

# 股價泡沫預測與投資人行為偏誤：粗集理論之運用

## Stock Bubble Prediction and Investor Behavior Bias: Application of Rough Set Theory

葉智丞(Chih-Cheng Yeh)

嶺東科技大學國際企業系副教授

張顯榮(Hia-Jong Teoh)

嶺東科技大學資訊科技系副教授

朱馨蕙\*(Hsing-Hui Chu)

嶺東科技大學會計資訊系助理教授

### 摘要

泡沫在人類的經濟發展史上始終存在，因此，泡沫的形成原因，一直是學者專家們研究的焦點，台灣股市30多年來歷經多次的股價泡沫，本研究從投資人的偏誤行為角度採用粗集理論進行分析。研究期間為1988年7月至2019年12月，採計動能、從眾、反從眾、投資人情緒、週轉率、融資餘額、貨幣供給與領先指標等9項變數。實證結果發現，當以9種變數的訓練樣本為基礎，第13條決策規則，即貨幣供給和融資餘額屬性數同時為1，預測泡沫期間的準確性最高；當以測試樣本為基礎，發現股價泡沫預測正確率為90%；在進一步的研究中，以4種約簡屬性變數為基礎，第7條決策規則，即動能、融資餘額與週轉率屬性數同時為1時有較高的準確性；當再以測試樣本為基礎，發現股價泡沫預測正確率為82.2%。研究結果對判斷目前股價是否泡沫，有很高的參考性。

**關鍵詞:** 股價泡沫; 投資人情緒; 從眾行為; 動能策略; 過度自信

---

\*聯絡作者: 嶺東科技大學會計資訊系。地址：台中市嶺東路一號。  
Tel：04-23892088 分機 3662。Email：teoh@teamail.ltu.edu.tw。

## ABSTRACT

Bubbles have always existed in the history of human economic development. Therefore, the reasons for the formation of bubbles have always been the focus of research by scholars and experts. Taiwan's stock market has experienced multiple stock price bubbles for more than 30 years. This study adopts a rough approach from the perspective of investor bias Set theory for analysis. The study period was from July 1988 to December 2019, and 9 variables including kinetic energy, herd, anti-herd, investor sentiment, turnover rate, financing balance, money supply and leading indicators were adopted. The empirical results found that when the training samples of 9 variables are used as the basis, the 13th decision rule is found, that is, the number of attributes of money supply and financing balance is 1, and the accuracy of predicting the bubble period is the highest; The correct rate of stock price bubble prediction is 90%; in further research, based on 4 reduced attribute variables, it is found that the seventh decision rule, that is, when the number of attributes of kinetic energy, financing balance and turnover rate is 1, there is a higher Accuracy: When based on test samples, it is found that the correct rate of stock price bubble prediction is 82.2%. The research results have a high reference value for judging whether the current stock price is bubble.

**Keywords:** Bubble; Sentiment; Herd behavior; Momentum; Overconfidence

## 壹、緒論

自從1630年代，荷蘭的鬱金香泡沫之後，泡沫在人類的經濟發展史上始終存在。例如，台灣股市投機熱潮在1990年2月份達到12682點的高峰，之後短短8個月，股市崩跌至2482點，跌幅達80.4%，不可謂不重矣。即使在2000年的科技泡沫，跌幅也是很可觀，台股在2000年2月指數10393點創新高點，然後每下愈況，在2001年9月來到最低3411點，跌幅67.1%，災情慘烈。由於股價泡沫結束後，市場經濟往往得經過很長一段時間的休養生息才得以恢復。因此，泡沫的形成原因，歷年來一直是學者專家們矚目的焦點。

那麼，泡沫現象是如何產生的？學者們研究發現認為是與新概念或新產品的出現或技術進步有關，一開始在媒體的報導下漸漸吸引投資人的注意，之後隨著股價逐漸上漲，參與投資風潮的民眾踴躍投入(Kindleberger and Aliber, 2005)，這時當經驗不足的投資人(inexperienced traders)參與市場交易的比例越來越高時，股價隨機性泡沫更是容易發生(Caginalp, Porter and Smith, 2001; Dufwenberg, Lindqvist and Moore, 2005)。顯而易見，假若市場投資人情緒變得高漲，理性成份降低，使得市場上經驗不足的投資人比重逐漸升高，原本遵照理性法則來決定投資決策的投資人交易比重相對較低，這種比例失調很容易讓市場價格失控。

更嚴重的是，經驗不足的投資人很容易相信最近的價格會依然上揚並且持續到未來(Byrne, 2005)，產生所謂的代表性偏誤(resentativeness bias)，進而前仆後繼的投入市場。Fisher and Statman (2002) 觀察在 1990s 年代末期科技泡沫，發現有很多投資人雖然相信

股價已漲過頭，但是他們仍然預期價格將會繼續上升。Ellison and Fudenberg (1993) 認為股票市場是一個人們互動與相互影響的地方，股價的決定在本質上就是一個在密切的社會互動脈絡中，對價格形成共識的一種社會過程。那麼，在社會互動的過程當中，一旦某個樂觀展望獲得共識，形成一定程度的氣氛時，會引發更多投資人加入，當一波又一波的跟隨者紛紛投入其中，泡沫可能就因此形成了。

Hirshleifer and Luo (2001) 指出人們有遵從或仿照他人決策判斷的行為傾向，由於正向回饋交易策略是以買進贏家股賣出輸家股。因此，動能現象很容易會引起從眾行為。Best (2005) 進一步推論動能現象與市場泡沫應該有很密切的關連。Boswijk, Hommes and Manzan(2007)將投資人區分成兩群：基本面交易者(fundamentalists)與市場跟隨者(trend followers)，發現 1990s 末科技泡沫期間幾乎所有投資人是市場跟隨者。Hwang and Salmon (2008) 證實高昂的市場情緒，會提高從眾行為，Siegel (1992)針對 1987 年 10 月美國股市崩盤進行研究，發現投資人情緒與股市崩盤前後的市場報酬之間有高度關連性。特別的是，在市場情緒與動能策略影響下，從眾行為的程度若越來越強，市場偏離真實價值的力量也會越來越大，這時套利者不但沒有矯正市場偏誤的功能，反而讓套利者陷入了市場情緒的旋渦之中，使得市場要回復真實的理性價位越發困難，反映了，市場投資人情緒、動能效應、從眾行為與市場泡沫有者很密切的關聯。

另一種心理偏誤是，投資人很容易因為自己的成功經驗而產生過度自信。Gervais and Odean (2001) 發現投資人會因為其經歷了高報酬而被強化了投資自信； Shiller (2000)認為泡沫的本質就是一種回饋的效果，當股價上升，激勵投資人的購買需求，回饋機制因而產生作用，投資人對未來看法會過於自信，從而增加對股票的需求，使市場趨勢產生正向的效果，將價格進一步地推升。Caginalp, Porter and Smith,(2001)進一步指出，有過度自信的投資人會企圖借錢去買更多的股票，使得在市場泡沫期間時經常與投資人的融資行為有關。國內學者，李春安、羅進水、蘇永裕等 (2006) 發現總體經濟及投資人情緒會顯著影響到追漲或殺跌策略，且三者之間有相互的影響關係。因此投資人的過度自信、融資行為與總體經濟指標，可能也是觀察市場價格是否有泡沫化的重要指標。

檢驗市場價格是否泡沫化?是一種資料探勘的過程。波蘭的 Zdzislaw Pawlak 於 1982 年提出粗集理論(Rough Set Theory)，粗集理論是資料探勘中的一種方法，最初是用來處理有關不確定性、模糊性和粗糙性資料的數學工具。後來經過不斷發展學習，目前粗集理論在人工智慧、資訊系統分析和知識擷取等方面有較佳的實務應用，也可以被用來擷取多維度屬性資料間的關聯性，所以近年來已逐漸被應用於資料探勘領域。另一方面，因應資訊科技的進步，面對龐大的數據庫，粗集理論可依據自身資訊去處理不精確、不確定與不完全的數據，其簡單且易於操作的模式，能夠降低決策難度，提高決策精度。

賴家瑞與溫坤禮(2005)認為粗集方法有以下幾個優點：因為不需要任何預備的或額外的關於數據之資訊或假設，同時可以適用於屬性(qualitative) 及屬量(quantitative)的因數變數分析。能夠由經驗數據中擷取出隱藏於其中的現象，並且以符合自然語法的方式來表達這些決策規則。這些所擷取出的決策規則，代表由數據資料庫中，經過消除多餘或無用的訊息後，再包含於數據資料中。與其他方法相較，模型的決策規則容易理解；

對使用者而言，不需額外對最終模型的參數(parameter)進行解讀。

過去文獻上對於檢定市場股價是否泡沫，有多種檢定泡沫的統計方法，如方差界限法(variance bound)；Kalman 過濾法；規格測試檢定法(Specification Test)；Monte-Carlo 模擬；內生泡沫法 (intrinsic bubbles)等。然而上述檢定方法，大致上是透過股價與股利的關係，來確認是否泡沫已形成，但是股利的發布有資訊的落後性，恐怕無法及時得知當前是否處於股價泡沫期間？從而喪失政策防微杜漸的時效性與預警的功能。有鑑於此，本文嘗試透過粗集理論的實證方法，來進一步檢視股市泡沫與投資人的行為偏誤的關聯性。研究結果發現，18條決策規則中，第13條決策規則，即貨幣供給和融資餘額屬性數同時為1，預測泡沫期間的準確性最高；當以測試樣本為基礎，發現股價泡沫預測正確率為90%；另外，以4種約簡屬性變數為基礎，8條決策規則中，第7條決策規則，即動能、融資餘額與週轉率屬性數同時為1時有較高的準確性；當再以測試樣本為基礎，發現股價泡沫預測正確率為82.2%。

本文的結構如下，第壹部分說明研究的背景與動機，第貳部分粗集理論的介紹與運用；第參部分敘述變數的衡量、資料來源與研究方法；第肆部分呈現實證結果；第伍部分總結。

## 貳、粗集理論

### 一、粗集理論

粗集理論(Rough Set Theory)為波蘭的 Zdzislaw Pawlak 於 1982 年所提出，最初為用來處理有關不確定性、模糊性和粗糙性資料的數學工具。在不斷發展學習後，目前粗集理論能處理各種數據，包括不完整的數據與大量的數據，並且可以從數據庫中來擷取知識，進而產生決策規則。特別的是，也可以用來做屬性值化簡以及尋找資料的隱藏樣本。近年來，逐漸被應用於資料探勘領域。

粗集方法是假設所分析的物件(objects set: 樣本集)本身，即隱含著知識(knowledge)，知識被認為是一種物件的分類能力，它是根據對象的特徵將其分門別類。粗集理論的主要目的在於從各物件(樣本)及其對應的屬性因數(attribute)所組成的資訊系統(information system)中，擷取出足以描述各物件在何種屬性條件下，能夠被分類的規則(rules)組。因此，描述物件分類的規則組即為知識的表達型式。(賴家瑞、溫坤禮，2005)

利用粗集理論處理資料的最大優點，在於不需要處理相關資料的附加資訊，就能容易應用處理，粗集理論主要用來處理下列相關問題領域：

- (1) 知識分析與簡化：能針對複雜的資訊做處理，以獲得簡化並有規則性的知識。
- (2) 不確定資料性驗證：粗集理論對於資料在明確性方面的判定，可透過近似值的概念進行分析。
- (3) 不確定、不嚴謹的知識表達：粗集理論利用近似值與等價類別的概念，可表達模糊或不確定性高的資料。
- (4) 規則學習與決策推導：藉由粗集理論的演算方法，可獲得關聯規則，並用以推

導其他資料，進而進行決策推導。

- (5) 資料屬性相關分析：粗集理論在處理與分析的過程中，皆以資料紀錄的屬性進行條件類別和決策類別的判定，因此對於資料屬性的相關分析研究，極有幫助。
- (6) 近似模型之分類：完成規則組的建立後，可以利用此規則組來對原始資料表進行建模與分類測試，以瞭解此規則組對於資料的解釋能力及適用性，在分類處理上有不錯的成效。

在粗集理論中，知識是以對象（目標、物件）的“分類”能力為基礎，對象包含真實的東西、狀態、抽象的概念、過程、時刻等，而知識是由兩部份所構成：一是從我們感到興趣的領域中，依照各種分類模式所構成的族群，這些可以提供明顯的事實；二是由明顯的知識所衍生出不明確事實的推理能力。最後，將所有相類似對象的集合稱為基本要素（elementary），並據此形成知識的基本單元（atom），基本要素的任何聯集（union）可為明確（crisp），即精確（precise）集合，否則就是約略（rough）—即不精確或模糊集合。

依照此定義，每個約略的集合都會有邊界線（boundary-line）要素（element），換言之，除了明確集合明顯無邊界線要素存在外，其餘集合即是無法將這些要素確定分類為集合的隸屬（member）或其餘集（complement），因此這些位於邊界上的案例便無法將所獲得的知識作適當分類，因此粗集理論可以被視為是一種處理模糊觀念的數學模型。另一方面，在粗集理論方法中，任何模糊的觀念都是藉由一組概念上，可視為精確的「下界近似（lower approximation）」與「上界近似（upper approximation）」來辨別：下界近似係由所有“確定”屬於該觀念的對象所組成；上界近似則包含“可能”屬於該觀念的所有對象所組成，因此，“近似”遂構成了粗集理論的兩個基本操作，如圖 1 所示。

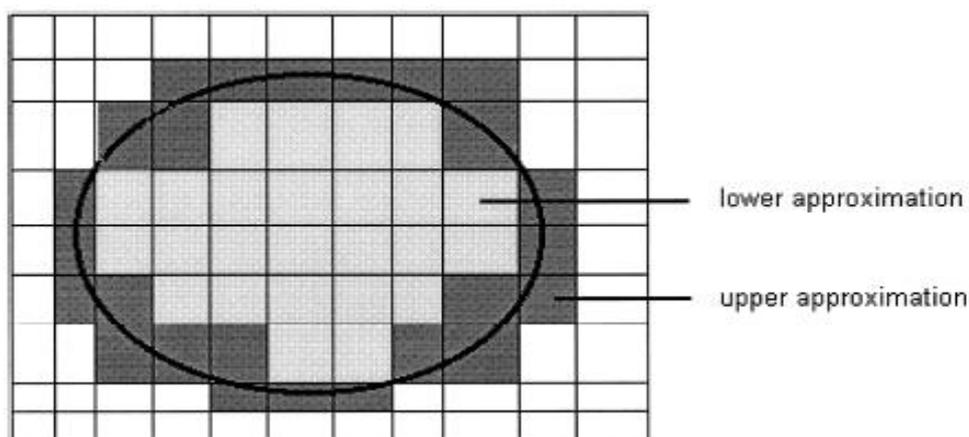


圖 1 下界近似與上界近似圖例（資料來源）

粗集理論的知識表達系統，稱作資訊系統，其數學表示式為： $S = \{U, Q, V, f\}$ ，其中  $U$  為一個有限的非空集合，表示資料庫中所有的記錄， $U = \{x_1, x_2, \dots, x_n\}$ ，稱為全集，可以將知識的概念加以整理； $Q$  表示資料庫中全部的屬性（Attributes）， $Q = \{q_1, q_2, \dots, q_m\}$ ； $V$

表示屬性組成的集合； $V = \cup q \in Q$ ； $f$  表示屬性和記錄的函數，例如  $f(x, q)$  表示記錄  $x$  關於屬性  $q$  的值。在一集中任意一個屬性  $q$ ，若其中某兩筆記錄  $x_i$  與  $x_j$  對於屬性  $q$  的值都相同，即可稱  $x_i$  與  $x_j$  為等價關係 (Equivalence Relation)。

表 1 為粗集理論對於知識擷取之關聯及預測的例子，表中為八筆有關汽車資料的記錄，分別為  $T_1$  至  $T_8$ ，每筆資料紀錄有四個屬性，分別為：Door，Size，Cyl，Mileage。粗集理論利用決策價值類表 (Decision Equivalent Classes) 與條件等值類別 (Condition Equivalent Classes)，來計算關聯式法則之信賴度。

表 1 粗集理論應用實例

| Tuple-Id       | Door | Size | Cyl | Mileage |
|----------------|------|------|-----|---------|
| T <sub>1</sub> | 2    | Camp | 4   | High    |
| T <sub>2</sub> | 4    | Sub  | 6   | Low     |
| T <sub>3</sub> | 4    | Camp | 4   | High    |
| T <sub>4</sub> | 2    | Camp | 6   | Low     |
| T <sub>5</sub> | 4    | Camp | 4   | Low     |
| T <sub>6</sub> | 4    | Camp | 4   | High    |
| T <sub>7</sub> | 4    | Sub  | 6   | Low     |
| T <sub>8</sub> | 2    | Sub  | 6   | Low     |

資料來源：<http://www.cis.drexel.edu/faculty/thu>

那麼粗集理論的資訊系統， $S = \{U, Q, V, f\}$ ，可表示如下：

$$U = \{T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8\}$$

$$Q = \{\text{Door, Size, Cyl, Mileage}\}$$

$$V = \{2, 4, \text{Camp, Sub}, 4, 6, \text{High, Low}\}$$

$$f(T_1, \text{Door})=2; (T_1, \text{Size})=\text{Camp}; f(T_1, \text{Cyl})=4; f(T_1, \text{Mileage})=\text{High}$$

根據上述四個屬性值，其中選擇 Mileage 屬性視為決策屬性，可表示為： $[\text{Mileage}] = \{[\text{Mileage}=\text{High}], [\text{Mileage}=\text{Low}]\}$ ；根據 Mileage 決策屬性值內容可分別表示為： $[\text{Mileage}=\text{Low}] = \{T_2, T_4, T_5, T_7, T_8\}$ ； $[\text{Mileage}=\text{High}] = \{T_1, T_3, T_6\}$ 。而若以 Door，Size，Cyl 三個屬性值的內容分析，則又可表示為 Door, Size, Cyl =  $\{\{T_1\}, \{T_2, T_7\}, \{T_3, T_5, T_6\}, \{T_4\}, \{T_8\}\}$

## 二、屬性選取

在粗集理論中，還有一個特別注意的運算，就是找出縮減屬性與核心屬性。就是將原始表格中的一個或多個屬性刪除後，其剩下的屬性仍然可達到原先屬性的分類效果，而被刪除的屬性稱為不必要的屬性，而剩餘的屬性稱為必要的屬性，使得變數可以縮減，以提高分析能力。

以  $U$  為物件的有限集合， $R$  為屬性的有限集合，而  $U/\text{IND}(R)$  為使用  $R$  來將  $U$  進行

分類的結果。  $U=\{P_1, P_2, P_3, P_4, P_5, P_6\}$  ,  $R=\{Headache, Muscle-pain, Temperature\}$  ,  $U/IND(R)=\{\{P_1\}, \{P_2, P_5\}, \{P_3\}, \{P_4\}, \{P_6\}\}$  。

若現在進行屬性的刪減,例如,考慮從屬性的有限集合  $R$  中拿掉屬性  $Headache$  時,會產生的結果為:  $U/IND(R-\{Headache\})=\{\{P_1\}, \{P_2, P_5\}, \{P_3, P_6\}, \{P_4\}\}$  ,再比較  $U/IND(R)$ 與  $U/IND(R-\{Headache\})$ 時,可以發現  $U/IND(R) \neq U/IND(R-\{Headache\})$  ,也就是說不考慮屬性  $Headache$  時,其對  $U$  所產生的分類,並無法得到利用  $R$  來對  $U$  所產生分類,因此可以得知不能刪除屬性  $Headache$  ,亦即屬性  $Headache$  為必要的屬性。

另一方面,若考慮從屬性的有限集合  $R$  中移除屬性  $Muscle-pain$  時,  $U/IND(R-\{Muscle-pain\})=\{\{P_1\}, \{P_2, P_5\}, \{P_3\}, \{P_4\}, \{P_6\}\}$  ,再比較  $U/IND(R)$ 與  $U/IND(R-\{Muscle-pain\})$  ,可以發現  $U/IND(R)=U/IND(R-\{Muscle-pain\})$  ,這代表可以將屬性  $Muscle-pain$  刪除之。因為將其刪除後,仍然可以達到原先利用屬性的有限集合  $R$  的分類效果,所以屬性  $Muscle-pain$  為不必要的屬性。

### 三、 規則產生

經由上界近似集合與下界近似集合、屬性化簡及資料離散化等處理過程後,就可以得到決策表。利用決策表可以產生決策規則以作為分類的依照,決策規則的集合,以下列形式來表示:

$$a_{ki} \Rightarrow d_j$$

其中  $a_{ki}$  表示屬性  $a_k$  的值為  $i$  ,  $d_j$  表示決策屬性  $d$  的值為  $j$  。

### 四、 決策支援應用決策規則

令  $S = (U, Q, V, f)$  是一個資訊系統。  $\forall P \subseteq Q$  , 公式  $For(P)$  可由數對  $(a, v)$  及邏輯連接符號  $\wedge, \vee, \sim$  來建立,這裡的  $a \in P$  and  $v \in Va$  。一個在  $S$  中的決策規則,可表示成  $\phi \rightarrow \Psi$  , 並讀成如果  $\phi$  則  $\Psi$  , 這裡  $\phi \in For(C)$  ,  $\Psi \in For(D)$  ,  $C$  和  $D$  分別為條件屬性和決策屬性;  $\phi$  和  $\Psi$  分別稱為條件規則和決策規則。數值  $suppS(\phi, \Psi) = card(\|\phi \wedge \Psi\| S)$  將稱作在  $S$  中規則  $\phi \rightarrow \Psi$  的支持(support)。  $\|\phi \wedge \Psi\| S$  的意義是指  $\|\phi\| S \cap \|\Psi\| S$  , 這裡的  $\|\phi\| S$  被定義為  $\{x \in U : \forall a \in C, v \in Va, a(v) = x\}$  , 而  $\|\Psi\| S$  被定義為  $\{x \in U : \forall a \in D, v \in Va, a(v) = x\}$  。

係數  $cerS(\phi, \Psi) = card(\|\phi \wedge \Psi\| S) / card(\|\phi\| S)$  稱為在  $S$  裡的規則  $\phi \rightarrow \Psi$  的信心係數(certainly factor)。明顯地,  $cerS(\phi, \Psi) = 1$  若且唯若  $\phi \rightarrow \Psi$  為真的。若  $cerS(\phi, \Psi) = 1$  , 則  $\phi \rightarrow \Psi$  將稱為確定決策規則(certain decision rule); 假如  $0 < cerS(\phi, \Psi) < 1$  則決策規則將參考成一個不確定的決策規則(uncertain decision rule)。數值  $suppS(\phi, \Psi) / card(U)$  將稱為決策規則  $\phi \rightarrow \Psi$  的強度(strength)。

## 參、研究方法

### 一、研究期間與樣本

本文研究期間為 1988 年 7 月至 2019 年 12 月，共計 378 個月頻率資料。本文以該段期間的上市公司為研究樣本，排除因停止交易或暫時下市導致資料缺漏的公司，所有變數與個股報酬率之資料均為月資料，並依月資料建立動能、投資人情緒、從眾、反從眾、過度自信、貨幣供給、領先指標與融資餘額等變數，相關的變數資料均取自於台灣證券交易所與台灣經濟新報資料庫(TEJ)。

### 二、操作型變數

本研究的應變數為泡沫期間 Y，自變數包括投資人情緒、從眾、反從眾、動能、過度自信、融資餘額、貨幣供給與領先指標等 8 個變數。各項變數定義如下：

(1) 泡沫期間: 本文將泡沫期間變數設為虛擬變數，泡沫期間設為 1，反之為 0。泡沫期間定義為泡沫期間股價高點出現的當月、次月以及前 1 個月，共 3 個月。台灣股市 30 多年來歷經多次的股價泡沫，如 1990s 的投機泡沫，股價指數高點日期為 1990 年 2 月；1997s 的科技泡沫，股價指數高點日期為 1997 年 8 月；2000 年的網路泡沫，股價指數高點日期為 2000 年 2 月；2007 年的美國次級房貸泡沫，股價指數高點日期為 2007 年 10 月。

(2) 投資人的情緒: Baker and Wurgler (2006) 採用六個情緒成份變數，即封閉型基金折價 (CEFD)、股市週轉率 (TURN)、IPO 數量(NIPO)、IPO 上市蜜月期報酬 (RIPO)、權益增資 (S)、及現金股利率 (DP) 等市場變數，依主成份分析法，萃取出投資人每月的情緒指標，計算出每月的 B&W 投資人情緒指數。本文依 Baker and Wurgler (2006) 的方法，以台灣股市上述 6 項變數，建立台灣股市投資人情緒指數。如下式：

$$SENTIMENT_t = 0.4142 * CEFD_{t-1} + 0.6344 * TURN_t - 0.2880 * NIPO_{t-1} + 0.4368 * RIPO_{t-1} - 0.0104 * S_{t-1} - 0.3900 * DP_t$$

(3) 從眾行為與非從眾行為: 在衡量市場投資人的從眾行為，有 Christie and Huang (1995) 以個股報酬及市場報酬率的 cross-sectional standard deviation (CSSD)；Chang, Cheng, and Khorral (2000) 以個股報酬及市場報酬率的 cross-sectional absolute deviation (CSAD)；及 Hwang and Salmon (2004) 以個股  $\beta_{i,t}$  橫斷面標準差來衡量從眾行為。本文從眾行為的衡量指標，主要採用 Hwang and Salmon (2004) 的方法來計算每月的從眾指標；非從眾指標的衡量方面則延續羅進水、李春安 (2009) 的處理方法與定義，以追隨市場報酬為從眾行為以及以不追隨市場報酬為非從眾行為。

(4) 動能策略: 本文動能變數的衡量，乃依 Jegadeesh and Titman (1993) 的方法。衡量方法為在每個月將上市公司過去期 6 個月的累積報酬，由小到大排序分成 10 群，最高的為贏家投資組合，最低的為輸家投資組合，同時買進贏家並賣出輸家投資組合，建立動能策略，採 overlapping holding periods 與等權法計算累積報酬，買入並持有 6 個月

的累積報酬。

(5) 過度自信: Gervais and Odean(2001)認為要驗證過度自信假設，可以利用成交量變化來分析，Statman Thorley and Vorkink(2004) 發現市場總體週轉率與落後期的市場報酬呈顯著的正向關係，顯示投資人在經歷了市場的良好報酬後，過度自信程度上升，以致做出更積極的交易行為，使得股市成交量上升。Caginalp, Porter and Smith,(2001)指出有過度自信的投資人會企圖借錢去買更多的股票。因此，本文以每月的市場週轉率以及融資餘額來衡量投資人的過度自信程度。

(6)總體經濟變數: 李春安、羅進水、蘇永裕等 (2006) 發現總體經濟也會顯著影響到投資人的追漲或殺跌策略，因此，本文以每月的市場總值、貨幣供給與領先指標來衡量總體經濟變數。

## 肆、實證結果

### 一、各項變數屬性與極大極小值

表 2 為動能、從眾、反從眾、投資人情緒、週轉率、融資餘額、貨幣供給與領先指標等 9 個變數的屬性，以及在 1988 年 7 月至 2019 年 12 月的極大與極小值。其中動能、從眾、反從眾、投資人情緒與貨幣供給的最小值都小於 0，週轉率、融資餘額與領先指標的最小值大於 0。從眾的值介於-0.902~0.555 之間，反從眾的值介於-0.81~0.328 之間。

表 2 變數的資訊集

| 屬性因子 | 變數名稱 | 變數英文名稱       | 屬性資訊  | 最小值     | 最大值       |
|------|------|--------------|-------|---------|-----------|
| 1    | 動能   | Momentum     | 數值，連續 | -9.823  | 10.831    |
| 2    | 從眾   | Follower     | 數值，連續 | -0.902  | 0.555     |
| 3    | 反從眾  | Against      | 數值，連續 | -0.81   | 0.328     |
| 4    | 情緒   | Emotion      | 數值，連續 | -18.911 | 97.884    |
| 5    | 貨幣供給 | M1B          | 數值，連續 | -10.28  | 32.05     |
| 6    | 融資餘額 | Mar          | 數值，連續 | 6462715 | 579802444 |
| 7    | 週轉率  | Tur          | 數值，連續 | 7.12    | 78.99     |
| 8    | 領先指標 | Lead         | 數值，連續 | 46.3    | 107.5     |
| 9    | 市值   | Market value | 數值，連續 | 0.994   | 3.745     |

### 二、變數的離散判別分類

利用粗集理論的分析方法，首先要把連續的資料轉換成離散資料，才能進行分析。離散化的方法本研究採用 MEPA 判別法(minimize entropy principle approach)，此法可在大量高維的連續數據中，計算資訊含量以便衡量資料的落點，再根據每一筆資訊含量在資料落點的左右方，再決定離散屬性。變數的離散判別分類如表 3，例如第 1 筆動能數

據 1.893，超過 0.51，則屬性數被歸類為 1；第 2 筆動能數據 0.823，超過 0.7，則屬性數被歸類為 0；第 4 筆動能數據 4.632，屬性數直接被歸類為 1；第 5 筆動能數據 2.594，超過 0.65，則屬性數被歸類為 1。本研究透過上述的動作，把 9 項變數的原始值資料，判別每一變數每一筆的屬性數。

表 4 則為變數的屬性數判別結果，如動能的第 1~5 比數據，屬性數被判為 1、0、1、1、1；從眾的第 1~5 比數據，屬性數被判為 1、1、1、1、1。

表 3 變數的離散判別分類

| no. | 屬性<br>原始值 | 動能   |      | 從眾      |     |     | ... | 市值    |     |     |
|-----|-----------|------|------|---------|-----|-----|-----|-------|-----|-----|
|     |           | 屬性數  |      | 原始<br>值 | 屬性數 |     |     | 原始值   | 屬性數 |     |
|     |           | 0    | 1    |         | 0   | 1   |     |       | 0   | 1   |
| 1   | 1.893     | 0.49 | 0.51 | 0.415   | 0   | 1   | ... | 1.852 | 1   | 0   |
| 2   | 0.823     | 0.7  | 0.3  | 0.419   | 0   | 1   | ... | 2.252 | 1   | 0   |
| 3   | 1.864     | 0.49 | 0.51 | 0.434   | 0   | 1   | ... | 2.557 | 1   | 0   |
| 4   | 4.632     | 0    | 1    | 0.439   | 0   | 1   | ... | 1.918 | 1   | 0   |
| 5   | 2.594     | 0.35 | 0.65 | 0.471   | 0   | 1   | ... | 2.169 | 1   | 0   |
| ... | ...       | ...  | ...  | ...     | ... | ... | ... | ...   | ... | ... |

表 4 變數的屬性數判別結果

| 樣本編號 | 動能  | 從眾  | .... | 市值  |
|------|-----|-----|------|-----|
| 1    | 1   | 1   |      | 0   |
| 2    | 0   | 1   |      | 0   |
| 3    | 1   | 1   |      | 0   |
| 4    | 1   | 1   |      | 0   |
| 5    | 1   | 1   |      | 0   |
| ...  | ... | ... | ...  | ... |

#### 四、9項變數分類規則模型之實證結果

當經由上界近似集合與下界近似集合及資料離散化等處理過程後，就可以得到決策表。利用決策表可以產生決策規則的集合，見表 5。為了測試決策規則的效果，將 378 筆資料以隨機的方式抽取，分為 277 筆資料的訓練樣本與 101 筆資料的測試樣本。

實證結果顯示，訓練樣本產生 13 條規則。第一條規則表示符合最多的樣本數，共有 73 筆樣本符合第一條規則。本研究最關心的泡沫期間，設定為  $Y1=1$ 。在表 5 中，13 條決策規則中出現  $Y1=1$ ，為第 4、9、10 與 13 條規則。其中，第 4 條規則在 29 筆樣本中出現 3 筆  $Y1=1$ ，預測正確率為 10%；第 9 條規則在 4 筆樣本中出現 1 筆  $Y1=1$ ，預測正確率為 25%；第 10 條規則在 4 筆樣本中出現 2 筆  $Y1=1$ ，預測正確率為 50%；第 13 條規則在 1 筆樣本中出現 1 筆  $Y1=1$ ，預測正確率為 100%。顯然第 13 條決策規則，即貨幣供給(m1b)和融資餘額(mar)屬性數同時為 1，預測泡沫期間的準確性最高。其次，第 10 條決策規則，動能(mom)、融資餘額(mar)與週轉率(tur) 屬性數同時為 1 以及貨幣供給(m1b) 屬性數為 0 也有很高的準確性。

表 5 以 9 項變數推演出的決策規則

| (1-13) | Match | Decision rules  |
|--------|-------|---|
| 1      | 73    | $(mom=0)\&(s=0)\Rightarrow\{y1=\{0[73]\}\}$                         |
| 2      | 66    | $(s=0)\&(mar=0)\Rightarrow\{y1=\{0[66]\}\}$                         |
| 3      | 46    | $(tur=0)\Rightarrow\{y1=\{0[46]\}\}$                                |
| 4      | 29    | $(s=1)\&(m1b=0)\&(mar=0)\&(tur=1)\Rightarrow\{y1=\{0[26],1[3]\}\}$  |
| 5      | 19    | $(mom=0)\&(mar=1)\Rightarrow\{y1=\{0[19]\}\}$                       |
| 6      | 18    | $(m1b=1)\&(mar=0)\Rightarrow\{y1=\{0[18]\}\}$                       |
| 7      | 10    | $(mom=0)\&(m1b=1)\Rightarrow\{y1=\{0[10]\}\}$                       |
| 8      | 5     | $(s=1)\&(m1b=1)\Rightarrow\{y1=\{0[5]\}\}$                          |
| 9      | 4     | $(mom=1)\&(s=1)\&(m1b=0)\Rightarrow\{y1=\{0[3],1[1]\}\}$            |
| 10     | 4     | $(mom=1)\&(m1b=0)\&(mar=1)\&(tur=1)\Rightarrow\{y1=\{0[2],1[2]\}\}$ |
| 11     | 1     | $(tra=0)\Rightarrow\{y1=\{0[1]\}\}$                                 |
| 12     | 1     | $(s=1)\&(mar=1)\Rightarrow\{y1=\{0[1]\}\}$                          |
| 13     | 1     | $(m1b=1)\&(mar=1)\Rightarrow\{y1=\{1[1]\}\}$                        |

接著，我們再將 101 筆測試樣本依照上述推演出的規則進行分類。表 6 為股價泡沫的預測準確率，表中顯示，當實際值為沒有股價泡沫時，95 筆預測值中，有 86 筆預測正確，9 筆預測錯誤，正確率為 90.5%；當實際值為有股價泡沫時，6 筆預測值中，有 6 筆預測正確，正確率為 100%。整體而言，總共有 101 筆資料 91 筆預測正確，有 9

筆預測錯誤，正確率為 90%。

表 6 股價泡沫的預測準確率

|     |            | 預測值        |           | 實際<br>筆數 | 個別<br>準確率 |
|-----|------------|------------|-----------|----------|-----------|
|     |            | 沒有股價<br>泡沫 | 有股價<br>泡沫 |          |           |
| 實際值 | 沒有股價<br>泡沫 | 86         | 9         | 95       | 0.905     |
|     | 有股價<br>泡沫  | 0          | 6         | 6        | 1         |

測試樣本:101  
準確率: 90.0

### 五、變數屬性縮減後分類規則模型之實證結果

粗集理論的分析方法中，有一個特點，可以找出多餘的屬性，從而刪除多餘的屬性，最後以約簡屬性與核心屬性來進行分析。此部份的運算可藉由辨識矩陣來加以判別，並且利用布林運算求取出約簡屬性與核心屬性，使得變數可以縮減，以提高模型的分析能力。

本研究變數，原本有動能、從眾、反從眾、投資人情緒、週轉率、融資餘額、貨幣供給與領先指標等 9 個變數，經由布林運算求取出 4 種約簡屬性變數，分別為動能、投資人情緒、週轉率與融資餘額等。再以表 5 的方法產生決策規則的集合。相同的，為了測試決策規則的效果，將 378 筆資料以隨機的方式抽取，分 277 筆資料為訓練樣本與 101 筆資料為測試樣本。

當以動能、投資人情緒、週轉率與融資餘額等 4 種約簡屬性變數進行實證，結果顯示訓練樣本產生 8 條規則。第一條規則表示符合最多的樣本數，共有 73 筆樣本符合第一條規則。泡沫期間依然設定為  $Y1=1$ ，表 7 中，8 條決策規則中出現  $Y1=1$ ，為第 4、6 與 7 條規則。其中，第 4 條規則在 34 筆樣本中出現 3 筆  $Y1=1$ ，預測正確率為 8.8%；第 6 條規則在 5 筆樣本中出現 1 筆  $Y1=1$ ，預測正確率為 20%；第 7 條規則在 5 筆樣本中出現 3 筆  $Y1=1$ ，預測正確率為 60%。顯然第 7 條決策規則，即動能(mom)、融資餘額(mar)與週轉率(tur) 屬性數同時為 1 有較高的準確性。

接著，我們再將 101 筆測試樣本依照上述推演出的規則進行分類。表 8 為縮減變數對股價泡沫的預測準確率，表中顯示，當實際值為沒有股價泡沫時，95 筆預測值中，有 81 筆預測正確，14 筆預測錯誤，正確率為 85.2%；當實際值為有股價泡沫時，6 筆預測值中，有 2 筆預測正確，4 筆預測錯誤，正確率為 33.3%。整體而言，總共有 101 筆資料 83 筆預測正確，有 18 筆預測錯誤，正確率為 82.2%。

表 7 以 4 項縮減變數推演出的決策規則

| (1-8) | Match | Decision rules                                 |
|-------|-------|--|
| 1     | 73    | (mom=0)&(emotion=0)=>(y1={0[73]})              |
| 2     | 66    | (emotion=0)&(mar=0)=>(y1={0[66]})              |
| 3     | 46    | (tur=0)=>(y1={0[46]})                          |
| 4     | 34    | (emotion=1)&(mar=0)&(tur=1)=>(y1={0[31],1[3]}) |
| 5     | 19    | (mom=0)&(mar=1)=>(y1={0[19]})                  |
| 6     | 5     | (mom=1)&(emotion=1)=>(y1={0[4],1[1]})          |
| 7     | 5     | (mom=1)&(mar=1)&(tur=1)=>(y1={0[2],1[3]})      |
| 8     | 1     | (emotion=1)&(mar=1)=>(y1={0[1]})               |

表 8 縮減變數對股價泡沫的預測準確率

|         |            | 預測值        |           |           | 準確率   |
|---------|------------|------------|-----------|-----------|-------|
|         |            | 沒有股價<br>泡沫 | 有股價<br>泡沫 | 實際<br>觀察值 |       |
| 實際<br>值 | 沒有股價<br>泡沫 | 81         | 14        | 95        | 0.852 |
|         | 有股價<br>泡沫  | 4          | 2         | 6         | 0.333 |

測試樣本:101

準確率:0.822

## 伍、結論

泡沫在人類的經濟發展史上始終存在，因此，泡沫的形成原因，一直是學者專家們研究的焦點。過去傳統財務學的兩大理論效率市場假說(EMH)與資本資產定價模型(CAPM)，支配了現代財務學領域達數十年，主要的假說依據是人是理性的，既不受情緒影響也沒有心理干擾。但是，近年來行為財務學興起，假說人並非是理性的，學者們陸陸續續發現了投資人有很多偏誤的心理與行為，包括投資人情緒、從眾行為與過度自信等。這些投資人的偏誤行為是否可以被用來預測股價泡沫?是個有趣的研究議題。

我們的研究期間為1988年7月至2019年12月，共計378個月頻率資料，採計動能、從眾、反從眾、投資人情緒、週轉率、融資餘額、貨幣供給與領先指標等9個變數，嘗試以粗集理論的方法來進行預測，為了測試決策規則的效果，將378筆資料以隨機的方式抽取，分為277筆資料的訓練樣本與101筆資料的測試樣本。

實證結果如下: (1)當9種變數的訓練樣本，經由上下界近似集合及資料離散化等處理

過程後得到決策表，再利用決策表產生13條決策規則。發現第13條規則，即貨幣供給和融資餘額屬性數同時為1，預測泡沫期間的準確性最高；當再將101筆測試樣本依照上述推演出的規則進行分類，發現股價泡沫預測正確率為90%。(2)當經由布林運算求取出4種約簡屬性變數，分別為動能、投資人情緒、週轉率與融資餘額等，再利用訓練樣本產生8條決策規則，其中第7條決策規則，即動能、融資餘額與週轉率屬性數同時為1有較高的準確性；再將101筆測試樣本，依照上述推演出的規則進行分類，發現股價泡沫預測正確率為82.2%。

台灣股市過去30多年來歷經多次的股價泡沫，未來是否會再次發生?或者目前已經是在發生中?相信會有很多投資大眾、學者專家與政府單位非常關心。本文的研究結果，對判斷目前股價是否泡沫，有很高的預測正確率，值得上述人士參考。

### 參考文獻

- 李春安、羅進水、蘇永裕 (2006)「動能策略報酬、投資人情緒與景氣循環之研究」，*財務金融學刊*，第 14 卷第 2 期，73-109。
- 羅進水、李春安 (2009)「從眾與非從眾行為對市場報酬衝擊之研究」，*管理評論*，第 28 卷第 3 期，21-41。
- 賴家瑞、溫坤禮 (2005)「粗糙集方法於授信審核之應用」，*計量管理期刊*，第 2 卷第 1 期，69-78。
- Baker, M. and J. Wurgler (2006). "Investor Sentiment and the Cross-Section of Stock Returns," *Journal of Finance*, Vol. 61, 1645-1680.
- Best, K. (2005). "Celebrity.com: Internet Finance and Frenzy at the Millennium," *Consumption, Markets and Culture*, 8.
- Boswijk, H.P., C.H. Hommes and S. Manzan (2007). "Behavioral Heterogeneity in Stock Prices," *Journal of Economic Dynamics & Control*, Vol. 31, 1938-1970.
- Byrne, K. (2005). "How Do Consumers Evaluate Risk in Financial Products?" *Journal of Financial Services Marketing*, Vol. 10, 21-36.
- Caginalp, G., D. Porter and V. Smith (2001). "Financial Bubbles: Excess Cash, Momentum, and Incomplete Information," *Journal of Psychology and Financial Markets*, Vol. 2, 80-99.
- Chang, J., W. Cheng and A. Khorana (2000), "An Examination of Herding Behavior in Equity Markets: An International Perspective," *Journal of Banking & Finance*, Vol. 24, 1651-1679.
- Christie, W. G. and R. D. Huang (1995). "Following the Pied Piper: Do Individual Returns Herd around the Market?" *Financial Analysts Journal*, Vol. 51, 31-37.
- Dufwenberg, M. ; T. Lindqvist and E. Moore (2005). "Bubbles and Experience: An Experiment," *American Economic Review*, Vol. 95, 1731-1737.
- Ellison, G. and D. Fudenberg (1993). "Rules of Thumb for Social Learning," *Journal of*

- Political Economy*, Vol. 104, 612-643.
- Fisher, K.L. and M. Statman (2002). "Blowing Bubbles," *Journal of Psychology and Financial Markets*, Vol. 3, 53-65.
- Gervais, S. and T. Odean (2001). "Learning to be Overconfident," *Review of Financial Studies*, Vol. 14, 1- 27.
- Hirshleifer, D. and G.Y. Luo (2001). "On the Survival of Overconfident Traders in a Competitive Security Market," *Journal of Financial Markets*, Vol. 4, 73-84.
- Hwang, S. and M. Salmon (2004). "Market Stress and Herding," *Journal of Empirical Finance*, Vol. 11, 585-616.
- Jegadeesh, N. and S. Titman (1993). "Returns to Buying Winners and Selling Losers: Implication for Market Efficiency," *Journal of Finance*, Vol. 48, 65-91.
- Kindleberger, C.P. and R.Z. Aliber (2005). "Manias, Panics, and Crashes: A History of Financial Crises." (5th Ed), *Palgrave Macmillan*.
- Nofsinger, J. (2003). "Social Mood and Financial Economics," *Journal of Behavioral Finance*. Vol. 6, 144-160.
- Shiller, R. J. (2000), "Irrational Exuberance," Princeton: Princeton University Press.
- Siegel, J. (1992), "Equity Risk Premia, Corporate Profit Forecasts, and Investor Sentiment around the Stock Market Crash of October 1987," *Journal of Business*, Vol. 65, 557-570.
- Statman, M., S. Thorley and K. Vorkink (2004). "Investor Overconfidence and Trading Volume," *Review of Financial Studies*, Vol. 19, 1531-1565.