國立政治大學金融學系研究所

碩士學位論文

10-K 財報情緒與多因子模型對超額報酬之影響：以美國股市為例

The Impact of Form 10-K Sentiment and Multi-Factor Model on Excess Return - Evidence from US Stock Market

國立政治大學 National Chengchi University

指導教授：林 士 貴 博士

江 彌 修 博士

研究生：蔡 承 恩 撰

中 華 民 國 一 百 零 九 年 一 月

DOI:10.6814/NCCU202000109

10-K 財報情緒與多因子模型對超額報酬之影響：以美國股市為例

學生：蔡承恩　　指導教授：林士貴 博士

江彌修 博士

政治大學金融學系

摘要

過去因子模型對於超額報酬之探討主要著重於公司營運與財務指標，然而，近年來隨著市場中的非結構化資料逐漸增加，如何透過分析各種不同的資料引入更多資訊成為重大議題，本研究為探討文字資料運用於財務領域的成效，使用情感分析技術 VADER，分析美國公司 10-K 財報中的文字資訊，計算出財報情緒分數，並驗證若以該分數做為建構投資組合之標準能否獲取超額報酬，並建立基於財報情緒分數之情緒因子，加入 Fama-French 三因子模型，使用此四因子模型針對美國股市進行超額報酬橫斷面變異進行分析。實證結果顯示使用財報情緒分數做為建立投資組合之標準在 Fama-French 三因子模型下確實能帶來異常報酬，尤其以負面情緒分數做為標準會有更明顯的趨勢，而更進一步使用基於財報情緒分數做為基準建構的情緒因子，更能使因子模型的解釋力上升。

關鍵字：情感分析、因子模型、VADER

DOI:10.6814/NCCU202000109

The Impact of Form 10-K Sentiment and Multi-Factor Model on Excess Return - Evidence from US Stock Market

Student：Tsai, Cheng-En Supervisor：Dr. Lin, Shih-Kuei

Dr. Chiang, Mi-Hsiu

Department of Money and Banking, National Chengchi University

ABSTRACT

In the past, the discussion of excess returns in factor models mainly focused on company operations and financial indicators. However, in recent years, as unstructured data has gradually increased, how to introduce more information through analysis of various data has become a major issue. In order to discuss the effectiveness of the application of text data in the financial field, we use VADER to analyze the text information in the 10-K financial reports of American companies, calculate the financial report sentiment score, and verify whether the score can be used as the criterion for constructing an investment portfolio and obtain excess returns. Further, we establish an emotional factor based on financial report sentiment scores and combine with Fama-French three-factor model, and use this four-factor model to analyze the cross-sectional variation of excess returns for the US stock market. Empirical results show that using financial report sentiment scores as the criterion for establishing portfolios can indeed bring abnormal returns under the Fama-French three-factor model. Especially with negative sentiment, we can see a clear trend with excess returns. And using the form 10-K sentiment based factor can improve the explanatory power of the multi-factor model.

Keywords: Sentiment analysis, Factor model, VADER

DOI:10.6814/NCCU202000109

目錄

第一章 緒論.....1
第一節 研究動機與背景.....1
第二節 研究目的.....2
第二章 文獻探討.....4
第一節 因子模型相關研究.....4
第二節 文字資料相關研究.....5
第三節 文本情感分析結合因子模型相關研究.....6
第三章 研究方法.....8
第一節 因子模型.....10
第二節 文本情感分析方法.....11
第三節 財報情緒因子與四因子模型.....15
第四章 資料來源與處理過程.....17
第一節 資料來源.....17
第二節 規模與帳面市價比因子.....17
第三節 10-K 財報與文字資料前處理.....18
第五章 實證分析結果.....20
第六章 結論.....37
參考文獻.....39
附錄.....41

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

表次

表 1　VADER 評分範例 13

表 2　依各情緒分數分五組之各投資組合超額報酬統計 22

表 3　依各情緒分數分五組之各投資組合 CAPM 模型統計量 25

表 4　依$Senti_{avg}$與$Senti_{Li}$分五組之各投資組合三因子模型統計量 28

表 5　依$Senti_{pos}$與$Senti_{neg}$分五組之各投資組合三因子模型統計量 29

表 6　依$Senti_{avg}$分五組之各投資組合四因子模型統計量 31

表 7　依$Senti_{pos}$與$Senti_{neg}$分五組之各投資組合四因子模型統計量 32

表 8　依公司規模與帳面市值比分二十五組之各投資組合PMN_{avg}四因子模型統計量 34

表 9　各因子之相關係數 36

表 10　依公司規模與帳面市值比分二十五組之各投資組合PMN_{pos}四因子模型統計量 41

表 11　依公司規模與帳面市值比分二十五組之各投資組合PMN_{neg}四因子模型統計量 43

DOI:10.6814/NCCU202000109

圖次

圖 1　研究流程.....3

圖 2　實驗流程.....9

國立政治大學
National Chengchi University

DOI:10.6814/NCCU202000109

第一章 緒論

第一節 研究動機與背景

如何有效的進行投資一直以來都是資本市場亟欲追尋的問題，自 Fama-French 三因子模型問世以降，產學業界皆不斷的在尋求能夠獲取超額報酬的投資組合與風險因子，隨著自然語言處理 (Nature language processing) 的技術成熟，研究也逐漸拓展至非結構化資料。

大數據時代資料分析所面臨的難處並不只有其所佔用的儲存空間極大，在由不同裝置與行為產生的非結構化資料的面前傳統統計量化工具顯得窒礙難行，近年來隨著資料探勘 (Data mining) 的研究逐步發展，我們發現了許多利用機器學習與統計方法來從非結構資料中萃取資訊的方法。又因為社群網路的普及，讓使用者能夠在網路上發表自己的言論與意見，使得以文字資料為目標的以自然語言處理與文本情感分析 (Sentiment analysis 或稱意見挖掘 opinion mining) 成為近年來最熱門的議題，應用於在財務金融領域如股價預測、總體經濟分析更是越發火熱。

所謂的大數據並不只是譁眾取寵的流行詞彙，更是我們正在面臨的難題，該如何在極為龐大且複雜的資料中汲取出珍貴的資訊並結合傳統分析方法運用在財務分析與投資組合管理也是目前學術、實務界所重視的議題。

DOI:10.6814/NCCU202000109

第二節 研究目的

Markowitz (1952)開創了現代投資組合理論，其以均質-方差分析的數學框架探討風險趨避投資人的最佳投資組合，並得出了通過投資組合可以有效降低風險的結論，然而模型中的每檔股票的風險、報酬率與相關係數之期望值在實作上並不容易估計，也使得實際使用該模型建構投資組合相對困難。

基於投資組合理論之上，Sharpe (1964)與 Lintner (1965)等人提出了資本資產定價模型(Capital Asset Pricing Model, CAPM)，CAPM 獲得了極大的成功，該論文主要強調個股投資報酬與市場風險之關係，也就是市場風險貼水 β，儘管實務上有許多市場異象(Market anomaly)無法使用該理論解釋(Fama and French, 2004)，但因其模型的簡約性仍常常被投資者做為建構衡量投資績效的方法。

針對其無法解釋的市場異象，有許多學者做出更進一步的研究，批評者認為 CAPM 模型過於簡化，忽略了某些風險因子，其中 Fama and French (1993)基於其於前一年做出的研究(Fama and French, 1992)發展出了規模因子與帳面市值比因子，並將其結合市場風險因子進而建立了著名的三因子模型。

然而，上述研究的目標皆為公司的會計項目並未考慮到財報中的非結構化文字資訊，事實上文字資料在股市的價格預測中，不論是基於字典的模型又或者是統計學習為主的模型都有研究指出文字資料對於價格預測是有助益(Gentzkow, Kelly, and Taddy, 2019)。

本研究將導入自然語言處理中的文本情感分析從公司的 10-K 財報的文字資料中萃取資訊並建立情緒因子，並將其引入多因子模型，驗證加入文字資料是否能夠捕捉文字中被傳統因子模型忽略的資訊，並驗證其是否能夠有效增強因子模型的解釋力。

本研究分為六章，第一章為研究動機與目的，第二章分三節分別探討因子模

DOI:10.6814/NCCU202000109

型、文字探勘、結合文本情感分析與財務分析相關之研究，第三章詳述本研究使用到的研究方法與實驗過程，第四章敘述資料來源與處理過程，第五章為實證分析，最後第六章為結論，研究流程如下圖。

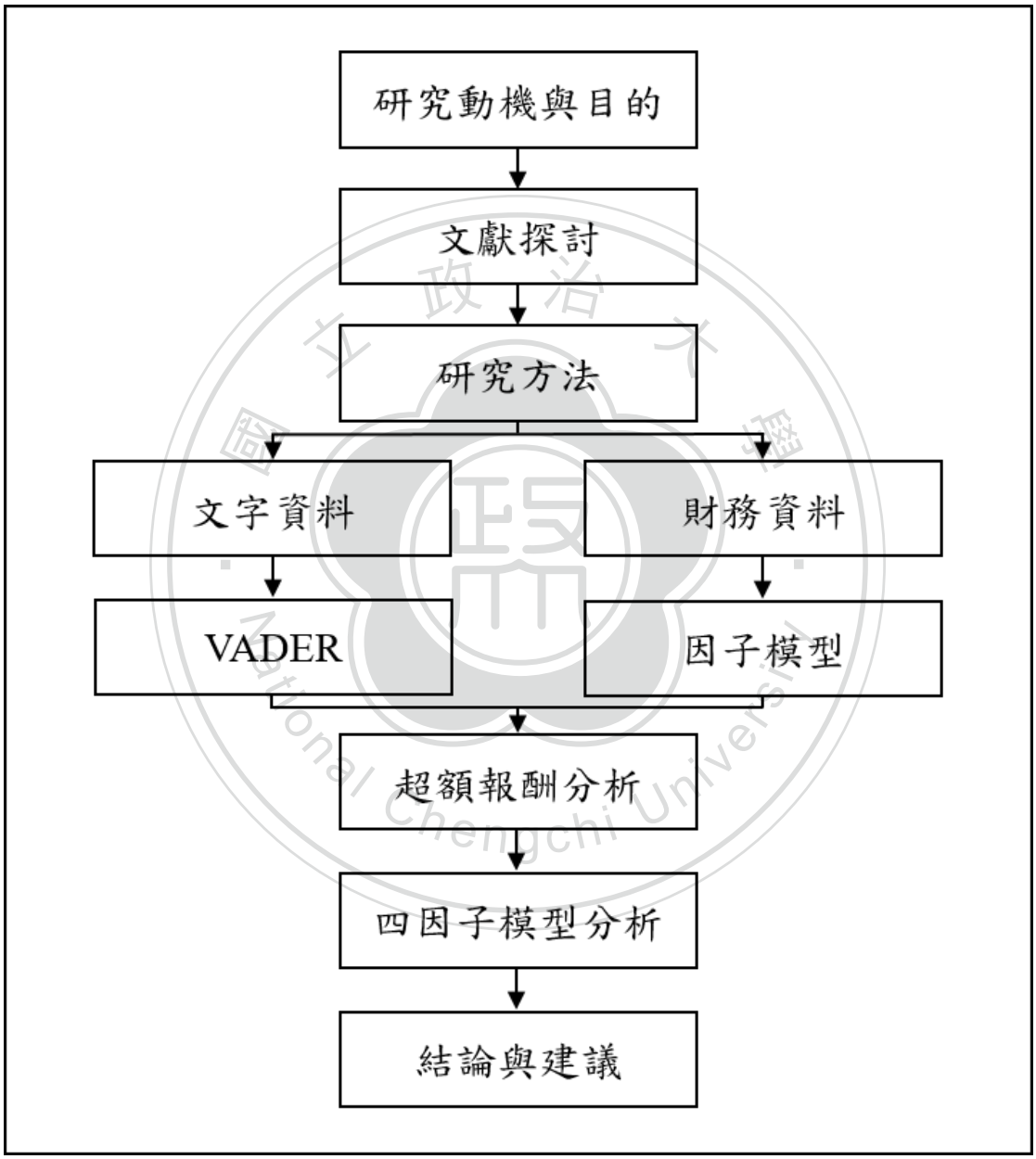

圖 1　研究流程

DOI:10.6814/NCCU202000109

第二章 文獻探討

第一節 因子模型相關研究

自資本資產定價模型 CAPM (Lintner, 1965; Sharpe, 1964)問世後，不論在學界或是實務界都有廣泛的使用，其影響不言而喻，然而此模型在許多時候並無法解釋市場異象(Schwert, 2003)，因此後續也有許多研究試圖找出對股票報酬橫斷面變異(Cross-sectional variation)具有解釋力的因子。

Banz (1981)指出在 NYSE 的上市公司中存在規模效應，透過線性模型觀察公司的市值大小與市場上平均公司市值大小的差距與投資報酬率之關係，發現小公司的風險調整後收益率大於大公司，在 Reinganum (1981)的研究中使用 NYSE 與美國證交所的上市公司做為樣本也顯示相同的結果，CAPM 模型在預測小公司的報酬率時會有低估的現象，但他們並未將規模效應的來源歸咎於特定原因，而是認為模型中的偏誤來自被忽略(Misspecified)的因子。

Ball (1978)的研究發現若公司發布之盈餘超乎預期則公司股價會有超額報酬，反之則會有負面效果，得出每股盈餘可被視為預期報酬的代理變數，且此變數在定價模型中應是解釋超額報酬的重要變數，Basu (1983)延續該研究重新驗證每股盈餘與公司規模與美國股票市場中的超額報酬的關係，實證結果顯示，高每股盈餘的公司的平均報酬率會高於低每股盈餘的公司，此一結果就算是在去除規模效應後仍然存在。

Fama and French (1992)使用公司規模、槓桿率、本益比、帳面市值比等公司資訊對其回報率之關係，發現帳面市值比、規模與市場超額報酬對投資組合超額報酬有顯著且穩固的關係，次年依其結果對美國證券市場進行實證，將資料依所選因子分組並使用橫斷面的迴歸模型，證實了美國股市確實存在規模與帳面市值比效應，此即著名的 Fama-French 三因子模型(Fama and French, 1993)。

DOI:10.6814/NCCU202000109

儘管 Fama-French 三因子模型成功的解釋了當時市場上許多的市場異象，但仍舊是有些市場異象並未被捕捉，此後仍有許多學者試圖嘗試捕捉新的因子並加入模型中，如 Carhart (1997)針對動量效應的市場異象分析，利用前一年基金的績效分組建立動量因子，並使用橫斷面回歸分析，發現加入動量因子能夠提升模型解釋力，有較高四因子的基金也會有較高的α。

第二節 文字資料相關研究

使用文字來預測股市並不是一個近年來才有的概念，早在 20 世紀初就有學者嘗試利用前華爾街日報主編對市場新聞的看法作為投資或放空道瓊指數的評判標準，在當時此一投資策略並未成功擊敗大盤，反而是較大盤差 3.5%(Cowles 3rd, 1933)，在這篇研究中投資策略極大的依賴金融專業人士對文字資訊的判斷，且較難將其中主觀的因素剃除，然而受限於當時的科技，在當時若要使用新聞資訊作為財務預測的變數，是必須要依賴專業判斷的。

在近代資訊科技逐漸成熟，文字資料逐漸擺脫必須仰賴於專業人士的判斷的限制，Tetlock (2007)使用新聞資料預測股市報酬，以預先設定好的字詞分類辭典 Harvard Psychosociological Dictionary(Stone, Dunphy, and Smith, 1966)針對華爾街日報的專欄「Abreast of Market」計算出字詞頻率並使用主成分分析建構出負面情緒因子，加入 VAR 模型中嘗試預測道瓊指數的日報酬率，發現此因子能夠有效預測股市的下跌，利用該因子的極端值也能發現與交易量之關係，此篇論文成功的利用電腦的計算力將文字資訊與財務預測做連結，為後續研究帶來了極大的影響。

Loughran and McDonald (2011)則進一步檢視 Harvard Psychosociological Dictionary 是否適用於財務領域，發現該字典中有 73.8%的負面字詞是錯誤分類的，如 Depreciation、cost、tax、liability 等，這些字詞在財務上只是敘述性的辭

DOI:10.6814/NCCU202000109

彙並不能代表負面情緒，因此他們針對金融上常用的字詞建立了新的字典，並利用該字典所產生的情緒因子加入 Tetlock (2007)中的因子作為控制變數，使用線性回歸預測股較超額報酬，實驗結果證明該字典能夠有效的提升模型的預測力。

Kothari, Li, and Short (2009)依產業別與資料的可用性選取美國 887 間上市公司，並使用 GI 針對其公司 10-K 財報、新聞做情感分析並利用因子模型研究文字情緒與公司風險的關係，結果顯示公司揭露其風險的語調影響揭露的訊息與資金成本的關係，在資料來源並非公司本身時，若語氣呈現悲觀則會增加公司的資金成本。

在 Campbell, Chen, Dhaliwal, Lu, and Steele (2014)的研究則更進一步顯示，公司的訊息揭露的悲觀程度與基於市場衡量的公司風險評估呈現正相關，此一研究並不止於非公司自主發布的資料，就算納入公司自行發布的財務資訊，也是有同樣的結果，這說明了就算隱匿公司風險對於公司經理人有利益，但因應法規經理人還是必須揭露風險，這使得公司自主發表的風險揭露是充滿著資訊的。

第三節 文本情感分析結合因子模型相關研究

Li (2006)則計算各公司的 10-K[1]報表中「Management Discussion and Analysis」的風險詞彙如 risk、risky、uncertainty 等的出現次數，並使用風險詞彙出現次數建構出不確定性情緒因子，發現不確定性情緒因子與隔一年度的收益呈現負相關，更將此因子結合三因子模型，有效地提升α超過 10%。

Loughran and McDonald (2011)同樣使用 10-K 報表中的「Management Discussion and Analysis」(MD&A)以及全文作為文字資料，但其使用研究中提出之字典計算詞彙出現次數並建構情緒因子，配合三因子模型(Fama and French,

[1]美國證券管理委員會規定公司須正式提交之財務報表

DOI:10.6814/NCCU202000109

1993)，建構高情緒減低情緒投資組合，儘管得出的α皆為正，但結果卻不顯著。

基於情感分析對於財報資料的研究方興未艾，然而目前結合因子模型與財報情緒因子的研究並未有統一之結論，主要困難之處在於如何有效且正確的計算文本的情感情分數，使用不同的分析模式計算財報情緒對於研究結果的影響甚大，本研究使用近年來廣為使用的 VADER 文本情感分析模型 (Hutto and Gilbert, 2014)做為文本情感分析工具。

VADER 除了使用正負面字典外，更對每個字詞有不同權重，也考慮了語法結構對於整體情緒之影響，對於文本情感分析的解構更為細膩，使得從文字資料萃取出之資訊更具代表性，然而 VADER 是以句子為層級做情感分析的工具，該如何在得出文章中所有句子的情感分數後進一步計算財報之分數亦是本研究之重點。

本研究將藉由 VADER 計算各公司年度財報的情感分數，並驗證財報情緒分數是否能創造超額報酬，並建構情感因子、建立更完善之基於財報情緒之多因子模型，完整研究流程與方法請見第三、第四章。

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

第三章 研究方法

本研究使用基於字典的文本情感分析方法VADER針對美國上市公司每年因應美國證券交易委員會規定所發布之10-K財務報表中「Management Discussion and Analysis」 (MD&A)文字資料建構出各公司各年度之財報情緒分數，再利用情感分數建構出投資組合並透過CAPM與Fama-French三因子模型驗證高低情感分數投資組合的超額報酬是否有顯著之差異。

其中建立投資組合之方式採用與Fama and French (1993)之方式，在每年六月底，將樣本中所有公司以其前一年度之財報情緒分數做排序，接著將公司依序分入投資組合中直到該投資組合達樣本中所有公司之總市值的五分之一，此即最高財報情緒分數之投資組合，接著再將之後的公司放入新的投資組合直到該投資組合達樣本中所有公司之總市值的五分之一，此即次高財報情緒分數之投資組合，重複五次後即可得到依財報情緒分數分組之五組投資組合，然而因事實上公司並不可切割一次放入兩投資組合，因此各投資組合市值僅約略等於總市值之五分之一。

因VADER是句子層級之文本情感分析工具，本研究中使用了三種不同方式總和財報中由VADER所產生之各句分數，分別是使用簡單平均作為總和方式的Senti_{avg}、計算財報中正面句子占總句數比例的Senti_{pos}、計算財報中負面句子占總句數比例的Senti_{neg}，在使用不同方式計算出分數後再使用三種方式分別建立投資組合，並利用CAPM與三因子模型觀察不同方式所計算出之財報情緒分數對於超額報酬之影響，為比較不同文本情感分析模型對於超額報酬之影響，本研究也使用了Li (2006)中提出之不確定性分數，本篇稱其為Senti_{Li}。

最後，為驗證加入財報情緒因子對因子模型之橫斷面變異解釋力，本文使用本文提出之三種分數分別建立財報情緒因子PMN_{avg}、PMN_{pos}、PMN_{neg}，並分別加入Fama-French三因子，檢視其解釋力。

DOI:10.6814/NCCU202000109

詳細研究方法之細節將分為三節在本章，分別介紹因子模型、文本情感分析方法、本篇提出之財報情緒因子與四因子模型，下圖為實驗流程圖。

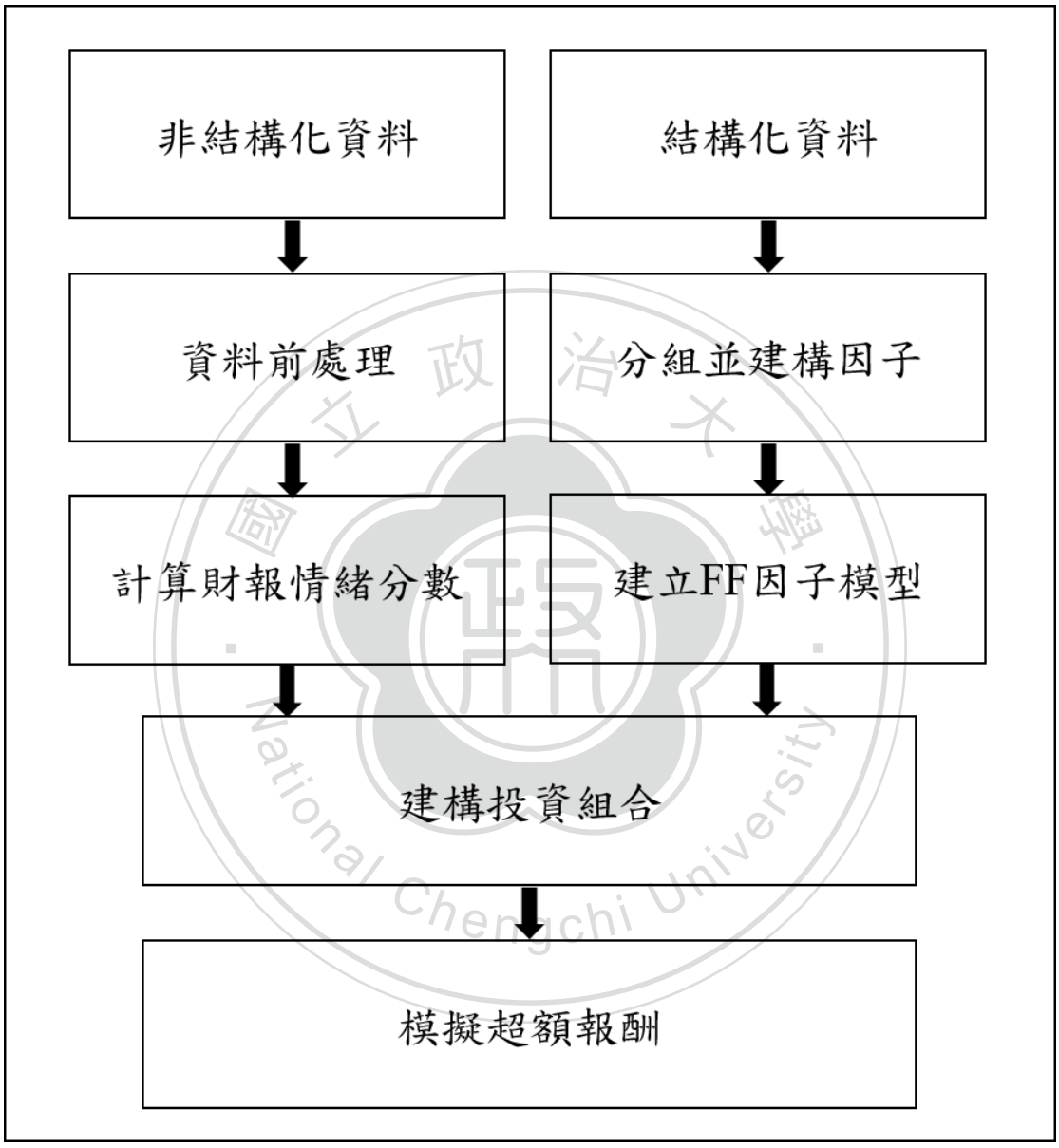

圖 2　實驗流程

DOI:10.6814/NCCU202000109

第一節 因子模型

本研究首先以 Black, Jensen, and Scholes (1972)提出之 CAPM 作為比較基準，觀察情緒因子是否會影響股價報酬，模型如下：

$$R_{it} - R_{ft} = \alpha_i + \beta_i\left(R_{mt} - R_{ft}\right) + \varepsilon_{it}$$

式中R_{it}為投資組合i在第t期之報酬率，R_{ft}為第t期之無風險利率，R_{mt}為市場投資組合在第t期之報酬率，其中$R_{it} - R_{ft}$為投資組合i在第t期之超額報酬，$R_{mt} - R_{ft}$為第t期之市場超額報酬即市場因子，而α_i為投資組合i在模型中之截距項，也就是著名的 Jensen's alpha，常作為衡量投資組合績效的指標。本篇論文中R_{ft}使用一月期之美國國庫券利率，R_{mt}使用所有可用之股票的價值平均(Value-weighted average)。

Fama-French 三因子模型(Fama and French, 1993)在單因子模型之上加入了規模與帳面市值比兩個因子，使得模型充分的解釋橫斷面變異，也更能有效的詮釋市場異象，在國內外學者業界也廣泛使用，因此本研究將基於三因子模型加入情緒因子來判別情緒因子的解釋力，三因子模型如下：

$$R_{it} - R_{ft} = \alpha_i + \beta_i\left(R_{mt} - R_{ft}\right) + s_i SMB_t + h_i HML_t + \varepsilon_{it}$$

上式中的 SMB 以及 HML 為規模以及帳面市值比因子，建構方式如下：首先於樣本期間每個月底，將所有淨值為正的股票依照公司規模分為大規模 B 與小規模 S 兩組；接著，將所有淨值為正的股票依照淨價市值比分為三組，前 30%為最高 H，其次 40%為中間 M，最末 30%則是最低 S ；將以上兩步驟得出之投資組合求交集得出 BH、BM、BL、SH、SM、SL 六個投資組合，並計算出次月各投資組合之市值加權投資報酬；利用小規模投資組合 SH、SM、SL 的市值加權平均報酬之總和減去大規模投資組合 BH、BM、BL 的市值加權平均報酬之總和後再除以 3，即規模因子 SMB；利用高帳面市值比投資組合 BH、SH 的市值

DOI:10.6814/NCCU202000109

加權平均報酬之總和減去低帳面市值比投資組合 BL、SL 的市值加權平均報酬之總和後再除以 2，即淨價市值比因子 HML。

第二節 文本情感分析方法

文字資料與數值資料在處理上最大的差別就在於其資料維度極高，假設僅以英文中最常出現的 1000 字作為字典，一則 30 字的短文就有1000^{30}種可能性，已經與宇宙中所有的原子總數相仿，因此使用文字作為資料時有很大一部份的重點在於如何將文字資料降維至經濟、科技尚可處理的程度又不失其中的資訊，處理文字資料的步驟可以概略拆解為以下三步(Gentzkow et al., 2019)：第一步，將未處理之文字資料$\mathcal{D}$以數值陣列$\mathcal{C}$呈現；第二步，將$\mathcal{C}$映射至未知結果$\mathcal{V}$之估計$\hat{\mathcal{V}}$；最後使用所估計之$\hat{\mathcal{V}}$做後續的敘述性或因果分析。

文本情感分析所要做的就是找出適當的模型來估計文章中的情感$\mathcal{V}$，得到估計$\hat{\mathcal{V}}$後再進行後續的研究，分析方法可依照模型的訓練方式分為基於預訓練(Pre-trained)字典的非監督式方法，如：LIWC (Pennebaker, Francis, and Booth, 2001), ANEW (Bradley and Lang, 1999), GI (Stone et al., 1966)、VADER(Hutto and Gilbert, 2014)等；另一種方法則是透過機器學習方法從已標記過情感或語意的資料中萃取出特徵，如：SVM (Support vector machine)、Naïve Bayes、Decision Tree 等。

基於字典的方法必須仰賴於先前專業的研究，且字典的泛用性往往不高，因在不同的領域下同一字詞可能有不同的意思，如 Cost 在一般情況下會被視為是負面的，但在財務領域 Cost 則只是一個描述公司成本的敘述，但若成功的取得一套優良的字典，其運行效率將遠高於基於機器學習之方法。而基於監督式學習之方法雖然可以達到較高的泛用性與精準度，但往往依賴於非常大的文本資料與極大量高效能的處理器、記憶體，囿於經濟與現實考量，本研究將使用基於字典的情感分析模型來處理文字資料。

DOI:10.6814/NCCU202000109

Hutto and Gilbert (2014) 所提出的 VADER 是目前非常受歡迎的基於字典的情感分析工具，他們首先統整 LIWC (Pennebaker et al., 2001), ANEW (Bradley and Lang, 1999), GI (Stone et al., 1966)等廣為使用的情感字典作為基礎加入網路上經常使用的表情符號、俗語與縮寫，並將詞彙利用 Amazon Mechanical Turk 群眾外包系統委託英文能力經認可之一般大眾對於其所統整的字詞作情緒評分，再利用文法結構調整分數。為了驗證此字典模型的效果，作者使用經過人工驗證、標註情緒過的推特短文、電影評價、科技產品評論以及新聞評論，作為欲預測之目標，此一模型簡單、快速且所得出之效果能擊敗當時最常使用的七種的文字情感模型，且具高泛用性，在作者實驗中使用的所有領域皆然。

VADER 相較於基於機器學習的文本情感分析不需要大量的訓練資料，運算速度也較快，相較於同屬基於字典的情感分析工具，其泛用性也較高，在社群網站的貼文與留言的情感分析甚至能達到人類的程度，本篇論文使用 Python 的 NLTK 套件來實踐，實作細節可拆為以下幾步驟[2]：首先，爬取 10-K 財報中的 Item 7. Management's Discussion and Analysis of Financial Condition and Results of Operations (簡稱 MD&A) ；接著，將使用 VADER 計算 MD&A 中各句的情感分數Sent_i，細節如下：首先，去除句子中不必要之標點符號，並將各單字分開放入待處理之清單；按順序查詢單字是否存在於 VADER 的字典中，若無，將該字詞的情感分數設定為 0，並跳到下一單字，若有則按字典中該字詞的情感極性與強度設定分數，並繼續以下步驟；將全大寫的字詞的分數做調整，檢查單詞前是否存在加強語氣(Intensifiers or degree adverbs)若有則調整分數，檢查單詞前面是否存在否定詞(Negation)，如有則調整分數；接著，加總句中所有單詞之分數並檢查句尾是否存在強調語氣之標點符號，如有則調整分數；最後，依句子長度不同，所得出之分數將會有很大差異，因此須對分數進行標準化，根據 Hutto 本人在網

[2] 其中各部份調整分數皆根據 Hutto and Gilbert (2014)所提出之論文中所做出之實證結果調整。

DOI:10.6814/NCCU202000109

路上之說法[3]，使用者當然可依其使用上的需求更改標準化之方法，但因模型的效能問題其建議可依 Hutto and Gilbert (2014)論文中之參數進行標準化，即可得出不錯的結果，因此本篇亦採用其所提出之參數作為標準化係數，在標準化後即可得到該句之情感分數$Score_i$，以下為一範例，其中S_i為字典中 good 之分數、I_j為 so 之加強語氣係數、E_k為驚嘆號之加強係數：

表 1　VADER 評分範例

步驟	獲得結果
輸入資料	“The service here is so not good.”
一、將字串拆成單字	[‘The’, ‘service’, ‘here’, ‘is’, ‘not’, ‘so’, ‘good’, ‘!’]
二、依字典給分	$[0, 0, 0, 0, 0, 0,\ S_i, 0]$
三、加強語氣	$[0, 0, 0, 0, 0, 0,\ I_j * S_i, 0]$
四、否定詞	$[0, 0, 0, 0, 0, 0,\ -I_j * S_i, 0]$
五、總和	$-I_j * S_i$
六、強調語氣之符號	$Score_{raw} = -I_j * S_i - E_k$
七、標準化	$Score_i = \frac{Score_{raw}}{\sqrt{{Score_{raw}}^2 + 15}}$

Chengchi Uni

然而，該 VADER 所得出的分數$Score_i$為一句句子的情感分數，該如何有效地建立一篇財報的情感分數正是本文的研究重點，以下我們將使用三種不同方式計算財報情緒分數，探討不同方式對於超額報酬的影響。

一、平均情緒分數：在計算出文章中各句之評分$Score_i$後取簡單平均，此項評分方式所計算出的分數代表的意義較為直觀，即一篇文章的平均情緒，預期若公司財報的平均情緒分數較佳會有更好的超額報酬。

[3] https://github.com/nltk/nltk/issues/1490

DOI:10.6814/NCCU202000109

$$\text{Senti}_{avg} = \frac{\sum_{i=1}^{N_{sentence}} Score_i}{N_{sentence}}$$

$N_{sentence}$：一篇文章中的總句數

二、除 VADER 外，本文亦使用 Li (2006)中提出的不確定性分數作為比較基準，該需計算文章中出現不確定性詞彙的次數NR_t，為避免真數為 0 的情況會先將不確定性詞彙出現次數加一，再將其取對數即得到一文的不確定性分數，公式如下：

$$\text{Senti}_{Li} = \log(1 + NR_t)$$

NR_t：文章中出現不確定性詞彙的次數

三、正面句子比例情緒分數：考慮到語法結構可能出現否定句的可能性，若以平均做加總可能會產生正副相抵的效果並會產生偏誤，因此本方法計算財報中正面句子占總句數之比例做為財報情緒分數，預期若公司財報的正面句子比例情緒分數較高則該投資組合超額報酬較高。

$$\text{Senti}_{pos} = \frac{N_{Positive\ sentence}}{N_{sentence}}$$

$N_{Positive\ sentence}$：文中正面句子出現次數

$N_{Sentence}$ ：文章總句數

四、負面句子比例情緒分數：在文本情感分析運用於財務的領域中，有許多研究都指出文章以負面的情緒影響較大，因此建構以負面句數占總句數之比例做為情緒分數，預期若公司財報的負面句子比例分數較高則超額報酬較低。

$$\text{Senti}_{neg} = \frac{N_{Negative\ sentence}}{N_{sentence}}$$

$N_{Negative\ sentence}$：文中負面句子出現次數

$N_{Sentence}$ ：文章總句數

DOI:10.6814/NCCU202000109

其中正負面句子按照 Hutto and Gilbert (2014)所建議以 0.05 作為門檻，若句子情感分數Senti_i大於 0.05 為正面，小於-0.05 則為負面，介於 0.05 與-0.05 的句子視為中性句。

第三節 財報情緒因子與四因子模型

以往的因子模型主要都著重於公司財務表現方面的探索，然而在公司財報中同樣重要的文字部分卻在解釋股價時被忽略，本研究在計算好各公司年度 10-K 財報情緒分數(Senti_{avg}、Senti_{Li}、Senti_{pos}、Senti_{neg})後，建立財報情緒因子，嘗試探索以不同方式建構之財報情緒分數是否對於投資組合之橫斷面變異具有解釋力。

財報情緒因子建立方法如下：在建立各公司年度財報情緒分數後依照情緒分數的高低將所有股票分為市值約略相等的 5 組投資組合，並使用單因子與三因子模型驗證其 Alpha 是否顯著，再使用財報情緒分數建立情緒因子，建立方法與建立帳面市值比因子相仿，細節如下：首先計算公司各年度之 10-K 財報情緒分數；於樣本內每年 7 月依所計算出之情緒分數排序，前 30%為高情緒投資組合 H，次 40%為中情緒投資組合 M，最後 30%為低情緒投資組合 L，再與規模因子之投資組合取交集，建立 BH、BM、BL、SH、SM、SL 六個投資組合，並計算出次月各投資組合之市值加權投資報酬；利用高情緒投資組合 BH、SH 的市值加權平均報酬之總和減去低情緒投資組合 BL、SL 的市值加權平均報酬之總和後再除以 2，即淨價市值比因子，使用不同財報情緒分數Senti_{avg}、Senti_{pos}、Senti_{neg}即可建立PMN_{avg}、PMN_{pos}、PMN_{neg}三種財報情緒因子。

在計算出情緒因子後，將其加入原先使用之三因子模型，即本研究所使用之基於 10-K 情緒之四因子模型，模型如下：

DOI:10.6814/NCCU202000109

$$R_{it} - R_{ft} = \alpha_i + \beta_i\left(R_{mt} - R_{ft}\right) + s_i SMB_t + h_i HML_t + p_i PMN_t + \varepsilon_{it}$$

本次研究因Senti_{Li}在初步驗證情緒分數能造成異常報酬時不顯著，最終僅使用三種不同方式Senti_{avg}、Senti_{pos}、Senti_{neg}建立情緒因子，並使用四因子模型對超額報酬橫斷面變異進行分析，期能改善因子模型之解釋力，並比較不同方式所得出結果之差異。

國立政治大學
National Chengchi University

DOI:10.6814/NCCU202000109

第四章 資料來源與處理過程

第一節 資料來源

本研究使用自 2005 年至 2018 年底在紐約證券交易所(New York Stock Exchange, NYSE)以及那斯達克股票交易所(National Association of Securities Dealers Automated Quotations, NASDAQ)上市的公司作為研究標的，使用 CRSP 資料庫抓取各公司的每月報酬與價格，公司淨值則使用 Compustat 資料庫抓取相關項目並計算出其淨值，財報中的文字部分則使用美國證券交易委員會(United States Securities and Exchange Commission, 簡稱 SEC)提供之 Electronic Data Gathering, Analysis, and Retrieval System (簡稱 EDGAR) 抓取其公司之 10-K 財報(Form 10-K)再透過資料前處理取得本研究與使用之文本資料，以下兩節將分別介紹公司規模、帳面市值比與非結構化文字資料之處理方法。

第二節 規模與帳面市價比因子

為建構規模因子與帳面市值比因子，使用 CRSP 資料庫抓取各公司的每月報酬與價格，無風險利率使用一月期之美國國庫券利率，市場報酬率使用 CRSP Value-Weighted Return。

公司淨值資料部分，使用 Compustat 資料庫中抓取直接相對應之項目，計算公司淨值時若投資人權益證券(SEQ)可行則直接採用，若無資料則使用普通股(CEQ)與特殊股(PSTK)相加，若 SEQ 與 CEQ 資料皆無，則使用總資產(AT)扣除總負債(LT)、少數股東權益(MIB)、特殊股(PSTKRV、PSTKL 或 PSTK 端看何者可用)，得到公司淨值後再除以流通在外之普通股(CSHPRI)即求得每股淨值。

根據 Fama and French (1993)，公司財報發布與財務年度結束日之間有數月差距，因此在計算各月的帳面市值比時所採用之每股淨值需遞延 6 個月，舉例來

DOI:10.6814/NCCU202000109

說，若一公司之財務年度結束於 12 月 31 日，其財報可能於次年三月到六月間揭露，因此在本研究在使用財報中的資料如公司淨值、10-K 中的文字資料等皆遞延 6 個月。

在計算帳面市值比的部分，本以每月更新股價的方式來計算淨價市值比，因財務報表有遞延性的問題，因子模型中計算淨價市值比實際上是以前期的變數來預測當期變數，在 Asness and Frazzini (2013)研究中的實證分析可見，以此種方式計算是對帳面市值比較好的預測。

第三節 10-K 財報與文字資料前處理

美國公司財報 10-K 是由 SEC 規定公司必須每年財務年度結束後一定期間必須提交之財務報表，與一般財務報表不同之處在於其具固定格式，公司必須按該格式交付給 SEC 並置於資料庫中供投資人與社會大眾檢視。

從 EDGAR 所下載之財報儲存格式為 txt 文字檔，其中亦包含了以 ASCII 碼呈現之圖片、表格以及因編碼解碼不同所造成的亂碼，因此在進行文本情感分析前必須先將資料中與本次研究非相關的部分先行去除，才能接續後續模型進行評分。

文字資料清洗部分，本研究使用 Python 內建的 Regular expression operation (簡稱 re)模組，搜尋並標註欲刪除之標籤如：<type>zip、<type>graphic、<type>excel 等，並標註其標籤之結尾</document>，接著將標註的區塊刪除；在刪除以上區塊後，同樣使用 re 模組搜尋並標記因編碼解碼方式不同而產生的亂碼並將其替換成原本應有的字符如：「&AMP」 → 「&」 、 「&APOS;」 → 「'」等；最後將資料中原本帶有的 HTML 標籤如： 、<P>等刪除，即完成初步的清洗。

接著同樣使用 re 模組，標記其中出現 Item 7. Management's Discussion and

DOI:10.6814/NCCU202000109

Analysis 以及出現 item8 financial statements 的位置，在這兩位置中間即為本研究的文字資料標的：MD&A。

然而，為了使用 VADER 對 MD&A 中的句子進行情感分析，需對上步驟中所抓取出之文字資料進行進一步的資料清洗，首先，將其中的換行符號與特殊空格替換成空格符號，以避免斷句時造成錯誤斷句。而因 VADER 中數字並不包含在字典中，若將其留在句中，VADER 會將其視為中立字眼，此會造成偏誤，因此移除 MD&A 中所有數字，同理特殊符號亦會造成偏誤，因此刪除，即完成 MD&A 的清理。

在完成清理後，將 MD&A 使用句號「.」做分句後使用 VADER 對句子做評價，其中若一句中字數小於 3，則將其結果剃除，再利用第三章所提出的三種方法(Senti_{avg}、Senti_{pos}、Senti_{neg})將整篇財報中所有句子的分數整合即可得到該財報的財報情緒分數。

國立政治大學
National Chengchi University

DOI:10.6814/NCCU202000109

第五章 實證分析結果

本章將使用利用 10-K 財報所建構的各情緒分數作為標準排序，將股票分為市值大致相等的五個投資組合，以$Senti_{avg}$為例，在每月初將各股票利用$Senti_{avg}$分數由低至高排序，依序將股票加入投資組合直到該投資組合之規模大於等於整體規模五分之一就開始建構下一投資組合，直到所有股票皆放入投資組合中，即可求得五個市值大約相等的投資組合，因各股實際上並不可切割，因此在分入投資組合時會造成投資組合市值大於或小與五分之一，並觀察高低情緒投資組合之間的差異。

本章中各表中所列出之統計檢定，其係數旁之***表示在 1%信心水準下顯著；**表示在 5%信心水準顯著；*表示在 10%信心水準下顯著。

首先針對利用不同分數分組之投資組合做敘述統計，在表 2 中分別列出使用$Senti_{avg}$、$Senti_{Li}$、$Senti_{pos}$、$Senti_{neg}$分組之各投資組合的敘述統計：其中 N 為每月平均持有股數、Avg. Score 為投資組合中所使用之情感分數的平均、Avg. Excess Return 為超額報酬平均、t(ER)為平均超額報酬顯著異於零之 t 統計量。

表 2 左上方，$Senti_{avg}$較低的投資組合中平均超額報酬小於較高的投資組合之超額報酬，而平均每月持有股票支數在低$Senti_{avg}$的投資組合中較大，即在低$Senti_{avg}$的投資組合中公司規模較小。

表 2 右上方中則是使用$Senti_{Li}$作為標準所建構之投資組合，雖然所建構出之投資組合超額報酬皆顯著異於零，但卻較難在分數高低的情況下判斷出投資報酬率之高低，平均每月持有股票之數在不同分數的投資組合中相差亦不大。

表 2 左下方，$Senti_{pos}$較高的投資組合相較於較低的投資組合有較高的平均超額報酬，然而在最高$Senti_{pos}$的投資組合中並未有最高的平均超額報酬，而平均每月持有股票支數有隨著$Senti_{pos}$增高而遞增的現象，即在高$Senti_{pos}$投資組

DOI:10.6814/NCCU202000109

合中的公司平均而言市值規模較大。

表 2 右下方則可看出$Senti_{neg}$越高的投資組合有較低的超額報酬，而平均每月持有股票支數相較於使用中其他分數分組的分布平均，即各投資組合中的公司規模並未能看出較大的差異，而在超額報酬部分，使用$Senti_{neg}$分組的投資組合所顯示出的趨勢也較明顯。

表 2 顯示出使用不同文本情感分析模型方法所建構出之投資組合會有很大的差異如以$Senti_{avg}$與以$Senti_{Li}$分組之投資組合的超額報酬趨勢不同，而在使用 VADER 做為情感分析模型的情況下，使用$Senti_{avg}$與$Senti_{pos}$所建構出的投資組合相仿在正面情緒的投資組合會有較高的報酬、負面情緒的投資組合會有較低的超額報酬，但並不是單調上升，而是在中間的投資組合會有較高的超額報酬，且在較高分數的投資組合中公司數量偏少，即公司規模較大。而使用$Senti_{neg}$同樣可看出正面情緒的投資組合會有較高的超額報酬、負面情緒的投資組合有較低的超額報酬，且其超額報酬上升的趨勢較明顯，而不同情緒之投資組合的公司規模並未有太明顯的差距。

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

表 2　依各情緒分數分五組之各投資組合超額報酬統計

	使用$Senti_{avg}$分組					使用$Senti_{Li}$分組				
	低	2	3	4	高	低	2	3	4	高
N	509.8	460.1	332.6	296.2	224.4	255.2	391.8	391.8	391.9	392.3
Avg. Score	0.1210	0.1684	0.1930	0.2171	0.2606	1.3155	2.258	2.7439	3.2969	4.4908
Avg. Excess Return	1.0784**	1.2086***	1.5895***	1.3859***	1.3644***	1.3764***	1.3962***	1.6494***	1.2873***	1.0780***
t(ER)	2.4830	3.1449	4.2597	4.1194	4.0897	3.7090	3.6916	4.4264	3.3941	3.0672
	使用$Senti_{pos}$分組					使用$Senti_{neg}$分組				
	低	2	3	4	高	低	2	3	4	高
N	625.1	405.3	328.5	248	216.1	426.1	358.3	364	353.5	320.9
Avg. Score	0.4120	0.4726	0.5038	0.5353	0.5880	0.0891	0.1187	0.1389	0.1625	0.2131
Avg. Excess Return	1.1083***	1.3053***	1.2303***	1.5581***	1.3675***	1.4539***	1.4563***	1.4383***	1.2676***	0.9138**
t(ER)	2.7378	3.2535	3.1609	4.7016	4.0482	4.0795	4.3702	4.1546	3.2408	2.0697

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

為比較不同分組方式對超額報酬之影響和對市場超額報酬敏感度之關係，表 3 使用 CAPM 單因子模型對各投資組合超額報酬進行迴歸，公式如下：

$$R_{it} - R_{ft} = \alpha_i + \beta_i\left(R_{mt} - R_{ft}\right) + \varepsilon_{it}$$

表 3 左上方使用$Senti_{avg}$作為建構投資組合之標準，各投資組合的市場因子迴歸係數 β 在 0.8808 至 1.1550 之間且皆顯著異於零，且在低$Senti_{avg}$的投資組合中 β 較大，也就是低$Senti_{avg}$投資組合所面臨之市場風險大於高$Senti_{avg}$的投資組合，在異常報酬α的部分，高$Senti_{avg}$的投資組合會有正向的異常報酬，低$Senti_{avg}$則為負向的異常報酬。

表 3 右上方則使用$Senti_{Li}$作為建構投資組合之標準，各投資組合的市場因子迴歸係數 β 在 0.9603 至 1.0360 之間且皆顯著異於零，然而在高低$Senti_{Li}$投資組合中並無遞增或遞減之關係，在異常報酬α則有三組投資組合未顯著異於零。

表 3 左下方中使用$Senti_{pos}$作為建構投資組合之標準，得出結果與使用$Senti_{avg}$作為標準的投資組合類似，在市場因子迴歸係數 β 的部分皆顯著異於零，且低$Senti_{pos}$所承受的市場風險高於高$Senti_{pos}$的投資組合，異常報酬α在高$Senti_{pos}$的投資組合有正向的異常報酬，低$Senti_{pos}$的則有負向的異常報酬，而在$Senti_{pos}$較為中立的投資組合中則是不顯著異於零。

表 3 右下方中使用$Senti_{neg}$作為建構投資組合之標準，與使用$Senti_{pos}$及$Senti_{avg}$較為不同的是$Senti_{neg}$分數較高表示情緒越負面，從表 8 中各投資組合之市場因子迴歸係數皆顯著異於零，且在高$Senti_{neg}$的投資組合中較大，低$Senti_{neg}$的投資組合中較小，異常報酬的則是在高$Senti_{neg}$的投資組合中有負向的異常報酬，低$Senti_{neg}$投資組合中則為正向。

在表 3 中可看出使用財報情緒分數做為建構投資組合之標準在單因子模型下確實能創造異常報酬，在正面情緒投資組合中皆有正向的異常報酬，且市場風險亦低於負面情緒投資組合，此與實驗前的假設相符，而使用$Senti_{Li}$做為標準的

DOI:10.6814/NCCU202000109

投資組合與本研究所提出之另外三種方法所產生的投資組合的結果差異較大，比較不易。單因子模型無法有效的解釋超額報酬的橫斷面變異，因此下段將使用三因子模型排除投資組合中可能具有的規模效應與帳面市值比效應做更進一步的檢測。

國立政治大學
National Chengchi University

DOI:10.6814/NCCU202000109

表 3　依各情緒分數分五組之各投資組合 CAPM 模型統計量

	使用$Senti_{avg}$分組					使用$Senti_{Li}$分組				
	低	2	3	4	高	低	2	3	4	高
α	-0.0045*** (-3.2172)	-0.0017* (-1.8542)	0.0026** (2.5656)	0.0019** (2.0081)	0.0020* (1.7393)	0.0009 (0.6634)	0.0005 (0.4741)	0.0031*** (3.3081)	-0.0008 (-0.9921)	-0.0019** (-2.4932)
β	1.1550*** (37.7282)	1.0446*** (51.0347)	1.0083*** (46.1787)	0.9071*** (44.7336)	0.8808*** (35.1047)	0.9730*** (33.0282)	1.0166*** (42.8233)	1.0102*** (48.7255)	1.0360*** (56.4020)	0.9603*** (56.8156)
Adj. R	90.5181	94.5868	93.4663	93.0668	89.2056	87.9728	92.4817	94.0924	95.5244	95.5865
	使用$Senti_{pos}$分組					使用$Senti_{neg}$分組				
	低	2	3	4	高	低	2	3	4	高
α	-0.0033*** (-2.9546)	-0.0012 (-1.0853)	-0.0016 (-1.4650)	0.0040*** (3.5648)	0.0016* (1.7308)	0.0019* (1.8412)	0.0029** (2.5854)	0.0021** (2.0602)	-0.0013 (-1.3216)	-0.0063*** (-4.2073)
β	1.0903*** (44.0231)	1.0796*** (43.5047)	1.0492*** (44.5212)	0.8769*** (35.8366)	0.9092*** (43.5832)	0.9591*** (43.5458)	0.8820*** (35.9552)	0.9283*** (41.5272)	1.0592*** (47.8256)	1.1673*** (35.5594)
Adj. R	92.8573	92.6984	93.0051	89.5969	92.7228	92.7112	89.6583	92.0428	93.8816	89.4511

DOI:10.6814/NCCU202000109

表 4、表 5 使用 Fama-French 三因子模型對各投資組合超額報酬進行迴歸，公式如下：

$$R_{it} - R_{ft} = \alpha_i + \beta_i\left(R_{mt} - R_{ft}\right) + s_i SMB_t + h_i HML_t + \varepsilon_{it}$$

表 4 左方使用$Senti_{avg}$做為標準建立投資組合，投資組合異常報酬與單因子分析時相仿，在低$Senti_{avg}$投資組合中有負向的異常報酬，但在最高$Senti_{avg}$的投資組合中三因子模型下其異常報酬卻不顯著，市場風險因子也與單因子模型下相似，在高$Senti_{avg}$投資組合中其所承擔之市場風險較低，低$Senti_{avg}$投資組合中則較高，最大及最小市場風險因子係數之間差距與從單因子 0.2742 降至 0.2437，而規模因子與帳面市值比因子並未有明顯趨勢，且在大多數投資組合中不顯著，而平均調整後 R 平方則從單因子的 92.1687 略微提升至 92.2998。

表 4 右方使用$Senti_{Li}$做為標準建立投資組合，從表中可看出在最高$Senti_{Li}$的投資組合裡有負向的異常報酬，然而正向的異常報酬卻僅出現於第三組投資組合，其餘皆不顯著，平均調整後 R 平方從單因子的 93.1316 微幅提升至 93.5773，而在市場風險因子、規模因子與帳面市值比因子中未能發現明顯的規律。

表 5 左邊使用$Senti_{pos}$做為標準建立投資組合，同樣的在低$Senti_{pos}$投資組合中有負向的異常報酬，但在最高$Senti_{pos}$的投資組合中三因子模型下其異常報酬卻不顯著，市場風險因子在低$Senti_{pos}$投資組合較高、高$Senti_{pos}$投資組合則較低，最大及最小市場風險因子係數之間差距與從單因子 0.1811 降至 0.1732，規模與市價比因子在大多數投資組合中皆不顯著，平均調整後 R 平方從單因子的 92.1761 微幅提升至 92.3514。

表 5 右邊則使用$Senti_{neg}$做為標準建立投資組合，異常報酬在高$Senti_{neg}$投資組合有負向報酬，在低$Senti_{neg}$投資組合則有正向報酬，市場風險因子在低$Senti_{neg}$投資組合較低、高$Senti_{neg}$投資組合則較高，最大及最小市場風險因子係數之間差距與從單因子 0.2853 降至 0.2582，規模與市價比因子在大多數投資

DOI:10.6814/NCCU202000109

組合中皆不顯著，平均調整後 R 平方從單因子的 91.5490 微幅提升至 91.5491。

使用三因子模型後，藉由情緒分數所建立之投資組合仍能創造異常報酬，然而，在各投資組合中規模與帳面市值比因子皆未能有效解釋超額報酬橫斷面變異，雖然平均而言提升了調整後 R 平方，但提升幅度並不大，且市場風險因子仍與 1 差距較遠，此與 Fama and French (1993)所做出結果不太相同。

比較不同方式所建構出之投資組合，可發現以$Senti_{neg}$為標準之投資組合在最高與最低的投資組合時，其異常報酬之趨勢仍能較好的被判別，而$Senti_{Li}$則是與本研究所提出之三種方法差別甚巨，因此在下一段建立情緒因子時，將不使用$Senti_{Li}$，僅比較$Senti_{avg}$、$Senti_{pos}$、$Senti_{neg}$不同方式建立情緒因子是否能增加因子模型對於超額報酬橫斷面之解釋力與不同評分方式是否有不同的效果。

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

表 4　依$Senti_{avg}$與$Senti_{Li}$分五組之各投資組合三因子模型統計量

	使用$Senti_{avg}$分組					使用$Senti_{Li}$分組				
	低	2	3	4	高	低	2	3	4	高
α	-0.0044*** (-3.1294)	-0.0014 (-1.4899)	0.0023** (2.3252)	0.0019** (2.1048)	0.0017 (1.4735)	0.0014 (1.0251)	0.0012 (1.1842)	0.0030*** (3.1950)	-0.0009 (-1.0632)	-0.0020** (-2.5811)
β	1.1442*** (34.3970)	1.0173*** (47.1660)	1.0203*** (43.0852)	0.9138*** (42.3486)	0.9005*** (33.2529)	0.9345*** (30.1465)	0.9692*** (40.6019)	1.0070*** (45.6278)	1.0500*** (54.3628)	0.9644*** (52.3899)
s	0.0732 (1.1633)	0.1273*** (3.1195)	-0.0403 (-0.9006)	-0.0784* (-1.9198)	-0.0825 (-1.6106)	0.1949*** (3.3234)	0.2049*** (4.5376)	0.0658 (1.5752)	-0.1077*** (-2.9467)	-0.0064 (-0.1837)
h	-0.0273 (-0.5245)	0.0283 (0.8397)	-0.0390 (-1.0549)	0.0734** (2.1788)	-0.0359 (-0.8479)	0.0129 (0.2673)	0.0762** (2.0434)	-0.0827** (-2.4003)	0.0575* (1.9052)	-0.0256 (-0.8910)
Adj. R	90.4889	94.8935	93.4697	93.3305	89.3165	88.6848	93.5420	94.3122	95.7954	95.5522

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

表 5　依$Senti_{pos}$與$Senti_{neg}$分五組之各投資組合三因子模型統計量

	使用$Senti_{pos}$分組					使用$Senti_{neg}$分組				
	低	2	3	4	高	低	2	3	4	高
α	-0.0029*** (-2.6154)	-0.0012 (-1.0855)	-0.0016 (-1.4571)	0.0038*** (3.5081)	0.0014 (1.4934)	0.0021** (2.1334)	0.0027** (2.3955)	0.0020* (1.9077)	-0.0015 (-1.4335)	-0.0060*** (-3.9790)
β	1.0604*** (40.3394)	1.0723*** (39.9246)	1.0476*** (40.7427)	0.8991*** (34.8846)	0.9191*** (40.6656)	0.9482*** (40.1330)	0.8915*** (33.3952)	0.9315*** (38.3549)	1.0669*** (44.2217)	1.1497*** (32.2435)
s	0.1341*** (2.6972)	0.0661 (1.3011)	0.0191 (0.3923)	-0.1508*** (-3.0942)	-0.0229 (-0.5349)	0.0073 (0.1625)	-0.0305 (-0.6041)	0.0117 (0.2540)	-0.0260 (-0.5701)	0.0725 (1.0743)
h	0.0390 (0.9502)	-0.0464 (-1.1068)	-0.0179 (-0.4460)	0.0573 (1.4227)	-0.0497 (-1.4078)	0.0846** (2.2931)	-0.0332 (-0.7966)	-0.0476 (-1.2543)	-0.0242 (-0.6426)	0.0345 (0.6198)
Adj. R	93.1626	92.7337	92.9252	90.1933	92.7421	92.8735	89.5935	92.0212	93.8315	89.4259

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

表 6 與表 7 使用本研究所提出之四因子模型對各投資組合超額報酬進行迴歸，公式如下：

$$R_{it} - R_{ft} = \alpha_i + \beta_i(R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + p_i PMN_t + \varepsilon_{it}$$

各表中皆使用其分組方式使用之評分方式之建立情緒因子，如表十三中的PMN_t在建立時所使用之財報情緒分數評分方式即使用$Senti_{avg}$。

表 6 中使用$Senti_{avg}$做為標準建立投資組合，可看出異常報酬皆不顯著，而最大及最小市場風險因子係數之間差距降至 0.0781，規模與市價比因子在大多數投資組合中皆不顯著，情緒因子在高$Senti_{avg}$投資組合為負向，低$Senti_{avg}$投資組合為正向，平均調整後 R 平方從三因子的 92.2998 提升至 94.9223。

表 7 左方使用$Senti_{pos}$做為標準建立投資組合，異常報酬僅在第三與第四投資組合顯著，而最大及最小市場風險因子係數之間差距降至 0.1136，規模與市價比因子在大多數投資組合中皆不顯著，情緒因子在高$Senti_{pos}$投資組合為負向，低$Senti_{pos}$投資組合為正向，平均調整後 R 平方從三因子的 92.3514 提升至 94.2576。

表 7 右方使用$Senti_{neg}$做為標準建立投資組合，異常報酬大多數不顯著，僅在最高$Senti_{neg}$投資組合顯著，而最大及最小市場風險因子係數之間差距降至 0.1130，規模與市價比因子在大多數投資組合中皆不顯著，情緒因子在高$Senti_{neg}$投資組合為負向，低$Senti_{neg}$投資組合為正向，平均調整後 R 平方從三因子的 91.5491 提升至 93.9330。

使用本研究所建構之四因子模型明顯能有效提升模型對於超額報酬橫斷面變異之解釋力，使得在大多數投資組合中的異常報酬皆不顯著異於零，也使市場風險因子之係數較收斂至 1，此結果與先前對於因子模型之相關研究結果相符，而在規模與帳面市值比因子中絕大多數仍不顯著，主因推測為以情感分數建立的投資組合中較無特定規模或帳面市值比的形成，如在高情緒分數投資組合中公司

DOI:10.6814/NCCU202000109

規模、帳面市值比並不會較大或小，同樣的在低情緒分數投資組合也不會有規模、帳面市值比效應。而加入情感因子後平均調整後 R 平方亦有明顯的提升，下部分則再次驗證，在使用不同分組方式的情況下情緒因子是否能提升模型解釋力。

表 6　依$Senti_{avg}$分五組之各投資組合四因子模型統計量

	使用$Senti_{avg}$分組				
	低	2	3	4	高
α	-0.0006 (-0.6389)	-0.0007 (-0.7434)	0.0017 (1.6452)	0.0005 (0.5719)	-0.0011 (-1.2919)
β	1.0071*** (45.9555)	0.9934*** (43.2062)	1.0432*** (41.0887)	0.9651*** (46.1710)	0.9997*** (47.9388)
s	-0.0241 (-0.6245)	0.1103*** (2.7240)	-0.0241 (-0.5387)	-0.0419 (-1.1378)	-0.0121 (-0.3290)
h	-0.0208 (-0.6600)	0.0294 (0.8908)	-0.0401 (-1.0997)	0.0710** (2.3652)	-0.0406 (-1.3539)
p	-0.6983*** (-15.9018)	-0.1216*** (-2.6398)	0.1166** (2.2920)	0.2617*** (6.2470)	0.5054*** (12.0937)
Adj. R	96.5099	95.0941	93.6545	94.7086	94.6446

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

表 7　依$Senti_{pos}$與$Senti_{neg}$分五組之各投資組合四因子模型統計量

	使用$Senti_{pos}$分組					使用$Senti_{neg}$分組				
	低	2	3	4	高	低	2	3	4	高
α	-0.0006 (-0.8744)	-0.0008 (-0.6955)	-0.0021* (-1.8731)	0.0024** (2.4645)	-0.0002 (-0.2538)	-0.0005 (-0.8219)	0.0014 (1.2681)	0.0007 (0.6973)	0.0000 (0.0289)	-0.0026** (-2.1893)
β	0.9766*** (54.9016)	1.0564*** (37.4627)	1.0652*** (39.6066)	0.9516*** (39.9615)	0.9789*** (53.0688)	1.0309*** (64.7351)	0.9316*** (34.5728)	0.9705*** (39.9520)	1.0208*** (43.6412)	1.0446*** (36.2003)
s	0.0830** (2.5909)	0.0564 (1.1114)	0.0299 (0.6169)	-0.1189*** (-2.7726)	0.0136 (0.4099)	0.0186 (0.6582)	-0.0250 (-0.5234)	0.0170 (0.3950)	-0.0323 (-0.7797)	0.0581 (1.1348)
h	0.0046 (0.1738)	-0.0529 (-1.2656)	-0.0106 (-0.2669)	0.0788** (2.2310)	-0.0251 (-0.9189)	0.0873*** (3.7465)	-0.0319 (-0.8089)	-0.0463 (-1.3023)	-0.0257 (-0.7513)	0.0311 (0.7360)
p	-0.5824*** (-14.5388)	-0.1102* (-1.7355)	0.1228** (2.0280)	0.3643*** (6.7935)	0.4153*** (9.9978)	-0.4454*** (-14.8664)	-0.2162*** (-4.2643)	-0.2103*** (-4.6026)	0.2485*** (5.6479)	0.5665*** (10.4352)
Adj. R	97.1989	92.8324	93.0729	92.5097	95.6741	97.1573	90.6893	92.9902	94.9089	93.9194

DOI:10.6814/NCCU202000109

表 8 中仿效 Fama and French (1993)使用規模與帳面市值比做排序建構出二維投資組合，並使用本研究所提出之四因子模型進行迴歸，因使用三種不同方式所建構之因子對於模型解釋力差異並不大，此處僅以使用基於$Senti_{avg}$所建構的PMN_{avg}為例，其餘兩組計算方式所建構之依公司規模與帳面市值比分二十五組之各投資組合四因子模型統計量置於附錄，迴歸公式如下：

$$R_{it} - R_{ft} = \alpha_i + \beta_i(R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + p_i PMN_t + \varepsilon_{it}$$

從表 8 中可看出超額報酬橫斷面變異並未被充分解釋，25 個投資組合中有 15 個異常報酬顯著異於零，而在高帳面市值比的投資組合中其所承擔之市場風險高於低帳面市值比之投資組合，高規模的投資組合所承擔之市場風險亦高於低規模之投資組合，而規模因子與帳面市值比因子之係數皆在其較高與較低之投資組合中顯著，與先前的研究相符，而情感因子則是在 25 個投資組合中的 13 個顯著，相較於未加入情感因子之三因子模型，其平均調整後 R 平方從 73.6832 提升至 74.6774。

使用不同情感分數計算方式所建構出之情感因子所驗證出之結果差距並不大，在此僅以基於$Senti_{avg}$所建構的為例，可以在表 9 中看見，不同情感分數所計算出之情感因子呈高度相關，而與規模、帳面市值比因子皆為低度相關，因此分別加入原先之三因子模型應不致造成線性重合的問題，而使用在模型分組也較能排除規模與帳面市值比的影響，有效地創造超額報酬。

DOI:10.6814/NCCU202000109

表 8　依公司規模與帳面市值比分二十五組之各投資組合PMN_{avg}四因子模型統計量

	帳面市值比分組					帳面市值比分組				
	α					t(α)				
規模分組	低	2	3	4	高	低	2	3	4	高
低	0.0123***	0.0013	-0.0034***	-0.0104***	-0.0244***	(10.2096)	(1.4213)	(-2.9096)	(-7.2501)	(-6.8571)
2	0.0054***	0.0002	-0.0049***	-0.0093***	-0.0170***	(4.4691)	(0.1447)	(-2.7874)	(-3.3856)	(-3.9581)
3	0.0055***	0.0016	-0.0039**	-0.0048*	-0.0070	(3.5440)	(1.1124)	(-2.1530)	(-1.9367)	(-1.2497)
4	0.0053**	-0.0006	-0.0044	-0.0093**	-0.0344***	(2.3694)	(-0.2899)	(-1.5751)	(-2.1812)	(-2.9050)
高	0.0016	-0.0037	-0.0112**	-0.0059	-0.0016	(0.5066)	(-1.1294)	(-2.0939)	(-0.8057)	(-0.2257)
	β					t(β)				
低	1.0326***	0.9739***	1.0549***	1.1456***	1.3312***	(34.9621)	(41.7724)	(36.8167)	(32.4103)	(15.1976)
2	0.9750***	1.0134***	1.0204***	1.0414***	1.2623***	(32.8005)	(30.4854)	(23.6832)	(15.4007)	(11.9299)
3	0.9124***	0.9752***	0.9696***	0.9103***	0.8435***	(23.9763)	(27.3066)	(21.9574)	(15.0009)	(6.0242)
4	0.8161***	0.9408***	1.0244***	1.1151***	1.2686***	(14.8541)	(18.4498)	(14.9746)	(10.5761)	(3.6309)
高	1.0698***	1.0545***	1.3604***	1.2002***	1.3088***	(14.1771)	(13.1112)	(10.3954)	(6.4764)	(5.8450)
	s					t(s)				
低	0.6237***	0.5385***	0.4969***	0.6379***	0.7998***	(11.9882)	(13.1124)	(9.8463)	(10.2451)	(5.1839)
2	0.1945***	0.1410**	0.1300*	0.0948	0.2289	(3.7145)	(2.4080)	(1.7134)	(0.7958)	(1.2292)
3	0.0370	-0.0700	-0.1528*	-0.0703	0.0076	(0.5513)	(-1.1133)	(-1.9755)	(-0.6587)	(0.0315)
4	-0.2177**	-0.0478	-0.2584**	-0.2386	0.3280	(-2.2491)	(-0.5327)	(-2.1657)	(-1.2748)	(0.6580)
高	-0.3722***	-0.3898***	-0.4015*	-1.2057***	-0.5862**	(-2.8001)	(-2.7559)	(-1.7538)	(-3.5814)	(-2.6556)

DOI:10.6814/NCCU202000109

	h					t(h)				
低	-0.1961***	0.1214***	0.2989***	0.4887***	1.0706***	(-4.6227)	(3.6244)	(7.2641)	(9.6251)	(8.5094)
2	-0.2419***	0.0297	0.0848	0.3959***	0.6787***	(-5.6665)	(0.6225)	(1.3702)	(4.0762)	(4.4331)
3	-0.2441***	-0.0844	-0.0547	0.1296	0.4106**	(-4.4655)	(-1.6458)	(-0.8563)	(1.4763)	(2.1215)
4	-0.1971**	0.0296	0.3585***	0.3612**	0.7520	(-2.4975)	(0.4044)	(3.6530)	(2.1283)	(1.4863)
高	-0.4099***	0.1215	0.8059***	0.5623**	-2.0915***	(-3.7818)	(1.0453)	(3.9147)	(2.3743)	(-4.8241)
	p					t(p)				
低	0.1497**	0.0133	0.0362	0.1231*	-0.1778	(2.5295)	(0.2851)	(0.6302)	(1.7375)	(-1.0129)
2	0.1065*	-0.0766	-0.2970***	0.1697	-0.7355***	(1.7879)	(-1.1494)	(-3.4401)	(1.2522)	(-3.0210)
3	0.0864	0.1225*	-0.2796***	-0.3495**	-0.2372	(1.1328)	(1.7118)	(-2.7508)	(-2.5031)	(-0.7425)
4	-0.2230**	0.1617	0.0501	-0.4459*	-2.0103***	(-2.0251)	(1.5827)	(0.3180)	(-1.7348)	(-2.8310)
高	0.4208***	0.4260**	0.3396	-1.0785***	-0.3091	(2.7832)	(2.2973)	(1.0307)	(-2.7602)	(-0.6478)
	R									
低	93.2873	95.5749	94.3729	93.3050	79.9902					
2	90.9430	90.5142	86.5262	70.8009	68.5098					
3	83.4189	86.2982	82.3892	71.1971	34.7704					
4	67.3816	74.3627	68.7226	62.4549	37.6587					
高	58.6470	57.1711	66.6393	65.1913	86.8079					

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

表 9　各因子之相關係數

	HML	SMB	PMN_{avg}	PMN_{neg}	PMN_{pos}
HML	1				
SMB	0.1360	1			
PMN_{avg}	-0.0848	-0.2973	1		
PMN_{neg}	0.0805	0.1629	-0.8227	1	
PMN_{pos}	-0.1614	-0.2421	0.8420	-0.7015	1

國立政治大學
National Chengchi University

DOI:10.6814/NCCU202000109

第六章 結論

本研究以情感分析技術分析美國公司 10-K 財報中的文字資訊，基於文章中句子的正負面情緒計算財報情緒分數、建構情緒因子並結合 Fama-French 三因子模型建立基於財報情緒之四因子模型，探討財報情緒是否能創造異常報酬、加入情緒因子之四因子模型是否能增強因子模型的超額報酬橫斷面變異解釋力。

根據實證結果，使用財報情緒確實能夠建立具有異常報酬的投資組合，且在正面情緒之投資組合中其市場風險較低，未來可嘗試使用財報情緒做為投資組合建構標準之一，並驗證是否能夠帶來更好的收益。而比較不同計算文章情緒分數之方法後，發現同樣是基於 VADER 的方法，Senti_{avg}、Senti_{pos}、Senti_{neg}在正面情緒的投資組合皆有較高的超額報酬，負面情緒的投資組合皆有較低的超額報酬，但在使用Senti_{neg}做為分組標準之投資組合中，其趨勢更加明顯。而若使用不同的文本情感分析方法如 Li (2006)則所得出之投資組合超額報酬趨勢與本研究提出之方式則有較明顯之差異，難以比較優劣。

在使用各公司年報所計算出之財報情緒分數來建構情感因子加入三因子模型後，各投資組合之異常報酬確實被消除，且平均調整後 R 平方亦有明顯上升的跡象，使用不同投資組合建構方式比較同樣有平均調整後 R 平方上升的現象，可見情緒因子確實能夠提升因子模型的超額報酬橫斷面變異解釋力，但仍並未能完全解釋。

本研究針對情緒因子對於超額報酬橫斷面變異之解釋力的影響得到的結果為正面的，但同時也發現其他問題可供後續研究：

1. 在財報前處理時，因各公司在處理 10-K 時的不同作法，造成資料的不完整，如有些公司因其公司年報已針對公司的表現與財務狀況作出評論，因此在 10-K 中的 MD&A 部分寫下請參照本公司的年報，此舉造成文本資料中有缺漏

DOI:10.6814/NCCU202000109

值，未來如欲研究 10-K 財報時，如何將資料前處理之細節處理細緻會是一大重點。

2. VADER 為句子層級的文本情感分析工具，本研究使用 VADER 評分後又再次對其分數做後續處理，雖然根據本研究結果證實不同後續處理方法影響並不大，但未來如能以文章層級的分析工具做情感分數評估應能產生更好的結果，如在一文中的首段通常代表著整篇文章最重要處，以文章層級的分析工具應能考量此效果並得出更精確之分析。

3. 使用不同文本情感分析方法對文章分析會產生不同的資訊，如 VADER 為正負面情感、Li 為文章不確定性，未來研究或可結合不同文本情感分析方法產生更進一步的資訊再進一步對股市進行分析。

國立政治大學
National Chengchi University

DOI:10.6814/NCCU202000109

參考文獻

Asness, C., and Frazzini, A. (2013). The devil in HML's details. *The Journal of Portfolio Management, 39*(4), 49-68.

Ball, R. (1978). Anomalies in relationships between securities' yields and yield-surrogates. *Journal of Financial Economics, 6*(2-3), 103-126.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics, 9*(1), 3-18.

Basu, S. (1983). The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. *Journal of Financial Economics, 12*(1), 129-156.

Black, F., Jensen, M. C., and Scholes, M. (1972). The capital asset pricing model: Some empirical tests. *Studies in the theory of capital markets, 81*(3), 79-121.

Bradley, M. M., and Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*.

Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H.-m., and Steele, L. B. (2014). The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies, 19*(1), 396-455. doi:10.1007/s11142-013-9258-3

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance, 52*(1), 57-82. doi:10.1111/j.1540-6261.1997.tb03808.x

Cowles 3rd, A. (1933). Can stock market forecasters forecast? *Econometrica: Journal of the Econometric Society*, 309-324.

Fama, E. F., and French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance, 47*(2), 427-465. doi:10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56. doi:10.1016/0304-405X(93)90023-5

Fama, E. F., and French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives, 18*(3), 25-46. doi:10.1257/0895330042162430

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as Data. *Journal of Economic Literature, 57*(3), 535-574. doi:10.1257/jel.20181020

DOI:10.6814/NCCU202000109

Hutto, C. J., and Gilbert, E. (2014). *Vader: A parsimonious rule-based model for sentiment analysis of social media text.* Paper presented at the Eighth international AAAI conference on weblogs and social media.

Kothari, S. P., Li, X., and Short, J. E. (2009). The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis. *The Accounting Review, 84*(5), 1639-1670. Retrieved from www.jstor.org/stable/27784235

Li, F. (2006). Do Stock Market Investors Understand the Risk Sentiment of Corporate Annual Reports? *SSRN Electronic Journal*. doi:10.2139/ssrn.898181

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics, 47*(1), 13-37. doi:10.2307/1924119

Loughran, T., and McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance, 66*(1), 35-65.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77-91.

Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates, 71*(2001), 2001.

Reinganum, M. R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics, 9*(1), 19-46.

Schwert, G. W. (2003). Chapter 15 Anomalies and market efficiency. In *Handbook of the Economics of Finance* (Vol. 1, pp. 939-974): Elsevier.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425-442.

Stone, P. J., Dunphy, D. C., and Smith, M. S. (1966). The general inquirer: A computer approach to content analysis.

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance, 62*(3), 1139-1168. doi:10.1111/j.1540-6261.2007.01232.x

DOI:10.6814/NCCU202000109

附錄

表 10　依公司規模與帳面市值比分二十五組之各投資組合PMN_{pos}四因子模型統計量

	帳面市值比分組					帳面市值比分組				
	α					t(α)				
規模分組	低	2	3	4	高	低	2	3	4	高
低	0.0124***	0.0013	-0.0034***	-0.0104***	-0.0251***	(10.5167)	(1.3705)	(-2.9571)	(-7.4046)	(-7.1458)
2	0.0054***	-0.0001	-0.0059***	-0.0088***	-0.0172***	(4.5640)	(-0.0974)	(-3.3066)	(-3.2451)	(-4.0332)
3	0.0051***	0.0019	-0.0042**	-0.0050**	-0.0078	(3.4139)	(1.3134)	(-2.3483)	(-2.0368)	(-1.3969)
4	0.0056**	0.0009	-0.0036	-0.0107**	-0.0299**	(2.5966)	(0.4487)	(-1.3061)	(-2.4701)	(-2.3999)
高	0.0018	-0.0028	-0.0104*	-0.0099	-0.0007	(0.5864)	(-0.8574)	(-1.9440)	(-1.2632)	(-0.1016)
	β					t(β)				
低	1.0288***	0.9766***	1.0551***	1.1462***	1.3567***	(35.9741)	(43.1565)	(37.9226)	(33.5842)	(15.8853)
2	0.9761***	1.0249***	1.0549***	1.0246***	1.2656***	(34.0603)	(31.5871)	(24.4606)	(15.5360)	(12.0949)
3	0.9270***	0.9661***	0.9837***	0.9175***	0.8776***	(25.5644)	(27.7182)	(22.2149)	(15.2008)	(6.3103)
4	0.8038***	0.8861***	0.9961***	1.1600***	1.1613***	(15.3938)	(17.8494)	(14.7068)	(10.6913)	(3.1841)
高	1.0653***	1.0232***	1.3431***	1.4001***	1.3239***	(14.6784)	(12.6655)	(10.1847)	(7.3463)	(6.3070)
	s					t(s)				
低	0.6184***	0.5398***	0.4964***	0.6358***	0.8189***	(12.0121)	(13.2525)	(9.9103)	(10.3488)	(5.3261)
2	0.1931***	0.1495**	0.1569**	0.0812	0.2287	(3.7426)	(2.5594)	(2.0210)	(0.6837)	(1.2327)
3	0.0442	-0.078	-0.1446*	-0.0662	0.0194	(0.6766)	(-1.2434)	(-1.8464)	(-0.6186)	(0.0800)

DOI:10.6814/NCCU202000109

4	-0.2208**	-0.0844	-0.2779**	-0.2100	0.4688	(-2.3486)	(-0.9442)	(-2.3361)	(-1.1116)	(0.8983)
高	-0.3832***	-0.4121***	-0.4330*	-1.1505***	-0.6256**	(-2.9330)	(-2.8786)	(-1.8865)	(-3.1872)	(-3.0621)
			h					t(h)		
低	-0.1842***	0.1237***	0.3023***	0.5000***	1.0651***	(-4.3426)	(3.6834)	(7.3236)	(9.8742)	(8.4059)
2	-0.2319***	0.0275	0.0723	0.4042***	0.6424***	(-5.4541)	(0.5722)	(1.1295)	(4.1316)	(4.1772)
3	-0.2303***	-0.0772	-0.0609	0.1157	0.4092**	(-4.2812)	(-1.4924)	(-0.9346)	(1.3045)	(2.1016)
4	-0.2222***	0.0217	0.3446***	0.3649**	0.6115	(-2.8677)	(0.2946)	(3.4929)	(2.1179)	(1.1583)
高	-0.3739***	0.1249	0.7770***	0.4567*	-2.0877***	(-3.4729)	(1.0531)	(3.7305)	(1.8232)	(-5.4096)
			p					t(p)		
低	0.1772***	0.0365	0.0512	0.1718**	-0.0651	(2.7519)	(0.7157)	(0.8170)	(2.2359)	(-0.3386)
2	0.1530**	-0.0248	-0.1654*	0.1141	-0.8963***	(2.3708)	(-0.3401)	(-1.7028)	(0.7686)	(-3.1521)
3	0.2192***	0.1035	-0.2404**	-0.3808**	-0.0601	(2.6839)	(1.3190)	(-1.9947)	(-2.3218)	(-0.1621)
4	-0.3891***	-0.1593	-0.1557	-0.1772	-1.9439**	(-3.3090)	(-1.4254)	(-0.8440)	(-0.5729)	(-2.1748)
高	0.5421***	0.2846	0.1576	-0.3185	-0.6625	(3.3171)	(1.2926)	(0.3985)	(-0.6846)	(-1.0135)
			Adj. R							
低	93.3390	95.5880	94.3833	93.3934	79.8645					
2	91.0888	90.4354	85.7122	70.6049	68.6808					
3	84.0639	86.1870	81.9504	71.0226	34.4965					
4	68.8140	74.2802	68.8604	61.5425	32.9756					
高	59.5104	56.0571	66.2812	59.9483	87.6958					

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109

表 11　依公司規模與帳面市值比分二十五組之各投資組合PMN_{neg}四因子模型統計量

	帳面市值比分組					帳面市值比分組				
	α					t(α)				
規模分組	低	2	3	4	高	低	2	3	4	高
低	0.0120***	0.0010	-0.0034***	-0.0099***	-0.0237***	(10.154)	(1.1127)	(-2.9040)	(-6.7965)	(-6.7049)
2	0.0051***	0.0000	-0.0054***	-0.0073***	-0.0173***	(4.3042)	(0.0009)	(-2.9998)	(-2.6720)	(-3.9514)
3	0.0050***	0.0020	-0.0036**	-0.0037	-0.0048	(3.2736)	(1.3388)	(-1.9940)	(-1.5493)	(-0.8703)
4	0.0041*	0.0001	-0.0053*	-0.0087**	-0.0342***	(1.8075)	(0.0345)	(-1.9001)	(-2.0576)	(-2.7792)
高	0.0030	-0.0039	-0.0108*	-0.0039	-0.0020	(0.9703)	(-1.1837)	(-1.9700)	(-0.5057)	(-0.2601)
	β					t(β)				
低	1.0355***	0.9828***	1.0538***	1.1255***	1.3146***	(36.2022)	(43.2567)	(37.4802)	(32.1296)	(15.3980)
2	0.9801***	1.0214***	1.0430***	0.9765***	1.2848***	(33.9979)	(31.2134)	(24.1772)	(14.7322)	(12.0194)
3	0.9250***	0.9614***	0.9682***	0.8826***	0.7640***	(25.1111)	(27.2272)	(22.0930)	(15.0065)	(5.4785)
4	0.8588***	0.9158***	1.0524***	1.0984***	1.2376***	(15.7097)	(18.1579)	(15.5229)	(10.5711)	(3.3997)
高	1.0128***	1.0528***	1.3472***	1.1781***	1.3057***	(13.3715)	(13.0966)	(10.1887)	(6.0409)	(5.2089)
	s					t(s)				
低	0.6073***	0.5382***	0.4927***	0.6213***	0.8175***	(11.9732)	(13.3600)	(9.8833)	(10.0027)	(5.4005)
2	0.1832***	0.1507**	0.1665**	0.0668	0.3024	(3.5842)	(2.5973)	(2.1771)	(0.5685)	(1.6241)
3	0.0290	-0.0857	-0.1289*	-0.0396	0.0582	(0.4435)	(-1.3688)	(-1.6900)	(-0.3865)	(0.2434)
4	-0.1868*	-0.0695	-0.2608**	-0.1784	0.6618	(-1.9266)	(-0.7766)	(-2.2250)	(-0.9672)	(1.2122)
高	-0.4273***	-0.4313***	-0.4543**	-1.0585***	-0.5791*	(-3.1815)	(-3.0821)	(-2.0186)	(-3.1123)	(-2.2555)

DOI:10.6814/NCCU202000109

	h						t(h)				
低	-0.1937***	0.1219***	0.2995***	0.4899***	1.0672***		(-4.6256)	(3.6647)	(7.2768)	(9.5553)	(8.5399)
2	-0.2401***	0.0288	0.0809	0.3965***	0.6960***		(-5.6896)	(0.6009)	(1.2808)	(4.0862)	(4.5026)
3	-0.2423***	-0.0830	-0.0491	0.1326	0.4286**		(-4.4937)	(-1.6049)	(-0.7727)	(1.5593)	(2.2436)
4	-0.1992**	0.0313	0.3614***	0.3646**	0.6966		(-2.4893)	(0.4245)	(3.7160)	(2.1692)	(1.3206)
高	-0.4051***	0.1129	0.7747***	0.6524**	-2.0234***		(-3.6540)	(0.9724)	(3.7880)	(2.6248)	(-4.6212)
	p						t(p)				
低	-0.1737***	-0.0618	-0.0323	-0.0215	0.2773*		(-3.2280)	(-1.4462)	(-0.6116)	(-0.3268)	(1.7265)
2	-0.1402**	0.0379	0.1922**	0.1705	0.5397**		(-2.5854)	(0.6163)	(2.3685)	(1.3673)	(2.4198)
3	-0.1594**	-0.0551	0.2683***	0.4806***	0.5850**		(-2.3002)	(-0.8300)	(2.9371)	(3.9167)	(2.0313)
4	0.0059	-0.0366	-0.2052	0.5197**	1.4842**		(0.0569)	(-0.3853)	(-1.4524)	(2.2463)	(2.0777)
高	-0.1381	-0.3793**	-0.1433	1.0752**	0.1529		(-0.9695)	(-2.2574)	(-0.5066)	(2.5620)	(0.3431)
	Adj. R										
低	93.4610	95.6353	94.3720	93.1707	80.2545						
2	91.1513	90.4527	85.9693	70.8608	67.7986						
3	83.8611	86.0874	82.5148	72.8907	36.5763						
4	66.4598	73.9465	69.1709	63.1195	32.3368						
高	56.7184	57.1162	66.3203	64.5058	86.3173						

國立政治大學 National Chengchi University

DOI:10.6814/NCCU202000109