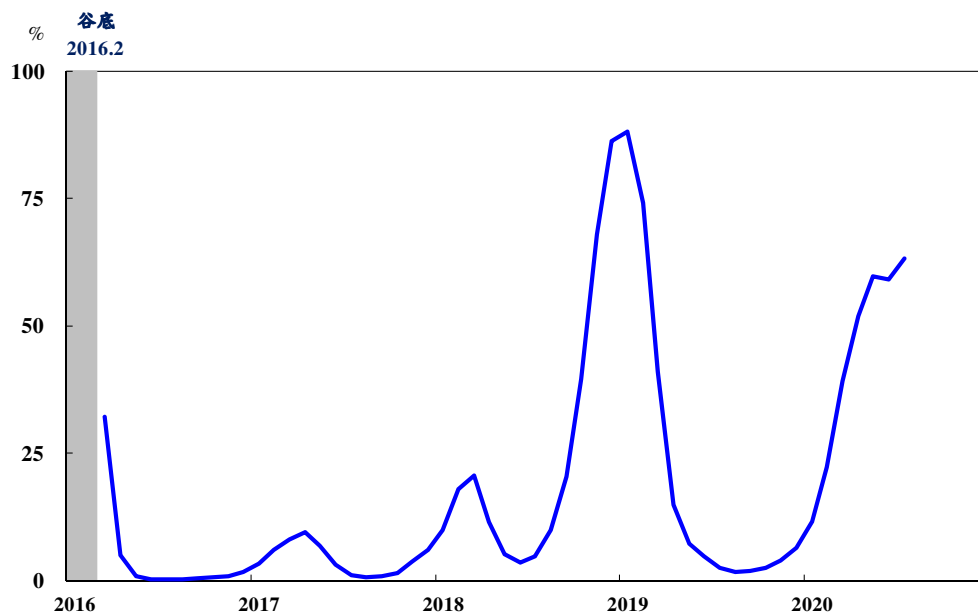
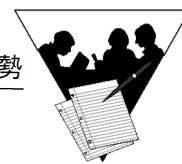


圖 14 近期景氣走勢（直接預測）



2020 年 1 月起景氣收縮機率上升，顯示武漢肺炎疫情對我國經濟之衝擊，2020 年 4 月景氣收縮機率達到最高後下降，則反應我國防疫措施得宜，經濟活動開始復甦，惟截至 2020 年 7 月，景氣收縮機率仍偏高。

### 三、實證分析探討

#### (一) 資料期間

本研究分別採用 1982 年資料集（自 1982 年 1 月起之 44 個變數）、2001 年資料集（自 2001 年 1 月起之 89 個變數）進行分析，1982 年資料集之特色為時間數列較長且變數較少，2001 年資料集特色則為時間數列較短且變數較多。

若比較兩者預測表現，發現使用 1982 年資料集較為穩定，不同參數設定之 AUC 最大值與最小值差距較小，且 1982 年資料集之多數預測表現較好，故實務應用上，建議使用時間數列較長的 1982 年資料集進行分析。（詳表 13）

表 13 模型預測能力

單位：%

| 資料集       | 研究方法      | AUC<br>最小值 | AUC<br>最大值 | AUC<br>全距 |
|-----------|-----------|------------|------------|-----------|
| 1982 年資料集 | Probit 模型 | 71.1       | 79.3       | 8.2       |
|           | 彈性網路      | 83.2       | 90.0       | 6.8       |
|           | 隨機森林      | 85.3       | 92.9       | 7.6       |
| 2001 年資料集 | Probit 模型 | 52.1       | 78.1       | 26.0      |
|           | 彈性網路      | 71.2       | 91.6       | 20.4      |
|           | 隨機森林      | 72.2       | 90.5       | 18.3      |

## (二) 選擇變數

本研究採用模擬即時分析，每月均加入最新資料，重新使用 LASSO 迴歸選擇變數，此部分列出 1982 年資料集、2001 年資料集之最佳模型選取之重要變數，並計算變數選取比率。

### 1. 1982 年資料集

1982 年資料集最佳模型採用之變數集合為包含落後 6 個月變數，若列出 1999 年 1 月至 2016 年 2 月之驗證資料期間，變數選取比率超過 50% 之重要變數，發現領先指標之製造業營業氣候測驗點，同時指標之製造業銷售量指數，落後指標之失業率、金融業隔夜拆款利率、製造業單位產出勞動成本指數皆為重要變數。

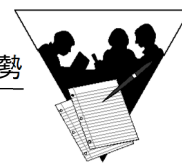
除此之外，如金融面之存款貨幣日平均，勞動面之勞動力、工業及服務業總工時，稅賦面之貨物稅、證券交易稅，物價面之核心消費者物價指數、消費者物價指數，貿易面之純貿易條件等，皆為選取比率較高之變數，可納入未來景氣指標構成項目檢討參考。(詳表 14)

### 2. 2001 年資料集

2001 年資料集最佳模型採用之變數集合為包含落後 2 個月變數，若列出 2009 年 3 月至 2016 年 2 月之驗證資料期間，變數選取比率超過 50% 之重要變數，則發現領先指標之建築物開工樓地板面積、落後指標之製造業單位產出勞動成本指數皆為重要變數。

除此之外，勞動面之求才利用率、平均失業週數，金融面之貨幣總計數 M2 日平均、消費者貸款，信心面之消費者信心指數、未來半年投資股票時機，工商業狀態之公司動態-新設立家數





等，皆為選取比率較高之變數，亦可納入未來景氣指標構成項目檢討參考。(詳表 15)

表 14 變數選取比率-1982 年資料集

單位：%

| 編號 | 變數名稱       | 落後期 | 選取比率  | 編號 | 變數名稱            | 落後期 | 選取比率 |
|----|------------|-----|-------|----|-----------------|-----|------|
| 1  | 存款貨幣日平均    | 0   | 100.0 | 13 | 短期票券市場 - 期末發行餘額 | 2   | 68.4 |
| 2  | 勞動力        | 5   | 100.0 | 14 | 消費者物價指數         | 6   | 65.0 |
| 3  | 失業率        | 0   | 100.0 | 15 | 準貨幣日平均          | 2   | 65.0 |
| 4  | 製造業營業氣候測驗點 | 0   | 100.0 | 16 | 製造業生產價值指數       | 0   | 62.6 |
| 5  | 貨物稅        | 6   | 97.6  | 17 | 娛樂稅             | 1   | 62.6 |
| 6  | 核心消費者物價指數  | 3   | 97.1  | 18 | 短期票券市場 - 期末發行餘額 | 1   | 55.8 |
| 7  | 消費者物價指數    | 0   | 94.7  | 19 | 工業及服務業加班工時      | 0   | 54.9 |
| 8  | 金融業隔夜拆款利率  | 5   | 92.2  | 20 | 海關進口值(美元)       | 4   | 53.9 |
| 9  | 工業及服務業總工時  | 0   | 83.0  | 21 | 製造業單位產出勞動成本指數   | 6   | 53.9 |
| 10 | 純貿易條件      | 2   | 80.1  | 22 | 製造業內銷價值         | 0   | 53.4 |
| 11 | 證券交易稅      | 2   | 78.6  | 23 | 工業及服務業經常性薪資     | 0   | 52.4 |
| 12 | 製造業銷售量指數   | 0   | 68.9  | 24 | 存款貨幣日平均         | 1   | 51.9 |

表 15 變數選取比率-2001 年資料集

單位：%

| 編號 | 變數名稱               | 落後期 | 選取比率  | 編號 | 變數名稱       | 落後期 | 選取比率 |
|----|--------------------|-----|-------|----|------------|-----|------|
| 1  | 製造業單位產出勞動成本指數      | 1   | 100.0 | 7  | 建築物開工樓地板面積 | 1   | 58.3 |
| 2  | 求才利用率              | 2   | 100.0 | 8  | 求才利用率      | 1   | 58.3 |
| 3  | 消費者信心指數            | 0   | 98.8  | 9  | 公司動態-新設立家數 | 0   | 56.0 |
| 4  | M2 日平均             | 2   | 89.3  | 10 | M2 日平均     | 1   | 56.0 |
| 5  | 消費者貸款              | 0   | 81.0  | 11 | 平均失業週數     | 0   | 54.8 |
| 6  | 消費者信心指數-未來半年投資股票時機 | 0   | 64.3  |    |            |     |      |

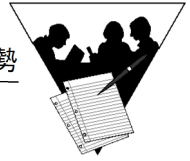
## 伍、結論與建議

### 一、結論

本研究運用大量經濟變數及其落後期，預測未來 1 個月景氣收縮機率，採用每月新增資料方式，進行模擬即時分析。將訓練資料以 LASSO 迴歸挑選變數後，分別使用 Probit 模型、彈性網路、隨機森林建立模型，再以驗證資料預測未來 1 個月景氣收縮機率，使用 AUC 評估預測表現。

由於各經濟變數起始月份不同，本研究設定 1982 年 1 月至 2020 年 7 月之 44 個變數為 1982 年資料集，2001 年 1 月至 2020 年 7 月之 89 個變數為 2001 年資料集進行分析，主要結論分述如下：

- (一) 隨機森林、彈性網路之預測表現較好。觀察各模型之 AUC，1982 年資料集之隨機森林均超過八成五，彈性網路超過八成，Probit 模型則未達八成；2001 年資料集之隨機森林及彈性網路皆為七成至九成，Probit 模型則未達八成。
- (二) 1982 年資料集最佳模型為隨機森林（變數集合包含落後 6 個月變數、挑選的變數比率為 1/4、進行 1500 次隨機抽樣），AUC 可達 92.9%；2001 年資料集最佳模型為彈性網路（變數集合包含落後 2 個月變數、彈性網路  $\alpha$  值為 0），AUC 可達 91.6%，且兩種最佳模型預測之景氣收縮機率，皆能對應我國景氣峰谷認定結果。
- (三) 比較 1982 年資料集、2001 年資料集之預測表現，發現 1982 年資料集較為穩定，不同參數設定之 AUC 最大值與最小值差距較小，且 1982 年資料集多數預測表現較好。實務應用上，建議使用時間數列較長的 1982 年資料集進行預測。

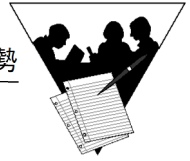


- (四) 觀察臺灣第 14 次景氣循環後之景氣走勢，2018 年 3 月起景氣收縮機率走揚，反應美中貿易紛爭爆發，同年 12 月景氣收縮機率達到最高後下降，則反應美中領導人於 G20 舉行會談，雙方暫停新的貿易措施之成果；2020 年 1 月起景氣收縮機率上升，顯示武漢肺炎疫情對我國經濟之衝擊，同年 4 月景氣收縮機率達到最高後開始下降，則反應我國防疫措施得宜，經濟活動開始復甦所致。

## 二、建議

- (一) 本研究使用 LASSO 迴歸選擇變數，若列出最佳模型選取之重要變數，發現部分為景氣指標構成項目，其餘如存款貨幣日平均、勞動力、貨物稅、證券交易稅、工業及服務業總工時等，可列入景氣指標構成項目檢討參考。
- (二) 可考慮其他方法對大量變數進行篩選或萃取資訊，如計算與領先指標、同時指標之相關係數，挑選相關性較高之變數，或使用 Stock and Watson (2002) 提出之 EM 演算法估算共同因子，或運用主成分分析 (principal component analysis, PCA)、偏最小平方法 (partial least squares, PLS) 等維度縮減 (dimension reduction) 方法萃取變數資訊。
- (三) 本研究變數來源為政府或民間機構發布之統計或調查結果，惟數位時代之資料種類更為豐富，如文字、影像、搜尋資料等，各界亦開始使用相關資料進行分析，如日本經濟產業省與野村證券合作編製「SNS×AI 景況感指數」，擷取社群網站之經濟相關推文，透過分析文字情緒，掌握當前景氣狀態，建議未來可納入更多種類資料進行分析。

(四) 我國綜合考量生產、消費、就業、貿易等各部門變化，以景氣領先、同時、落後指標衡量總體經濟之波動，惟各產業之景氣循環態樣有所差異，如零售、餐飲業與內需消費市場有關，半導體產業則受國際外需市場影響，建議未來可運用計量方法，分析各產業之景氣循環與關聯性，以掌握產業之景氣走勢。



## 參考文獻

1. 周大森 (2019), 「運用大量數據認定景氣轉折點」, 國家發展委員會委託研究計畫。
2. 徐士勛、黃裕烈、徐之強、張瑞雲 (2020), 「經濟趨勢調查之應用研究」, 國家發展委員會委託研究計畫。
3. 黃裕烈 (2019), 「景氣監測預警系統之研究」, 國家發展委員會委託研究計畫。
4. Abberger, K., Graff, M., Campelo Jr, A., Lemos Gouveia, A. C., Müller, O., & Sturm, J. E. (2020). The Global Economic Barometers: Composite indicators for the world economy. *KOF Working Papers*, 471.
5. Berge, T. J. (2015). Predicting recessions with leading indicators: Model averaging and selection over the business cycle. *Journal of Forecasting*, 34(6), 455-471.
6. Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
7. Bry, G. and C. Boschan (1971). Cyclical Analysis of Time Series: Procedures and Computer Programs, New York, NBER.
8. Burns, A. F., & Mitchell, W. C. (1946). Measuring Business Cycles. NBER Books.
9. Chauvet, M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switching. *International economic review*, 969-996.
10. Chauvet, M., & Hamilton, J. D. (2006). Dating business cycle turning points. *Contributions to Economic Analysis*, 276, 1-54.
11. Chauvet, M., & Piger, J. (2008). A comparison of the real-time performance of business cycle dating methods. *Journal of Business & Economic Statistics*, 26(1), 42-49.
12. Chauvet, M., & Potter, S. (2005). Forecasting recessions using the yield curve. *Journal of Forecasting*, 24(2), 77-103.
13. Fossati, S. (2016). Dating US business cycles with macro factors. *Studies in Nonlinear Dynamics & Econometrics*, 20(5), 529-547.
14. Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1), 1.
15. Giusto, A., & Piger, J. (2017). Identifying business cycle turning points in real time with vector quantization. *International Journal of Forecasting*, 33(1), 174-184.
16. Greiner, M., Pfeiffer, D., & Smith, R. D. (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Preventive*

- veterinary medicine*, 45(1-2), 23-41.
17. Harding, D. and A. Pagan (2002). Dissecting the Cycle: A Methodological Investigation. *Journal of Monetary Economics*, 49, 365-381.
  18. Hanck, C., Arnold, M., Gerber, A., & Schmelzer, M. (2019). Introduction to Econometrics with R. *Obtenido de <https://www.econometrics-with-r.org/ITER.pdf>*.
  19. Hodrick, R. J., & Prescott, E. C. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, credit, and Banking*, 1-16.
  20. Jung, J. K., Patnam, M., & Ter-Martirosyan, A. (2018). *An Algorithmic Crystal Ball: Forecasts-based on Machine Learning*. International Monetary Fund.
  21. Kauppi, H., & Saikkonen, P. (2008). Predicting US recessions with dynamic binary response models. *The Review of Economics and Statistics*, 90(4), 777-791.
  22. Ng, S. (2013). *Boosting recessions*. *Canadian Journal of Economics*, forthcoming.
  23. Rudebusch, G. D., & Williams, J. C. (2009). Forecasting recessions: the puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics*, 27(4), 492-503.
  24. Smeekes, S., & Wijler, E. (2018). Macroeconomic forecasting using penalized regression methods. *International journal of forecasting*, 34(3), 408-430.
  25. Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2), 147-162.
  26. Stock, J. H., & Watson, M. W. (2014). Estimating turning points using large data sets. *Journal of Econometrics*, 178, 368-381.
  27. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
  28. Tiffin, A. (2016). Seeing in the dark: a machine-learning approach to nowcasting in Lebanon. *IMF Working Paper*.
  29. Pereira, J. M., Basto, M., & da Silva, A. F. (2016). The logistic lasso and ridge regression in predicting corporate failure. *Procedia Economics and Finance*, 39, 634-641.
  30. Woloszko, N. (2020). Adaptive Trees: a new approach to economic forecasting. *OECD Working Papers*.
  31. Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, 67(2), 301-320.