

國立政治大學金融學系

碩士學位論文

基於打分法之多因子選股策略建構：

中國 A 股市場實證分析

**Constructing a Multi-factor Investing Strategy Based on Scoring**

**Method: An Empirical Analysis of China A-share market**



指導教授：林建秀 博士

研究生：胡維維 撰

中華民國一百一十年六月

## 謝辭

時光荏苒，在政大的研究所生涯即將畫上句點，在此我想感謝各位師長以及同學給予的幫助與關懷，讓我在異鄉求學之路上感受到有如家人般的親切。

首先，我想感謝指導教授林建秀老師，感謝老師在論文研究方向以及思路提供的寶貴建議，當研究過程遇到瓶頸時，老師您總是耐心地與我討論解決方案，每次討論結束後聽到老師鼓勵的話語，總能讓我充滿幹勁。同時也非常感謝兩位口試委員，政大金融系廖四郎教授與臺北大學經濟系程智男教授，兩位在口試當天針對論文不足之處提供指導意見的同時，也傳授了一些實務上的寶貴經驗，相信這些經驗能在本人未來的職涯上提供幫助。

最後也非常感謝政大以及政大的各位師長、學長姐、同學。感謝學校在疫情期間及時提供遠程授課方案協助我們順利完成學業，感謝系上各位老師的教導讓我在政大金融所兩年的學業中學習到了許多金融領域的新知識，尤其是金融工程以及金融科技相關的課程，在教學內容上都緊跟金融趨勢，令人受益匪淺。同時感謝金融系同儕在兩年裏的陪伴以及勉勵，希望大家未來都有美好的前程。

最後的最後也想說聲抱歉，因為疫情以及一些個人原因選擇提前返家，非常遺憾沒來得及與政大的同學還有老師們做正式的告別，兩年裏非常感謝大家的陪伴以及鼓勵，期待下次相見！

# 基於打分法之多因子選股策略建構：

## 中國 A 股市場實證分析

國立政治大學金融系

### 摘要

本研究以 2006 年至 2020 年的中國 A 股市場為對象構建打分法下的多因子選股策略，首先以中國 A 股市場股災發生前的數據為樣本，在規模類、估值類、質量類、成長類等常見的 10 大類因子中選擇了 18 個因子對其有效性進行識別，經過橫斷面 t 檢定、IC-IR 檢定法、相關性檢定、因子分層回測一系列步驟對各類因子進行有效性、穩健型、獨立性以及一致性進行分析，最終篩選出 6 個風格因子，分別為總市值對數、賬面市值比、每股收益成長率、非流動性指標、過去一個月股價報酬、過去一個月 BETA 值。

識別出有效因子後，再基於打分法對六因子進行隨機組合，發現在單因子的基礎上不斷加入其餘因子可以在一定程度上提升選股效果，樣本內績效最佳的投資策略為由六因子共同排序打分所構建的策略，投組的夏普比率與信息比率分別達到了 1.1767 以及 1.3061。為檢驗六因子打分法的穩定性，將該策略應用於股災後的數據進行回測檢驗，雖然結果顯示投組在整體上的表現仍優於基準指數滬深 300，但在 2017-2018 年間的超額收益有所下降，最後透過投組歸因分析發現由於因子風格轉變以及行業輪動等原因造成策略發生階段性的失效。

**關鍵詞：**中國 A 股市場、因子有效性、多因子模型、歸因分析

# **Constructing a Multi-factor Investing Strategy Based on Scoring Method : An Empirical Analysis of China A- share market**

Department of Money and Banking, National Chengchi University

## **Abstract**

This paper takes China's A-share market from 2006 to 2020 as the object to construct a multi-factor investing strategy under the scoring method. First, taking the data before the 2015 China stock market crash as the sample, selected 18 factors from 10 common categories, such as size, valuation, quality, and growth to identify their effectiveness. Through a series of steps of cross-sectional T-test, IC-IR test, correlation test, and factor stratification test, the effectiveness, robustness, independence, and consistency of various factors were analyzed, and finally, 6 style factors were screened out: logarithm of total market capitalization, book-to-market ratio, EPS growth rate, Amihud ILLIQ, a stock price return of the past month, BETA of the past month.

After the effective factors are identified, these six factors are randomly combined based on the scoring method. It is found that adding other factors based on a single factor can improve the effect of stock selection to a certain extent. The investment strategy with the best performance in the sample is the six-factor model, the Sharpe ratio and Information ratio of the portfolio reached 1.1767 and 1.3061. This strategy was applied to the market after the 2015 crash to test the stability of the six-factor scoring method. Although the results showed that the portfolio's overall performance was still better than the benchmark index CSI 300, the excess return declined from 2017 to 2018. Finally, through the attribution analysis, we can find that the strategy fails periodically due to factor style and industry rotation change.

**Keywords: China A-share market, factor validity, multi-factor model, attribution analysis**

# 目次

第一章 緒論.....	1
1.1 研究動機：.....	1
1.2 研究目的：.....	2
1.3 論文結構安排：.....	3
第二章、文獻回顧.....	4
2.1 多因子理論：.....	4
2.2 風格因子文獻探討：.....	4
第三章 多因子選股方法.....	11
3.1 風格因子的選取：.....	11
3.2 去極值與標準化：.....	11
3.3 有效因子識別：.....	12
3.4 多因子選股模型介紹：.....	18
第四章 實證分析.....	21
4.1 實證數據描述：.....	21
4.2 有效因子識別：.....	22
4.3 多因子隨機組合績效.....	33
4.4 樣本外多因子組合績效.....	38
第五章 投資組合歸因分析.....	39
5.1 因子風格的轉變：.....	39
5.2 因子歸因分析：.....	41
5.3 Brinson 歸因分析：.....	43
第六章 結論與展望.....	46
6.1 研究結論：.....	46
6.2 論文不足與展望：.....	48
參考文獻.....	50

## 表目錄

表 3-1：風格因子表 .....	11
表 3-2：斯皮爾曼相關係數示意表 .....	15
表 3-3：交互式打分法示意表 .....	20
表 4-1：風格因子有效性檢驗結果 .....	22
表 4-2：規模因子分組回測績效 .....	25
表 4-3：估值因子分組回測績效 .....	27
表 4-4：成長因子分組回測績效 .....	28
表 4-5：流動性因子分組回測績效 .....	29
表 4-8：單因子隨機組合投組績效 .....	33
表 4-9：雙因子隨機組合投組績效 .....	33
表 4-10：三因子隨機組合投組績效 .....	34
表 4-11：四因子隨機組合投組績效 .....	35
表 4-12：五因子隨機組合投組績效 .....	36
表 4-13：六因子隨機組合投組績效 .....	37
表 4-14：樣本內外打分法回測績效 .....	38
表 5-1：投組重倉行業 .....	44
表 5-2：基準指數沖倉行業 .....	44

## 圖目錄

圖 1：論文結構安排.....	3
圖 3-1：排序打分法三模型 .....	18
圖 4-1：風格因子相關性矩陣 .....	24
圖 4-2：規模因子分組回測累積淨值 .....	25
圖 4-3：估值因子分組回測累積淨值 .....	26
圖 4-4：成長因子分組回測累積淨值 .....	28
圖 4-5：流動性因子分組回測累積淨值 .....	29
圖 4-6：短期反轉因子分組回測累積淨值 .....	30
圖 4-7：波動度因子分組回測累積淨值 .....	32
圖 4-7：波動度因子分組回測累積淨值 .....	32
圖 4-9：樣本外打分法回測累計淨值 .....	38
圖 4-8：樣本內打分法回測累計淨值 .....	38
圖 4-10：規模因子分組回測累積淨值（樣本外） .....	39
圖 4-11：估值因子分組回測累積淨值（樣本外） .....	39
圖 4-14：短期反轉因子分組回測累積淨值（樣本外） .....	40
圖 4-15：流動性因子分組回測累積淨值（樣本外） .....	40
圖 4-12：成長因子分組回測累積淨值（樣本外） .....	40
圖 4-13：波動度因子分組回測累積淨值（樣本外） .....	40
圖 5-1：因子歸因分析結果 .....	42
圖 5-2：Brinson 歸因分析結果.....	44

# 第一章 緒論

## 1.1 研究動機：

在眾多的資產種類中，股票投資為大多數人主要的投資管道，然而在廣大的股票市場中，要如何從中選取正確的投資標的，擬定適當投資策略以獲取報酬是投資學的一大學問。在股票投資中，我們通常使用某種或是多種指標對股票池進行篩選，進而將這些被選出來的股票構建成投資組合，這些用於選股的指標一般被稱為因子，用於描述某類股票所共有的特徵，如果基於該因子所構建的投組能帶來穩定的收益，則說明該因子具有有效的選股能力。特徵因子一直以來都是構建股票投資組合的重要基石，從美國學者威廉·夏普與 1964 年提出資本資產模型定價理論（CAPM）以來，學者們就致力於找到可以解釋資產報酬的有效因子，同時在對因子進行探勘的過程中也逐漸發現了一些市場異象無法被傳統資產定價模型所解釋，例如：Banz（1981）提出的規模效應，Rosenberg, Reid and Lanstein（1985）提出的價值股效應，Amihud（1986）提出的“非流動性風險溢酬”等。自從 1990 年代 Fama French 三因子模型誕生以來，以多因子為媒介進行選股的模型層出不窮，資本市場也出現了以“seeking alpha“，”Smart Beta“為目標的因子投資組合的策略，這些策略目前也被廣泛應用於基金市場。

多因子模型的構建過程可以大致分為兩個階段：一是有效因子的選取，即判斷具有何種特徵的股票的可獲得最佳的報酬率，二是多因子的合成，即通過組合多個有效因子構建績效表現最佳的投資組合，其中涉及多因子合成模型以及各個因子以及個股之間的權重配置。



在現今大數據以及科技化的時代背景下，各類因子投資策略上往往能精準發掘過往投資組合內推動回報的因素，廣大投資人也可以通過瞭解因子的運作方式，在降低風險的同時賺取投資回報。本文旨在模擬投資人在 A 股市場上的投資情形，基於過往文獻對於有效選股因子的探討，試著從規模類、估值類、質量類、成長類、杠杆類、股票交易流動性類、動量類、波動度類、股利分配類以及分析師情緒類這 10 大類因子中找到能幫助投資人提升選股收益率的因子，再通過隨機組合的方式測試多因子組合相對於單因子的表現，找到可以使投組績效達到最大的多因子投資策略，並希望能在實證研究中為投資人帶來投資上的啟示及建議。

## 1.2 研究目的：

本研究的目的是在於從規模類、估值類、質量類、成長類等 10 大類因子中尋找適用於中國 A 股市場的有效因子，並通過多因子隨機組合的方式找到最佳的投資策略，使得選出的有效因子所組建的投組能為投資人創造額外的投資報酬，且其報酬表現經由風險調整後仍能優於市場的績效表現。同時，通過比較中國 2015 年股災發生前後市場上投組的績效表現來探討各因子表現的持續性以及有效性。

研究目的如下：

1. 過往文獻中提到的有效選股因子是否適用於中國 A 股市場，這些有效因子的選股表現如何？
2. 通過多因子的組合方式是否可以在提升投組風格多元性的同時提升因子的選股績效？
3. 通過觀察中國股災前後有效因子表現，來檢視因子的有效性及持續性，判斷因子是否在風格上是否發生了改變以及探討可能發生的緣由為何？

### 1.3 論文結構安排：

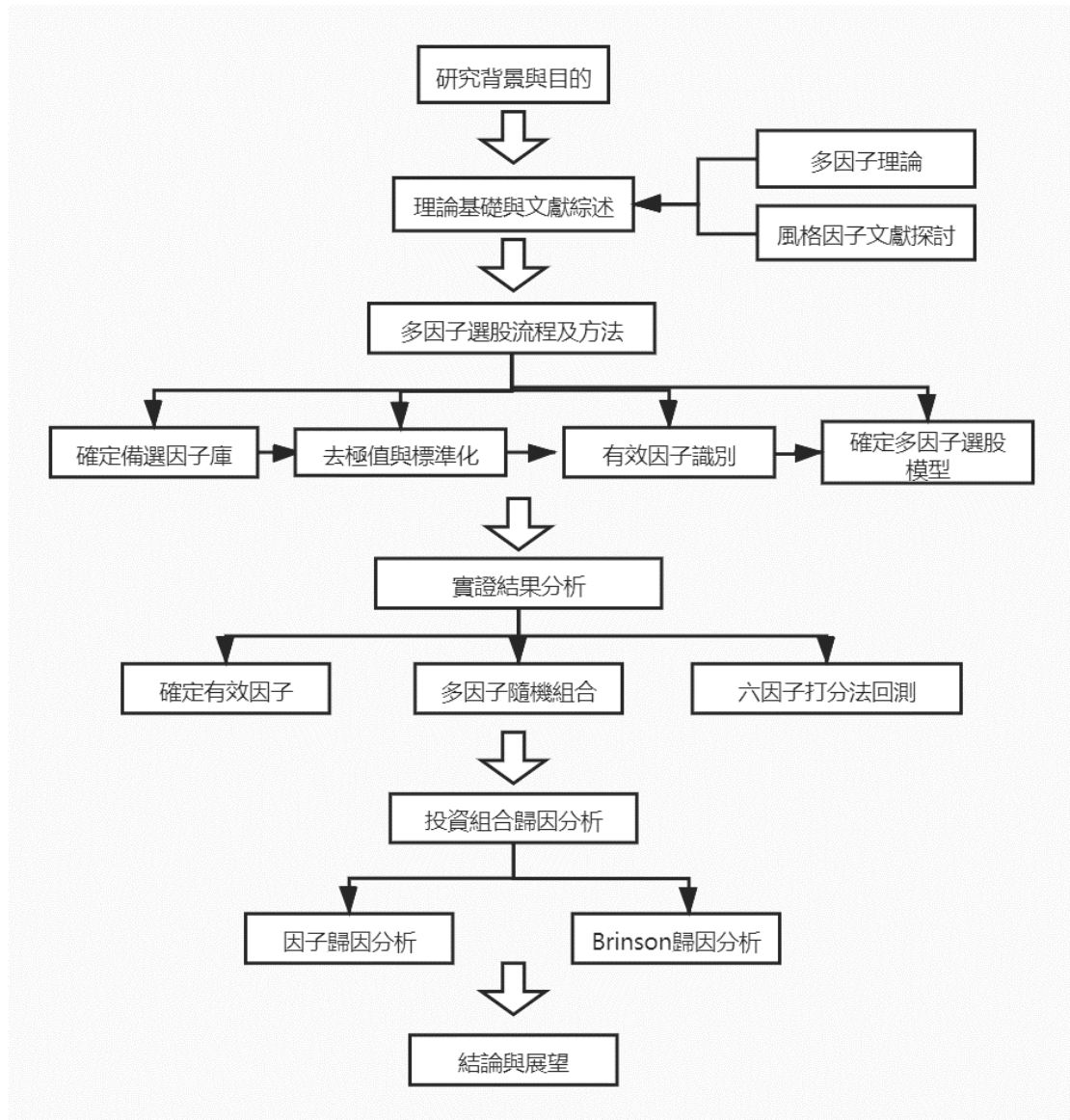


圖 1：論文結構安排

## 第二章、文獻回顧

### 2.1 多因子理論：

多因子選股理論最早要追溯到馬科維茲提出的均值-方差理論，該理論首次從量化的角度探討投資收益與風險之間此消彼長的關係，提出最佳投組在於達到一定風險水平下的收益最大化。在馬科維茲的理論基礎上，夏普以及利特納等學者提出了資本資產定價模型（CAPM），該模型認為個股只受到市場整體行情影響出現同漲同跌的情況，因此引入了系統性風險  $\beta$  以衡量投組與市場之間的關聯性。

CAPM 被提出以後，投資人發現市場上除了存在“可被解釋”的  $\beta$  因子之外，仍存在某些來源於股票共同特徵的因子無法被模型所解釋。因此在此基礎上，羅斯提出了資產套利定價理論（APT），以無套利理論構建股價收益率與多因子之間的綫性模型，為日後各類多因子模型奠定了堅實的理論基礎。以因子來源分類大致可以將其分為總體經濟因子、行業因子以及風格因子，其中總經主要源於地區經濟增長因素，例如：GDP 增長率、通貨膨脹；行業因子來自於能代表行業共同特徵的因素，例如：行業類型、行業景氣周期性；風格因子則來自股票或公司自身的鮮明特徵。多因子模型在跟隨整體市場系統性風險的基礎上，更加強調特定風格因子的暴露程度，重視優化選股以及擇時投資，以獲取高於傳統指數的投資回報。

### 2.2 風格因子文獻探討：

以下將從幾類經典的風格因子出發，探討中外文獻對於風格因子的勘探歷程。

## 1.規模因子：

規模因子最早由 Banz (1981) 提出，該研究以美國紐約證券交易所股票為研究對象，發現美股市場在過去至少 40 年的時間裏都存在“小規模效應”，Reimganum (1981) 也於同年也提出了同樣的觀點。大部分學者嘗試從風險溢酬的角度解釋此“市場異象”，例如 Arbel 與 Strebel 於 1982 年提出的“被忽略公司效應”：由於券商分析師與專業投資人更傾向於關注於規模較大的公司而忽視了小公司的資訊揭露，導致市場投資人低估了小盤股的績效表現。Chan, Chen(1991) 指出公司的小規模特徵作為低效能、高杠杆的代理變數承擔了由外部經濟動態變化所帶來的高風險，因此在理論上會取得更高的風險補償。

中國股市同樣也存在明顯的小市值效應。朱世武、鄭淳 (2003) 以 1995 年到 2002 年的中國上市公司數據為樣本，按照股票的流通市值對其進行排序分組，發現投組收益率隨著規模的增加而呈現遞減的效果，文中解釋了中國股數出現“小市值效應”的緣由，例如：市場不成熟導致投資人高度投機的特徵，小盤股的殼資源價值以及國家的政策扶持等。陳慰、張兵 (2018) 使用 1996 至 2016 年所有上市公司的數據對中國股市“小市值效應”再次進行探討，研究發現在控制住流動性、特徵波動度、博彩型偏好等 8 種定價因素後小市值效應仍然顯著有效，使用大小盤對沖的策略可以帶來每月 2.32% 的收益。

## 2.公司財務因子：

基本面投資法著重於分析公司的財務因子，按照特定面向又可以為估值類、盈利類、杠杆類、成長類等，這些因子通常會被定期反映在公司的財報上。以估值類因子為例，B/M ratio、E/M ratio 相對估值指標反映了公司賬面、盈利相對於市價的比值，相當於標準化後財務數據，投資人可以依據相對價格指標的大小

對公司價值進行比較分析。關於估值類因子的選股效果最早要追溯到 Basu (1977) 對於 P/E ratio 的選股實證研究，而 B/M ratio 的有效性則發掘於 Rosenberg, Reid and Lanstein (1985)，二者都證明瞭價值股優於成長股的表現，B/M 因子與規模因子於 1992 年被納入 Fama French 三因子模型成為資產定價模型中最重要的因子之一。在中國股市的研究上，肖軍，徐信忠 (2004) 以中國滬深 A 股股票為例進行價值反轉的投資策略，通過買入 B/M 較高的股票的確可以獲得超額報酬，說明 A 股市場同樣存在“價值效應”。

作為價值投資法的先驅，Benjamin Graham 認為估值因子應搭配質量因子做輔助才能找到真正意義上被低估的高品質股票，因此除了估值因子之外，公司的盈利能力、財務結構、成長能力等都同等重要。首先是盈利類因子，盈利因子選股效能主要體現在公司的經營能力，公司當前獲利能力良好說明公司擁有穩定的資金來源，能為投資人創造更多投資價值。Chen, Novy-Marx, and Zhang (2010) 根據托賓 Q 理論中企業市場價值以及重置成本的關係，認為高盈利而少投資的公司會有更高的預期收益率，因此在 CAPM 的基礎上加入投資因子 I/A 及盈利因子 ROA 構建新三因子模型，該模型在 1972 到 2009 年的美國股市上解釋能力超過傳統的 Fama-French 三因子模型，以至於 Fama, French (2015) 也在此基礎上納入了投資以及盈利因子，形成後來的 Fama-French 五因子。

除了公司的盈利能力，同樣需要綜合考量財務上安全性以及收益上的成長性。公司經營狀況不佳會導致企業的投資收益小於負債利息支出，令企業陷入財務危機，因此普遍上認為融資杠杆較高的公司在財務上有一定的風險性。而在收益上保持高成長的公司通常在投資人看來在未來會有更好的發展潛能。中國股市在質量因子的研究方面，李斌、馮佳捷 (2019) 從公司盈利性、成長性、安全性以及股

利分配的特徵出發，考察了公司質量因子對 A 股股票定價的影響，發現高質量的股票在風險調整後呈現最佳的收益表現。

### 3.流動性因子：

流動性反映了股票在市場上交易的活絡程度，理論上，當股票流動性水平較高時，投資人通常能以較為合理的價格對股票進行及時的交易，同時交易成本也相對較低，更受市場投資人的青睞。而學者在過去對於流動性水平與收益率的研究中却發現二者呈現負向相關性，針對流動性帶來的市場異象，Amihud and Mendelson(1986)以買賣價差與股票報酬之間的關係為例，首次提出了“非流動性補償”的概念，即當股票的流動性水平較低時，股票的價格更容易受到交易量的影響，且因為交易量少導致交易成本上升，投資人將面臨較高的流動性風險，因此為了補償流動性風險帶來的影響要求更高的預期報酬。同樣，Haugen and Baker(1996)以股票換手率、交易量等流動性相關指標對股票進行排序分組，發現流動性因子與股票之間存在負相關性。Amihud(2002)通過價量結合的方式計算單位成交量下的絕對報酬變動，從而構建出 Amihud 非流動性因子，用橫斷面資料再次證明瞭非流動性越高的股票收益率更高。在中國股市，流動性因子的研究也展現出了相同的效果，王春峰、韓冬、蔣祥林（2002）使用 ILLIQ 因子為度量，分別使用截面回歸以及時間序列回歸法驗證了 ILLIQ 因子有效性。蘇冬蔚、麥元勛（2004）使用換手率作為流動性衡量因子證明 A 股市場存在顯著的非流動性溢價現象，認為這是由於換手率與交易成本之間存在正向關聯性導致投資人對流動較差的股票的預期報酬更高。孔東民（2006）通過將 ILLIQ 因子納入 CAPM 形成 LCAPM 模型發現加入非流動性因子的資產定價模型能更好地解釋資產報

酬率，同時在對公司規模進行控制之後結果依然穩健，說明非流動性因子在中國股市有重要的經濟意涵。

#### 4.短期反轉效應：

相對於財務相關的基本面因子，股價信息總能最及時地被廣大投資人所接收，更真實地反映市場行情以及市場投資人情緒，因此投資人也更傾向於通過當前獲得的股價信息對未來股價變化做出預判。基於股價信息的因子主要體現在歷史收益率上，根據其效果又分為“慣性效應”以及“反轉效應”。其中“反轉效應”最早由 DeBondt and Thaler (1985) 提出，通過買入過去表現最差的投組能創造相對於市場每年平均 19.6% 的超額報酬，而買入過去表現最佳的投組的報酬則低於市場 5%，作者認為這是投資人的“過度反應心理”導致過去表現較差的股票被相對低估而過去表現較佳的股票被過分高估，而當市場出現修正後，市場就出現反轉現象。而“慣性效應”的發現源自於 Jegedeesh and Titman (1993)，由於股票信息的“反應不足”，其投組表現出與反轉效應相反的“強者恆強”現象，由其構建的動量因子在 Carhart (1997) 被加入 Fama-French 三因子模型形成經典的 Carhart 四因子模型。

同樣，中國股市對於股價信息構建的動量/反轉策略也有相應的研究，但大多數研究發現中國股市的確存在短期反轉現象，慣性現象則並不明顯。王永宏、趙學軍 (2001) 以股票的 N 個形成月的累計超額收益為選股因子，通過“買輸賣贏”以及“買贏賣輸”分別構建慣性策略以及反轉策略，發現中國股市存在明顯的“過度反應”現象，但並不存在“反應不足”的現象。彭慧文 (2013) 以 1997 年至 2010 年的滬深股市數據為例，在排除了非同步交易、買賣報價反彈等干擾因素後，仍然存在明顯的反轉效應。同時，文中也呈現了換手率與非流動性指標對短期反轉

效應的影響，結果顯示反轉效應隨著非流動性的提高而增加，而換手率則沒有明顯影響。

## 5.波動度：

股價波動度指標的表現形式有許多，以 Barra<sup>1</sup>波動度衡量因子為例，其中包含了系統性波動度因子以及剩餘波動度，系統性波動度因子即常見的 beta 係數，beta 係數用於衡量股價收益率相對於基準指數的波動程度，從風險溢酬的角度來看，若 beta 係數越高則說明系統風險越高，理論上應獲得更高的收益率。但 Blitz, D., & Van Vliet, P.(2007)提出了“波動度異象”：由於低波動的股票存在一定的抗周期性以及逆市場性因此在長期表現上優於高波動股票。楊雨晴（2020）以中國股市為研究對象，發現股票收益率與系統性風險呈正向關係，而與特質波動度呈反向關係，長期而言整體上則是呈現低波效應。

## 6.股利類：

股票股利相關特徵通常使用股利率作為衡量指標，例如 Barra 股利類因子指標就包含了股利率以及預期股利率，股利率是股息與股價之間的相對比率，同樣可以作為一種估值指標以衡量公司價值。股息率越高說明該股票本身更具收益性，尤其對於保守投資人更具投資價值。熊海斌，楊帆（2013）通過將股息率由大到小分為五個投組，發現中國股市中股息率與股票收益率呈現明顯的正向關係。

## 7.分析師情緒：

分析師情緒相關的因子包含許多，例如分析師預測類因子以及分析師評級變化等。券商通常會定期公布分析師預測報告為投資人做股價預測與股票推薦，因

---

<sup>1</sup> Barra 公司作為因子研究的先鋒，由美國學者 Barr Rosenberg 創立并專注於多因子量化投資以及風險管理，2004 年 MSCI 公司對其進行併購形成 MSCI Barra，致力於創建因子投資方面的探索以及革新。



此分析師情緒可能會影響股價走勢。蔡慶豐，楊侃（2013）以中國 A 股為研究對象，發現分析師對評級調整動機主要來自於公司外部的無形信息，但相對而言，分析師以及投資人更加關注基本面等有形信息。胡昌生，高玉森（2018）以研報數為研究代理變數對股票進行分組，發現分析師薦股力度與短期溢價水平之間呈現正向關係，而隨著時間推移則表現出反轉效果。



## 第三章 多因子選股方法

### 3.1 風格因子的選取：

表 3-1：風格因子表

序號	風格	因子	因子描述
1	市值類	LN (MV)	總市值對數
		LN(MVOSD)	流通市值對數
2	估值類	B/M	賬面市值比
		E/M	益本比
3	質量類	ROA	總資產報酬率
		ATO	資產周轉率
4	成長類	EPS_G	每股盈餘同比成長率
		SALES_G	營業收入同比成長率
5	杠杆類	MLEV	市場杠杆
		DTOA	資產負債率
6	流動性類	ILLIQ	Amihud 非流動性因子
		STO	換手率
7	短期反轉類	STR_1M	過去一個月股價報酬
		STR_3M	過去三個月股價報酬
8	波動度類	BETA_1M	由 CAPM 公式推導出的最近一個月的 BETA 值
		BETA_1Y	由 CAPM 公式推導出的最近一年的 BETA 值
9	股利類	DDY	股息率
10	分析師預測	SCORE	證券公司分析師對股票的評級均分，其中買入、增持、中性、減持，賣出的分數分別為 1, 2, 3, 4, 5

### 3.2 去極值與標準化：

因子的數據採集完畢之後需要對因子進行進一步的清洗以及標準化的工作，以便後續進行各因子之間的比較以及有效因子識別的步驟。

#### ● 去極值：

爲了避免極端數據的存在對因子後續分析造成干擾，需要對各因子的原始數值進行去極值的處理。本文選用了中位數去極值法來進行去極化的操作，該方法

為標準差法的改進，使用數值中位數以及平均絕對偏差(MAD)來替換均值以及標準差，令結果更加穩健。該方法先通過計算各數值與中位數的絕對差值的算數平均數得到平均絕對偏差 (MAD)，再通過中位數以及 MAD 來決定數值的極值範圍，其中業界通常選用基於平均絕對偏差 3 倍標準差的範圍來作為上下界，具體公式如下：

$$MAD = \frac{1}{n} \sum_{i=1}^n |x_i - \text{median}(x_i)| \quad (3-1)$$

$$x'_i = \begin{cases} x_{\text{median}} + 3MAD & \text{if } x_i > x_{\text{median}} + 3MAD \\ x_{\text{median}} - 3MAD & \text{if } x_i < x_{\text{median}} - 3MAD \\ x_i & \text{if } x_{\text{median}} - 3MAD < x_i < x_{\text{median}} + 3MAD \end{cases} \quad (3-2)$$

- 標準化：

因為各因子的單位以及數值大小之間存在差異無法直接比較以及分析，因此需要使用 z-score 的方式對去極值後的數據進行標準化，各類因子數據經由 Z-score 的方法處理後的將符合平均值為 0，標準差為 1 的標準常態分配，使得數據更加集中的同時也消除因各因子的數值結構不同對結果造成的影響。

$$x'_i = \frac{x_i - \mu}{\sigma} \sim N(0,1) \quad (3-3)$$

### 3.3 有效因子識別：

雖然在主觀邏輯上 18 個因子與股票收益率之間都存在一定的關聯性。但各個因子，尤其是同類因子之間在選股效能上一定會有優劣之分，因此需要引入單因子有效性分析，從客觀的統計角度來驗證各因子與未來收益率之間的關係，並找到確實能提升未來選股收益的特徵性因子。因子的選股效能可通過以下四個步驟來檢驗其有效性、穩健型、獨立性以及一致性。

- 單因子回歸 t 檢定：

首先，業界最常用的因子有效性檢驗方法為因子與收益率之間的回歸分析，通過回歸係數 t 檢定來檢驗各個因子與收益率之間是否存在顯著的綫性關係。為在一定程度上消除行業對因子有效性的影響，本文將行業虛擬變量矩陣納入考量。回歸模型如下：

$$R_i^{T+1} = \sum_j \beta_j^T * Industry_{j,i}^T + \beta_F^T * Factor_i^T + \mu_i^T \quad (3-4)$$

$R_i^{T+1}$ : 股票 i 在第 T+1 期的收益率

$Industry_{j,i}^T$ : 股票 i 在第 T 期第 j 個行業因子<sup>2</sup>上的暴露度（屬於該行業則為 1，否則為 0）

$Factor_i^T$ : 股票 i 在第 T 期某單因子上的暴露度

$\beta^T$ : 第 T 期第 j 個行業因子或者因子 Factor 的回歸係數，通過每個月回歸得到

$\mu_i^T$ : 股票 i 在第 T 期的殘差值

對於每一期的單因子截面回歸都可以得到其對應的因子回歸係數以及 t 統計量，因此通過整理所有期間上的橫截面回歸 t 檢定情況可以從整體上對因子的有效性做評判。各期單因子回歸係數的 t 檢定方式如下：

$$\begin{cases} H_0 : \beta_F = 0 \\ H_1 : \beta_F \neq 0 \end{cases}$$

$$t = \frac{\beta_F - 0}{\sigma / \sqrt{N-1}} \quad (3-5)$$

t: 單因子係數的 t 統計量

$\beta_F^T$ : 單因子回歸係數

$\sigma$ : 單因子係數的標準差

N: 樣本個數

1) t 值絕對值的均值：針對每一期橫截面數據，我們可以得到各個單因子的 t 統計量，通過比較 t 值絕對值均值是否大於 2 來初步判定因子與收益率之間是否顯著相關。

<sup>2</sup> 行業分類標準採用申萬一級行業分類標準，從采掘業到綜合業務共 28 個基本行業類別。

2) t 值中絕對值大於 2 的占比：爲了防止極值對數據結果造成影響，仍需比較各期 t 值絕對值大於 2 的比率來協助評判，若比率大於 30%，則認同該因子具有有效性。

● 單因子 IC 與 IR 值：

由於 t 檢定的顯著性只能單純表現因子值對未來收益率的影響程度，無法反映因子與收益率之間是否存在持續的關聯性，因此引入因子 IC 以及 IR 值以反映因子對於下一期收益率預測的穩定性。IC 值在本質上是計算因子與下一期收益率的相關係數，IR 值則是進一步表現因子在獲取收益能力的穩定程度，相應的公式如下：

$$IC^T = \text{corr}(R^{T+1}, \text{Factor}^T) \quad (3-6)$$

$$IR = \text{mean}(IC) / \text{std}(IC) \quad (3-7)$$

其中 IC 值的計算方法有兩種，一是 Normal IC 法，只是單純的計算因子與下一期收益率的相關性係數，一般直接使用皮爾森 (pearson) 相關係數的公式來計算。二是 Rank IC (秩相關係數)，先將每一期的因子以及收益率按照值的大小進行排序，再使用相應公式計算兩者序列之間的相關係數，通常使用斯皮爾曼 (spearman) 相關係數的公式來計算。兩種方法的具體計算方式如下：

方法一、皮爾森 (pearson) 相關係數：

$$\rho_{(\text{pearson})} = \frac{\text{cov}(R_i^{T+1}, \text{Factor}_i^T)}{\sigma(R_i^{T+1}) * \sigma(\text{Factor}_i^T)} \quad (3-8)$$

其中  $\text{cov}(R_i^{T+1}, \text{Factor}_i^T)$  表示各期因子值與下一期的收益率之間的共變異數， $\sigma(R_i^{T+1})$  與  $\sigma(\text{Factor}_i^T)$  分別代表各期因子值的標準差以及下一期收益率的標準差。

方法二、斯皮爾曼 (spearman) 相關係數：

以下表的數據為例，首先，對原始兩個變量 $R_i^{T+1}$ ， $Factor_i^T$ 進行結對形成 n 組  $(X_i, Y_i)$ 按照各自值的大小進行排序，然後記錄兩個變量的排序情況記為  $(X'_i, Y'_i)$ ，再計算出兩個變量排序差值的平方記為 $d_i^2 = (X'_i - Y'_i)^2$ 。

表 3-2：斯皮爾曼相關係數示意表

變量 $X_i$	變量 $X_i$ 的排序	變量 $Y_i$	變量 $Y_i$ 的排序	排序差值的平方
30	3	70	1	$d_1^2 = (3 - 1)^2 = 4$
20	4	20	3	$d_2^2 = (4 - 3)^2 = 1$
4	5	10	4	$d_3^2 = (5 - 4)^2 = 1$
40	2	5	5	$d_3^2 = (2 - 5)^2 = 9$
60	1	30	2	$d_4^2 = (1 - 2)^2 = 1$

最後通過下列的公式計算出每期的相應的斯皮爾曼相關係數：

$$\rho(\text{spearman}) = \frac{1 - 6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (3-9)$$

在本文中， $X_i$ 代表每一期因子值， $Y_i$ 代表下一期股票收益率，因此 $d_i$ 代表每一期因子排序與下一期收益率排序的差異值， $n$ 則表示排序列的總數。

雖然兩種方法的理論基礎相同，但後者在使用過程中受異常數據的影響較小，適用範圍更廣，且考慮後續將使用因子排序打分的方式進行選股，因此本文實證中採用 Rank IC 法進行評估，並通過下列的方式進行因子的篩選：

1) IC 值均值：針對每一期的數據，我們可以計算得到各個因子與下一期收益率之間的 IC 值，再通過比較 IC 值的平均數是否大於 0.03 來評判因子是否具有有效性。

2) IR 值：通過 IR 值公式 (3-7) 計算得到各因子所對應的 IR 值大小，根據業界常用 IR 值標準：比較因子 IR 值是否大於 0.3 來評判因子與收益率序列之間是否存在穩定的相關性。

● 因子相關性分析：

在候選因子較多的情況下，對於經濟意涵相似、相關性較高的同類因子，可採取因子之間的有效性比較刪除冗餘因子。而各大類的因子之間，雖然經濟意涵不同，但考慮到因相關程度過高而選出相同的股票，也應儘量保證因子之間的獨立性以最大程度地發揮多因子的選股能力，此處相關性係數的計算同樣採用斯皮爾曼法。

1) 首先，按照每月數據計算出各因子值之間的相關係數矩陣：

$$\rho^t = \begin{bmatrix} 1 & \rho_{12}^t & \dots & \rho_{1k}^t \\ \rho_{21}^t & 1 & \dots & \rho_{2k}^t \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k1}^t & \rho_{k2}^t & \dots & 1 \end{bmatrix} \quad (3-10)$$

2) 然後對各期的相關性係數數據做平均得到因子之間平均相關性矩陣：

$$\bar{\rho}_{ij} = \frac{1}{N} \sum_{t=1}^N \rho_{ij}^t \quad (3-11)$$

$$\bar{\rho} = \begin{bmatrix} 1 & \bar{\rho}_{12} & \dots & \bar{\rho}_{k1} \\ \bar{\rho}_{21} & 1 & \dots & \bar{\rho}_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\rho}_{k1} & \bar{\rho}_{k2} & \dots & 1 \end{bmatrix} \quad (3-12)$$

● 單因子分組回測：

單因子分組回測的目的在於檢視各因子的實際選股效果在時間序列上是否具有一致性，作為檢驗因子有效性的輔助步驟，相較於統計檢定更為直觀。我們以每月為回測頻率，每個月底都依照各個因子值的大小對股票進行排序，再按排序結果將股票均分為 5 組，1 到 5 組分別表示因子值最大的前 20% 支股票到最小的後 20% 支股票，各組內個股採等權重方式持有到下個月，計算出各個組別所對應的下一期收益率，重複此選股操作即可呈現 5 個投組在樣本內的實際績效表現。除了各投組的終淨值圖像之外，同樣計算下列各類績效指標來評判各組在風險以及收益上的表現：

1) **年化收益率**：通過將回測期間投組的累計收益率轉換成以年為單位表示的集合平均收益率來表現投組在投資期間整體的報酬率情況，年化收益率越高表示投組在收益上的表現越好。

$$R_{annualized} = (1 + R_{cumulative})^{\frac{12}{n}} - 1 \quad (3-13)$$

其中  $P_{cumulative}$  代表投組累計收益率， $n$  代表投資期數。

2) **勝率**：通過計算回測期間投組收益率打敗基準指數的比率來評判投組相對收益情況，勝率越高說明投組相對於基準指數的表現更佳，此處的基準指數採用滬深 300 指數。

$$Odd = (\text{投組報酬大於基準指數報酬的次數}) / \text{總投資期數} \quad (3-14)$$

3) **年化收益波動度**：通過計算年化後的投資組合收益率標準差來衡量投組整體報酬的風險程度，通常年化收益波動度越高代表收益的不確定性越高，投資風險越高。

$$\sigma_p = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (R_p - \bar{R}_p)^2} * \sqrt{12} \quad (3-15)$$

其中  $P_p$  代表投組各期的報酬率， $n$  代表投資期數。

4) **最大回撤率**：通過計算投組每期淨值相對過去最高點的最大回落幅度，衡量投組在投資期間發生最大虧損程度。最大回撤率越大代表投資所遭受的最大虧損越大，投資人很有可能因為承受不住風險而提前結束投資。

$$Max\ Drawdown = Max(1 - \frac{P_t}{P_{Max}}) \quad (3-16)$$

其中  $P_t$  代表各期投資的累計淨值， $P_{Max}$  代表  $t$  時點前投組發生的最大累計淨值。

5) **夏普比率**：表示承擔一單位風險報酬下，投組相對於無風險利率的超額表現，夏普比率越高表示投組在風險調整後的收益率越高，投組整體的績效表現越好。

$$Sharp\ Ratio = \frac{R_p - R_f}{\sigma_{ap}} \quad (3-17)$$



其中 $R_p$ 與 $R_f$ 分別代表投組的年化收益率以及無風險利率，無風險利率采用十年期國債之年化收益率， $\sigma_p$ 代表年化收益率波動度。

6) 信息比率：計算了投組相對於基準指數在風險調整後的收益，在一定程度上反映了投組相對於基準的主動選股能力。通常信息比率越高代表投組相對基準的績效表現更好。

$$\text{Information Ratio} = \frac{R_p - R_B}{\sigma(R_p - R_B)} \quad (3-18)$$

其中 $R_p$ 與 $R_B$ 分別代表投組的年化收益率以及基準年化收益率，基準指數采用滬深 300 指數， $\sigma(R_p - R_B)$ 為投組的追蹤誤差，即投組與基準指數收益率差值之年化標準差。

### 3.4 多因子選股模型介紹：

多因子選股的主要思路是透過結合有效性因子的選股效能創造出能提升整體選股績效的策略方法，目前常見的模型方法包含排序打分法、回歸法以及近年來流行的機器學習選股模型。機器學習與回歸法的本質都是通過統計、模擬預測下一期的股票收益率進行選股，算法相對複雜，普通投資人在實際應用中也存在技術上的局限性。因此，本文采用業界最常用的排序打分法進行實作，相關的方法主要有以下三種：

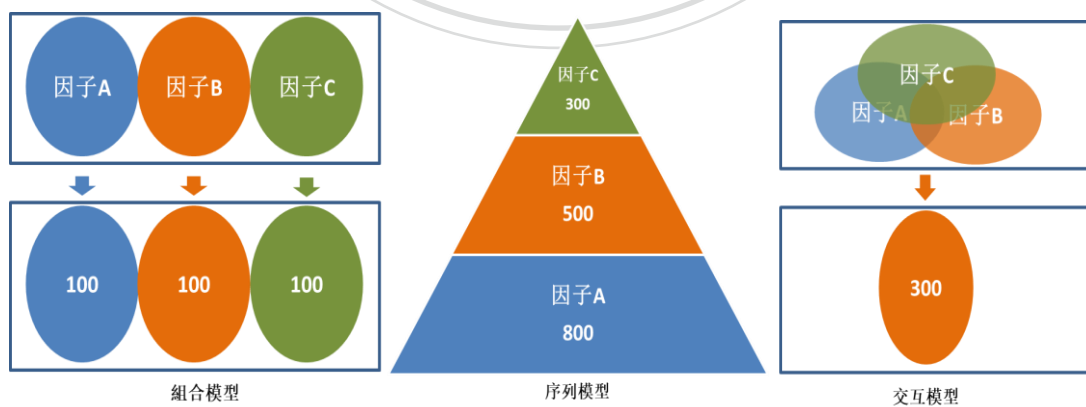


圖 3-1：排序打分法三模型

- 組合模型：

如圖 3-1 左圖所示，組合模型的概念相對簡單，針對每一期的數據，運用單因子的選股效果在各個因子上篩選出收益表現較好的股票，再將這些股票組合成一個投資組合進行投資。

- 序列模型：

序列模型又稱多重因子選股法，使用各個因子對股票池進行層層篩選，得到最終的股票並形成投資組合。構建序列模型的難點主要在於因子篩選的先後順序，為了保證因子的信息能夠最有效地用於選股，因子的篩選順序一般取決於各因子的重要性程度。例如，構建投組前可以對通過有效性檢驗的候選因子進行績效評分以決定因子的主次順序，或是通過遍歷因子的選股情況或依據選股經驗與邏輯來決定。

- 交互模型：

交互模型通過多因子的綜合評分對備選股票進行排序，依據得分高低從股票池中篩選出排序靠前的目標股票進行倉位配置以構建投資組合，具體的操作方法如下：

1. 首先，按照因子有效性檢驗步驟中因子與收益率之間相關性方向對各類因子進行排序。以對數市值為例，按照 IC 值的結果可以判斷出對數市值與下一期的收益率呈現負相關，因此在每個換倉日，根據該因子值的大小對股票進行倒序排序，即因子值越小的股票市值因子的分數越高。相反，若因子 IC 值顯示該因子與收益率呈現正向關係，則對該因子進行正序排序，即因子值越大的股票在該因子上的得分越高。

2. 將上一步得到的各因子分數進行等權加總，得到各個股票的多因子綜合評分。

3.在每個換倉日都按照因子綜合評分對股票池中的股票從大到小進行排序，並從中篩選出得分最高的股票合成投資組合作為投資標的。

下表以三檔股票進行三因子排序打分法為例，展示了交互模型下的打分規則：

表 3-3：交互式打分法示意表

	股票 1	股票 2	股票 3
因子 A (-)	30	20	10
因子 A 排序得分	1	2	3
因子 B (+)	20	30	40
因子 B 排序得分	1	2	3
因子 C (+)	80	60	10
因子 C 排序得分	3	2	1
因子綜合得分 (各因子等權重)	$(1+1+3) \times \frac{1}{3} = \frac{5}{3}$	$(2+2+2) \times \frac{1}{3} = \frac{6}{3}$	$(3+3+1) \times \frac{1}{3} = \frac{7}{3}$
選股次序	3	2	1

\*表中各因子後的符號表示該因子與下一期收益率之間的相關性方向，若為“-”表示下一期收益率與因子之間呈負相關，備選股票將按因子值大小以倒序排序，因子值越大排序計分越小，“+”則相反。

雖然以上三種模型都具有各自的邏輯以及特點，但在應用上同樣有一定的局限性，例如在第一種方法之下選出的各個股票並不能反映所有因子的特性，若考慮因子有效性的差異，可能投組的表現不如單個因子選出的投組表現好。而第二種方法在操作流程上會相對複雜，除了因子篩選的先後順序需要個人的主觀判斷之外，也較難設定各層篩股的比例，尤其是當股票少因子多的情況下這種方法不太實用。相較而言，交互模型既無需考慮因子篩選的先後順序，也能充分展現因子的綜合選股效能，在操作流程上也相對便捷，基於分析對比，本文實證中采用交互式模型對股票進行篩選建模，並從原有的股票池中選出因子綜合評分最高的 300 支股票以等權重的方式形成投組。

## 第四章 實證分析

### 4.1 實證數據描述：

- 數據來源：

數據均來源於 CSMAR（國泰安中國經濟金融研究數據庫）

- 數據樣本範圍：

股票池的選擇為中國 A 股全體股票，其中刪除了當月為 ST、ST\* 的股票，ST 與 ST\* 是證交所對出現公司財務狀況或者運營商出現異常的上市公司股票交易進行特別處理的標志，提醒投資人這些股票的公司未來可能有退市的風險。在保證整體樣本數據足夠的條件下，為了保證選到的股票能在市場上能基本正常交易，對樣本中上一期的交易天數小於 10 天的股票進行刪除，認為上一期交易天數過少的股票缺乏正常的交易量以及流動性，不適合進入下一期的股票池。

- 採樣期間：

採樣期間為 2005-01-01 到 2020-12-31，其中樣本內的投組測試期間為 2006-01 到 2014-12，樣本外的測試期為 2016-03 到 2020-12。選擇刪除 2015-01 到 2016-02 的數據是因為這段期間中國股市發生了史上最嚴重的“2015 股災”事件，期間中國 A 股市場遭遇了前所未有的暴漲暴跌的情形：先是從進入 2015 年到 2015 年 6 月 12 日的五個半月裏，上證指數與深成指分別驟升了 60% 和 122%，隨之而來的是股市的連續性暴跌，在 2016 年年初前兩個月又迎來新一輪的驟跌，甚至兩度觸發熔斷機制。因此，為了避免這一階段由於外界事件以及人為操弄股市

對因子選股帶來的影響，決定刪除這一期間的股票數據，並以一這段期間為樣本內外數據的分界線，來觀察 2015 年股災前後 A 股市場的多因子選股效果。

#### 4.2 有效因子識別：

表 4-1：風格因子有效性檢驗結果

風格大類	因子	t 均值	t ≥2 的比率	IC 均值	IR 值	是否通過檢驗
市值類	LN(MV)	6.4017	80.56%	-0.0656	-0.3483	是
	LN(MVOSD)	6.3974	79.63%	-0.0500	-0.2549	否
估值類	B/M	4.1602	65.74%	0.0543	0.3549	是
	E/M	4.4048	74.07%	0.0357	0.2447	否
質量類	ROA	4.7010	74.07%	0.0012	0.0077	否
	ATO	1.8549	38.89%	0.0069	0.0919	否
成長類	EPS_G	2.0111	37.96%	0.0336	0.5414	是
	SALES_G	1.5862	26.85%	0.0187	0.3222	否
杠杆類	MLEV	3.1795	56.48%	-0.0045	-0.0321	否
	DTOA	3.7659	56.48%	-0.0082	-0.0585	否
流動性類	ILLIQ	5.3982	79.63%	0.0883	0.5492	是
	STO	6.1862	80.56%	-0.0712	-0.4225	是
短期反轉類	STR_1M	5.1053	79.63%	-0.0683	-0.5053	是
	STR_3M	5.1830	69.44%	-0.0639	-0.4175	是
波動度類	beta_1Y	4.2424	66.67%	-0.0119	-0.0825	否
	beta_1M	4.1559	70.37%	0.0314	0.3104	是
股利類	DDY	3.5593	69.90%	0.0322	0.2785	否
分析師預測	SCORE	2.8780	59.26%	-0.0084	-0.0886	否

通過上表我們可以大致判斷出各個因子的選股效能。首先，可以發現大部分因子的回歸係數 t 值的絕對數均值都大於 2，且整體大於 2 的比率較高，說明大部分因子對股票收益率具有顯著的解釋作用，其中市值類對於股票收益率的解釋效果最顯著。然而，與公司的營收相關的營收成長以及資產周轉率因子在 t 值上均未達標，說明雖然市場投資人會在一定程度上關注營收相關的財務因子，也更看好運營能力越好的企業，但因 A 股市場上市公司眾多且市場尚不成熟，單從營收表現上無法判斷其未來真實的表現，因此這一類因子的有效性相對較低。

雖然許多因子的 t 檢定都達到標準，但從再各類因子的 IC、IR 值的表現來看却存在明顯的參差，這說明部分因子的有效性並不穩定。從大類來看表現最佳的因子類別為流動性類，這類的因子與股價報酬率保持著穩定的高相關性，說明流動性因子所表現的特性在 A 股市場是長期穩定存在的。然而，大多數不滿足條件的因子都屬公司財務類，這在一定程度上說明瞭財務類因子在 A 股上的選股效能是較弱的，一方面因為這些因子都屬季數據，一般只有公布財報的特定期間才會給公司股價帶來一定的影響，因而整體上與股票報酬率的相關性程度較低，另一方面可能數據庫本身在數據采集上存在滯後性。

根據因子有效性的檢驗結果，可以初步篩選出 6 大類因子風格，分別為：市值類、估值類、成長類、流動性類、短期反轉類以及波動度類，而其中流動性以及短期反轉類的兩個因子都通過了因子有效性的檢定，但因為因子之間計算方式相似，大類因子之間具有相似的經濟意涵以及較高的相關性程度，例如近三個月報酬率與近一個月報酬之間的相關性就達到了 0.77，因此選擇在其中挑選出有效性最佳的因子作為其風格的代表。通過有效性的結果，我們選擇保留代表流動性類的 ILLIQ 因子以及代表短期反轉類的 STR\_1M。

根據上述步驟，我們選出了代表不同風格的六個因子，雖然這些因子在邏輯上均代表不同的經濟意涵，但為了避免各因子之間過於相似而導致選出的股票存在趨同性，我們通過各因子之間的相關性分析來檢視大類因子之間是否具備獨立的選股效能。一般而言，若變數之間的相關係數絕對值低於 0.3 表示兩者之間為低度綫性相關，處於 0.3~0.7 之間，說明變數之間存在中等相關性，而大於 0.7 則代表因子之間存在較強的相關性。以下是每期各因子之間的平均相關性形成的矩陣：

	LN(MV)	B/M	EPS_G	ILLIQ	STR_1M	BETA_1M
LN(MV)	1	0.25	0.45	-0.54	0.024	0.031
B/M	0.25	1	-0.12	-0.0014	-0.33	0.0056
EPS_G	0.45	-0.12	1	-0.19	0.23	-0.015
ILLIQ	-0.54	-0.0014	-0.19	1	0.024	-0.021
STR_1M	0.024	-0.33	0.23	0.024	1	0.0049
BETA_1M	-0.022	0.0083	-0.015	-0.023	-0.036	1

圖 4-1：風格因子相關性矩陣

從圖 4-1 顯示的結果可以發現除了 ILLIQ 與 LN(MV)之間的相關係數等於-0.54，屬中等相關之外，其餘因子之間均呈現低度相關。結合第二章文獻探討的部分我們也瞭解到，通常 A 股大市值股為國有控股以及知名企業，相對於小市值股的流動性更高，因此在價格漲跌上並不會受到交易額變化導致的大幅波動，ILLIQ 值較低。然而，大市值高流動的情況並不能完全反映市場狀況，一些小市值公司因為“高成長、低股價”的特性也受到了眾多散戶的青睞，相較於一些成熟大公司而言交易活躍度更高，因此二者並不能完全解釋彼此，我們通過相關性檢定也印證了這點。因此，根據相關性分析，得到最終的六大因子，分別是：代表規模的 LN(MV),代表估值的 B/M，代表成長的 EPS\_G，代表流動性的 ILLIQ，代表短期反轉的 STR\_1M 以及代表波動度的 BETA\_1M。

接下來，根據第五章描述的步驟，通過各個因子對所有股票進行排序以及分層，初步探究各個因子下的各層投組的選股績效表現：

## 1.LN (MV)：

$$\text{LN (MV)} = \text{LN (總市值)}$$

市值因子作為經典的 Fama 三因子體現了上市公司的規模特徵，在絕大多數市場上都存在“小市值效應”，即小市值公司的股價收益率普遍大於大公司的現象，為使數據表現更加平穩，本文參考 Barra 風格因子計算方式，採用總市值取對數作為衡量市值因子的指標。

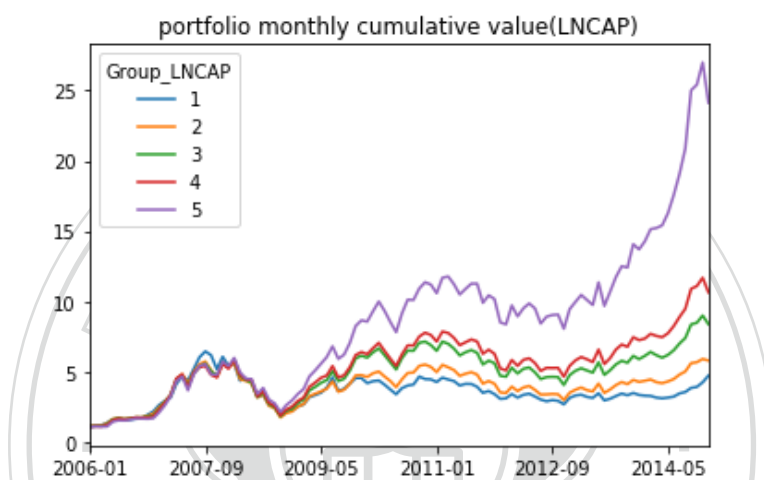


圖 4-2：規模因子分組回測累積淨值

表 4-2：規模因子分組回測績效

組別	年化收益率	勝率	年化收益 波動度	最大回撤 比率	夏普比率	信息比率
1	18.79%	53.70%	34.59%	70.65%	0.4368	0.2796
2	21.55%	53.70%	36.81%	69.58%	0.4855	0.2763
3	26.64%	56.48%	37.95%	68.89%	0.6050	0.4541
4	30.05%	57.40%	39.29%	67.75%	0.6711	0.5443
5	42.44%	63.88%	39.86%	64.55%	0.9724	0.9350

通過圖 4-2，我們可以發現，經由市值因子對股票由大到小分成 1-5 組之後，各層的投資組合之間出現了較明顯的差異性，除了 2008 年以前出現大市值股的表现優於其他組別之外<sup>3</sup>，對數市值最小的投資組合的累計終值始終是所有投組

<sup>3</sup> 這段期間并未出現“小市值效應”是由於中國人民銀行在 2007 年期間曾多次調升存款基準利率造成市場資金流出股市，上市公司尤其是小公司的籌資成本上升，導致小公司的績效表現普遍低於大公司。



中最高，同時呈現出一定的次序效果，進一步證實該因子與下一期的股價報酬之間存在穩定的相關性。再經表 4-2 具體比較各組的績效指標可以看出，隨著對數市值的減小，年化收益率呈現單調遞增，且市值最小組別的年化收益率甚至超過市值最大組別的兩倍，說明 A 股市場存在明顯的“小市值效應”。同時也可以發現波動度也隨著市值的減小而增大，說明投資小市值股的風險更大，而最大回撤則呈現一定反向效果。最後，從風險調整後的績效表現來看，小市值投組的表現即使在調整過波動度以及追蹤誤差之後仍能明顯超過大市值投組的表現。整體上來看，對數市值這個因子無論在有效性還是穩定性上的表現都是非常令人滿意的，因此後續在構造多因子組合時將選取對數總市值比較小的股票。

## 2. B/M:

$$B/M = \text{公司賬面價值} / \text{市值}$$

B/M 作為經典的估值指標，通過公司賬面價值與市值的比值，體現了股東實際擁有的公司淨值相對於目前的股東認為公司的投資價值高低。B/M 較高的股票被成為價值股，相對於 B/M 較低的成長股在收益率上一般會有更好的表現。本文 B/M 採用市淨率的倒數來表示。

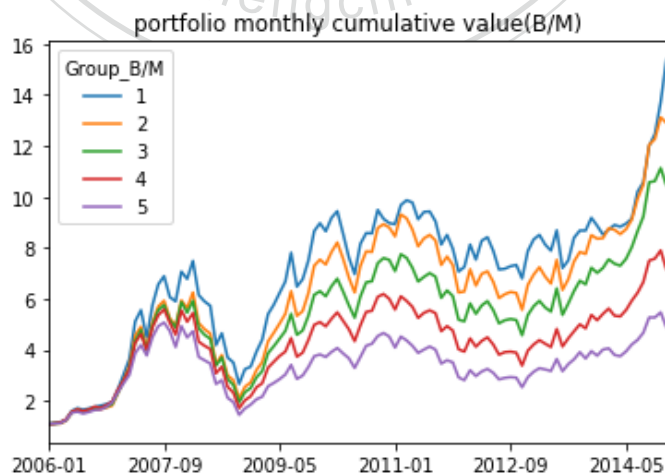


圖 4-3：估值因子分組回測累積淨值

表 4-3：估值因子分組回測績效

組別	年化收益率	勝率	年化收益波動度	最大回撤比率	夏普比率	信息比率
1	35.49%	64.81%	38.88%	67.39%	0.8182	0.9861
2	32.86%	57.40%	38.53%	66.56%	0.7573	0.7629
3	29.79%	56.48%	37.48%	67.23%	0.6966	0.6132
4	24.43%	57.40%	36.90%	69.41%	0.5623	0.3752
5	19.26%	56.48%	35.27%	71.40%	0.4417	0.1451

從圖 4-3 可以觀察到，除了在 2014 年期間因為整體股市處於急速攀升階段造成成長股投組淨值趕上價值股之外，整體上 B/M 因子表現從第一組到第五組依次呈現單調遞減的趨勢，且區分度較明顯。再從各組的績效指標上來看，年化報酬率以及年度勝率的部分隨著 B/M 的上升而呈現明顯的上升趨勢，說明整體上 A 股存在價值股表現優於成長股的特性。在風險層面上看，年化波動度隨著 B/M 的上升而上升，而最大回撤比率上則不存在一定的趨勢性，因此整體上各組風險層面上不存在明顯的差異性。最後是風險下的投組報酬，夏普比率以及信息比率與 B/M 均呈現明顯的正向關係，且最高分別達到了 0.82 以及 0.99，說明 A 股市場中價值股的績效表現是非常優秀的，因此後續多因子的構建過程中會選擇 B/M 較高的股票。

### 3. EPS\_G：

$$EPS\_G = (\text{EPS 本期金額} - \text{EPS 上年同期金額}) / (\text{EPS 上年同期金額})$$

每股收益增長率作為個股成長能力的指標代表，反映了股東每股所分得的淨收益的相對增長。為去除季節性因素的影響，採取同年同季比較的方式，同時預期每股盈餘若相較於上一年度有顯著的成長的公司將在未來有更好的發展的潛能，則能為股東創造穩定的收益。

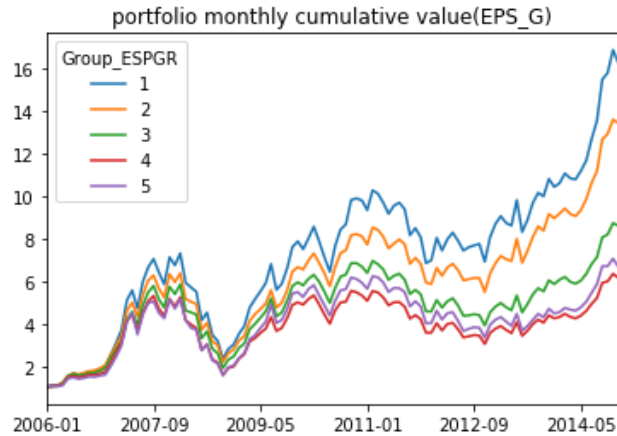


圖 4-4：成長因子分組回測累積淨值

表 4-4：成長因子分組回測績效

組別	年化收益率	勝率	年化收益波動度	最大回撤比率	夏普比率	信息比率
1	36.30%	63.88%	38.18%	67.39%	0.8544	0.9760
2	33.45%	63.88%	35.14%	64.84%	0.8472	0.8801
3	27.06%	56.48%	36.00%	66.67%	0.6494	0.5538
4	22.45%	52.78%	37.36%	69.29%	0.5024	0.3008
5	23.55%	54.62%	39.76%	68.46%	0.4997	0.3193

將所有股票按照 EPS 成長率從大到小將其分為一到五個投資組合，通過圖 4-4 可以觀察到，隨著該因子的上升，投組累積淨值呈現依次上升的趨勢，且這種正向關係在樣本期間相對穩定，同時，我們通過 2011 年到 2014 年的數據可以觀察到 EPS\_G 較高的前兩個投組終值隨著時間不斷與其他三組拉開距離。再從表 6 中具體的績效指標上分析，在年化收益率、年化勝率的部分基本與因子大小保持正向關係，而在風險部分，年化收益率波動度以及最大回撤比率之間不存在線性關係，且差距不太明顯，最後在夏普比率以及信息比率上都呈現了該 EPS\_G 較高的投組收益率在調整過風險過後仍然能保持其優勢，因此後續構建投組的過程中會選擇 EPS\_G 較高的股票。

4. ILLIQ：

$$ILLIQ = \frac{1}{N} \sum_{i=1}^N \frac{|R_i|}{dvol_i}$$

其中 $R_i$ 和 $dvol_i$ 分別代表股票在當月的第 $i$ 個交易日的每日漲幅和成交金額， $N$ 代表當月的有效交易天數。

ILLIQ 是由 Amihud(2002)提出的衡量股票價量關係的流動性因子，反映了一檔股票在觀察期內單位成交量下價格波動度的平均情況。一般而言，ILLIQ 越大代表一單位成交額對該股票的價格衝擊越大，說明該股票在市場上的流動性較差，因此根據其性質該指標又被稱為“非流動性因子”。根據資本市場“風險補償”的定價原理，投資人希望從“非流動性”較高的股票上獲得收益上的補償以彌補非流動性所帶來的風險，因此 ILLIQ 因子越高的股票股價收益率更高。

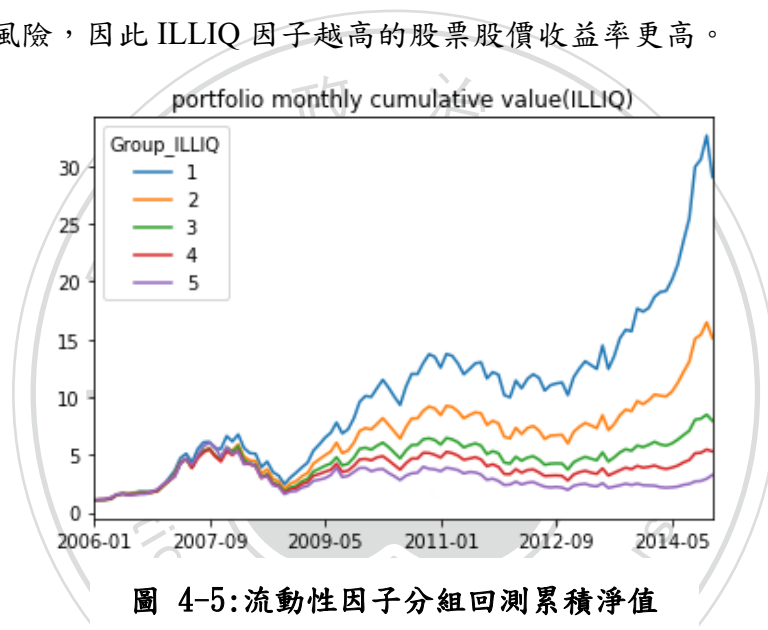


圖 4-5:流動性因子分組回測累積淨值

表 4-5：流動性因子分組回測績效

組別	年化收益率	勝率	年化收益 波動度	最大回撤 比率	夏普比率	信息比率
1	45.40%	65.74%	38.11%	63.66%	1.0947	1.2023
2	35.22%	61.11%	37.71%	65.48%	0.8364	0.8767
3	25.90%	56.48%	37.21%	68.70%	0.5972	0.4834
4	20.36%	54.62%	37.47%	69.27%	0.4452	0.2381
5	14.03%	49.07%	37.27%	73.55%	0.2777	-0.1786

將所有股票按照 ILLIQ 的值從大到小將其分為一到五個投資組合,通過圖 4-5 我們可以觀察到，整體上各個投組與累積淨值之間存在著穩定的綫性關係，ILLIQ 越大的組別報酬越高，各組之間的區分度也較為明顯。同時，我們仔細觀

察也會發現在 2007 年期間，ILLIQ 較低的投組表現出現過一段短期間的上升，這個情況與對數市值因子類似，相對於市值對數，ILLIQ 並沒有出現 ILLIQ 最小組別的績效超過 ILLIQ 最大組別的現象，這也側面說明該因子相對市值效果更佳。再從具體的績效指標上分析，各組的年化報酬率以及年度勝率呈現均勻上升的情況，風險調整過的投組報酬也與因子大小呈現明顯的正向關係，且夏普比率、信息比率最高分別達到 1.09 和 1.20，與最低的組別差距明顯，因此可以說明 ILLIQ 是所有因子中區分度最高且選股效果最好的因子。根據分析結果，後續多因子選股過程中會選擇 ILLIQ 值較大的股票進入投資組合。

#### 5. STR\_1M：

STR\_1M=過去一個月股價報酬率

根據第二章的文獻探討我們瞭解到 A 股市場存在較明顯的“短期反轉”現象，因為投資人的“過度反應”行爲，過去在股價收益率上表現較差的股票在未來的績效表現上呈現出一定程度的逆轉效果。結合過往文獻的探索結果以及操作的便捷性，本文採用業界常規用的過去一段期間的股價報酬率作為“短期反轉”的指標。

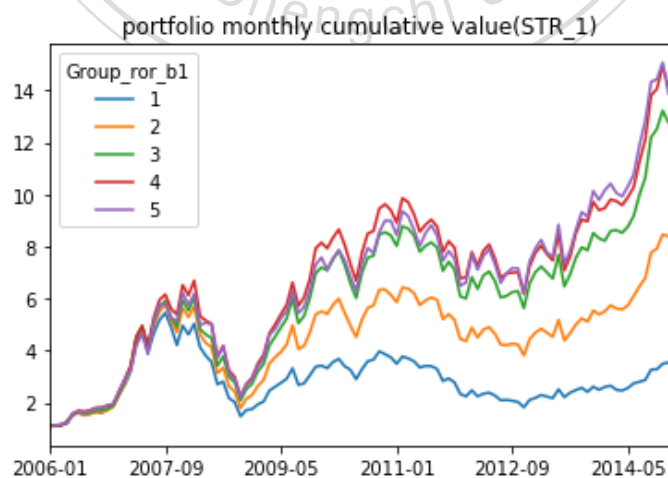


圖 4-6：短期反轉因子分組回測累積淨值

表 4-6：短期反轉因子分組回測績效

組別	年化收益率	勝率	年化收益波動度	最大回撤比率	夏普比率	信息比率
1	15.08%	52.77%	36.25%	73.03%	0.3145	-0.0519
2	26.66%	55.55%	36.82%	69.14%	0.6241	0.5276
3	32.71%	61.11%	37.25%	65.56%	0.7774	0.7314
4	34.25%	62.03%	38.94%	66.52%	0.7851	0.7799
5	33.95%	62.96%	37.82%	64.69%	0.8004	0.8060

將所有股票按照 STR\_1M 的值從大到小將其分為一到五個投資組合，通過圖 4-6 我們可以觀察到，短期反轉因子整體上與股票累積淨值呈反向關係，尤其是從 2012 年開始，第五組的投組報酬逐漸超過第四組，因子的分組效果出現真正的次序關係，且前兩組的績效表現明顯優於其他組別。再從具體的績效指標上看，年化報酬率基本隨著因子的下降而上升，最高的組別為第四組，這是因為在 2012 年之前第四組的表現一直優於其他組別，而在年化勝率、年化收益波動度以及風險調整後的收益率上來看，各組表現基本呈現一定的單調性，因此，STR\_1M 最小的組別整體表現上是超過其他組別的，後續選股的過程中將會選擇 STR\_1M 較小的股票進入投組。

#### 6. BETA\_1M:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it}$$

$$\hat{\beta}_i = \frac{\sum_{t=1}^T (R_{it} - \bar{R}_{it})(R_{mt} - \bar{R}_{mt})}{\sum_{t=1}^T (R_{mt} - \bar{R}_{mt})^2}$$

其中  $\hat{\beta}_i$  的計算採用月頻日數據， $R_{it}$  為每日股票的股價報酬率， $R_{mt}$  為每日市場指數的報酬率，使用滬深綜合市場指數收益率來表示。

Beta 來源於最經典的 CAPM 模型，反映了個股股價報酬相對於市場報酬的波動率，屬系統性風險。BETA 的變化在很大程度上其實與基準市場指數相關，根據 CAPM 的公式，當市場處於上漲階段時，BETA 越大的股票表現越好，但市

場處於下跌階段時，BETA 越大的股票則表現越差，由此看來，BETA 在市場行情變化時並不具有

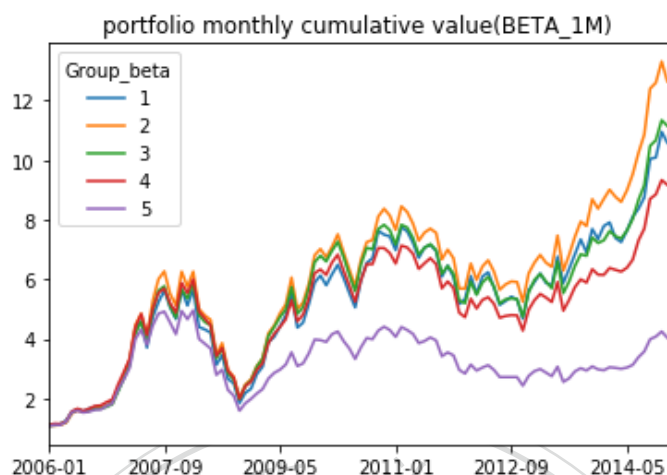


圖 4-7：波動度因子分組回測累積淨值

表 4-7：波動度因子分組回測績效

組別	年化收益率	年化勝率	年化收益 波動度	最大回撤 比率	夏普比率	信息比率
1	29.97%	62.04%	39.67%	67.31%	0.6627	0.6345
2	32.56%	59.26%	39.22%	67.62%	0.7364	0.7235
3	30.73%	58.33%	37.58%	66.21%	0.7198	0.6801
4	27.92%	61.11%	35.72%	66.83%	0.6786	0.5796
5	16.78%	50.00%	33.70%	67.94%	0.3887	0.0341

當我們將股票按照 BETA\_1M 從大到小將其分為一到五個投資組合，通過圖 4-7 我們可以觀察到，在 A 股市場上，一個月的 BETA 因子整體上與股票報酬率是呈正向關係的，尤其是第二組的績效常年都優於其他組別，即使是在 2008 年熊市期間也保有一定的效果。再結合具體的績效指標來看，年化收益率、年化勝率方面，前兩組的年化收益率是顯著大於後兩組的，當然風險的部分也相對較高，但從風險調整後的績效表現夏普比率、信息比來看，仍然優於其他組別。雖然整體來看 BETA 因子的表現並沒有像其他因子那樣具有清晰的區分度，但結合因子有效性的各方檢驗結果以及出於選股因子多元化的目的，仍然將該因子納入多因子選股的考量，且在後續選股過程中將納入 BETA\_1M 較大的股票進入投組。雖

然 BETA 因子的實證表現沒有想像中優秀，但後續會採用不同個數因子組合的形式構建最佳的多因子模型，因此如果後續結果中若該因子沒能提升綜合選股效果仍然有被刪除的可能。

### 4.3 多因子隨機組合績效

目前已經通過因子的有效性分析確定了最終的因子庫，接下來從第二個研究目的出發，通過這 6 個因子之間進行互相組合的方式構建出 63 個投組，並對比 63 個投資之間的績效表現找到最佳的投資組合。按照第三章介紹的打分法，每期會根據各因子的表現對樣本池中的股票進行打分，選出綜合分數排名前 300 的股票等權重構建投資組合。實證結果如下表：

表 4-8：單因子隨機組合投組績效

單因子	投組終值	年化收益率	年化收益波動度	夏普比率	信息比率
ILLIQ	31.11	0.4651	0.3786	1.1313	1.1683
LV(MV)	26.25	0.4377	0.3986	1.0057	0.9945
EPS_G	16.12	0.3619	0.3713	0.8755	1.0019
B/M	14.36	0.3445	0.3879	0.7933	0.9367
STR_1M	14.32	0.3441	0.3903	0.7873	0.7586
BETA	11.55	0.3124	0.4013	0.6868	0.6958

首先是單因子的表現情況，通過表 10 顯示的結果來看，六因子中選股效果最佳的是 ILLIQ 因子，這與因子分層回測部分的數據結果相同，該因子下構建的投組不論在年化收益率還是年化收益波動度方面都優於其他因子，夏普比率、信息比率的表現也都大於 1。

表 4-9：雙因子隨機組合投組績效

雙因子	投組終值	年化收益率	年化收益波動度	夏普比率	信息比率
ILLIQ, LN(MV)	28.08	0.4485	0.3916	1.0512	1.0617
ILLIQ, EPS_G	26.19	0.4374	0.3715	1.0783	1.1545



ILLIQ,B/M	25.5	0.4331	0.3878	1.022	1.1245
ILLIQ,STR	32.26	0.4711	0.3880	1.1194	1.1974
ILLIQ, BETA	28.89	0.4531	0.3925	1.0605	1.1191
LN(MV),EPS_G	21.59	0.4069	0.3874	0.9553	0.9792
LN(MV),B/M	23.06	0.4172	0.3955	0.9619	1.0174
LN(MV),STR_1M	27.31	0.4441	0.4024	1.0121	1.0468
LN(MV),BETA	24.31	0.4255	0.4036	0.963	0.9759
EPS_G,B/M	19.45	0.3906	0.3802	0.9306	1.1973
EPS_G,STR_1M	18.66	0.3842	0.3788	0.9172	1.0071
EPS_G,BETA	17.04	0.3703	0.3877	0.8603	1.0089
B/M,STR_1M	16.21	0.3628	0.3879	0.8404	0.9228
B/M,BETA	17.05	0.3704	0.3981	0.8379	1.0457
STR_1M,BETA	16.30	0.3636	0.4001	0.8168	0.839

從雙因子的表現來看，ILLIQ 與 STR\_1M 組合打分的選股效果是所有投組中表現最佳的，不論是在收益還是風險或是風險調整下的績效表現均優於其他雙因子大分下構建出的投組，同時對比 ILLIQ 單因子以及 ILLIQ 與 STR\_1M 雙因子打分下的投資組合，發現 STR\_1M 的加入確實提升了整體的選股效果，年化收益率以及信息比率分別從 46.51%和 1.1683 上升到了 47.11%和 1.1974，然而其他雙因子的組合打分並沒有出現這種提升的效果。

表 4-10：三因子隨機組合投組績效

三因子	投組終值	年化收益率	年化收益波動度	夏普比率	信息比率
ILLIQ,LN(MV),ESP_G	29.95	0.4590	0.3866	1.0920	1.1375
ILLIQ,LN(MV),B/M	31.35	0.4664	0.3947	1.0885	1.1661
ILLIQ,LN(MV),STR_1M	34.75	0.4833	0.3955	1.1289	1.1931
ILLIQ,LN(MV),BETA	33.50	0.4772	0.3964	1.1110	1.1533
ILLIQ,EPS_G,B/M	27.07	0.4427	0.3820	1.0625	1.2030
ILLIQ,EPS_G,STR_1M	31.51	0.4672	0.3851	1.1178	1.2280
ILLIQ,EPS_G,BETA	27.52	0.4453	0.3862	1.0578	1.1599

ILLIQ,B/M,STR_1M	29.71	0.4577	0.3908	1.0769	1.1983
ILLIQ,B/M,BETA	27.59	0.4457	0.3942	1.0374	1.1626
ILLIQ,STR_1M,BETA	30.32	0.4610	0.3989	1.0635	1.1366
LN(MV),EPS_G,B/M	24.61	0.4275	0.3918	0.9971	1.0978
LN(MV),EPS_G,STR_1M	27.83	0.4471	0.3958	1.0366	1.1032
LN(MV),EPS_G,BETA	24.08	0.4240	0.3968	0.9757	1.0343
LN(MV),B/M,STR_1M	26.53	0.4394	0.3987	1.0097	1.0859
LN(MV),B/M, BETA	25.44	0.4327	0.4002	0.9892	1.0734
LN(MV),STR_1M, BETA	26.19	0.4374	0.4106	0.9756	1.0124
EPS_G,B/M,STR_1M	21.19	0.4039	0.3844	0.9550	1.1447
EPS_G,B/M,BETA	20.50	0.3988	0.3906	0.9268	1.2161
EPS_G,STR_1M,BETA	19.04	0.3873	0.3931	0.8916	0.9919
B/M,STR_1M,BETA	18.98	0.3869	0.3998	0.8757	1.0168

三因子的組合共有 20 組，若從風險調整後的績效表現來看，綜合表現最好的組別為 ILLIQ、STR\_1M、EPS\_G 三因子綜合打分構建出的投資組合。相較於雙因子最優的投組而言，雖然其中有兩組的年化收益率超過了雙因子最佳投組，但由於波動度較大的原因，風險調整後的績效並沒有帶來改善，而由 ILLIQ、STR\_1M、EPS\_G 三因子構建的投資組合雖然年化報酬率略低，但考慮波動度後，整體的績效有相對提升。

表 4-11：四因子隨機組合投組績效

四因子	投組終值	年化收益率	年化收益波動度	夏普比率	信息比率
ILLIQ,LN(MV),EPS_G, B/M	30.89	0.464	0.3904	1.0942	1.187
ILLIQ,LN(MV),EPS_G, STR_1M	38.46	0.5001	0.3941	1.1754	1.2701
ILLIQ,LN(MV),EPS_G,BETA	32.58	0.4727	0.3924	1.1107	1.1699
ILLIQ,LN(MV),B/M,STR_1M	35.6	0.4872	0.398	1.1316	1.2316
ILLIQ,LN(MV),B/M,BETA	33.25	0.476	0.3987	1.1016	1.1846

ILLIQ, LN(MV), STR_1M, BETA	38	0.4981	0.4032	1.1442	1.2213
ILLIQ, EPS_G, B/M, STR_1M	33.62	0.4778	0.3891	1.1332	1.308
ILLIQ, EPS_G, B/M, BETA	31.48	0.4671	0.3915	1.0992	1.2889
ILLIQ, EPS_G, STR_1M, BETA	33.68	0.4781	0.393	1.1229	1.2529
ILLIQ, B/M, STR_1M, BETA	31.23	0.4658	0.4006	1.071	1.2054
LN(MV), EPS_G, B/M, STR_1M	28.31	0.4499	0.3976	1.0389	1.1559
LN(MV), EPS_G, B/M, BETA	29.51	0.4566	0.3976	1.0558	1.2069
LN(MV), EPS_G, STR_1M, BETA	28.33	0.45	0.403	1.0253	1.1059
LN(MV), B/M, STR_1M, BETA	28.67	0.4519	0.406	1.0224	1.1206
EPS_G, B/M, STR_1M, BETA	22.91	0.4162	0.3947	0.9613	1.1917

由四因子共同打分構建的投組中，由 ILLIQ、STR\_1M、EPS\_G、LN(MV) 四因子構建的投資組合的績效表現是所有投組中表現最佳，其年化報酬率達到了 50.01%，相對雙因子以及三因子的投組都有了一定的提升，同時投組年化收益波動度的部分也隨著 LN(MV) 因子的加入出現上升，但綜合考慮收益-風險的情況下，該因子在夏普比率以及信息比率上都有一定的提升，分別為 1.1754 以及 1.2701。

表 4-12：五因子隨機組合投組績效

五因子	投組終值	年化收益率	年化收益波動度	夏普比率	信息比率
ILLIQ, LN(MV), EPS_G, B/M, STR_1M	36.82	0.4928	0.3942	1.1568	1.2769
ILLIQ, LN(MV), EPS_G, B/M, BETA	35.54	0.487	0.3966	1.1351	1.255
ILLIQ, LN(MV), EPS_G, STR_1M, BETA	39.86	0.506	0.401	1.1701	1.2742
ILLIQ, LN(MV), B/M, STR_1M, BETA	39.1	0.5028	0.4031	1.156	1.2689
ILLIQ, EPS_G, B/M, STR_1M, BETA	33.96	0.4795	0.3947	1.1216	1.2976

LN(MV),EPS_G,B/M, STR_1M, BETA	32.24	0.471	0.3999	1.0857	1.2356
-----------------------------------	-------	-------	--------	--------	--------

由五因子的共同打分構建的投組中，ILLIQ、STR\_1M、EPS\_G、LN(MV)、BETA 構建的投組表現最佳，年化收益率達到 50.6%，但是波動度也相應地上升到了 40.1%，夏普比率以及信息比率為 1.1701 和 1.2742，雖然年化收益率的部分超過了四因子中的最佳投組，但是風險調整後的收益率並沒有相應地提升，說明 BETA 因子的加入能有效地提升投組收益率，但與此同時也增加了投組收益的風險，因此 BETA 因子的加入有待考慮。

表 4-13：六因子隨機組合投組績效

六因子	投組終值	年化收益率	年化收益波動度	夏普比率	信息比率
ILLIQ, LN(MV), EPS_G, B/M, STR_1M, BETA	40.07	0.5069	0.3995	1.1767	1.3061

最後是六因子共同打分下構建的投資組合，通過上表可以觀察到，相對於五因子中最佳投組，隨著最後一個 B/M 因子的加入，整體投組的年化收益率得到了進一步的提升，同時年化波動度方面也呈現了一定程度的下降，因此風險調整後的收益率得到了整體上的提升，夏普比率以及信息比率分別為 1.1767 以及 1.3061。雖然在五因子的部分，BETA 因子的加入並沒有給原來的投組帶來更好的選股效果，但再通過對比五因子以及六因子各投組的結果會發現，BETA 因子的加入却為原本只有 ILLIQ、STR\_1M、EPS\_G、LN(MV)、B/M 的投組改善了選股效能，年化收益率從原本的 47.95% 上升到了 50.69%，說明六因子綜合打分之下，各個因子之間的交互作用力的確提升了選股效能。總結以上的實證結果，最終可以確定由六因子構建的投組策略是所有因子組合中的最佳方案。

#### 4.4 樣本外多因子組合績效

通過樣本內數據的實證結果，已經確定六因子的選股模型是最佳的，在預期該模型能持續帶來選股效能的基礎上，我們使用該模型對樣本外數據進行進一步的回測分析，通過對比樣本內外的實證結果來評價該策略的有效性以及穩定性。

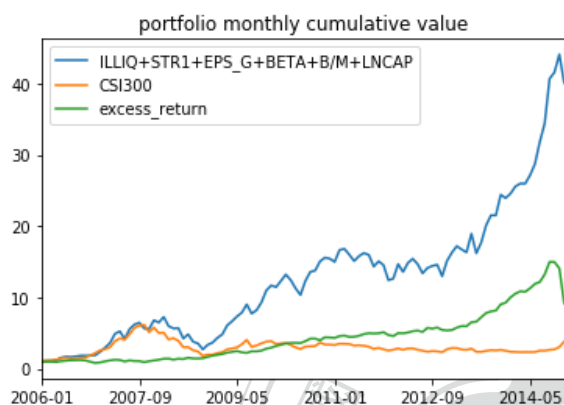


圖 4-8：樣本內打分法回測累計淨值

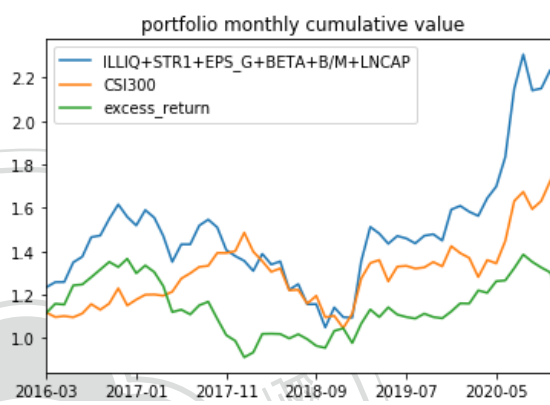


圖 4-9：樣本外打分法回測累計淨值

表 4-14：樣本內外打分法回測績效

2006-01--2014-12	2016-03--2020-12
年化收益率: [0.5069]	年化收益率: [0.1693]
年化收益率波動度: [0.3995]	年化收益率波動度: [0.2345]
夏普比率: [1.1767]	夏普比率: [0.5651]
信息比率: [1.3061]	信息比率: [0.2463]

上圖表分別呈現了樣本內外的實證結果，整體來看六因子綜合打分法模型下的投組期末淨值表現均超過基準滬深 300 指數。再通過對比可以發現，樣本內除了 2007 年期間曾短暫出現基準指數超過投組的現象之外，整體上投組績效均超過基準指數，且呈現較大的差異性。相較於樣本內績效優異的投組表現，樣本外投組的績效則相對一般，年化收益率從 50.69% 下降到了 16.93%，夏普比率以及信息比率也大幅下降，分別從 1.1767 和 1.3061 下降到了 0.5651 和 0.2463，尤其是從 2017-11 月到 2018 年-12 月這段期間，投組模型表現出現低於基準指數的跡象，超額累積淨值部分開始出現大幅下滑。

## 第五章 投資組合歸因分析

整體投資流程分為投前、投中以及投後管理這三個步驟，在本文中，投前主要進行因子有效性分析，找到具有穩定選股效能的因子，並使用多因子隨機組合的形式對樣本內數據進行回測找到最佳的投組策略。而投中的部分是在樣本外進行回測，對比分析因子以及策略在樣本內外的表現情況，最後投後管理的部分則是在投資之後，通過對投資組合進行績效歸因，找到投組獲利或者虧損的來源。

在對投組績效進行歸因分析前，我們需要對樣本外因子表現以及該階段市場發生的事件進行初步瞭解。為了更直觀地觀察各風格因子隨時間的變化情況，我們使用因子分組回測的方式對因子的表現進行以下分析，通過觀察可以發現大多因子都隨時間發生了較為明顯的轉變。

### 5.1 因子風格的轉變：

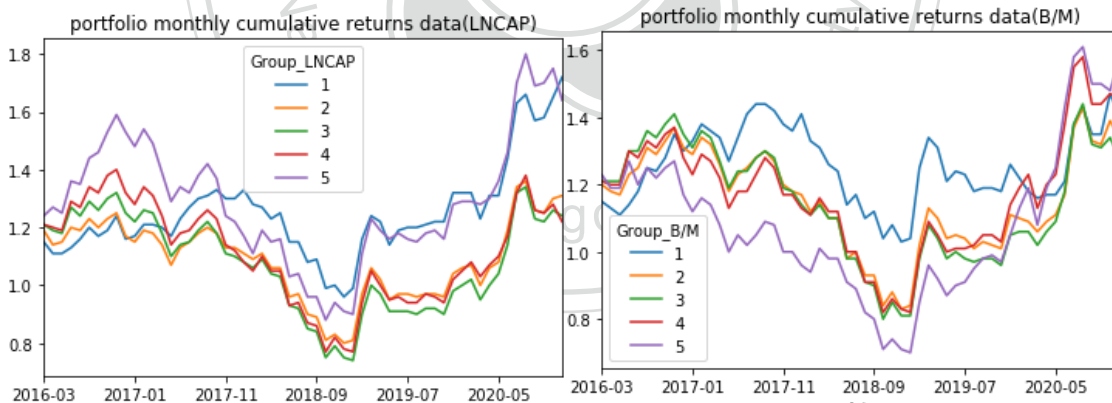


圖 4-10：規模因子分組回測累積淨值  
(樣本外)

圖 4-11：估值因子分組回測累積淨值  
(樣本外)

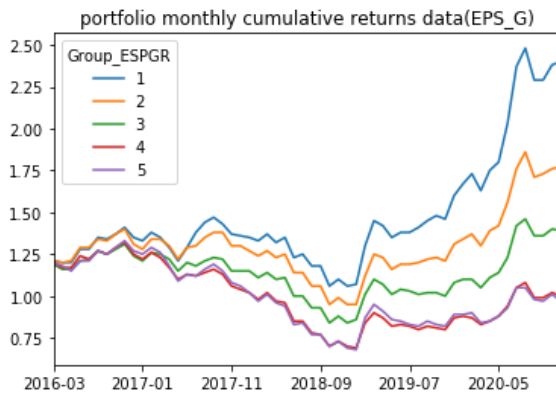


圖 4-12：成長因子分組回測累積淨值  
(樣本外)

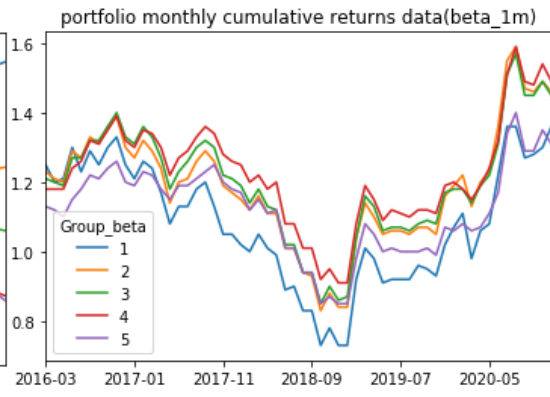


圖 4-13：波動度因子分組回測累積淨值  
(樣本外)

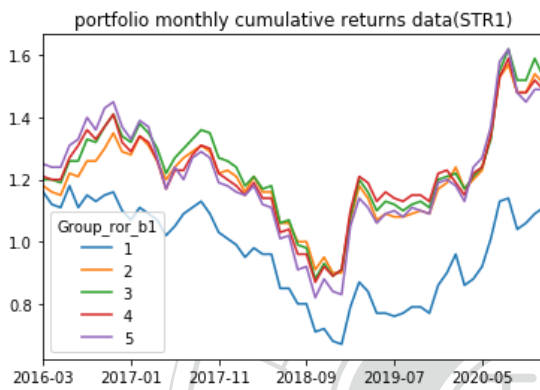


圖 4-14：短期反轉因子分組回測累積  
淨值 (樣本外)

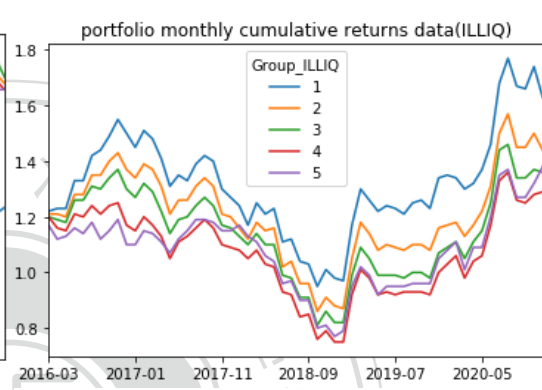


圖 4-15：流動性因子分組回測累積淨值  
(樣本外)

通過各類風格因子在樣本外期間的表現情況可以發現，除了 EPS\_G 以及 ILLIQ 因子分層後整體上仍保持其原有特性之外，其餘因子的風格都發生了較大的改變。首先是市值因子，通過圖 4-10 我們可以觀察到，在 2017 年期間，市值最大的投組累積淨值逐漸上升，並於 2017 年 11 月超過了市值最小的組別，且在 2019 年往後的期間裏也一直表現出與小市值股相當的績效。再結合 2017 年的市場狀態，這段期間裏主要有兩大影響因素：一是 A 股市場的註冊制的深化改革以及退市制度的不斷完善導致 A 股市場“小盤股殼價值溢價”<sup>4</sup>逐漸消退，二是中國金融監督管理部門為防範股災重蹈覆轍，於 2017 年出臺了《資管新規》以整

<sup>4</sup> 未上市公司為避開 IPO 流程，通過買入小公司進行控股的行為稱為“借殼上市”(reverse merger)，Lee et al. (2017)指出，冗長的 IPO 機制導致中國股市存在明顯的“小盤股殼價值溢價”現象。

治資管行業中的杠杆、套利現象，行規的推出使得市場資金有所縮減（市場資金主要集中於機構投資人），並集中於 A 股市場“大而美”的優質股上，因此出現了一定程度“大強小弱”的風格分化現象。再來是 B/M，從圖 4-11 可以看到，2015 股災以及 2020 中美貿易戰過後的一段期間裏該因子出現了反向效果，這可能是經歷熊市後整體市場處於恢復狀態，在所有股票都處於急劇上漲階段時，投資人就容易忽略 B/M 的效果。B/M 效果直到 2017 年開始才逐漸得到恢復，結合 2017 到 2019 年期間的市場狀態，B/M 風格變化與“大市值效應”引起的因素大致相同，加上中美貿易戰給股市帶來的影響，這段期間 A 股市場的價值股板塊的確得到了投資人的重點關注。再來是 STR\_1M 以及 BETA 因子，通過上圖可以瞭解到二者也在 2017 年到 2018 年期間發生了較大的風格改變，除了第一組之外，其都出現了一定程度的反轉效果，說明在市場整體投資情緒表現悲觀的情況下，投資人對持續下跌以及波動度較大的股票並沒有報以希望。

整體而言，可以判斷出上述 6 因子中有 4 個因子在 2017 年以後的 A 股市場上發生了風格轉變，同時投資人也逐漸偏好藍籌、權重以及白馬股這些相對優質的股票，而不太關注市場上的冷門股，因此樣本內有效的六因子模型到了樣本外並沒有達到預期的效果。

接下來使用業界常用的兩大方法對樣本外投組進行歸因分析：一是因子歸因分析，通過回歸分析將投組的收益率映射到各個因子上以觀察各因子對投組的貢獻程度；二是 Brinson 歸因分析，該分析涉及到行業因子，將投組在各行業上的配置以及收益情況與基準指數在各行業的情況做比較，探討超額報酬的主要來源。

## 5.2 因子歸因分析：

對於一個投組來說，因子歸因分析通常採用多因子回歸模型：



$$R^P(t) = \sum_{k=1}^n X_K^P(t) * f_k + u^P$$

實證分析以年頻率的數據進行以上回歸，其中因變數為 $R^P(t)$ 代表 t 期投組報酬率，而自變數 $X_K^P(t)$ 為 t 期投組中各公司 K 因子的平均值。通過回歸可以得到該年度各個的因子回歸係數 $f_k$ ，因此在 t 期，投組在因子 K 上的因子的貢獻值為：

$$RC_K(t) = X_K^P(t) * f_k$$

K 因子每期對於投組的貢獻占比為：

$$RCR_K = \frac{|X_K^P(t) * f_k|}{\sum_{k=1}^n |X_K^P(t) * f_k|}$$

以下是各因子每年平均貢獻占比的實證結果：

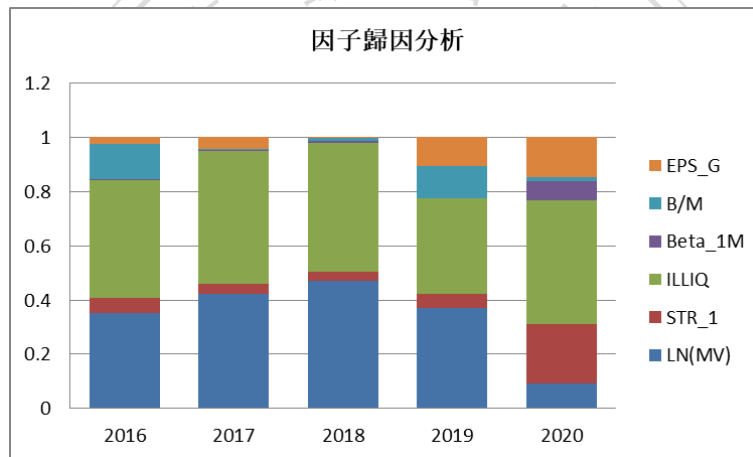


圖 5-1：因子歸因分析結果

通過圖 5-1 可以瞭解到，六因子綜合打分法構建的投組在樣本外的績效表現主要歸因於 ILLIQ 以及 LN(MV) 這兩個因子，五年的平均貢獻度分別為 34.12% 以及 44.15%，而績效歸因效果最弱的因子 Beta\_1M 的平均貢獻度只有 1.79%，這與我們之前在樣本內因子有效性檢驗的結果較符合：ILLIQ、LN(MV) 兩個因子在有效性方面都優於其他因子，而 Beta\_1M 在有效性方面則明顯劣於其他因子。尤其是 2017-2018 期間，市場上符合所有因子條件的股票很少，由六因子綜合打分選出的股票在剩餘因子上並不具有明顯的效果。再結合圖 4-9 以及 4-10 來看，

在 2017 到 2018 期間貢獻程度較大的 LN (MV) 因子在此期間發生了明顯的反轉效果，相對於由藍籌股構成的基準指數，按照“小市值效應”所構建的打分規則使得投組挑選到的股票表現更弱，這可能是造成投組績效相較基準指數較差的原因之一。

### 5.3 Brinson 歸因分析：

由於存在行業板塊輪動效果，A 股市場上時常會出現熱門行業帶動整個板塊集體上漲的情況，因此在對投組進行歸因分析上，我們需要通過對比投組與基準指數在選股配置以及行業配置上的表現來檢視投組的表現。根據 Brinson 的績效歸因理論，我們可以大致將投組的超額收益分為以下兩個方面：

$$\left\{ \begin{array}{l} \text{行業配置能力 (Allocation): } \sum_{i=1}^j (w_j^P - w_j^B) r_j^B \\ \text{選股能力 (Selection): } \sum_{i=1}^j (r_j^P - r_j^B) w_j^B \end{array} \right.$$

其中  $w_j^P$  與  $w_j^B$  分別代表投資組合與基準指數（滬深 300 指數）在行業 j 上的權重，而  $r_j^P$  與  $r_j^B$  分別表示投資組合與基準指數在行業 j 上的收益率。因此根據上圖的公式所呈現的，投組與基準指數之間的行業權重差異乘上基準指數在該行業上的報酬率就代表行業配置上產生的超額收益；而投組與基準指數之間的收益率差異乘上基準指數在該行業上的權重則代表投組在選股能力上產生的超額收益。

以下是 Brinson 歸因分析的實證結果：

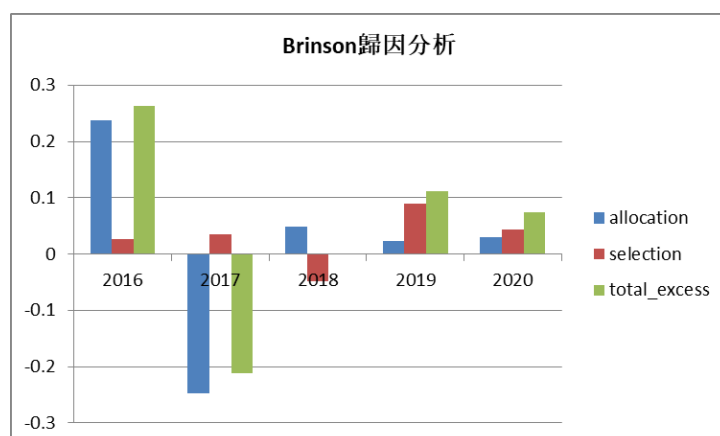


圖 5-2：Brinson 歸因分析結果

表 5-1：投組重倉行業

投組重倉行業					
排名	2016	2017	2018	2019	2020
1	化工	化工	化工	化工	化工
2	機械設備	機械設備	機械設備	機械設備	機械設備
3	醫藥生物	醫藥生物	建築裝飾	建築裝飾	公用事業

表 5-2：基準指數沖倉行業

滬深 300 指數重倉行業					
排名	2016	2017	2018	2019	2020
1	非銀金融	非銀金融	非銀金融	非銀金融	非銀金融
2	銀行	銀行	銀行	銀行	銀行
3	房地產	食品飲料	食品飲料	食品飲料	食品飲料

通過圖 5-2 我們可以觀察到 2016 年到 2017 年期間，主要是投組的行業配置能力主導投組的超額報酬，尤其是 2017 年投組發生的虧損主要源於行業配置問題。為了探討其中的緣由，我們從股市狀態進行分析，首先，可以通過圖 4-9 的回測結果觀察到 2017 年 11 月基準指數的累積淨值首次超過了投組，通過 5.1 節的分析，我們已經瞭解到這段期間 A 股市場的投資人偏好轉向了“藍籌股”，如表 5-1 以及 5-2 所示，投組重倉的行業主要為於化工、機械設備、醫藥生物，這些生產製造行業中的公司多為民營企業，而滬深 300 指數中的行業主要集中在金融

機構以及食品業等傳統類的產業，這些行業公司數量較少且多數為國有企業，因此投資風險更小，也更容易吸引大額資金。

從圖 5-2 可以瞭解到，六因子選股模型在整體選股能力上為正向表現，說明投組在各行業基本都能選到表現較優的股票，但 2018 年則出現例外結果，結合該期間樣本內的回測結果（見圖 4-9），可以發現投組與基準指數都出現了一定程度的下跌。這段期間 A 股市場受到了中美貿易戰的衝擊：美國總統川普於 2018 年 3 月發布公告，要求其貿易代表對中國進口的商品徵收關稅，根據美國的對華的關稅條例，貿易戰主要給中國出口產業以及製造業帶來較大壓力，出口產品包括：機械設備、紡織品、化工、金屬製品等。通過表 5-1 可以發現每年投組中的化工業以及機械設備行業權重占比分別穩居第一、第二名，這兩個行業作為中國出口美國的主要產業，在貿易戰期間一定會受到相應的影響，因此從 2018 年開始，雖然行業配置效果為正，但由行業分配帶來的超額報酬已經逐漸低於前兩年的效果；且 2018 年期間由於多因子模型下各行業選到的小型冷門股公司在貿易戰期間更容易受到重挫，所以投組在選股能力上並無獲得超額報酬。

## 第六章 結論與展望

### 6.1 研究結論：

本文從研究目的出發，在規模類、估值類、質量類、成長類等常見的 10 大類因子中選擇了 18 個因子對中國 A 股市場進行了實證研究。首先，通過業界常用的回歸 t 檢定以及 IC、IR 值對各個因子進行有效性與穩定性方面的檢定，再通過計算大類因子之間的相關矩陣檢驗因子之間的獨立性，最後依據各因子值對股票池進行分組回測檢驗了因子效果的單調性，最終篩選出 6 個風格因子，分別是規模因子：總市值對數 (LN(MV))、估值因子：賬面市值比 (B/M)、成長因子：每股收益成長率 (EPS\_G)、流動性因子：ILLIQ、短期反轉因子：過去一個月股價報酬 (STR\_1M)、波動度因子：過去一個月 BETA 值 (BETA\_1M)。

識別出有效因子後，再基於打分法對六因子進行隨機組合，發現在單因子的基礎上不斷加入其餘因子可以在一定程度上提升選股效果，且樣本內績效最佳的投資策略為由六因子共同排序打分所構建的策略，夏普比率與信息比率分別達到 1.1767 以及 1.3061。為了檢驗六因子打分法的穩定性，同樣將該策略應用於樣本外數據進行回測，雖然投組的累計淨值在整體上仍優於基準指數，但在 2017-2018 年間由於基準指數的略微強勢造成投組超額收益率出現大幅削減。

為了深入探討樣本外六因子投資策略失效的緣由，本文從因子風格轉變及其原因出發，使用因子分層回測的方式對樣本外數據進行檢視，發現除了 ILLIQ 以及 EPS\_G 因子之外，其餘因子在風格上均發生了較大轉變，例如：在 2017 到 2018 年期間，隨著 A 股市場法規制度的深化改革，由規模因子所帶來的“小市值”效應逐漸被“大市值”所取代，其他因子同樣在 2017 年度發生了不同程度的反轉

效果。最後，使用業界最常用的兩大方法對樣本外投組報酬進行歸因分析，一是風格因子歸因，通過各因子的年平均貢獻程度可以發現投組的報酬主要來自於規模因子以及流動性因子，結合這兩個因子在 2017 到 2018 年發生的反轉效果，以原先選股邏輯構建的投組策略的確不能展現較優的選股效能，因此造成投組績效被基準指數反超的現象。二是使用 Brinson 歸因分析，將投組與基準指數在各行業的配置以及報酬進行比較以探討投組績效的來源，結果顯示在 2017 期間，投組的相對虧損主要歸因於投組在行業上的配置情況，由於投組一直以來主要集中於化工以及機械設備等行業，相對於滬深 300 所持的傳統產業而言多為規模較小的民營企業，結合這段期間的投資人對指數“藍籌股”的追捧現象，投組在行業上配置能力稍遜一籌。而在 2018 年期間，中美貿易戰的發生導致中國進出口行業遭受的虧損首當其衝，機械設備、化工等作為中國重要的對外經貿產業遭受衝擊，因此該年度在行業配置上的優勢已經不若 2016 年明顯了，且投組在選股能力上首次出現虧損，說明在整體市場景氣下滑的情況下，由六因子綜合打分選出的股票相對市場缺乏抗跌能力。

根據以上的分析，我們可以結合研究目的得出以下結論以及投資建議：

- 1、過往文獻中提到的經典因子效果的確有體現於中國 A 股市場。例如“小市值效應”、“價值股效應”以及“非流動性溢價效應”等在 A 股市場都有一定程度的呈現，也說明這些經典因子，尤其是在 2015 年股災發生前都具有良好且穩定的選股效能。
- 2、通過 63 組多因子隨機組合，可以發現特定因子之間的交互作用的確可以提升整體的選股績效，尤其是在風險調整後的報酬率層面。本文中，由六因子綜合打分法構建的投組在夏普比率以及信息比率上都有更好的表現，說明多因子模型下，

投組不僅能體現多元化的風格特徵，在績效上也有更好的改善。因此，建議投資人在進行投資策略規劃時，不僅需要考量因子有效性，同時也要關注因子之間的交互作用，以達到收益-風險的最佳化。

3、通過觀察 A 股市場在經歷 2015 年股災前後的投組表現可以發現，雖然六因子在股災發生前的 9 年裏一直保持穩定的選股效能，風險調整後的報酬可以持續打敗大盤，但在股災後的期間裏，尤其是 2017 到 2018 年間，由於法規限制、經濟景氣等各方面因素造成大多數風格因子發生轉變，尤其像市值這類會顯著影響股票特徵的風格因子，一旦發生反轉效果將使投組報酬大幅下降。因此，建議投資人在使用過往表現較佳的有效因子構建投組時，需要格外注意市場動向以及因子效能的變化。

## 6.2 論文不足與展望：

1、本文在多因子打分法過程中採用的因子權重為等權重，屬靜態權重法，雖然該方法在操作上相對便捷，但無法體現各個因子有效性、穩定性、相關性等性質在時間序列上表現出的差異化，因此在未來的研究中可以通過動態調整因子權重的方式，例如對因子使用 IC、IR 滾動加權評分方式來提升各個因子在投組建構過程中的選股擇時能力。

2、考慮到中國 A 股市場股票衆多，本文選用滬深 300 指數作為基準以反映市場整體狀況，因此在模型的選股基礎上同樣選擇了 300 檔股票構建投組進行績效的比較。但在股票型基金等實際投資應用中，為考量基金管理成本通常會控制持倉股數於 30 至 50 檔，因此在未來研究中可以考慮模擬基金的運作形式進行選股流程，並使用不同的市場基準指數作為績效衡量之標杆來驗證選股模型的適用性。

3、在樣本期間的考量上，為探究股災發生前後各類因子在股市表現上的穩定性，本文使用中國 2015 年股災發生時期作為切割期間，對前後數據做樣本內外績效的分析比較。這雖然在一定程度上隔絕了股災對因子表現力的干擾，但在實際應用中存在一定的局限性，因此在後續的研究中可以考慮對股市狀態做區分，研究不同市場狀態下的各因子表現情況，並針對不同市場階段擬定合適的交易策略。

4、因為收集到的備選因子個數較少，本文並沒有使用機器學習模型進行選股上比較分析，但隨著目前機器學習模型在量化交易領域的開發以及測驗，由大數據驅動的算法越來越受到投資界的關注，在因子重要性識別以及股票收益率預測都具有重要貢獻作用，因此後續研究可以考量使用機器學習方法進行模型優化。





## 參考文獻

- [1] Amihud, Y. and H. Mendelson, (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2), 223-249.
- [2] Amihud, Y., (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31-56
- [3] Arbel, A., Strebel, P. (1982). The neglected and small firm effects. *Financial Review*, 17(4), 201-218.
- [4] Banz, Rolf W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1), 3-18.
- [5] Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient market hypothesis. *Journal of Finance*, 32(3), 663-682
- [6] Blitz, D., & Van Vliet, P. (2007). The volatility effect: lower risk without lower return. *Journal of Portfolio Management*, 34(1), 102-113
- [7] Chan, K.C. and Chen, N. (1991). Structural and return characteristics of small and large firms. *Journal of Finance*, 46(4), 1467-1484.
- [8] Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57-82.
- [9] Chen, L, Novy-Marx, R, Zhang, L. (2010). An alternative three-factor model. Working paper. Washington University in St. Louis.
- [10] De Bondt, W. F. M., R. H. Thaler. (1985). Does the stock market overreact?. *Journal of Finance*, 40(3), 793-805.
- [11] Fama, E. F. and K. R. French, (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2), 427-465.
- [12] Fama, E. F. and K. R. French. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.
- [13] Haugen, R., W. Baker, (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, 41(3), 401-439.
- [14] Jegadeesh, N. and S. Titman. (1993). Return to buying winner and selling losers: implication for stock market efficiency, *Journal of Finance*, 48(1), 65-91.
- [15] Lee, C. M. C., Y. Qu, and T. Shen (2017). Reverse mergers, shell value, and regulation risk in Chinese equity markets. Working paper. Stanford University.
- [16] Reinganum, Marc R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics*, 9(1), 19-46.
- [17] Rosenberg, B., Reid, K. and Lanstein, R. (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management*, 11(3), 9-16.

- [18] 孔東民.“流動性風險與資產定價：來自中國股市的證據”.*南方經濟*.2006(3). 91-107
- [19] 王永宏,趙學軍.“中國股市‘慣性策略’和‘反轉策略’的實證分析”.*經濟研究*.2001(6).56-89
- [20] 王春峰,韓冬,蔣祥林.“流動性與股票回報：基於上海股市的實證研究”.*經濟管理*.2002(24)58-67
- [21] 朱世武,鄭淳.“中國資本市場股權風險溢價研究”.*世界經濟*.2003(11).62-70
- [22] 李斌,馮佳捷.“中國股市的公司質量因子研究”.*管理評論*. 2019(3).14-26
- [23] 肖軍,徐信忠.“中國股市價值反轉投資策略有效實證研究”.*經濟研究*.2004(3).55-64
- [24] 林曉明.“華泰多因子模型體系初探”.*華泰證券研究報告*.(2016).
- [25] 胡昌生,高玉森.“分析師情緒會影響股票價格嗎?”.*金融研究*. 2013 (6).193-206
- [26] 柴宗澤.“證券流動專題研究(一)：非流動性因子 ILLIQ 的選股效力”.*安信證券研究報告*.(2013).
- [27] 郭少石.“基於多因子模型的量化選股研究”.*天津大學金融學系碩士論文*.(2016).
- [28] 彭慧文.“中國股市個股非流動性、換手率和短期反轉的關係”.*金融發展研究*.2013(9). 8-14
- [29] 費洋.“A 股市場多因選股量化模型構建及其檢驗”.*浙江大學金融學系碩士論文*.(2018).
- [30] 楊雨晴.“系統波動率還是特質波動率?”.*廈門大學金融學系碩士論文*.(2020).
- [31] 雷璇.“基於回歸和打分法的因子選股模型對比分析”.*大連理工大學金融學系碩士論文*.(2019).
- [32] 熊海斌,楊帆.“股息率與股票收益:基於中國股市經驗證據的研究”.*商業研究*. 2013 (8).151-156
- [33] 蔡慶豐,楊侃.“是誰在“捕風捉影”：機構投資者 VS 證券分析師”.*投資研究*. 2018 (1).193-206
- [34] 蘇冬蔚,麥元勛.“流動性與資產定價：基於我國股市資產換手率與預期收益的實證研究”.*經濟科學*.2004(2).95-105