

台灣房價下行風險評估及其可能應用*

鄭漢亮**

摘 要

台灣房地產法規嚴謹，且無造成美國次貸危機的證券化商品與次貸產品，房貸違約率不高並續創歷史新低，房市過熱時，政府、央行也會適時進場調控，穩定房市，要如同美國因房地產過度發展，而發生金融危機的可能性不高。綜觀，台灣過去的幾次金融危機，均多由信用過度擴張(如2005 Q3~2006 Q4雙卡風暴等)所致。惟房價下行風險或可能蘊含資產價格領先資訊，因此本研究除使用分量隨機森林模型建構房價下行風險，亦嘗試將其當做危機預警指標之一。實證結果顯示，分量隨機森林模型預測未來1季房價下行風險的表現優於傳統分量迴歸。此外，其在預測房價下行風險能力優於上行風險。另在樣本內外預測，房價下行風險或可做為危機有效早期預警指標之一。

* 本文初稿完成於民國112年8月，114年3月修正完稿。本文承蒙嚴副總裁宗大、吳局長懿娟、曹處長體仁、廖行務委員俊男、俞助理研究員欣榮、徐副研究員婉容與兩位匿名審稿人之悉心審閱，以及處內其他同仁給予寶貴意見，特此衷心謝忱。惟本文觀點純屬個人意見，與服務單位無關，若有任何疏漏或錯誤，概由作者負責。

** 作者為中央銀行經濟研究處研究員。

壹、前言

房屋不但是家庭擁有的重要資產，也是企業借款抵押品之一(Adrian et al., 2020)。房屋跌價會減損家庭淨資產，以及抵押品價值，而使信用大幅萎縮，並對銀行的投資組合、獲利能力與穩定性產生負面影響，進而衝擊經濟活動，且影響的時間可能比經濟衰退更為持久(Adrian et al., 2020)。長久以來，先進國家金融(銀行)危機發生之前，房價大多出現暴漲暴跌，2008年全球金融危機便是一個最好的例子，從房市崩跌蔓延至全面性的金融危機。因此，國外文獻已發現房價風險可幫助預測危機發生(參見IMF, 2019；Adrian et al., 2020)。由於這幾年許多先進國家房價同步上漲，房價下行風險增溫，與家庭槓桿率上升的情況下，預先掌握房市變動有助金融穩定以及政策形成(Adrian et al., 2020)，促使各國央行更加關注房價下行風險。

過去以來先進國家房價上漲後，急遽驟跌而阻礙經濟活動的情況屢見不鮮，若能找到與房價尾端分配關係緊密的金融經濟變數，分量迴歸或可成為預測房價下行風險的可行方式(IMF, 2019；Adrian et al., 2020)。若能準確預測房地產市場的脆弱性，可將之用於金融穩定監測，且總體審慎政策或可依據房價下行風險制定，減少房地產市場震盪(參見IMF, 2019；Galán and Rodríguez-

Moreno, 2020；Adrian et al., 2020)。

然而過去國內外關於房價的學術文獻多集中在估計房價基本面與泡沫，例如Mayer (2011)、Chen and Patel(2002)、Smith and Smith (2006)，少部分研究探討在房市不同狀態(繁榮與蕭條)之下，重要總體經濟變數對美國房價之影響，並進一步評估政策變動對房市狀態轉換之效果，如Nneji et al. (2013)等。近幾年才有國外文獻根據Adrian et al. (2019)，以分量模型評估房價下行風險(downside entropy)，並探討其對金融危機的預測力，如Galán and Rodríguez-Moreno (2020)、Adrian et al. (2020)等，並逐漸有研究將其拓展應用至通膨等下行風險評估，如Banerjee et al. (2020)。

國外文獻也發現房價風險評估可提供貨幣政策決策者經濟前景資訊。例如，Adrian et al. (2020)認為房價下行風險對經濟成長有重大影響；此外，張天惠、朱浩榜 (2020)則發現納入房地產與信用變數的金融情勢指數，相對於未考量前述變數，對於經濟衰退之預測能力表現更勝一籌。對於政策制定者而言，房價下行風險是一項重要且有用的資訊。

另一方面，台灣房地產市場法規嚴謹，且無造成美國次貸危機的證券化商品與次貸產品，加以房貸違約率不高並續創歷史新

低，房市過熱時，政府、央行也會適時進場調控，穩定房市，沒有美國房市暴漲暴跌的情況，要如同美國因房地產過度發展，而發生金融危機的可能性不高。綜觀，台灣過去的幾次金融危機，均由信用過度擴張(如2005 Q3~2006 Q4雙卡風暴等)所致。所以，尚無文獻檢視台灣房價下行風險與金融危機的關係。惟房價本身具有資產價格特性，可能隱含未來資訊(如前述之經濟成長動向)，且 Cheng and Chen (2021)研究發現台灣房價循環領先信用循環，房價下行風險或可提前反應部分金融失衡現象(如信用過度擴張等)。因此，本研究嘗試將房價下行風險做為金融危機預警指標之一。

此外，過去因計量工具的限制，致大量數據分析窒礙難行。直至隨機森林(Random Forest, RF)等機器學習模型問世，允許引入高維度資料(high-dimensional data)，並可排序解釋變數重要性，進行特徵變數篩選(features selection)，且能捕捉非線性關係的特質，使其表現往往較傳統線性時間序列迴歸模型更佳(Duan et al., 2022)，成為實證研究另一項選擇。相較其他Lasso等機器學習模型，Lasso無法處理嚴重的多重共線性，且當自變量高度相關，Lasso會隨機刪除其中一個變數；Lasso也只適用於線性分析，沒有辦法有效捕捉總體變數彼此間存在的非線性影響。由於總體經濟金融變數與房價變動之間可能為非線性關係，而隨機森林演算

法在變數篩選機制上更具優勢，將有助了解驅動房價上行或是下行風險之因素。

迄今為止，國內仍缺乏以分量隨機森林模型建構房價成長分配與下行風險之文獻^{註1}。有鑒於此，本文選取房地產市場供需、總體經濟、利率與貨幣供給、股票市場以及外部因素等共計47個變數，並使用Meinshausen (2006)分量隨機森林(quantile random forest, QRF)模型，以平均減少吉尼係數(Mean Decrease Gini)為標準，排序解釋變數相對重要性(relative important)，並參酌Dubey (2018)、Cysouw et al. (2021)建議，從大量數據中選取前10%重要變數，降低不相關變數的干擾。進一步，再根據Adrian et al. (2019)作法，以偏斜t分配(Skew-t Distribution)，建構房價成長分配、下行風險、預期損失(expected shortfall)與長期收益(expected longrise)等，並探討台灣房價下行風險可否做為預警指標之一。

實證結果顯示，由於分量隨機森林模型充分反映資料資訊，且能捕捉非線性關係，相較傳統分量模型更可捕捉房價成長脆弱性，樣本外未來1季與1年房價下行風險估計值與樣本內比較，幾乎沒有太大區別。此外，使用Quantile R^2 比較各分量預測表現，發現模型在房價下行風險預測力均高於上行風險；並且模型預測未來1年房價下行風險能力雖有下降，但其準確性仍優於中位數以及上行風險。

最後，本研究亦發現房價下行風險搭配其他總體變數，有助預測未來1年危機的發生。此外，模型較長期預測力較高，有可能是因為房價下行風險結合景氣與國際情勢波動等資訊，能更早的反應危機發生。

本文的其餘部分安排如下。首先，第

二節回顧房價下行風險與危機事件關聯之文獻；然後，第三節說明量化方法；第四節介紹房價風險估計的主要結論，包括樣本內外預測，以及房價下行風險對危機預測表現；最後一節為結論。

貳、文獻回顧

自Adrian et al. (2019)發表「Vulnerable Growth」乙文後，使用偏斜t分配，評估經濟成長下行風險的研究，如雨後春筍般湧現，並成為其他領域評估風險的主要工具，包括總體金融與匯率風險等，詳見Chicana and Nivin (2022)、Ferrara and Yapi (2022)。

近幾年先進國家等房價出現同步大漲現象，文獻研究紛紛將目光轉移至房地產市場，並以分量迴歸建構房價成長分配、下行風險(downside entropy或downside risk)，進而探討房價風險與危機預測之關聯。相關文獻包括：Adrian et al. (2020)使用分量迴歸，預測32個先進與新興市場經濟體未來房價成長下行風險。Adrian et al. (2020)也針對預防房價風險，提出政策建議。整體而言，總體審慎政策遏止房價下行風險最為有效。

Galán and Rodríguez-Moreno (2020)指出，分量迴歸風險評估係分析金融變數等對房價下行風險之影響，一實用架構與政策決策工具，且房價風險與金融穩定密切關聯。IMF(2019)認為分量迴歸是計算房價下行風

險的新方法，並且在金融危機預測模型加入房價風險指標，可增強模型預測能力，故房價下行風險是良好早期預警指標(early-warning indicator)。

另一方面，分量隨機森林模型雖尚未有應用至房地產風險研究，但已有大量運用至經濟風險的文章。如Duan et al. (2022)使用分量隨機森林模型，構建預測中國大陸系統風險(systemic risk)之資訊集，以評估未來總體經濟的下行風險，並發現分量隨機森林模型相對傳統分量迴歸，樣本外預測表現更為出色；Foltas and Pierdzioch (2022)探討1970年~2017年德國3個經濟研究機構發布經濟成長預測的正確性。他們使用分量隨機森林模型，以偏斜t分配，模擬各機構預測誤差的條件分配，發現樣本內外，均不支持強式效率預測(strongly efficient forecasts)，但無法拒絕弱式效率預測(weak efficiency of forecasts)。

國內房價相關研究方面，Chen and Patel(2002)發現家庭收入是長期決定台灣房價最重要的因素；短期內，股價與貨幣供

給等變數至關重要。至於房價風險評估方面的文獻，陳明吉與蔡怡純(2022)使用2階段ARCH模型並假設常態分配，評估台灣房價分配與下行風險；江明珠、李政峰、權清全(2011)假設房價報酬為常態分配，計算下行風險，或是以泰勒展開式近似厚尾分配，評估房價下行風險。但迄今為止，台灣尚未有使用分量隨機森林模型，建構房價成長分配

及下行風險之文章。

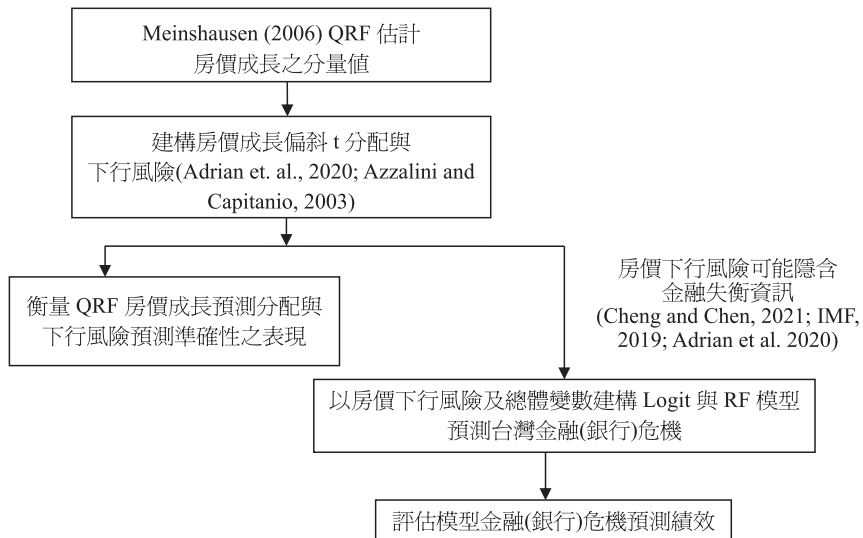
另也有國內研究進一步探討房價波動與變動的不對稱性，如使用T-GARCH模型、門檻誤差修正模型等，研究房地產價格向下波動不對稱性，藉以驗證不動產市場抗跌性，參見蔡怡純與陳明吉(2008)、蔡怡純、陳明吉(2013)。

參、模型設定

本研究實證步驟為，首先使用Meinshausen (2006)分量隨機森林模型篩選有用變數與估計房價成長 $\alpha = \{0.1, 0.25, 0.75, 0.9\}$ 之分量值。再依照Adrian et. al.(2020)、Azzalini and Capitanio (2003)做法，以平滑量

化函數，重建房價成長的偏斜t分配，估算下行風險，並進行樣本外預測分配(predicted distributions)績效與下行風險預測準確性評估等(如圖1所示)。

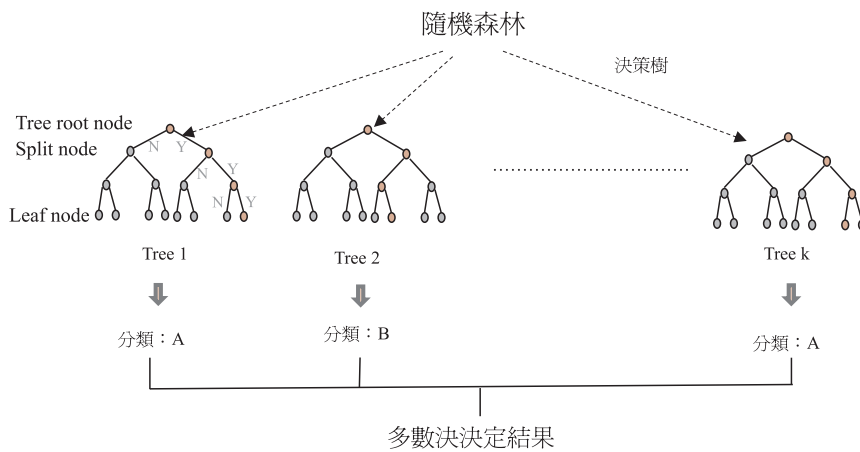
圖1 實證方法流程圖



其次，雖然台灣歷次金融危機與房市波動無關，但房價變動可能蘊含危機發生前的金融失衡資訊，如信用過度擴張等，因此以房價下行風險及總體變數等，建構傳統Logit與隨機森林模型預測台灣金融(銀行)危機，並評估其表現。

一、分量隨機森林

隨機森林(RF)模型集成(ensemble)多個決策樹(decision tree)，能有效降低模型的錯誤率、共線性等問題，常被運用在預測，且表現良好。以隨機森林處理分類問題為例，簡述如下，M表示訓練個數，n為變數數目：



1. 隨機抽樣M筆資料，形成一組新的訓練集。每次取出M筆樣本皆會重新放回母體，允許資料可重複。
2. 從選取的M筆資料中，建立決策樹。每筆資料皆含有n個觀察值，而每棵樹的節點大小(nodesize)，隨機選取m比例，共d個變數，用於做決策，且d小於n，並根據此d個變數，分割該節點，並對剩餘樣本作預測，評估誤差。
3. 以拔靴法(Bootstrap)隨機重新取樣產生新資料^{註2}。重複步驟(1)與(2)k次，

產生k棵決策樹。

4. 彙總所有決策樹的預測，以多數決方式決定最終結果。針對分類問題，以投票機制進行；迴歸則選擇平均值預測。

Meinshausen (2006)指出隨機森林除可預測條件平均值，亦可擴展至條件分量預測。過程簡單說明如下，給定 $X_t = x_t$ ，第 τ 分量值 $Q_{y_{t+h}|x_t}(\tau|x_t)$ 出現機率，正好等於 τ ，數學式可表示為：

$$Q_{y_{t+h}|x_t}(\tau|x_t) = \inf \{y_{t+h} : F(y_{t+h}|X_t = x_t) \geq \tau\}, \quad (1)$$

其中， \inf 為最大下界； y_{t+h} 為 $t+h$ 之房價年增率； X_t 係變數向量，包含常數項。而式(1)中的 $F(y_{t+h}|X_t = x_t)$ 可進一步寫為：

$$F(y_{t+h}|X_t = x_t) = P(Y_{t+h} \leq y_{t+h}|X_t = x_t) \\ = E(1_{\{Y_{t+h} \leq y_{t+h}\}}|X_t = x_t), \quad (2)$$

式(2)說明 $F(y_{t+h}|X_t = x_t)$ 估計式可透過加權 $1_{\{Y_{t+h} \leq y_{t+h}\}}$ 指示函數(indicator function)得到，如下：

$$\hat{F}(y_{t+h}|X_t = x_t) = \sum_{i=1}^d \omega_i(x) 1_{\{y_{t+h}^i < y_{t+h}\}}, \quad (3)$$

其中， $i = 1, \dots, d$ ，為觀察值的數量；權數 $\omega_i(x)$ 設定可參見Meinshausen (2006)，簡單說明如下：

$$\omega_i(x, \theta) = \frac{1_{\{x_t^i \in R_{l(x, \theta)}\}}}{\#\{j: x_t^j \in R_{l(x, \theta)}\}}, \\ \omega_i(x) = k^{-1} \sum_{s=1}^k \omega_i(x, \theta_s), \quad (4)$$

其中， $\#$ 代表次數； $l(x, \theta)$ 為決策樹的每一片葉子(terminal nodes or leaf)； $R_{l(x, \theta)}$ 為葉子 l 對應的 x ； θ 為隨機參數向量，決定如每個節點分割考慮的變數數目。

進一步由 $\hat{F}(y_{t+h}|X_t = x_t)$ 可推知 τ 分量估計值 $\hat{Q}_{y_{t+h}|x_t}(\tau|x_t)$ ，即 y_{t+h} 小於 $\hat{Q}_{y_{t+h}|x_t}(\tau|x_t)$ 的機率等於 τ ，如下式所示：

$$\hat{Q}_{y_{t+h}|x_t}(\tau|x_t) = \inf \{y_{t+h}: \\ \hat{F}(y_{t+h}|X_t = x_t) \geq \tau\}. \quad (5)$$

分量隨機森林使用大量變數進行訓練，若參數設定不得宜，樣本內外表現易不一致。面對前述情況，一種處理方式為調整決策樹的參數，但Dubey (2018)認為該方法可

能落於太任意(ad hoc)；另一種方式為對變數進行隨機森林訓練，並根據變數重要性選擇較小的子集合。例如Kuhn and Johnson (2013)發現隨機森林模型中，變數過多可能包含許多不相關變數的負面影響，建議刪除不相關變數，而非調整隨機森林參數；Zhang and Yang (2020)依重要性順序，遞迴刪減不重要變數，選擇最佳隨機森林模型；Dubey (2018)建議簡單地選取前10%重要性最高的變數，或者也可依據 Fang et al. (2018)，選取重要性高於平均數加上2個標準差的變數，重新做為解釋變數；此外，Cysouw et al. (2021)為了減輕隨機森林模型過度配適，以變異數分析(ANOVA)分析決定保留前10%變數，減少變數輸入；Abdellatif et al. (2022)以前10個重要變數為準則，使用隨機森林模型檢測心臟病以及病患存活率；Masarat et al. (2016)指出隨機森林模型最佳變數數量亦常使用 $\log_2(p+1)$ 為標準，其中， p 是解釋變數個數。

本研究使用分量隨機森林演算法產生的「平均減少吉尼係數」，衡量特定解釋變數被隨機置換時，模型誤差增加的程度，以評估變數之重要性，並依據Dubey (2018)、Cysouw et al. (2021)方式，從中選取重要性最高的前10%變數，降低不相關變數的影響

註3。

二、建構房價成長偏斜t分配與下行風險

過去研究多使用2階段ARCH模型並假設常態分配^{註4}，評估房價分配與下行風險，新興研究趨勢為使用分量迴歸建構分配，而該方法也逐漸擴散應用至各類風險估計，如通膨風險等。由於分量迴歸不需強加分配參數或iid假設，相較ARCH模型，在構建下行風險指標更加彈性。

在Meinshausen (2006)分量隨機森林模型估計得到房價成長 $\tau = \{0.1, 0.25, 0.75, 0.9\}$ 之分量估計值 $\hat{Q}_{y_{t+h}|x_t}(\tau|x_t)$ ，進一步仿照Azzalini and Capitanio (2003)以平滑量化函數，重建房價成長偏斜t分配(更多討論可見 Adrian et al., 2019)：

$$f(y; \mu, \sigma, \alpha, \nu) = \frac{2}{\sigma} \phi\left(\frac{y-\mu}{\sigma}; \nu\right) \Phi\left(\alpha \frac{y-\mu}{\sigma} \sqrt{\frac{\nu+1}{\nu+\frac{y-\mu}{\sigma}}}; \nu+1\right), \quad (6)$$

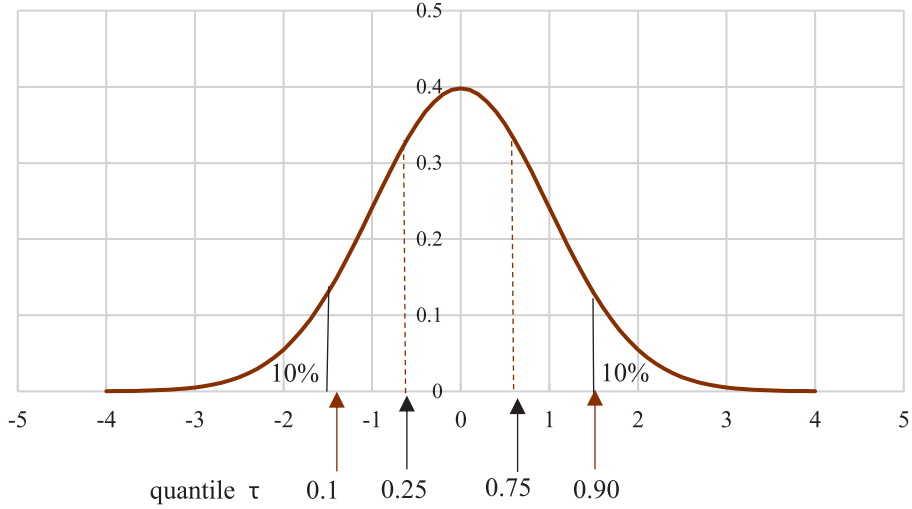
其中， $\phi(\cdot)$ 與 $\Phi(\cdot)$ 分別表示t分配的機率密度函數(probability density function, PDF)與累積機率函數(cumulative distribution function,

CDF)。t分配的 μ 、 σ 、 ν 以及 α 參數決定分配平均值(location)、規模(scale)、厚度(fatness)以及形狀(shape)，改變參數，可調整PDF偏斜效果。若 $\alpha=0$ 為傳統(traditional) t分配；在 $\alpha=0$ 與 $\nu=\infty$ 的情況，為平均值為 μ 、標準差為 σ 的高斯分配(Gaussian distribution)；當 $\nu=\infty$ 與 $\alpha \neq 0$ 時，為偏斜常態分配(skewed normal distribution)。偏斜t分配平滑化為在每次分量估計，選取偏斜t分配的4個參數 $\{\mu_t, \sigma_t, \alpha_t, \nu_t\}$ ，極小化0.1、0.25、0.75、0.9分量 $\hat{Q}_{y_{t+h}|x_t}(\tau|x_t)$ 值與偏斜t分配 $F^{-1}(\tau; \mu_t, \sigma_t, \alpha_t, \nu_t)$ 之間的差距平方項，如下式，進而再以前述4個參數建構偏斜t分配。

$$\{\hat{\mu}_{t+h}, \hat{\sigma}_{t+h}, \hat{\alpha}_{t+h}, \hat{\nu}_{t+h}\} = \underset{\mu, \sigma, \alpha, \nu}{\operatorname{argmin}} \sum_{\tau} (\hat{Q}_{y_{t+h}|x_t}(\tau|x_t) - F^{-1}(\tau; \mu, \sigma, \alpha, \nu))^2, \quad (7)$$

偏斜t分配若以均值 μ 為0與標準差 σ 為1的t分配(df=99)為例，0.1分量 $\hat{Q}_{y_{t+h}|x_t}(\tau|x_t)$ 值對應t分配累積機率達10%之處，如圖2所示，約莫是-1.66。所以當分量越低，越可反映尾端下行風險；反之，則為上行風險。

圖2 t分配與分量的關係



政策制定者除了關心房價趨勢，也會擔憂總體金融因素等帶來的下行或上行風險。依據Adrian et al. (2019)，總體金融造成的房價風險可藉由比較房價成長條件分配與無條件分配尾端差異得知，亦即衡量條件密度函數相對於非條件機率分配右尾與左尾的額外機率。Adrian et al. (2019)定義相對上行及下行風險值 Γ_t^U 與 Γ_t^D ，如下式：

$$\Gamma_{h,t}^D(\hat{\phi}_{y_{t+h}|x_t}; \hat{g}_{y_{t+h}}) = - \int_{-\infty}^{\Phi_{\tau_{t+h}|x_t}^{-1}(0.5|x_t)} (\log \hat{g}_{y_{t+h}}(y) - \log \hat{\phi}_{y_{t+h}}(y|x_t)) \hat{\phi}_{y_{t+h}|x_t}(y|x_t) dy, \quad (8)$$

$$\Gamma_{h,t}^U(\hat{\phi}_{y_{t+h}|x_t}; \hat{g}_{y_{t+h}}) = - \int_{\Phi_{\tau_{t+h}|x_t}^{-1}(0.5|x_t)}^{\infty} (\log \hat{g}_{y_{t+h}}(y) - \log \hat{\phi}_{y_{t+h}}(y|x_t)) \hat{\phi}_{y_{t+h}|x_t}(y|x_t) dy, \quad (9)$$

其中， $h=1, 4$ ； $\hat{g}_{y_{t+h}}$ 為利用偏斜t分配建構的房價成長非條件密度函數； $\hat{\phi}_{y_{t+h}}$ 則是條件密度函數； $\Phi_{\tau_{t+h}|x_t}^{-1}(0.5|x_t)$ 為條件中位數 (conditional median)。

Adrian et al. (2019)又稱 $\Gamma_{h,t}^U$ 與 $\Gamma_{h,t}^D$ 為上

行與下行熵(entropy)風險值。熵起初代表能量退化的指標，用於計算一個系統中的失序，當系統內越混亂，熵值越高，故熵又稱為亂度。後來，廣泛地被應用於風險測度及經濟領域。熵風險值為相對熵(relative entropy)組成，用來解釋兩個分配的離散程度，衡量不確定性。下行熵已被文獻應用如通膨、經濟成長等下行風險評估，如Banerjee et al. (2020)。

依據式(8)與式(9)，當條件分配機率集中左尾，而非條件分配集中在右尾，上行與下行熵風險值會分別呈現負值與正值；反之，當條件分配機率集中右尾，而非條件分配多落在左尾，上行與下行熵風險值會分別出現正值與負值。綜上，若條件分配比無條件分配賦予左尾更多的機率，下行風險上升， $\Gamma_{h,t}^D$ 會增加；同樣地，條件分配比無條件分配賦予右尾更高的機率，代表上行風險

增溫， $\Gamma_{h,t}^U$ 將上升。

另外，為了更完整衡量房價成長的極端值風險，本文亦參照Adrian et al. (2019)，計算下行風險期望值 (SF_{t+h}) 與上行風險期望值 (LR_{t+h})，定義如下：

$$SF_{t+h} = \frac{1}{\pi} \int_0^{\pi} \hat{F}_{y_{t+h}|x_t}^{-1}(\tau|x_t) d\tau, \quad (10)$$

$$LR_{t+h} = \frac{1}{\pi} \int_{1-\pi}^1 \hat{F}_{y_{t+h}|x_t}^{-1}(\tau|x_t) d\tau, \quad (11)$$

此處設定 $\pi = 0.1$ ，計算房價變動大於0.9(小於第0.1)分量的條件期望值 $LR_{t+h}(SF_{t+h})$ ，提供尾端分配資訊。下行風險期望值增加，預期損失可能上升；反之，上行風險期望值增加亦同，預期長期收益將提高。

三、房價成長預測分配績效與下行風險預測準確性

(一) 預測分配表現評比

分配預測樣本外表現評比過程，為將樣本期間分為I個樣本內，以及O個樣本外部分，採遞迴法(recursive scheme)，每次分析擴增1季樣本，模擬預測分配，直至資料結束。至於樣本外分配預測表現比較，Diebold et al. (1998)最早提出透過機率積分轉換(Probability integral transform, PIT)評估樣本外預測分配的可靠性(reliability)。

令分量模型預測小於實際值 y_{t+h} 的累積分配機率为：

$$\begin{aligned} \hat{\Psi}_t(y_{t+h}) &= \int_{-\infty}^{y_{t+h}} \hat{\phi}_{t+h}(\epsilon|\Theta) d\epsilon \\ &= \hat{\Phi}_{t+h}(y_{t+h}|\Theta), \end{aligned} \quad (12)$$

其中， Θ 為資訊集。將累積分配 $\hat{\Psi}_t(y_{t+h})$ 機率積分轉換為：

$$Z_{t+h} = \hat{\Psi}_t(y_{t+h})$$

若條件分配正確預測，與資料 y 的生成過程(data-generating process)一致，Diebold et al. (1998)證明累積分配經機率積分轉換， Z_{t+h} 會服從i.i.d. $U(0,1)$ ，藉此可評斷樣本外分配預測可靠性^{註5}。

當 Z_{t+h} 對應 τ 的累積分配圖形，越接近45度線(即均勻分配的累積分佈)，模型分配預測具有更高的可靠性。Rossi and Sekhposyan (2017)進一步透過計算樣本外預測 Z_{t+h} 累積分配的信賴區間，評估分配正確性。若累積分配位於信賴區間之內，則預測分配被正確設定。

$$\text{信賴區間} = \gamma \pm \kappa_{\alpha} \sqrt{O}, \quad (13)$$

其中， γ 是均勻分配的CDF， κ_{α} 為Rossi and Sekhposyan (2017)計算的臨界值，O是樣本外數量。

(二) 房價下行風險預測準確性

樣本外各分量預測準確性可用Koenker and Machado (1999)的 Quantile R^2 指標評比，如下示：

$$\text{Quantile } R^2 = 1 - \frac{\frac{1}{T} \sum_t [\rho_{\tau}(y_{t+h} - \hat{y}_{t+h} \hat{\beta}_{\tau})]}{\frac{1}{T} \sum_t [\rho_{\tau}(y_{t+h} - \hat{q}_{\tau})]}, \quad (14)$$

其中， ρ_{τ} 為分量迴歸損失函數^{註6}。上式分子為條件分量迴歸殘差，而分母則代表無條件分量迴歸殘差，Quantile R^2 越大，解釋變數改善模型配適度的程度越高。但若無

條件分量迴歸比條件分量迴歸提供更好的解釋力，Quantile R^2 為負值。下文分別檢驗 $\tau=0.9、0.5、0.1$ 的預測表現，以了解模型捕捉房價成長上行風險、中位數以及下行風險的情形。

四、金融(銀行)危機預警與預警績效評估

黃朝熙、黃裕烈、黃淑君、謝依珊、楊茜文(2014)指出2008年前台灣經歷過兩次金融風暴，分別為台灣本土金融風暴(1998 Q2~2001 Q4)、雙卡風暴(2005 Q3~2006 Q4)。另因2008年10月次貸風暴引發實施存款全額保障措施，李桐豪、江永裕(2009)將其定義為金融危機。綜合前述文獻，分別將

1998 Q2~2001 Q4、2005 Q3~2006 Q4、2008 Q3~2008 Q4定義為金融(銀行)危機事件發生期間^{註7}。

文獻已指出房價風險可改善金融危機預測(參見IMF, 2019; Adrian et al., 2020)，雖然不一定適用台灣。本研究仍嘗試以Logit迴歸及隨機森林，評估樣本內外包含房價成長下行風險值的預警模型，可否預測台灣金融(銀行)危機，並以接收者操作特徵曲線面積(Area Under the Receiver Operating Characteristic, AUROC)與二次機率分數(the quadratic probability score, QPS)做為表現評斷標準。Logit迴歸設定、AUROC以及QPS等說明詳見附錄。

肆、實證結果

一、資料來源

本研究的房價指數取自信義房屋的台灣房價指數^{註8}，並從TEJ等資料庫，依據房地產市場、總體經濟、利率與貨幣供給、股票市場以及外部因素等面向，選取47變數，詳參見附錄，加上房價成長自我一階落後項，共計48個解釋變數，可供分量隨機森林模型篩選。而全體樣本期間為1992年Q1~2022年Q1^{註9}。

值得注意的是，從表1房價年增率基本統計量觀察，平均值為4.59%；整體非條件

分配的偏態(skewness)為0.25，大於零，為正偏(positive skew)或右偏，即少數時候房價會出現高漲幅，此代表台灣房價變動具不對稱性。此外，房價年增率分配的峰度(kurtosis)小於3，分布較常態分配分散。

表1 房價年增率基本統計量

	房價年增率
平均值	4.59%
中位數	3.35%
極大值	20.95%
極小值	-8.27%
峰度	2.07
偏態	0.25

而之所以納入外部因素為候選解釋變數，其一，考量資本移動會影響全球資產價格共同變動，房價可能會隨證券投資及FDI等資本流入起伏。例如，證券投資資金流入會激勵股價上揚，民眾財富增加後，購屋需求可能會擴增(更多討論參見IMF, 2019; Kim and Yang, 2009a; Kim and Yang, 2009b; Taguchi et al., 2015)；其次，全球風險趨避傾向(global risk aversion)(通常以VIX指數做為代理變數)與影響全球資產價格共同變動的全球因子(global factor)密切相關，而且對於全球資本移動的走向也有重大影響力(Nier et al., 2014; Bergant et al., 2020)，故納入VIX指標捕捉全球風險趨避傾向。

至於危機預測，Logit樣本內迴歸以及隨機森林樣本外預測的解釋變數，除房價下行風險值，還包含景氣循環、實質有效匯率等控制變數。其中，GDP取自主計總處，並以CF濾波器萃取循環項；實質有效匯率則來自BIS。

二、房價條件分配之決定因素及其隱含資訊對預測分配之影響

本節首先以分量隨機森林訓練全樣本變數，預測未來1季房價成長分配，並依據Mean Decrease Gini準則，選取前10%重要性高的變數。除了房價成長1階自我落後項，還包括其他4個變數，分別為失業率、核發建造執照件數、核發建照執照樓板面積以及

就業人數。另一方面，隨機森林模型最佳變數數量亦常使用 $\log_2(p+1)$ 為標準，其中， p 是解釋變數個數，見Masarat et al. (2016)。依此標準，最佳變數數目約為5.62個，與前10%重要解釋變數數目共5個變數，差距不大，因此以前10%重要變數做為新變數集合。

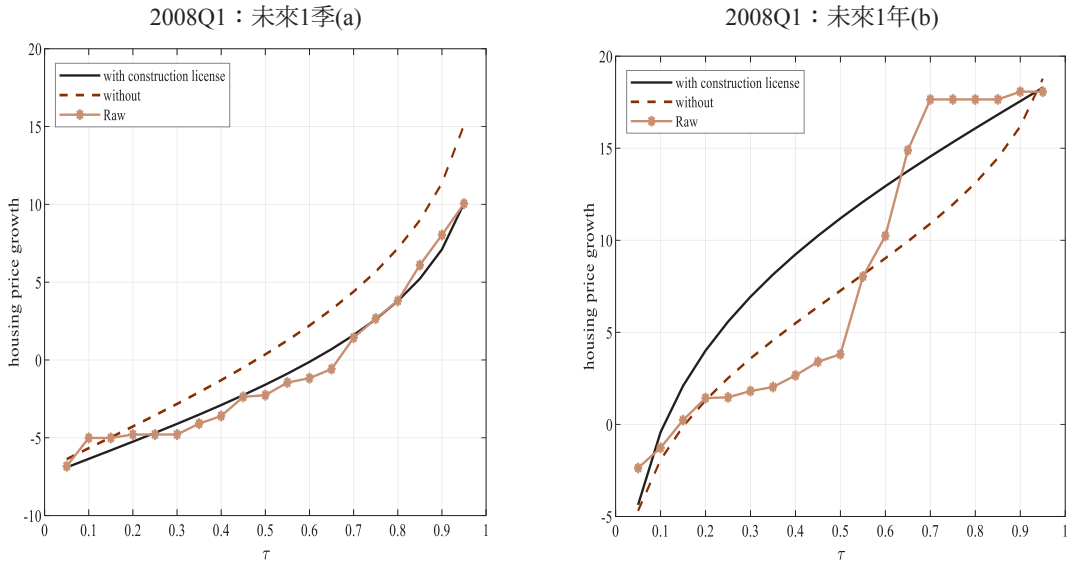
選取的變數中，失業率與就業人數可反映房地產需求，失業率越低以及就業人數越多，民眾所得增加，有助推升購房意願；核發建照執照件數與面積反映房屋供給，建商若看好未來房市前景，會積極購地推案，核發建照執照件數與面積將增加；反之則是放緩推案步調，建照數及建照面積將下滑(林左裕、程于芳, 2015)。綜上，核發建照執照件數與面積變化或可提早反應房市重大事件的來臨。

由於核發建照執照件數與面積反映建商對房地產市場前景的看法，可能隱含未來房價成長分配變動資訊。以下便以核發建造執照件數為例，模擬分析2008 Q1全球金融危機前以及2014 Q2房價起跌點，檢驗模型有無加入核發建造執照件數，能否提前預知房價成長分配變動。

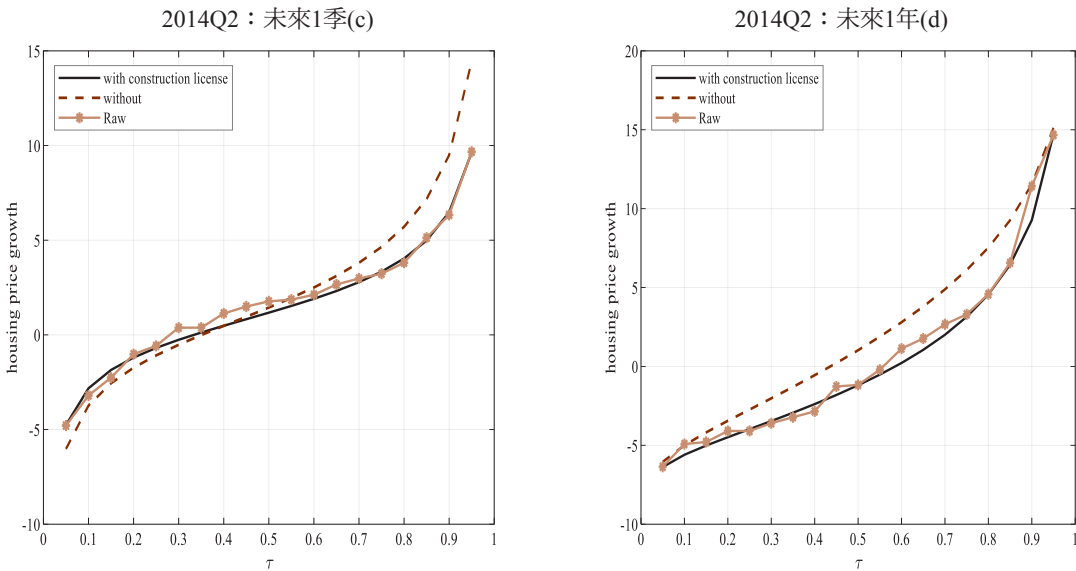
圖3的(a)與(c)、(b)與(d)分別表示站在時點2008Q1以及2014Q2，預測未來1季與1年房價成長條件分量之結果；其中，黑色實線與橘紅色虛線為分別代表，模型加入與不加入核發建造執照件數的條件預測分量^{註10}；

而圖3之中的(a)、(b)、(c)、(d)等4個子圖形下，所估算的實際房價成長條件分量值。裡的Raw曲線則為已擁有未來1季及1年資訊

圖3 條件分量預測(2008Q1與2014Q1)



註：對1季與1年的條件分量值估計；為便於比較，Raw為使用當時實際資料，以分量隨機森林估計的條件分量值。



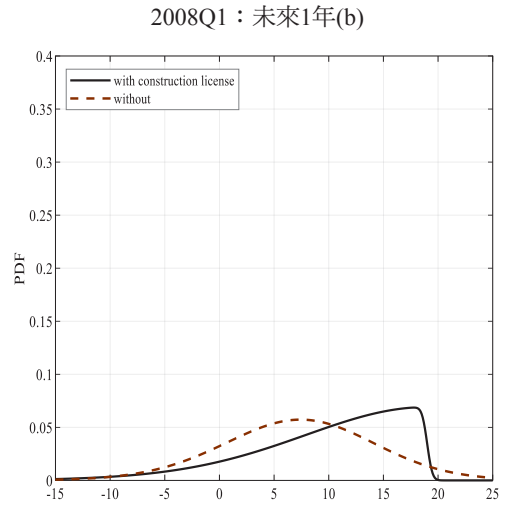
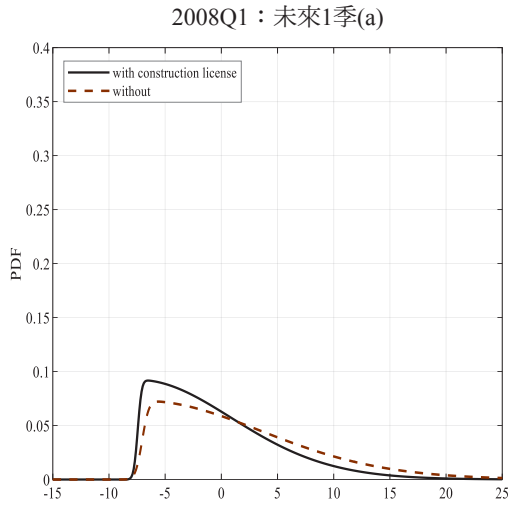
註：對1季與1年的條件分量值估計；Raw為使用當時實際資料，以分量隨機森林估計的條件分量值。

對照房價實際走勢，2008 Q2台灣房價在歷經全球金融危機後，走跌；另一方面，2014 Q3台灣房價亦出現跌勢。模型在納入核發建造執照件數之後，在2008 Q1以及2014 Q2對未來1季條件分量預測，明顯低於不考慮該變數時，特別是在右尾，並與Raw曲線(實際分量值)幾乎重合。即便只使用過去資料，模型仍提早預知了未來房價可能的下跌風險。綜上，核發建造執照件數能夠幫助適度先行反映房價成長下行風險；另一方面，台灣房價成長在金融危機後約1年的2009 Q1轉為回升；而相反地，2015 Q2則延續2014前1年房市疲軟態勢。模型在考量核發建造執照件數後，在2008 Q1對未來1年(即2009Q1)條件高分量預測值上移，與Raw曲線類似，顯示其或可幫助捕捉房價上漲轉折點；且在2014 Q2對未來1年(即2015Q2)的全

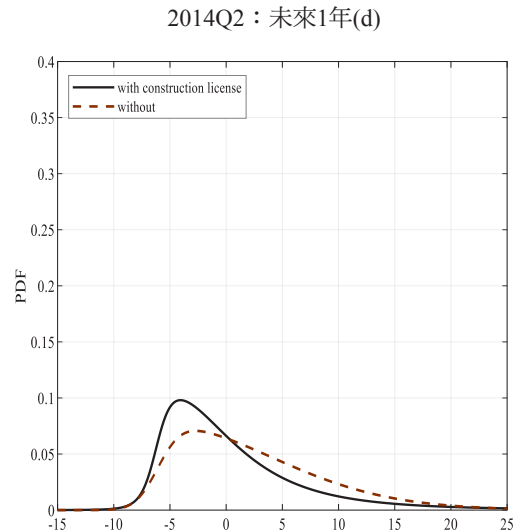
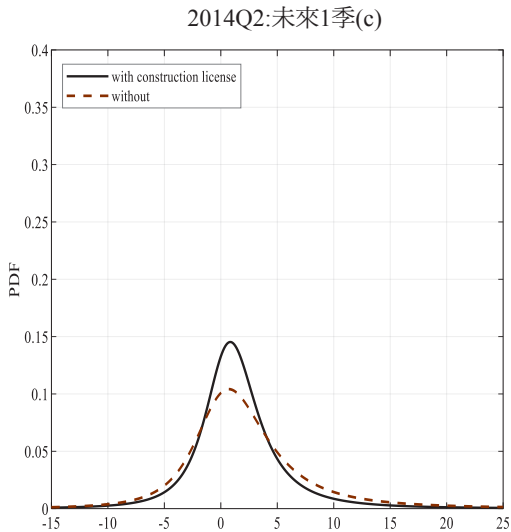
體分量預測值接近實際分量值，預測表現更勝一籌。

進一步，圖4為對應圖3，將不同預測時點之房價成長條件分量值，透過擬值方式轉換為t分配。從中可更清楚看出，當2008 Q1與2014 Q2之後未來1季(2008Q2與2014Q3)房價將出現大幅修正前，模型加入核發建造執照件數，未來1季條件分配預測更傾向左偏(例如在2008 Q1對未來1季的預測)，或者整體分配左移(例如在2014 Q2對未來1季的預測)，平均值降低。另一方面，在2008 Q1與2014 Q2對未來1年條件分配預測分別表現出，更大程度右偏與更高的平均值；以及分配左移與較低的平均值，大致符合2009Q1與2015Q2實際房價變化。總體而言，核發建造執照件數可改善模型預測房價成長分配變化，甚至是房市景氣轉折。

圖4 房價成長t分配預測(2008Q1與2014Q1)



註：以偏斜t分配建構。



註：以偏斜t分配建構。

三、全樣本分配預測

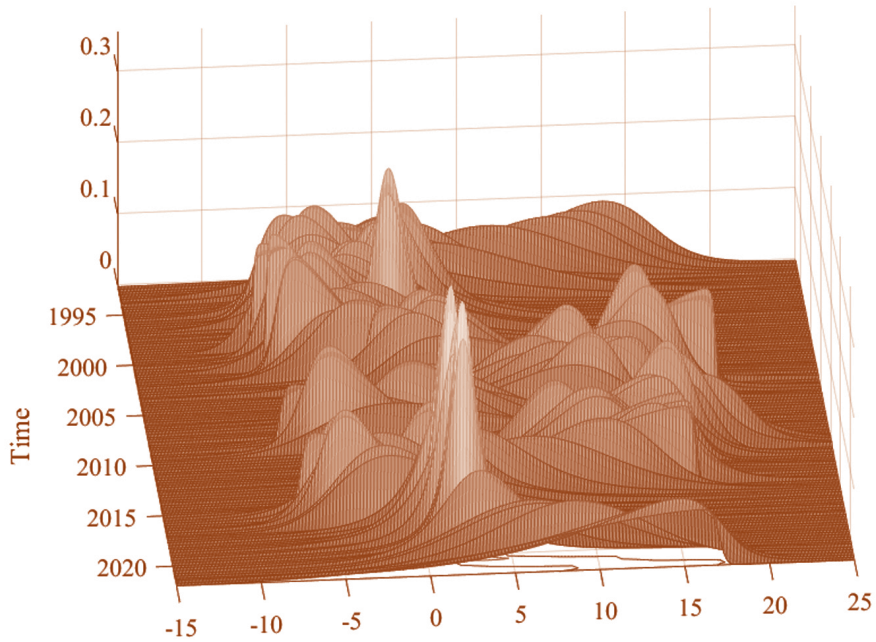
圖5與圖6為全樣本台灣歷年未來1季與1年房價成長預測分配，具有兩項特徵。

其一，分配中位數與左右尾表現出強烈的隨時間變化。例如，房市擴張階段，下行風險逐漸累積，條件分配傾向為左尾分配。

以最近一次房價上漲來看，條件分配明顯左傾；反之，在房價下跌時期，條件分配則多為右尾分配，房價可能在未來反彈。

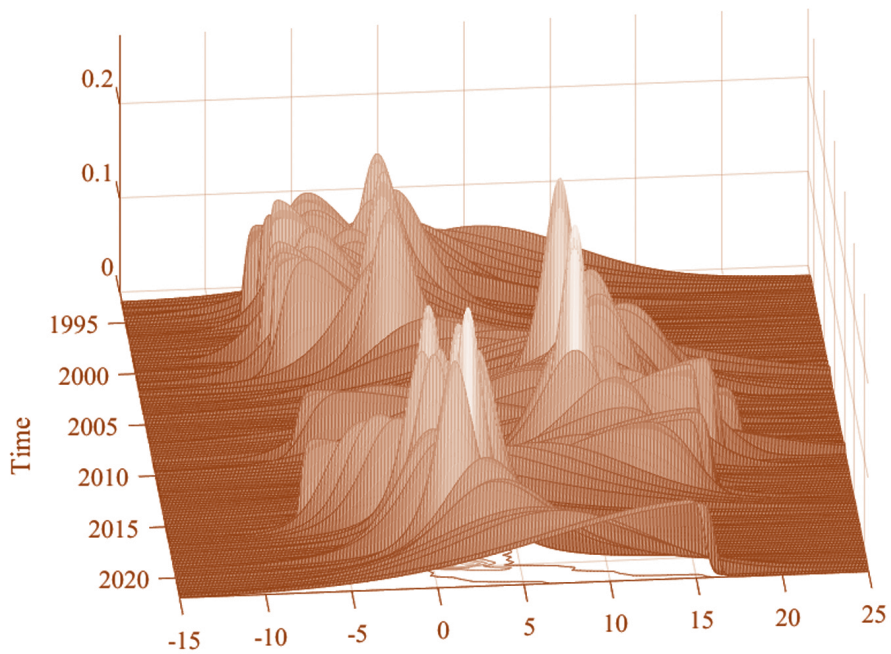
其次，房價成長條件分配動態變化具不對稱性，分配左尾極端值較集中；房價成長上行風險則較明顯波動，分布範圍較廣。

圖5 歷年未來1季房價成長預測分配



註：縱軸為時間；橫軸為分配對應的房價成長率(%)。
資料來源：自行計算

圖6 歷年未來1年房價成長預測分配



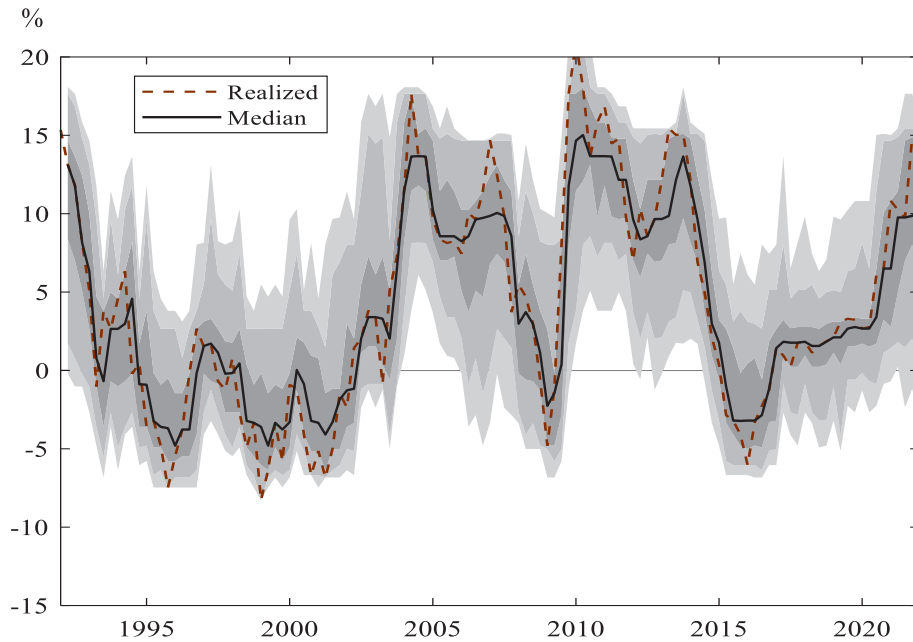
註：縱軸為時間；橫軸為分配對應的房價成長率(%)。
資料來源：自行計算

綜上，在房價在上漲過程，分配會逐漸往左傾斜，下行風險增加；反之，亦然。因此，房價持續上漲將累積潛在風險。

圖7進一步顯示未來1季房價成長的條件中位數以及0.05、0.1、0.25、0.75、0.9與0.95分量值估計，除2000年初房價上漲階段以及2008年後的房市反彈之外，低分量(例

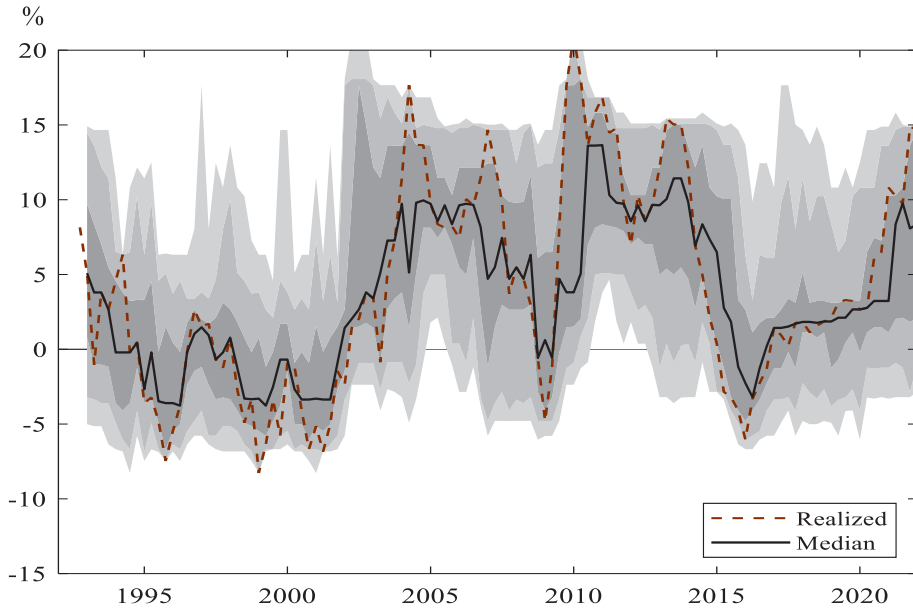
如0.05、0.1、0.25)數值分布集中；而高分量(例如0.75、0.9、0.95)分布範圍較廣，呈現不對稱。前述現象與圖5與圖6觀察一致；若進一步觀察未來1年條件房價成長高分量，因為預測期間越長，資訊量降低，造成不同分量分布區間擴大。

圖7 歷年房價成長各分量預測
未來1季



註：灰色色階分別為0.05、0.1、0.25、0.75、0.9與0.95分量估計值；黑色實線為0.5分量估計值；橘紅色虛線為實際值。

續圖7
未來1年



註：灰色色階分別為0.05、0.1、0.25、0.75、0.9與0.95分量估計值；黑色實線為0.5分量估計值；橘紅色虛線為實際值。

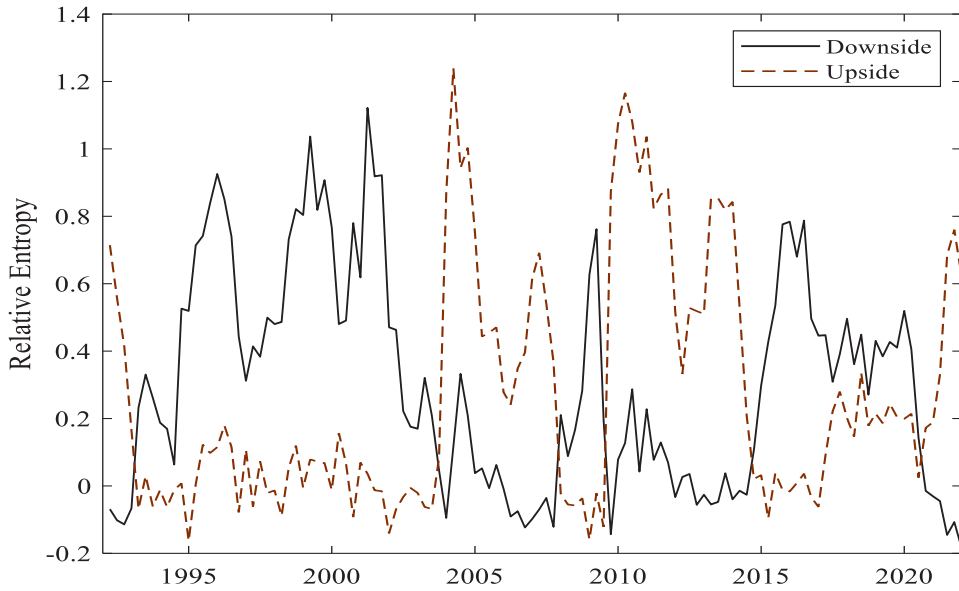
四、房價上行與下行風險

當房價下行風險較高時，且解釋變數蘊含額外資訊，條件分配的左尾機率將大於無條件分配(意味下行風險的相對熵(relative entropy)將上升)。同樣地，房價上行風險較高時，條件分配將比無條件分配給予右尾更高的機率(意味上行風險的相對熵(relative entropy)增加)。圖8縱軸為未來1季與1年的房價上行與下行風險值(即相對熵)變化走勢，

橫軸則為時間。以2022年期間趨勢為例，上行風險值為正，而下行風險值為負數，房價成長展現高度上行風險。

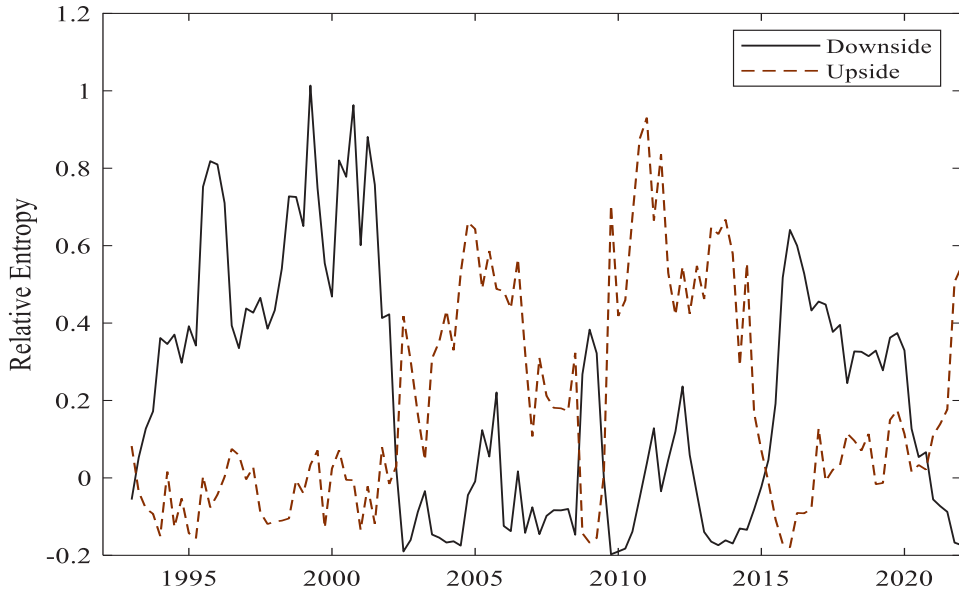
圖9上行(下行)風險期望值為計算房價年增率大於第90(小於第10)百分位的條件期望值，即為90%(10%)的預期長期收益(預期損失)。若下行風險期望值增加，預期損失上升；反之，預期長期收益可能提高。從圖9可發現，房價預期損失波動性相較預期長期收益明顯較低，具不對稱現象^{註11}。

圖8 房價風險變化(未來1季及1年)
房價下行與上行風險值(Γ_t^D 與 Γ_t^U)：未來1季



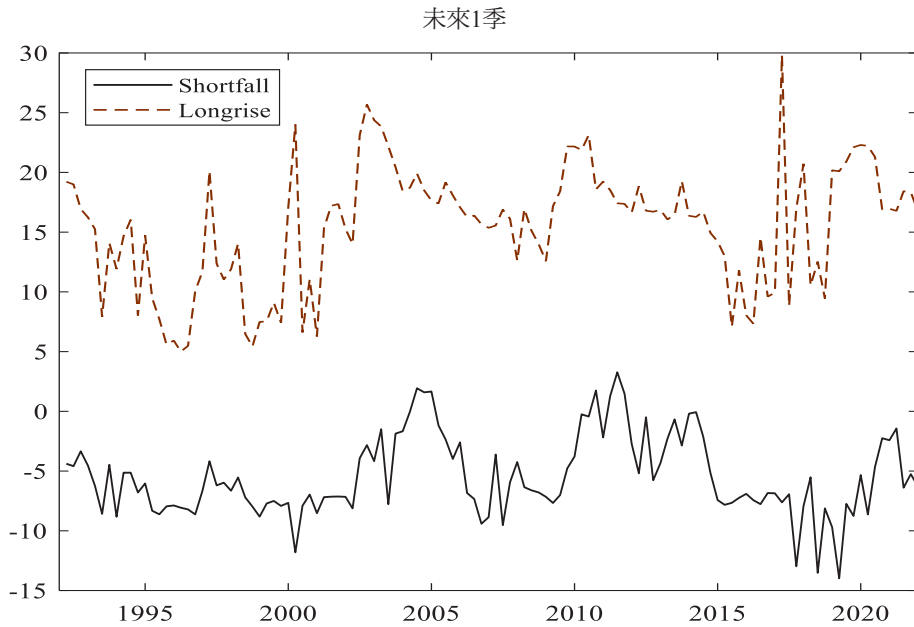
資料來源：自行計算

續圖8
房價下行與上行風險值(Γ_t^D 與 Γ_t^U)：未來1年



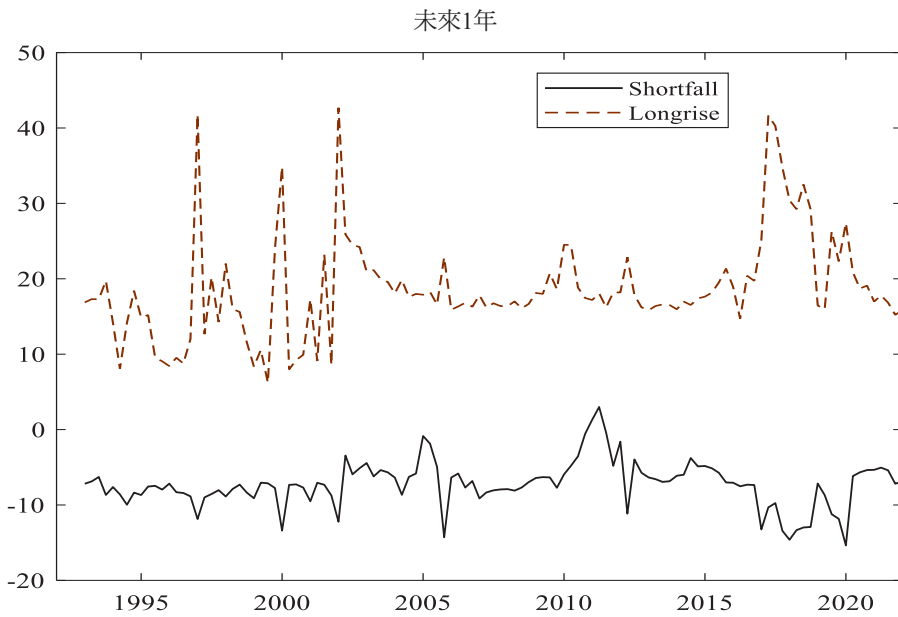
資料來源：自行計算

圖9 預期房價損失與長期收益(未來1季及1年)



註：Shortfall代表10%的預期損失；Longrise為90%的預期長期收益。
資料來源：自行計算

續圖9



註：Shortfall代表10%的預期損失；Longrise為90%的預期長期收益。
資料來源：自行計算

五、樣本外分配預測與績效評比

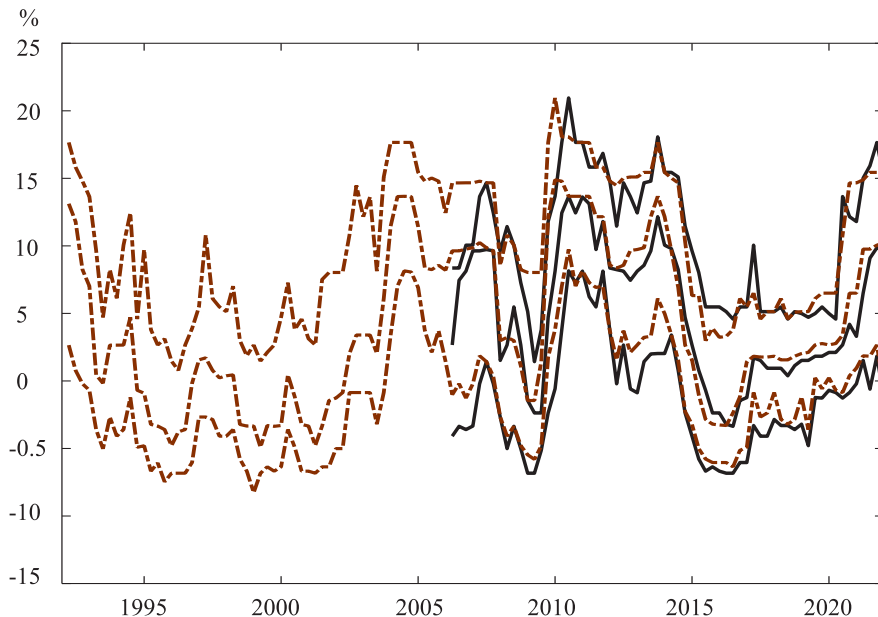
本節先使用全樣本進行分量隨機森林分類，建立各分量的預測值。再以2006 Q1為起點，使用遞迴方式，每次增加1筆樣本，估計未來1季及1年預測分配，直到樣本結束。換言之，在每次遞迴過程，進行分量隨機森林分類，除預測房價成長下行(0.1分量)與上行風險(0.9分量)以及中位數(0.5分量)，也建構偏斜t分配。

圖10黑色實線為樣本外各分量估計值，分別在2008年全球金融危機，與2011年歐債

危機，以及央行以總體審慎政策調控房市，奢侈稅與豪宅稅上路，2014年後出現較大幅修正；因此，模型遞迴估計樣本外各分量值時，雖然只使用過去資訊，惟仍可良好預測2008年與2014年等房市下跌事件對房價成長各分量值的影響。

此外，樣本外分量估計不論是未來1季或是1年大多與全樣本結果(橘紅色虛線)貼近，特別是在0.1分量估計值，模型較能發現房價成長脆弱性。反觀，模型在0.9分量預測能力相對較弱，且預測時間越長，估計值越明顯偏離全樣本結果。

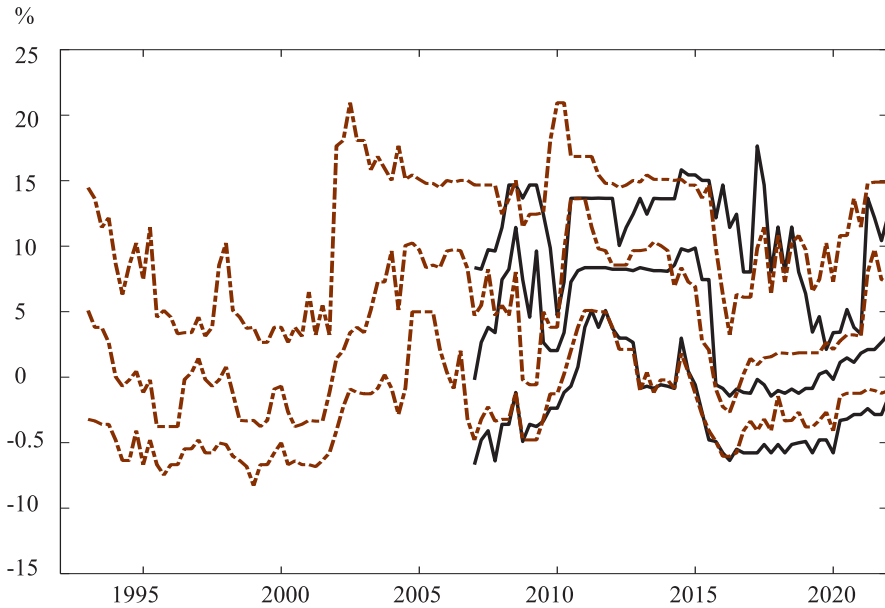
圖10 樣本內外主要分量值預測
未來1季



註：0.1、0.5與0.9分量估計值樣本內外預測；橘紅色虛線為全樣本估計，黑色實線為樣本外估計。

資料來源：自行計算

續圖10
未來1年



註：0.1、0.5與0.9分量估計值樣本內外預測；橘紅色虛線為全樣本估計，黑色實線為樣本外估計。

資料來源：自行計算

為評估各分量預測準確性，進一步使用 Quantile R^2 比較，表2顯示，模型在0.1分量預測力均高於0.9分量；並且雖然模型對未來1年各分量預測能力皆下降，但是於0.1分量預測準確性仍大於中位數以及0.9分量，模型對於較長期房價成長下行脆弱性預測之表現相對較佳。

最後，以機率積分轉換(PIT)檢驗樣本外分配預測可靠性。PIT衡量給定分量，累積分配越接近45度線，模型分配預測準確性就越高。

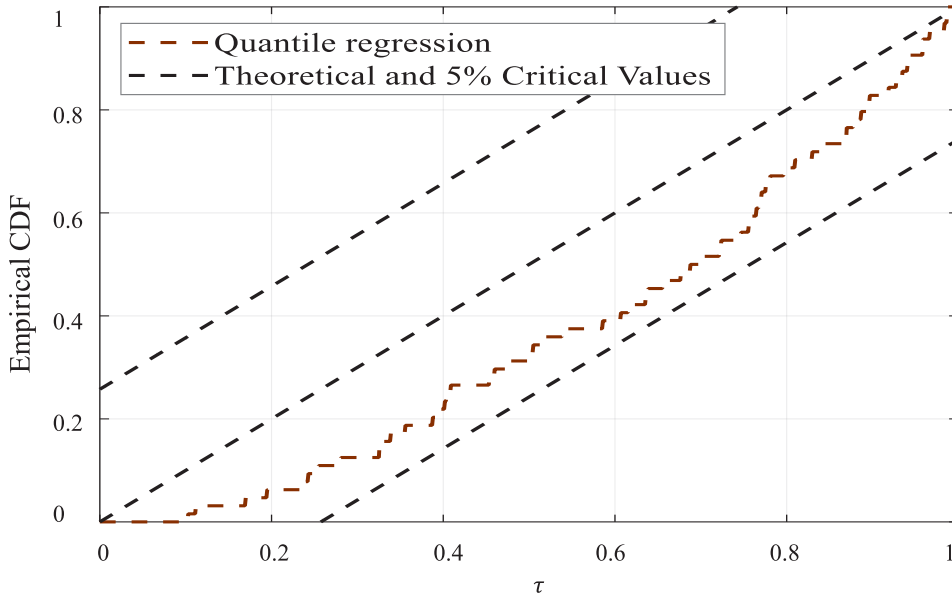
圖11中，PIT累積分配皆分布於信賴區間之內，並且未來1年低分量的預測累積分配更加靠近45度線，反映模型預測房價下行脆弱性的表現較優。整體來說，預測分配表現適當，且更能夠捕捉下行風險。

表2 Quantile R^2 -分量隨機森林模型

	$\tau=0.10$	$\tau=0.5$	$\tau=0.90$
$h=1$	0.7538	0.7428	0.2598
$h=4$	0.4781	0.3463	-0.5911

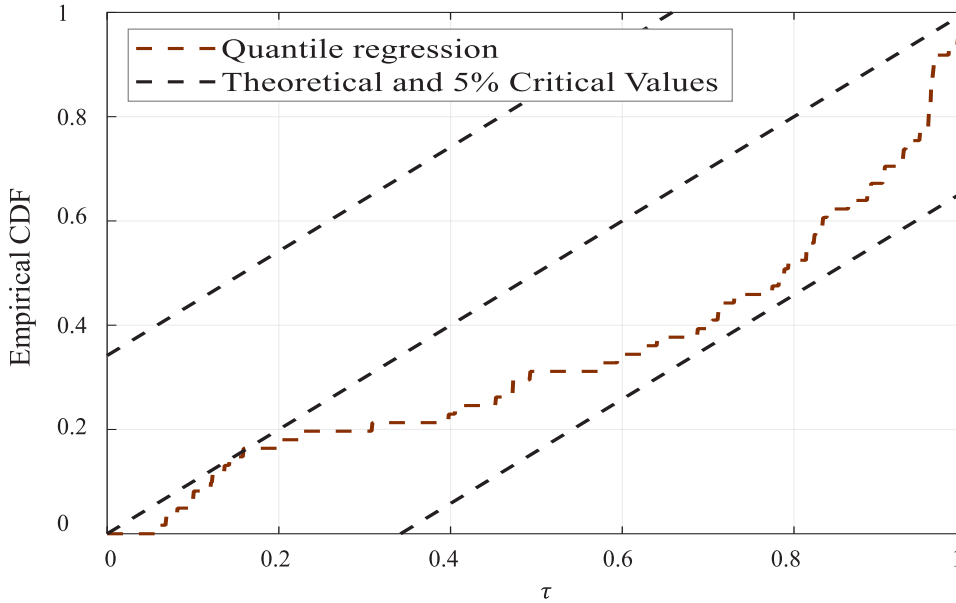
註：h代表預測長度；預測未來1季： $h=1$ ；預測未來4季： $h=4$ 。

圖11 樣本外分配預測準確性
未來1季



註：機率積分轉換累積分配的臨界值以Rossi and Sekhposyan (2017)方式計算。
資料來源：自行計算

續圖11
未來1年



註：機率積分轉換累積分配的臨界值以Rossi and Sekhposyan (2017)方式計算。
資料來源：自行計算

六、穩健性檢定-傳統分量迴歸

本節以傳統分量迴歸進行各分量樣本外預測(變數同樣包括房價1階自我落後項、失業率、核發建造執照件數、核發建照執照樓板面積以及就業人數)，並利用 Quantile R^2 評比績效。比較表3與表2，傳統分量迴歸模型相較於分量隨機森林，在未來1季房價成長低分量(0.1分量)預測的 R^2 較小，且對於未來1年房價成長低分量預測的 R^2 為負值，模型預測下行風險能力較弱。

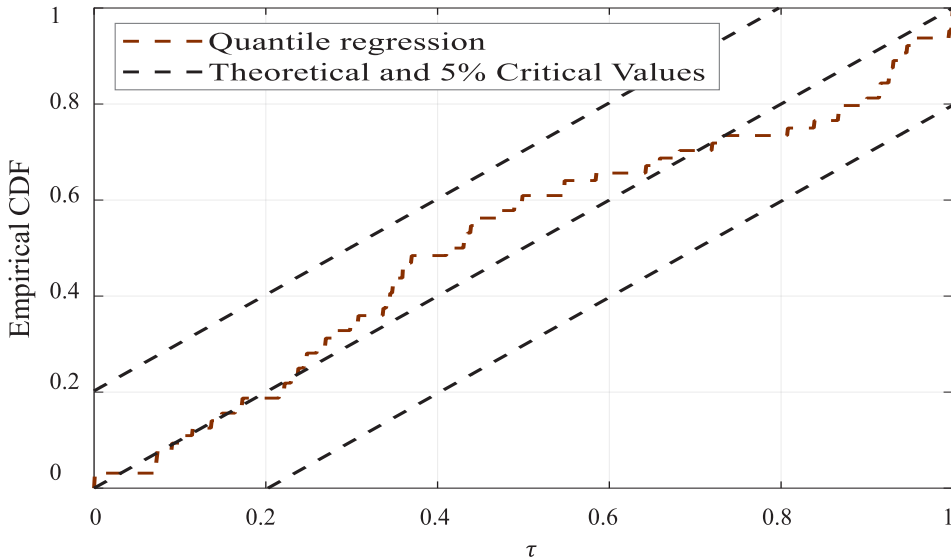
而圖12顯示，PIT累積分配均在信賴區間內，整體分配預測仍屬恰當；惟未來1年房價成長低分量預測的累積分配較偏離45度線，反映傳統分量模型預測較長期的房價下行風險表現相對較差。

表3 Quantile R^2 -傳統分量迴歸模型

	$\tau=0.10$	$\tau=0.5$	$\tau=0.90$
$h=1$	0.4232	0.9423	0.7929
$h=4$	-2.2742	0.8732	0.6701

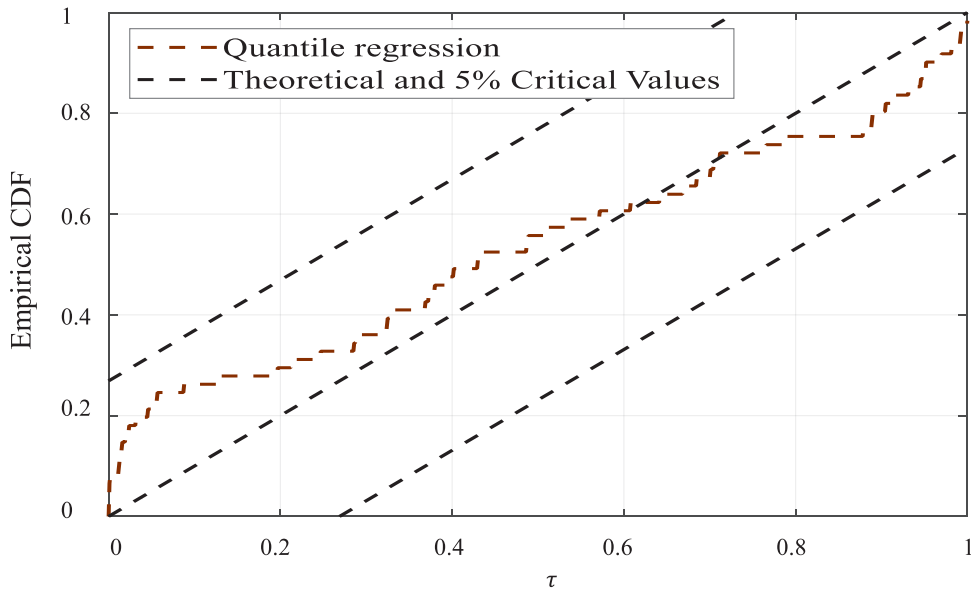
註：h代表預測長度；預測未來1季：h=1；預測未來4季：h=4。

圖12 樣本外分配預測準確性
未來1季



註：機率積分轉換累積分配的臨界值以Rossi and Sekhposyan (2017)方式計算。
資料來源：自行計算

續圖12
未來1年



註：機率積分轉換累積分配的臨界值以Rossi and Sekhposyan (2017)方式計算。
資料來源：自行計算

七、金融危機事件預警

本節參考Chen and Sviryzdenka (2021)，透過Logit模型觀察房價風險指標迴歸係數是否顯著，評估其是否可作為金融(銀行)危機的預警指標之一。此外，由於房價風險係由分量隨機森林模型計算，為求模型一致性，金融(銀行)危機樣本外預測以隨機森林模型進行，兩者均為機器學習模型^{註12}。

首先，進行傳統Logit模型迴歸的樣本內表現評估，並設定危機預測期間 $h = 1, 4$ (即未

來1季與1年)。表4顯示，模型引入總體經濟控制變數，如產出循環($GDP.Cycle_t$)控制如景氣、信用擴張等造成之影響，以及REER捕捉外部衝擊，如國際情勢波動等，未來1季與1年房價下行風險值($\Gamma_{1,t}^D$ 、 $\Gamma_{4,t}^D$)迴歸係數仍統計顯著，房價風險如同其他資產價格會提前反應金融(銀行)危機到來。整體而言，房價下行風險資訊可做為金融(銀行)危機預警指標之一，特別是未來1季Logit迴歸模型II McFadden (1977)的 pseudo R^2 最高。

表4 Logit模型迴歸樣本內配適

	$h=1$		$h=4$	
	模型I	模型II	模型I	模型II
constant	-1.26 (1.83)	1.85 (2.13)	-3.66 (2.21)	-2.22 (2.54)
$\Gamma_{1,t}^D$	3.14*** (0.92)	--	2.40*** (0.75)	--
$\Gamma_{4,t}^D$	--	4.31*** (1.06)	--	3.16*** (0.88)
$GDP.Cycle_t$	0.87*** (0.33)	0.79*** (0.32)	1.14*** (0.39)	1.13 (0.41)
REER	-0.013 (0.02)	-0.04*** (0.02)	0.01 (0.02)	-0.002 (0.02)
McFadden's pseudo R^2	0.15	0.21	0.15	0.19

註：* ** ***分別代表在10%, 5%, 1%顯著水準之下拒絕虛無假設；括弧內為標準誤。
資料來源：自行估計

緊接著，使用隨機森林模型，以遞迴方式進行金融(銀行)危機之樣本外預測。樣本內期間為1992Q1~2003Q1，預測範圍包括雙卡風暴(2005 Q3~2006 Q4)，以及2008年全球金融危機。表5實證結果顯示，當預測期間越長，則模型預測表現越佳，且所有AUROC值均大於0.5；而 QPS 相對穩定。模

型長期預測力較高，有可能是因為房價下行風險結合景氣與國際情勢波動等資料，能提供更早的危機資訊。

綜上，房價下行風險等變數建構的隨機森林模型最早可在1年前，便發出有用的危機警訊。

表5 金融危機樣本外預測

	模型I ($\Gamma_{1,t}^D$)		模型II ($\Gamma_{4,t}^D$)	
	$h=1$	$h=4$	$h=1$	$h=4$
AUROC	0.61	0.71	0.52	0.68
QPS	0.11	0.11	0.12	0.12

資料來源：自行估計

伍、結 論

過去以來，先進國家發生金融危機之前，房價大多先暴跌，如2008年的全球金融危機。近幾年許多先進國家的房價同步上漲，擴大房價下行風險並引起各界關注相關潛在後果。

近期國外許多研究開始使用Adrian et al. (2019)分量方法建構房價風險，並試圖探討其與危機的關聯，但台灣在此方面的研究仍相當匱乏。這可能係因為台灣房地產法規嚴謹，無造成美國次貸危機的證券化商品與次貸產品，且房貸違約率不高並續創歷史新低，房市過熱時，政府、央行也會適時進場調控，穩定房市，要如同美國因房地產過度發展，而發生金融危機的可能性不高。綜觀，台灣過去的幾次金融危機，均由信用過度擴張(如2005 Q3~2006 Q4雙卡風暴等)所致。惟房價下行風險或可能蘊含資產價格領先或台灣金融失衡資訊，因此本研究也嘗試將房價下行風險當做危機預警指標之一。

另一方面，分量隨機森林模型除能應用大量資料，亦可進一步篩選有用的特徵變數，加以具捕捉變數非線性關係的特質，預測表現可能較傳統時間序列模型更佳，成為實證研究新工具。本文蒐集房地產市場供需面向、總體經濟、利率與貨幣供給、股票市場以及外部因素等共47個變數，並以Meinshausen (2006)分量隨機森林模型建構

台灣房價成長分配及下行風險，以及探討其可能應用。截至目前為止，國內尚無以分量隨機森林模型分析房價成長分配等之相關文獻。

本研究主要結論如下：

1. 分量隨機森林模型可良好捕捉房價下行風險

分量隨機森林模型預測未來1季房價下行風險的表現優於傳統分量迴歸。此外，其在預測房價下行風險能力優於上行風險。

2. 房價下行風險或可做為危機預警指標之一

雖然台灣歷次金融危機與房價波動並無太大關連，但房價內含的領先資訊，在樣本內預測顯示，房價下行風險可做為危機有效早期預警指標之一；而樣本外預測期間越長，模型預測表現越佳，最早可在1年前發出警訊。

本研究亦提出幾點未來研究的可能方向：其一，本研究定義房價下行風險係以近期Adrian et al. (2019)提出的下行熵風險值衡量，未來可考量其他測量方式，例如風險價值(value-at-risk)等，並分析不同危機期間預測表現。或者使用因子模型縮減變數維度，再進行分量迴歸分析等。其次，可延伸本研究結果，探討總體審慎政策對於房價風險之抑制效果。

附 註

- (註1) 當房價風險與其他總體經濟變數結構關係複雜，或存在非線性關係(包括變數間潛在的非線性與交乘作用等)，傳統VAR線性模型將很難捕捉非線性傳遞管道與尾端風險資訊(相關討論見Duan et al., 2022)。本文目的為模擬房價成長分配與下行風險，針對變數之間與房價如何傳遞影響，並非研究重點。而分量隨機森林模型不受限於傳遞機制假設，更適合本文之研究目的。
- (註2) 台灣房價與總體變數均可能呈現自我相關，抽樣時或可考慮區塊抽樣(block bootstrap)。惟文獻顯示，即便拔靴法(bootstrap)抽樣考慮時間序列相關性，也不一定可以改善模型預測能力。例如，Cordeiro and Neves (2009)使用ETS對時間序列資料進行分解(分為趨勢項、季節性與殘差項)，然後以AR模型配適殘差，再利用篩網拔靴法(Sieve Bootstrap)，重製殘差，建構新的時間序列。整體而言，提升預測效果不佳，在季與月資料預測獲得些許改善，並且篩網拔靴法重建的時間序列，其預測能力往往不如原始序列。Bergmeir et al. (2016)對資料進行Box-Cox變換，將時間序列分解為趨勢、季節性與剩餘部分(remainder)，然後使用移動區塊拔靴法(moving block bootstrap)重製剩餘部分，再將趨勢與季節成分加回，建構新的數列。實證結果發現，因季與年頻率的樣本太短，預測改進效果不顯著，而月頻率序列較長，則可增進預測能力，故該方法較適合月資料。Goehry et al.(2021)考慮時間序列資料相關性，建立隨機森林模型，發現當最適區塊長度(block length)為最小季節成分(minimal seasonality)乘以2或3倍，以區塊拔靴法(block Bootstrap)重建時間序列，可顯著提高預測能力。區塊拔靴法目的是構建既能保持區塊間的依賴性，又能在一定程度上保持區塊獨立性，因此，在具季節性趨勢的資料下，最適區塊長度與季節成分相關。然而本研究序列多為年增率，已去除季節性，或不適用Goehry et al.(2021)的方法。再者，部分隨機森林應用實證研究均未考慮解釋變數自我相關的問題。如Foltas and Pierdzioch (2022)使用時間序列解釋變數，並以分量隨機森林模型與偏斜分配，模擬1970~2017德國各機構預測誤差之條件分量；或者是Lenza et al (2023)使用分量隨機森林，研究通膨分配等，均未考慮解釋變數自我相關。因此，本研究依據文獻做法，未使用區塊抽樣，相關改進研究或可待後續探討。
- (註3) 分量隨機森林模型參數設定方面，節點大小(nodesize)是控制樹(tree)生長的參數，分類問題可設節點數為1，而迴歸問題則設定為5，可產生良好的結果(Díaz-Uriarte and de Andrés, 2006；Goldstein et al., 2011)。本研究設定樣本內節點數為3，樣本外預測節點數為2；決策樹每個分叉的變數數量比例 m 為1/3；至於決策樹數量 k ，改變此參數，結果不會產生很大的差異，設定決策樹數量為1,000顆；另外，增加訓練樣本數 M 可提升模型正確率，但有可能會產生樣本內外表現不一致，降低樣本外預測能力，為了更好的樣本外預測，設定訓練樣本比例為30%。
- (註4) 例如假設房價成長為常態分配，利用ARCH模型估計條件變異數後，再計算下行風險等。
- (註5) PIT指連續隨機變數的累積分配函數 $Z=\Phi(y)$ 會服從均勻 $[0,1]$ 分配。簡單證明如下：累積分配函數 $\Phi(y)$ 介於0與1之間， $0<Z=\Phi(y)<1$ ， $z\in[0,1]$ 。由於 Z 小於 z 的機率為 $P(Z \leq z) = P(\Phi(y) \leq z) = P(\Phi^{-1}(\Phi(y)) \leq \Phi^{-1}(z)) = P(y \leq \Phi^{-1}(z)) = \Phi(\Phi^{-1}(z)) = z$ ，故 $Z \sim U(0,1)$ 。
- (註6) 損失函數 $\rho_{\tau}(u) = u(\tau - 1_{\{u < 0\}})$ 也稱為檢查函數(check function)，由 u 給予權重。其中， $1_{\{u < 0\}}$ 為指示函數。
- (註7) 每次金融危機發生的獨特性及原因可能大相逕庭，有些學者認為以統一標準認定危機具有侷限性，因此Caprio and Klingebiel (2003)除了使用事件觀察，也輔以人為判斷。透過文獻已有共識的金融危機事件，進行研究，將更妥適。
- (註8) 信義房價指數於1991年編製，自2009年起，信義房屋與政治大學商學院共同編製，其主要特點為選取公寓、華廈及電梯大樓，排除預售物件、排除偏差樣本，如工業住宅、屋齡偏高之成交物件等，相當於中古屋市場價格。
- (註9) 房價風險分析以高頻資料為基礎較佳，惟考量部分總體經濟變數為季資料以及台灣房價序列長度取得的限制，統一以季頻率進行分析。
- (註10) 模型其餘變數還包括落後1期的房價成長、失業率、核發建照執照樓板面積與就業人數。

- (註11) 此有可能是因房價下跌，屋主恐面臨資本損失，減低賣房意願，產生類似股票市場的「處置效應(disposition effect)」(林左裕、傅聖元, 2015)，而降低波動度。「處置效應」一詞源於行為財務學，指投資人傾向處置獲利的股票，持有價值下跌的股票。
- (註12) 隨機森林模型參數使用R軟體原始設定，包括針對分類問題，決策樹每個分叉的變數數量為 \sqrt{n} ；至於決策樹數量 k 則為500顆等。
- (註13) Chen and Svirydenka (2021)使用REER等控制變數，分析金融循環對金融(銀行)危機的預測力。

參考文獻

中文文獻

- 江明珠、李政峰、權清全(2011)，「台灣不動產市場的下方風險-以台灣四個縣市為例」，住宅學報，20(2)，頁1-24。
- 李桐豪、江永裕(2009)，「台灣金融危機領先指標之研究」，台灣金融服務業聯合總會委託計畫。
- 林左裕、傅聖元(2015)，「住宅市場處置效應之研究」，都市與計劃，42(3)，頁273-294。
- 林左裕、程于芳(2015)，「影響不動產市場之從眾行為與總體經濟因素之研究」，應用經濟論叢，95，頁61-99。
- 張天惠、朱浩榜(2020)，「台灣金融情勢與經濟預測」，經濟論文叢刊，48(1)，頁77-106。
- 陳明吉、蔡怡純(2022)，「建構台灣房價風險值量化評估模型」，央行季刊，44(4)，頁5-44。
- 黃朝熙、黃裕烈、黃淑君、謝依珊、楊茜文(2014)，「資產價格與信用循環對金融穩定之影響效果分析」，中央銀行季刊，36(4)，頁15-49。
- 蔡怡純、陳明吉(2008)。台北地區不動產價格波動之不對稱性探討。住宅學報，17(2)，頁1-11。
- 蔡怡純、陳明吉(2013)。房價之不對稱均衡調整：門檻誤差修正模型應用。臺灣土地研究，16(1)，頁37-58。

英文文獻

- Abdellatif, A., H. Abdellatef, J. Kanesan, C. O. Chow, J. H. Chuah, and H. M. Gheni (2022), "Improving the Heart Disease Detection and Patients' Survival Using Supervised Infinite Feature Selection and Improved Weighted Random Forest," *IEEE Access*, 10, 67363-67372.
- Adrian, T., A. Deghi, M. Katagiri, S. Shahid, and N. Valckx (2020), "Predicting Downside Risks to House Prices and Macro-Financial Stability," *IMF Working Papers* 2020/011.
- Adrian, T., N. Boyarchenko, and D. Giannone (2019), "Vulnerable Growth," *American Economic Review*, 109(4), 1263-89.
- Azzalini, A., and A. Capitanio (2003), "Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t-Distribution," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2), 367-89.
- Banerjee, R. N., J. Contreras, A. Mehrotra, and F. Zampolli (2020), "Inflation at Risk in Advanced and Emerging Economies," *BIS Working Papers* 883.
- Bergant, K., F. Grigoli, N. J. Hansen, and D. Sandri (2020), "Dampening Global Financial Shocks: Can Macroprudential Regulation Help (More than Capital Controls)," *CEPR Discussion Paper* 14948.
- Bergmeir, C., R. J. Hyndman, and J. M. Benítez (2016), "Bagging Exponential Smoothing Methods using STL Decomposition and

- Box-Cox Transformation,” *International Journal of Forecasting*, 32, 303–312.
- Bush, O., R. Guimarães, and H. Stremmel (2013), “Beyond the Credit Gap: Quantity and Price of Risk. Indicators for Macroprudential Policy,” Unpublished Manuscript.
- Caprio, G., and D. Klingebiel (2003), “Episodes of Systemic and Borderline Financial Crises.” World bank Research Dataset.
- Chen, M. C., and K. Patel (2002), “An Empirical Analysis of Determination of Housing Prices in the Taipei Area,” *Taiwan Economic Reviews*, 30(4), 563–595.
- Chen, S., and K. Sviryzdenka (2021), “Financial Cycles – Early Warning Indicators of Banking Crises?” Available at SSRN: <https://ssrn.com/abstract=3970194>.
- Cheng, H. L., and N. K. Chen (2021), “A study of Financial Cycles and the Macroeconomy in Taiwan,” *Empirical Economics*, 61(4), 1749-1778.
- Chicana, D., and R. Nivin (2022), “Evaluating Growth-at-Risk as A Tool for Monitoring Macro-financial Risks in the Peruvian Economy,” *IHEID Working Papers 07–2021*, Economics Section, The Graduate Institute of International Studies.
- Cordeiro, C., and M. M. Neves (2009), “Forecasting Time Series with Boot.EXPOS Procedure,” *Revstat - Statistical Journal*, 7(2), 135–149.
- Cysouw, M. C. F., B. H. E. Jansen, and T. van de Brug et al (2021), “Machine Learning-based Analysis of [18F]DCFPyL PET Radiomics for Risk Stratification in Primary Prostate Cancer,” *Eur J Nucl Med Mol Imaging*, 48(2), 340–349.
- Díaz-Uriarte, R., and S. Alvarez de Andrés (2006), “Gene Selection and Classification of Microarray Data using Random Forest,” *BMC Bioinformatics*, 7, 3.
- Diebold, F. X., T. A. Gunther, and A. S. Tay (1998), “Evaluating Density Forecasts with Applications to Financial Risk Management,” *International economic Review*, 39(4), 863–883.
- Duan, Y., J. W. Goodell, H. Li, and X. Li (2022), “Assessing Machine Learning for Forecasting Economic Risk: Evidence from an Expanded Chinese Financial Information Set,” *Finance Research Letters*, 46, 102273.
- Dubey, A. (2018), “Feature Selection Using Random Forest,” *Towards Data Science*.
- Fang, Y., P. Xu, J. Yang, and Y. Qin, (2018), “A Quantile Regression Forest Based Method to Predict Drug Response and Assess Prediction Reliability,” *PLOS ONE*, 13(10), 1–16.
- Ferrara, L., and J. Yapi (2022), “Measuring Exchange Rate Risks during Periods of Uncertainty,” *International Economics*, 170, 202–212.
- Foltas, A., and C. Pierdzioch, (2022), “On the Efficiency of German Growth Forecasts: An Empirical Analysis using Quantile Random Forests and Density Forecasts,” *Applied Economics Letters*, 29, 1644–1653.
- Galán, J. E., and M. R. Moreno (2020), “At-risk Measures and Financial Stability,” *Revista de Estabilidad Financiera*, Banco de España, issue Autumn.
- Goehry, B. (2019), “Random Forests for Time-dependent Processes,” HAL-01955331v2.
- Goldstein, B., E. Polley, and F. Briggs (2011), “Random Forests for Genetic Association Studies,” *Statistical Applications in Genetics and Molecular Biology*, 10, 32–32.
- IMF (2019), “Downside Risks to House Prices,” *IMF Global Financial Stability Report*, Chapter 2, April.
- Kim, S., and D. Y. Yang (2009a), “The Impact of Capital Inflows on Asset Prices in Emerging Asian Economies: Is Too Much Money Chasing Too Little Good?,” *Open Economies Review*, 22(2), 293–315.
- Kim, S., and D.Y. Yang (2009b), “Do Capital Inflows Matter to Asset Prices? The Case of Korea,” *Asian Economic Journal*, 23(3),

323–348.

- Koenker, R., and J. A. F. Machado (1999), “Goodness of Fit and Related Inference Processes for Quantile Regression,” *Journal of the American Statistical Association*, 94(448), 1296–1310.
- Kuhn, M., and K. Johnson (2013), *Applied Predictive Modeling*, Springer.
- Lenza, M., I. Moutachaker, and J. Paredes (2023), “Density Forecasts of Inflation: A Quantile Regression Forest Approach,” *ECB Working Paper Series* 2830.
- Masarat, S., S. Sharifian, and H. Taheri (2016), “Modified Parallel Random Forest for Intrusion Detection Systems,” *The Journal of Supercomputing*, 72(6), 2235–2258.
- Mayer, C. J. (2011), “Housing Bubbles: A Survey,” *Annual Review of Economics*, 3, 559–577.
- McFadden, D. (1977), “Quantitative methods for Analysing Travel Behavior of Individuals: Some Recent Developments,” *Cowles Foundation Discussion Papers* 474, Cowles Foundation for Research in Economics, Yale University.
- Meinshausen, N. (2006), “Quantile Regression Forests,” *Journal of Machine Learning Research*, 7, 983–999.
- Nier, E., T. S. Sedik, and T. Mondino (2014), “Gross Private Capital flows To Emerging Markets: Can the Global Financial Cycle Be Tamed?,” *IMF working paper* 14/196.
- Nneji, O., C. Brooks, and C. W. R. Ward (2013), “House Price Dynamics and Their Reaction to Macroeconomic Changes,” *Economic Modelling*, 32, 172–178.
- Rossi, B., and T. Sekhposyan (2017), “Alternative Tests for Correct Specification of Conditional Predictive Densities,” Unpublished.
- Smith, M. H., and G. Smith (2006), “Bubble, Bubble, Where’s the Housing Bubble?” *Brookings Papers on Economic Activity*, 2006(1), 1–67.
- Taguchi, H., S. Pravakar, and G. Nataraj (2015), “Capital Flows and Asset Prices: Empirical Evidence from Emerging and Developing Economies,” *International Economics*, 141, 1–14.
- Zhang, F., and X. Yang (2020), “Improving Land Cover Classification in an Urbanized Coastal Area by Random Forests: The Role of Variable Selection,” *Remote Sensing of Environment*, 251, 112105.

附錄一 傳統Logit迴歸模型設定與預警績效評比

Logit迴歸模型主要用於解釋二元變數的發生，當 $t+h$ 時間點爆發金融危機事件， $F_{t+h} = 1$ ，否則 $F_{t+h} = 0$ ，定義其機率分別為 p 與 $1-p$ ，如下式：

$$p = \text{prob}(F_{t+h} = 1) = \frac{e^{f(X)}}{1+e^{f(X)}}, \quad (1-1)$$

$$1 - p = \text{prob}(F_{t+h} = 0) = \frac{1}{1+e^{f(X)}}, \quad (1-2)$$

預測金融危機的解釋變數(X)參考Adrian et al. (2020)，註¹³ 包括未來1季或1年下行風險值 $\Gamma_{d,t}^D$ ，以及參考Chen and Svirydzenka (2021)，選取REER，體現全球金融風險的傳遞。另外，考慮經濟體質變差導致的金融壓力增加，加入景氣循環($GDP.Cycle_t$)等控制變數(Z_t)。

Logit迴歸式可表示如下：

$$P(X) = \ln \left[\frac{p}{1-p} \right] =$$

$$\beta_0 + \beta_1 \Gamma_{h,t}^D + Z_t + \varepsilon_t, h = 1, 4, \quad (1-3)$$

Logit迴歸樣本內配適度的優劣可由McFadden (1977)的pseudo R^2 判斷。

$$R_{McFadden}^2 = 1 - \frac{\log L_c}{\log L_{null}}, \quad (1-4)$$

其中， L_c 為模型最大概似函數值， L_{null} 代表模型只有截距項的最大概似函數值。

至於樣本外預測，由於房價風險指標係由分量隨機森林模型計算，為求模型一致性，金融危機樣本外預測以隨機森林模型進行。另一方面，模型樣本外預測金融(銀行)危機能力的高低，可透過計算預測值與實際值命中比例來評估。下表將預測情況分為：A正確預測金融危機發生；B為誤判，惟未發生金融危機；C是誤判，卻發生金融危機；D代表正確預測未發生金融危機。

各類預測情況

		真實值	
		$Y_t = 1$	$Y_t = 0$
預測值	$\hat{F}_t = 1$	A (正確預測)	B (誤判，未發生金融危機)
	$\hat{F}_t = 0$	C (誤判，發生金融危機)	D (正確預測未發生金融危機)

定義TPR為在所有金融危機樣本之中，被正確地判斷之比率；而在所有未發生金融危機樣本中，誤判為假警報的比率，稱為FPR。

$$TPR = \frac{A}{A+C}, \quad FPR = \frac{B}{B+D},$$

若模型預期危機將至，但未發生(B情況)，為型I錯誤(Type I error)；或預期危機未發生，但實際上發生(C情況)，則是型II錯誤(Type II error)。良好的金融危機預警除須能正確地預測危機出現時點，亦須減少假警報的發布(B情況)，以提供決策者有效資訊，

俾益及時採取因應措施。

此時，可使用接收者操作特徵曲線面積(AUROC)，評估金融危機預測的正確性。操作特徵曲線(ROC)將1-FPR當做X軸(介於1~0)，正確判斷TPR機率則置於Y軸(介於0~1)；ROC曲線下方的面積稱為AUC (Area Under The Curve)。相關應用可參考Bush et al. (2013)。

AUROC運作方式為當模型預測機率高於門檻值(threshold)，分類為危機事件發生，最適門檻值為使ROC曲線下方面積極大化。AUROC值介於0~1之間：AUROC =1，

為完美分類； $0.5 < \text{AUROC} < 1$ ，預測優於隨機猜測；AUROC= 0.5，與隨機猜測無異。AUROC<0.5，劣於隨機猜測。

此外，亦使用二次機率分數(QPS)評估模型捕捉金融危機的表現。QPS可由以下運算式計算。

$$QPS = 2/O \sum_{t=1}^O (P_t - F_t)^2, \quad (1-5)$$

其中， P_t 為時點 t 的危機預測機率； F_t 為虛擬變數(金融危機發生等於1，否則為0)；O代表樣本外預測數；QPS 介於[0, 1]之間，值越低，預測結果越佳。

附錄二

附表 分量隨機森林模型之變數定義及資料來源

	變數名稱	處理方式	資料來源
房地產市場	1. 核發建照執照樓板面積(平方公尺)-月	2	TEJ
	2. 核發使用執照樓板面積(平方公尺)-月	2	TEJ
	3. 核發建造執照件數-月	2	TEJ
	4. 本國銀行消費者貸款購置住宅貸款年增率-月	1	TEJ
	5. 營造工程總指數(105=100)-月	2	主計總處
總體經濟	6. 領先指標不含趨勢指數-月	2	TEJ
	7. 就業人數(千人)-月	2	TEJ
	8. 勞動參與率(百分比)-月	1	TEJ
	9. 經常性薪資(元)-各行業-月	2	TEJ
	10. 勞動生產力指數-工業(2016=100)-月	1	TEJ
	11. 美元兌新台幣匯率-月	2	TEJ
	12. 實質GDP(2016=100)-台幣(百萬)-季	2	TEJ
	13. 實質國民消費(2016=100)-台幣(百萬)-季	2	TEJ
	14. 台灣名目平均每人GDP-台幣(元)-季	2	TEJ
	15. 核心CPI年增率(百分比)-月	1	TEJ
	16. WPI-前六個變動率(2016=100)(百分比)-月	1	TEJ
	17. 外銷訂單金額總計-美元(百萬元)-月	2	TEJ
	18. 生產量指數-製造業(2016=100)-月	2	TEJ
	19. 出口年增率(美元)-月	1	TEJ
	20. 進口年增率(美元)-月	1	TEJ
	21. 失業率-月	1	TEJ
利率與貨幣供給	22. 商業本票-次級市場-1-30天-月	1	TEJ
	23. 商業本票-次級市場-31-90天-月	1	TEJ
	24. 金融業隔夜拆款利率-加權平均-月	1	TEJ
	25. 短期融通利率-月	1	TEJ
	26. 商業本票-初級市場-91-180天-月	1	TEJ
	27. 可轉讓定期存單-次級市場-1-90天-月	1	TEJ
	28. 可轉讓定期存單-次級市場-91-180天-月	1	TEJ
	29. 五大銀行-放款基準利率-月	1	TEJ
	30. 郵政存簿儲金利率-月	1	TEJ
	31. 存款加權平均利率-本國一般銀行-月	1	TEJ
	32. 貨幣總計數-M1B-月底-台幣(百萬)-月	2	TEJ
	33. 貨幣總計數-M2-月底-台幣(百萬)-月	2	TEJ
	34. 重貼現率-月	1	TEJ
	35. 擔保放款融通利率-月	1	TEJ
	36. 貨幣機構放款與投資(性質別)-月底餘額-月	2	TEJ
	股票市場	37. 殖利率-上市公司-大盤-月	1
38. 股價淨值比-上市公司-大盤-月		1	TEJ
39. 股票成交量週轉率-月		1	TEJ
40. 上市公司家數-月		2	主計總處
41. 上市股票總面額-月		2	主計總處
42. 上市股票總市值-月		2	主計總處
43. 上市股票總成交值-月		2	主計總處
44. 加權平均股價指數-月		2	主計總處
外部因素	45. VIX-月	1	CBOE
	46. BOP-直接投資-資產(百萬美元)-季	2	TEJ
	47. BOP-證券投資-資產(百萬美元)-季	2	TEJ

註：1為取月平均值，轉季資料；2為取年增率，或取月平均年增率，轉季資料。