DOI: 10.6277/TER.201909_47(3).0002

美國聯準會會議紀要的文字探勘與台灣經濟變數預測

黃裕烈・管中閔*

本文利用文字探勘的技巧, 從美國聯邦公開市場委員會 (Federal Open Market Committee, 簡稱 FOMC) 所發表的正式官方文件: minutes of the FOMC 中萃取出重要的資訊, 再利用情緒分析來判斷 FOMC 對三大經濟「使命」的正/負面看法, 並建立指標, 以預測台灣的財經相關變數。相較於之前的文獻, 本文統一 FOMC 常用的專門術語, 並且考慮了複合字的情況; 這可以避免後續在計算文字字數時, 產生重覆計算以及語意錯誤的情況。我們也利用統計模型 (MAP-PLSA) 將每一份文件中的句子依據 FOMC 的經濟使命分成三個主題, 再利用情緒分析技巧, 建立各類主題的指標數列。最後, 我們以迴歸模型來分析情緒指標與台灣相關財經變數之間的關聯性。

關鍵詞: 文字探勘, 美國聯邦公開市場委員, 主題模型, 情緒分析
JEL **分類代號**: C10, E30, O11

1 前言

文字探勘 (text mining) 是處理文字資料的技術, 可用來發掘文字檔案中

*作者分別爲國立清華大學計量財務與金融學系教授與國立台灣大學財務金融學系暨研究所教授, 國立台灣大學計量理論與應用研究中心特約研究員。黃裕烈爲通訊作者。作者感謝吳俊毅副研究員、徐之強以及徐士勛教授對本文初稿所提供的寶貴意見; 我們也特別感謝兩位匿名評審的意見, 使本文內容更加豐富和正確。此外, 本研究承蒙國家科學委員會與教育部計畫經費補助 (計畫編號: MOST-106-2014-H-007-002, MOST 107-3017-F-002-004; NTU-107L900202, NTU-107L900204), 在此一併致謝。文中若有任何錯誤, 當屬作者之責。

臺灣大學學術期刊論文

隱含的資訊, 本文將採用這種技術去分析美國聯邦公開市場委員會 (Federal Open Market Committee, 簡稱 FOMC) 的會議紀要 (minutes of the FOMC, 以下稱爲 Minutes)。FOMC 是美國聯邦準備系統 (Federal Reserve System) 下的貨幣政策決策單位, 每年召開 8 次會議, 並於會後發表正式的 Minutes。由於貨幣政策是影響美國經濟成長和物價穩定的關鍵因素, 而美國本身的經濟影響力又遍及全球, 所以 FOMC 的 Minutes 一向動見觀瞻。這些 Minutes 除了陳述貨幣政策的調整外, 還包括對未來美國經濟的展望, 所以也是全世界投資人觀察經濟動向的重要參考文件。由於文字內容並非數據資料, 難以量化方式分析, 所以過去外界只能設法解讀 Minutes 的文意, 嘗試從字裡行間揣摩出 FOMC 的態度。如今透過文字探勘的技術, 我們以更科學的方法來分析 Minutes 所包含的訊息。

文字探勘本屬於資訊工程或自然語言處理 (natural language processing) 領域, 但隨著各種演算法的快速演進, 文字資訊的處理變得更加迅速有效, 因此愈來愈多的研究領域開始運用文字探勘技術。以財經相關議題爲例, Lucca and Trebbi (2009), Berger, Haan, and Sturm (2011), Bernd and Neuenkirch (2013), Sadique et al. (2013) 等利用文字探勘技巧分析中央銀行的文件 (例如 Minutes 文件), 並利用這些資訊來協助預測總體變數; 這些文獻發現, Minuts 文件的資訊確實對總體變數的預測有所幫助。Doran, Peterson, and McKay (2012), Loughran and McDonald (2011), Loughran and McDonald (2014), and Loughran and McDonald (2016) 則利用此技術來處理企業合併, 訂價等財務問題。其中, Jubinskia and Tomljanovich (2013) 與 Rosa (2013) 的實證研究也發現, Minutes 文件對財務市場有顯著的影響效果。目前運用文字探勘的文獻, 多專注在如何將文字資訊轉換成數據資料的技術層面, 例如字數的計算以及指標的建立, 只有少數研究 (如前引 Loughran and McDonald 的數篇論文) 嘗試將財經相關的領域知識 (domain knowledge) 整合其中。

本文也將應用文字探勘來分析 Minutes, 但與過去相關文獻最大不同之處在於, 本文引入經濟相關的領域知識, 使文字探勘技術更具合理性。在作法上, 我們除了統一 FOMC 常用的專門術語, 也考慮了複合字 (compound words), 以避免後續在計算文字字數時, 產生重覆計算以及語意錯

臺灣大學圖書館期刊資料庫

誤的情況。此外, 我們利用自然語言處理的技術, 將文字資訊轉換成數據。更重要的是, 我們依據 FOMC 的三大經濟使命 (mandates), 用統計模型將每一份 Minutes 的句子 (sentence) 分成三個類別, 再應用情緒分析 (sentiment analysis) 分別建立這三大類別的指標; 據我們所知, 引入經濟相關的領域知識來區分文字檔案文句的類別, 然後建立個別情緒指標的作法, 是國、內外文獻中首次的嘗試。最後, 我們以迴歸模型來探討這些情緒指標與台灣相關財經變數之間的關聯性。

本文的實證結果發現, 在考量專門術語、複合字以及文句區分爲三大類別的情況下, 利用情緒分析所建構的各類指標可以呈現 Minutes 中更多樣的訊息。舉例而言, 我們發現每年第一次的 Minutes 文件中, FOMC 都會例行性地提到「穩定物價」這個使命。我們也發現, FOMC 在不同時期所關注的使命不盡相同, 而這些時期又和主席的任期有關。此外, 我們發現 FOMC 對三大經濟使命的正/負面看法跟美國的景氣循環有關聯。迴歸分析的結果顯示, 本文建構的各類指標對台灣的財經變數有解釋力以及預測能力。由於 FOMC 的決策應並不會考慮到台灣經濟, 此迴歸分析的結果也意謂這些指標可以在未來其他台灣經濟的研究中被用爲工具變數 (instrumental variables)。

本文的內容安排如下: 第 2 節概述文字資料的處理方式, 第 3 節則說明本文所採用的統計模型以及情緒分析方法的設定。第 4 節爲實證結果, 而最後一節爲總結。

2 資料處理

本文所欲分析的 Minutes 爲非結構性的 (unstructured) 文字資料, 必須適當處理後才能加以分析; 以下我們將說明資料整理的步驟與方式。我們自官方網站下載 Minutes,[1] 時間從 1993 至 2016 年 (共 24 年), 每年 8 份文件, 總共 192 個文件檔案, 文件格式均爲 HTML。

我們首先參考文獻上的作法, 將每份文件的開場敘述以及圖表註解刪除, 只保留主要的文字內容。其次, 我們統一文件中的專門術語。我們撰寫相關程式, 找出 95 個 FOMC 常用的專門術語, 用以取代文件中的相關

[1] 官方網址爲 https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm。

臺灣大學學術期刊資料庫

文字內容。舉例而言, Minutes 中有時以 “consumer price index”, 有時則以 “consumer price (CPI)” 或 “CPIs” 來表示消費者物價指數。若這些字詞不予以統一, 後續計算字數時可能會重覆計算, 例如會將 “consumer” 和 “price” 各分別計算一次, 並造成語意上的錯誤。因此, 消費者物價指數將統一由 “CPI” 取代。其他如 “reverse repurchase agreement”, “reverse repurchase (RRP)”, “reverse repurchase agreements (RRPs)” 等皆以 RRP 來取代, 而 “real personal consumption expenditures”, “real personal consumption expenditures (real PCE)” 等則以 “real PCE” 取代。

除了專門術語, 我們也處理複合字的問題。複合字如 “rate of return”, “labor force participation rate”, “trading volume”, “reserve requirement”, 以及前面提到的 “real PCE” 等, 必須適當處理, 否則亦會有重覆計算與語意錯誤的情形。我們利用文字探勘中的 N-Gram ($N \leq 12$) 技巧, 挑出 572 個 Minutes 中常用的複合字, 例如 “open market operations”, “basis points” 等。[2] 此外, 我們也撰寫程式從 Wikipedia 中挑出 30 萬個常用英文複合字, 然後跟 Minutes 文件比對, 從中再挑出 4,453 個英文複合字, 例如 “retail business”, “agricultural commodity” 等。最後, 我們找出 Minutes 文件中利用連字符號 (hyphen) 或是引號 (quotation mark) 產生的字詞, 共 848 個, 例如 “advanced economy”, “more positive than expected” 等。這些字詞, 加上之前的專門術語, 整合後共產生 5,770 個複合字; 我們後續的程式會將這些複合字視爲單一字詞。在前述資料處理的過程中, 我們也考量名詞的單複數以及字首大小寫, 並將這些專門術語及複合字統一以單數名詞表示。

我們將上述整理過後的文件視爲資料內容, 透過 Stanford University 的自然語言處理程式 coreNLP, 標記出文件中各字詞的詞性 (part of speech tagging; 例如名詞, 動詞等), 並且將不同詞性的字詞還原 (lemmatization) 爲單一字詞。[3] 最後, 利用 R 語言的 tm 套件建構所有文件的 document-term matrix (簡稱 DTM), 供後續模型使用。DTM 爲 $N \times M$ 維度的矩陣, 其中 N 爲文件數 (本文中 $N = 192$), M 爲字詞數 (本文中 $M = 9{,}311$), 故矩陣中第 i 列 (row) 第 j 行 (column) 的數字表示第 i 個文件中第 j 個

[2] 有關 N-Gram 的概念, 可以參考 Banchs (2012) 之說明。

[3] 將詞性還原是 coreNLP 套件的重要功能之一。舉例而言, 英文的 “go”, “goes”, “went” 與 “gone” 都屬同樣的語意, 因此該套件會將這些字詞還原成單數的 “go”。

臺灣大學經濟學術期刊資料庫

字詞被引用的次數 (以下稱爲詞頻)。DTM 矩陣中字詞爲所有文件所形成的集合, 包括 (但不限於) 專門術語與複合字, 但排除了 tm 套件所選定的停用詞 (stop words), 亦即出現頻率高, 但不影響文意的字詞, 如英文中的 “the”, “a”, “and” 等。

透過 DTM 的資料內容, 我們可以掌握 FOMC 的常用字詞及其詞頻。表 1 爲依據 DTM 內容彙整詞頻最高前 20 名的結果。其中, 第 2–3 行是依據所有字詞進行排序後所找出的高詞頻字詞。若只單看這些字詞, 我們可能無法清楚了解字詞所要表達之意; 這也是爲什麼我們在整理文字資料時, 必須考量專門術語與複合字。舉例來說, “growth” 在資料中出現多達 2,990 次, 這是排除了複合字 (如表中第 4 行的 “economic growth” 與 “real GDP growth”) 與專門術語的結果, 但我們無法只從這個字了解其在文件中的意義。相反地, 複合字如 “economic growth” 本身有其經濟意義, 因此其詞頻 (817 次) 就有了更豐富的意涵。

最後, 由於 FOMC 的三大經濟使命分別爲:「高度可持續就業」(maximum sustainable employment),「穩定物價」(stable prices), 以及「適度的長期利率」(moderate long-term interest rates), 因此表 1 中第 4–9 行也分別列出與這些使命相關的高詞頻字詞, 供後續模型使用。

3 模型設定與估計

建構 DTM 後, 我們將進行情緒分析。我們參考情緒關鍵字典, 如哈佛心理字典 Havard General Inquirer 或 Loughran and McDonald (2011) 所建構的財務相關字典, 藉此了解 FOMC 對每一份文件所抱持的態度爲正面或是負面; 參見 Sadique et al. (2013)。由於這些情緒關鍵字典已區分出正面字詞 (如 achieve, benefit 等) 或負面字詞 (如 abolish, bailout 等), 透過 DTM 我們可計算每一份文件的正/負面字詞數目, 並建構相關指標。然而與過去文獻中作法不同之處在於, 我們認爲每一份 Minutes 也都在陳述當時 FOMC 對三大經濟使命的看法, 因此進行情緒分析前, 有必要將每一份文件的內容先按這些使命區分成三大主題 (topics), 然後再針對各主題進行情緒分析。這樣的作法不僅可以揭露 FOMC 對個別主題所持之態度, 也可避免將主題合併分析時可能產生不同情緒互相抵銷的情況。

臺灣大學圖書館

表 1: FOMC 常用字詞與引用詞頻

排名	所有字詞 (不分類)	詞頻	就業相關字詞	詞頻	物價相關字詞	詞頻	利率與其它相關字詞	詞頻
1	increase	5113	economic activity	1454	inflation	3516	monetary policy	1354
2	continue	4513	labor market	902	price	1383	interest rate	776
3	remain	4058	economic growth	817	inflation expectation	1160	monetary	430
4	inflation	3516	unemployment rate	761	price stability	916	foreign exchange market	286
5	rise	3495	employment	735	energy price	657	monetary and financial condition	202
6	decline	3382	output	711	equity price	306	monetary aggregate	184
7	growth	2990	labor market condition	489	core inflation	287	foreign exchange value	163
8	expect	2151	maximum employment	339	consumer price	261	long term interest rate	105
9	rate	2150	real GDP	300	inflation compensation	249	market interest rate	103
10	policy	1987	labor cost	271	price inflation	234	foreign exchange	90
11	pace	1848	labor	229	inflation pressure	230	mortgage interest rate	85
12	time	1822	real GDP growth	221	commodity price	177	exchange	76
13	level	1796	unemployment	214	house price	174	monetary policy accommodation	55
14	suggest	1783	labor compensation	201	oil price	158	foreign monetary authority	52
15	report	1634	job	196	consumer price inflation	156	real interest rate	51
16	federal funds rate	1621	payroll employment	145	producer price	150	monetary growth	48
17	note	1612	civilian unemployment rate	136	inflationary pressure	140	market exchange rate	47
18	change	1593	employ	132	food price	127	monetary accommodation	41
19	reflect	1571	potential output	120	CPI	125	monetary base	33
20	measure	1557	labor market indicator	115	gasoline price	92	interest sensitive	29

臺灣大學學術期刊資料庫

3.1 PLSA 模型

我們採用 Hofmann (1999) 的 probability latent semantic analysis (簡稱 PLSA) 模型, 並加以修改, 以便將每份文件的內容區分爲三大主題。假設每一份 Minutes d_i, $i = 1, \cdots, N$ 是由 K 個不同主題 z_k, $k = 1, \cdots, K$ 所組成, 而每一個主題包含許多不同的字詞 w_j, $j = 1, \cdots, M$, 各主題之間允許有相同的字詞, 而這些字詞之集合 $W = \{w_1, \cdots, w_M\}$ 稱爲 bag of words。在本研究中, $K = 3$, 即 FOMC 的三大經濟使命, d_i, w_j 是可觀察的變數, 而 z_k 則是不可觀察的變數 (我們知道有三個主題, 但不知道每個字詞是屬於那一個特定主題)。PLSA 模型主要就是利用 expectation-maximization algorithm (簡稱 EM algorithm) 來計算

$$\theta = \left\{\mathbf{P}\left(z_k\right), \mathbf{P}\left(d_i/z_k\right), \mathbf{P}\left(w_j/z_k\right)\right\}, \quad \forall i, j, k,$$

的最大概似估計值 (maximum likelihood estimates, 簡稱 MLE)。

在不管文法與字詞順序等要求下, PLSA 模型假設任何一份文件或句子都是從 $\mathbf{W}$ 這集合中獨立抽樣出來的結果, 只是每一個字詞被抽出的頻率不同。舉例來說, 在2016年12月14日的 Minutes 文件中 (第三句) 提到:

> "The value of the dollar on foreign exchange markets rose, U.S. equity indexes increased considerably, and credit spreads on U.S. corporate bonds narrowed."

因此, PLSA 模型假設此句子是從 $\mathbf{W}$ 這集合中隨機抽取一些字詞出來組合而成, 該句子只是這些隨機取樣的實現值 (realization); 其中, "the", "on" 與 "U.S." 分別抽出2次, 其餘字詞抽出1次。在不考慮文法以及字詞的順序下, 這些抽出來的字詞與詞頻組成這一個句子。然而, 在我們的實際運算過程中, 我們考量了複合字, 且已排除停用詞, 並利用 coreNLP 套件將詞性還原, 所以 $\mathbf{W}$ 這集合中只會包含以下字詞:

> {value, dollar, foreign exchange market, rise, U.S., equity index, increase, considerably, credit spread, corporate bond, narrow},

臺灣人文與藝術期刊資料庫

而這些抽出來的字詞與詞頻則組成上述的例句。

PLSA 模型也假設在給定 z_k 的情況下, w_j 與 d_i 相互獨立: $\mathbf{P}(w_j/d_i, z_k) = \mathbf{P}(w_j/z_k)$。在上述這些假設條件下, 字詞 w_j 在 d_i 文件下發生的機率可表示成:

$$\mathbf{P}\left(d_i, w_j\right) = \sum_{k=1}^{K} \mathbf{P}\left(z_k\right) \mathbf{P}\left(d_i/z_k\right) \mathbf{P}\left(w_j/z_k\right),$$

而所有 Minutes 文件集合 (即 corpus) 的對數最大概似函數則爲:

$$\begin{aligned} \log L_1 &= \log \left[\prod_{i=1}^{N} \prod_{j+1}^{M} \mathbf{P}\left(d_i, w_j\right)^{n\left(d_i, w_j\right)} \right] \\ &= \sum_{i=1}^{N} \sum_{j=1}^{M} n\left(d_i, w_j\right) \log \left[\sum_{k=1}^{K} \mathbf{P}\left(z_k\right) \mathbf{P}\left(d_i/z_k\right) \mathbf{P}\left(w_j/z_k\right) \right], \end{aligned} \quad (1)$$

其中 $n(d_i, w_j)$ 爲在文件 d_i 下, 字詞 w_j 所發生的次數, 亦即 DTM 矩陣中第 i 列第 j 行的詞頻。從式 (1) 可看到, 此一概似函數爲多項分佈函數 (multinomial distribution)。由於 z_k 是無法觀察到的變數, 因此在極大化式 (1) 的目標下, Hofmann (1999) 導出 EM algorithm 分別爲:

E-Step:

$$\begin{aligned} \mathbf{P}\left(z_k/d_i, w_j\right) &= \frac{\mathbf{P}\left(z_k, w_j, d_i\right)}{\sum_{k=1}^{K} \mathbf{P}\left(z_k, w_j, d_i\right)} \\ &= \frac{\mathbf{P}\left(z_k\right) \mathbf{P}\left(d_i/z_k\right) \mathbf{P}\left(w_j/z_k\right)}{\sum_{k=1}^{K} \mathbf{P}\left(z_k\right) \mathbf{P}\left(d_i/z_k\right) \mathbf{P}\left(w_j/z_k\right)}; \end{aligned} \quad (2)$$

M-Step:

$$\mathbf{P}\left(w_j/z_k\right) \propto \sum_{i=1}^{N} n\left(d_i, w_j\right) \mathbf{P}\left(z_k/d_i, w_j\right);$$

$$\mathbf{P}\left(d_i/z_k\right) \propto \sum_{j=1}^{M} n\left(d_i, w_j\right) \mathbf{P}\left(z_k/d_i, w_j\right);$$

臺灣大學期刊資料庫

$$\mathbf{P}(z_k) \propto \sum_{i=1}^{N}\sum_{j=1}^{M} n(d_i, w_j)\,\mathbf{P}(z_k/d_i, w_j), \tag{3}$$

相關計算過程可參見 Hofmann (1999, pp. 50–51) 的內容。換言之, 只要給定起始值 $\{\mathbf{P}(z_k), \mathbf{P}(d_i/z_k), \mathbf{P}(w_j/z_k)\}^{(t=0)}$, 透過式 (2) 即可計算出事後機率 $\{\mathbf{P}(z_k/d_i, w_j)\}^{t=0}$; 給定此事後機率以及 $n(d_i, w_j)$ 結果, 透過式 (3) 即可計算出下一次的參數值 $\{\mathbf{P}(z_k), \mathbf{P}(d_i/z_k), \mathbf{P}(w_j/z_k)\}^{(t=1)}$, 如此可反覆計算直到收斂爲止。收斂結果即爲 $\{\mathbf{P}(z_k), \mathbf{P}(d_i/z_k), \mathbf{P}(w_j/z_k)\}$ 的 MLE 值。

利用 PLSA 模型雖可以計算出各主題發生的機率 $\mathbf{P}(z_k)$, 但我們很難判斷該主題是屬於 FOMC 那一個使命。這就類似主成份 (principal component) 或是因子分析 (factor analysis), 我們通常難以將這些主成份或是因子歸類成某一類特定變數。但實務上, 對於某些特定的字詞, 我們主觀即可判斷出該字詞是屬於那一個主題, 這樣的訊息應該善加運用。舉例來說, 當文件中的句子若出現 “economic activity” 以及 “labor market” 等字詞時, 該句子有比較大的機率是跟「高度可持續就業」這個使命有關; 又若出現 “inflation” 或是 “CPI” 時, 該句子則有比較大的機率是跟「穩定物價」有關。據此, 我們可將與「高度可持續就業」相關的字詞設定在第 1 類主題, 將與「穩定物價」相關的字詞設在第 2 類,「適度的長期利率」與其它部份則歸類在第 3 個主題, 透過 DTM 資料以及上述的設定估計出 MLE 結果, 便比較能判斷出該主題是屬於那一使命, 以解決上述的困擾。

3.2 MAP-PLSA 模型

爲了適度判斷相關字詞屬於那一個使命, 我們依循 Blei, Ng, and Jordan (2003) 的做法, 假設 $\theta = \{\mathbf{P}(z_k), \mathbf{P}(d_i/z_k), \mathbf{P}(w_j/z_k)\}$ 的先驗機率爲相互獨立的 Dirichlet 分佈。假設此一分佈的好處在於, 該分佈爲多項分佈函數 (即式 (1)) 的分配函數 (Dirichlet is a distribution over multinomial distributions), 其隨機樣本一定滿足多項分佈機率所要求的條件 (即大於 0, 加總爲 1 且爲多項分佈機率值), 並且在某些特殊的先驗參數 (deep parameters) 設定下, 所導出的最大概似估計式結果與 (2), (3) 的結果一樣; 換言

臺灣人文與社會科學資料庫

之, PLSA 模型只是此設定下的特例。[4] 更具體來說, 給定某一主題 $z_k = k$, 若

$$\left\{\mathbf{P}\left(w_j/z_k\right), \cdots, \mathbf{P}\left(w_M/z_k\right)\right\} \sim \text{Dirichlet } \left(\alpha_k\right),$$

則其機率密度函數可表示成:

$$f\left(\mathbf{P}\left(w_j/z_k\right), \cdots, \mathbf{P}\left(w_M/z_k\right); \alpha_k\right) = \frac{1}{\mathbf{B}\left(\alpha_k\right)} \prod_{j=1}^{M} \mathbf{P}\left(w_j/z_k\right)^{\alpha_{k,j}-1},$$

其中 $\alpha_k = \{\alpha_{k,1}, \cdots, \alpha_{k,M}\}$, $\alpha_{k,j} > 0 \; \forall j$ 爲 Dirichlet 分佈的先驗參數, $\mathbf{B}(\alpha_k)$ 爲 Beta 函數, 並且

$$\mathbf{P}\left(w_j/z_k\right) > 0, \quad \sum_{j=1}^{M} \mathbf{P}\left(w_j/z_k\right) = 1。$$

當 $\alpha_{k,j} = 1 \; \forall j$ 時, 該分佈等同於在 $M-1$ 維 open simplex 中的 uniform 分佈 (a uniform distribution over the open standard $M-1$ simplex)。這表示在此設定下, 我們並沒有對任何字詞有任何先驗假設, 每一個字詞在第 k 主題下發生的機率都相同 (此爲 uniform 分佈的特性)。但若我們將 $\alpha_{k,j}$ 先驗參數值設定比 1 大時, 表示我們有一些先驗認知, 認爲字詞 w_j 在第 k 主題下發生的機率比較大。

後續實證分析時, 我們會建構一個 $K \times M$ 的 $\alpha_{k,j}$ 先驗參數值矩陣, 該矩陣的第 k 列表示第 k 個主題, 而矩陣的行則與 DTM 矩陣中字詞的行相互對應。此外, 先驗參數矩陣的值大部份爲 1, 表示我們並沒有對這些字詞有任何先驗假設, 大部份的字詞在該主題下發生的機率相同。但我們對矩陣中第一列與表 1 就業相關字詞的 $\alpha_{1,j}$, 第 2 列與表 1 物價相關字詞的 $\alpha_{2,j}$, 以及第 3 列對應表 1 利率相關字詞的 $\alpha_{3,j}$ 均給予大於 1 的值。除了 $\{\mathbf{P}(w_1/z_k), \cdots, \mathbf{P}(w_M/z_k)\}$ 之外, 我們也假設:

$$\{\mathbf{P}\left(d_1/z_k\right), \cdots, \mathbf{P}\left(d_N/z_k\right)\} \sim \text{Dirichlet } \left(\beta_{k,1}, \cdots, \beta_{k,N}\right),$$

$$\{\mathbf{P}\left(z_1\right), \cdots, \mathbf{P}\left(z_K\right)\} \sim \text{Dirichlet } \left(\gamma_1, \cdots, \gamma_K\right),$$

[4] 假設 θ 的先驗分佈爲 Dirichlet 分佈有其優點, 也有它的缺點。如同貝氏估計 (Bayesian estimation) 的缺點一樣, 因每一個人對 θ 的主觀判斷可能不同, 所以任何先驗分佈的假設都可能會有錯誤設定的情況。

臺灣大學學術期刊論文資料庫

且在實證分析中, 我們設定 $\beta_{k,i} = \gamma_k = 1, \forall i, k$, 表示我們對文件以及主題都沒有任何的先驗假設。

依據這些假設以及獨立的條件, 先驗機率函數 $g(\theta)$ 便可表示成:

$$g(\theta) \propto \prod_{k=1}^{K} \left[\mathbf{P}(z_k)^{\gamma_k - 1} \prod_{j=1}^{M} \mathbf{P}\left(w_j/z_k\right)^{\alpha_{k,j}-1} \prod_{i=1}^{N} \mathbf{P}\left(d_i/z_k\right)^{\beta_{k,i}-1} \right],$$

而我們極大化的目標函數則成爲

$$\log L_2 = \log L_1 + \log[g(\theta)]。 \tag{4}$$

相較於 MLE 估計法, 文獻上稱此方法爲 maximum *a posteriori* method (簡稱 MAP); 請參考 Murphy (2012) 的討論。Chien and Wu (2008) 推導極大化式 (4) 的 EM algorithm, 其中 E-Step 與式 (2) 相同, 但

M-Step:

$$\begin{aligned}
\mathbf{P}\left(w_j/z_k\right) &\propto \left(\alpha_{k,j} - 1\right) + \sum_{i=1}^{N} n\left(d_i, w_j\right) \mathbf{P}\left(z_k/d_i, w_j\right); \\
\mathbf{P}\left(d_i/z_k\right) &\propto \left(\beta_{k,i} - 1\right) + \sum_{j=1}^{M} n\left(d_i, w_j\right) \mathbf{P}\left(z_k/d_i, w_j\right); \\
\mathbf{P}\left(z_k\right) &\propto \left(\gamma_k - 1\right) + \sum_{i=1}^{N}\sum_{j=1}^{M} n\left(d_i, w_j\right) \mathbf{P}\left(z_k/d_i, w_j\right)。
\end{aligned} \tag{5}$$

在這樣的架構下, 我們稱此模型爲 MAP-PLSA。很明顯地, 若我們對字詞, 文件與主題沒有任何先驗假設, $\alpha_{k,j} = \beta_{k,i} = \gamma_k = 1, \forall i, j, k$, 則式 (5) 會變成式 (3); 換言之, PLSA 模型只是 MAP-PLSA 模型的特例。

3.3 情緒分析

透過 MAP-PLSA 模型以及其對應的 EM-algorithm, 我們可估計出不同主題下各個字詞發生的機率 $\{\mathbf{P}(w_1/z_k), \cdots, \mathbf{P}(w_M/z_k)\}$。然而文字有其語意上的脈絡, 所以在理解文字內容時, 應該先從句子的意義切入, 而非僅機

臺灣大學圖書館期刊資料庫

械性的將個別字詞分開來判別。因此, 我們做的情緒分析將以句子爲單位, 先計算每一個句子屬於那一類主題的機率:

$$\begin{aligned}\mathbf{P}\left(z_k/\text{sentence}_{i,\ell}\right) &\propto \mathbf{P}\left(\text{sentence}_{i,\ell}/z_k\right)\mathbf{P}\left(z_k\right)\\ &= \left[\prod_{w_j\in\text{sentence}_{i,\ell}} \mathbf{P}\left(w_j/z_k\right)\right]\mathbf{P}\left(z_k\right),\end{aligned}$$

並取其最大值來判斷該句子是屬於那一類主題:

$$k_\ell^* := \arg\max_k \left\{\mathbf{P}\left(z_1/\text{sentence}_{i,\ell}\right),\cdots,\mathbf{P}\left(z_K/\text{sentence}_{i,\ell}\right)\right\},$$

其中 $\text{sentence}_{i,\ell}$ 代表在第 i 個文件中第 ℓ 個句子, 等式成立是因爲 $\{\mathbf{P}(w_j/z_k)\}\forall j,k$ 相互獨立的假設。換言之, 我們可將每一份文件以句子爲單位, 區分成三大主題。有了此一結果, 我們即可依據 Loughran and McDonald (2011) 的做法, 計算第 i 個文件中第 k 個主題的正/負面字詞總數: 正面字數$_{i,k}$ 和負面字數$_{i,k}$, 以及下述指標:

$$\begin{aligned}\text{Tone}_{i,k} = \frac{\text{正面字數}_{i,k}-\text{負面字數}_{i,k}}{\text{正面字數}_{i,k}+\text{負面字數}_{i,k}}\times 100,\\ k=1,\cdots,K;\ \ i=1,\cdots,N,\end{aligned}$$

當 $\text{Tone}_{i,k} > 0 (< 0)$ 時, 表示 FOMC 對第 k 個主題持正面 (負面) 態度。此外, 我們也以「正/負面字詞數差」當作指標:

$$\text{PN}_{i,k} = \text{正面字數}_{i,k} - \text{負面字數}_{i,k}。$$

爲了計算正/負字詞數, 我們參考了文獻上常用的正/負面詞庫: Loughran and McDonald (2011) 的財務詞庫, 以及 Bing Liu 教授所提供的正/負面詞庫。[5] 至此, 我們將原本文字資訊轉換成了清楚的數據。

[5]考量 Loughran and McDonald (2011) 的情緒字典主要是因爲此字典是針對經濟與財務相關的字詞進行情緒分析; 而考量 Bing Liu 的字典則是因爲此字典還包含一些複合字詞的情緒分類。文獻上尚有其它詞庫 (如 General Inquirer) 以及分類方式 (如分成增加以及減少); 相關的實證結果我們將列於附錄中。

臺灣大學經濟期刊

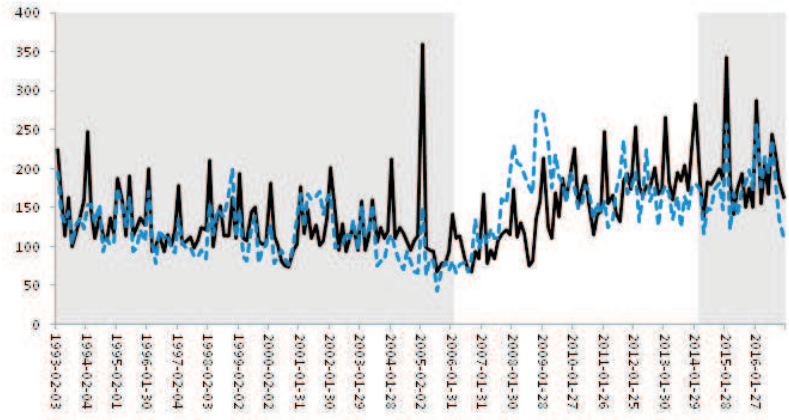


圖 1: 未區分主題下的正面 (實線) 與負面 (虛線) 字詞數, 圖中陰影部份用來區分不同 FED 主席的擔任期間

4 實證分析

4.1 未區分主題之情緒分析

首先, 我們依據第 2 節介紹的資料處理方式, 在不區分主題的情況下進行情緒分析。圖 1 是合併兩個文獻上常用的正/負面詞庫, 計算歷年各 Minutes 的正面 (實線) 與負面 (虛線) 字詞的詞頻圖。圖中第一個陰影時段 (資料開始至 2006 年 1 月) 爲 A. Greenspan 擔任聯邦準備系統主席 (以下簡稱 FED 主席) 的期間; 第二個陰影 (2014 年 2 月至今) 爲 J. Yellen 擔任 FED 主席的期間; 前述兩時段中間無陰影部分 (2006 年 2 月至 2014 年 1 月) 則爲 B. Bernanke 擔任主席的期間。

從圖 1 中不難看出, 正/負面字詞的次數會出現週期性的高點, 而這個高點均發生在每年第一次的 Minutes。此外, 2008–2009 年金融海嘯期間, 負面字詞的次數明顯增加。圖 2 爲「正/負面字詞數差」的數列, 其中陰影時段代表美國景氣循環的擴張期間。我們從此圖可以清楚看出, 在衰退期間, 負面的字詞數會比正面的字詞數來的多; 但在擴張期間, 多數 Minutes 的正面字詞數會比負面字詞數多。這些圖固然顯示了 FOMC 在景氣循環過程中的一些態度, 但我們卻很難從這些圖得知更多有關 FOMC 態度的

臺灣人口學資料庫

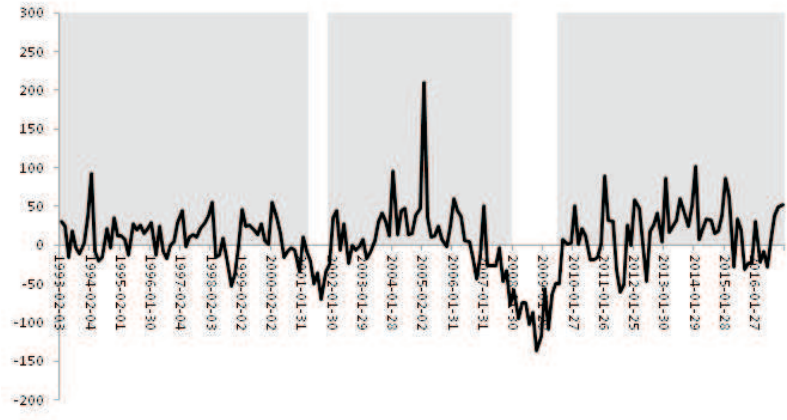


圖 2: 未區分主題下的正/負面字詞數差, 圖中陰影部份爲美國景氣循環擴張期間

訊息 (例如爲何高點均發生在第一次的 Minutes); 這正是傳統情緒分析的侷限之處。

4.2 區分主題下之情緒分析

我們依據第 3.2 節的 MAP-PLSA 模型並參考 Huang and Kuan (2017), 將歷年來的 Minutes 區分成「高度可持續就業」,「穩定物價」以及「適度的長期利率」三大主題進行情緒分析。圖 3 是不同主題下正面 (實線) 與負面 (虛線) 字詞的次數圖; 其中, 陰影部份主要是用來區分不同 FED 主席的擔任期間 (同圖 1), 而上圖爲「高度可持續就業」主題的結果, 中圖與下圖分別是「穩定物價」與「適度的長期利率」的結果。

從圖 3 可以看出, 在 A. Greenspan 擔任 FED 主席的期間, FOMC 多著重在「適度的長期利率」這個經濟使命, 所以這段期間中此一主題的正/負面字詞數特別多。B. Bernanke 擔任主席的前段期間, FOMC 關注的重點在於「穩定物價」, 但在其任期後段, FOMC 的關注重點則明顯轉爲「高度可持續就業」這個使命, 清楚顯示 FOMC 對雙重經濟使命 (dual mandate) 的態度。至於 J. Yellen 擔任主席後, FOMC 延續 Bernanke 後期的態度, 更關注「高度可持續就業」這個經濟使命。此外, 由圖 3 的中圖我們可看出,

臺灣大學 財務金融 資料庫

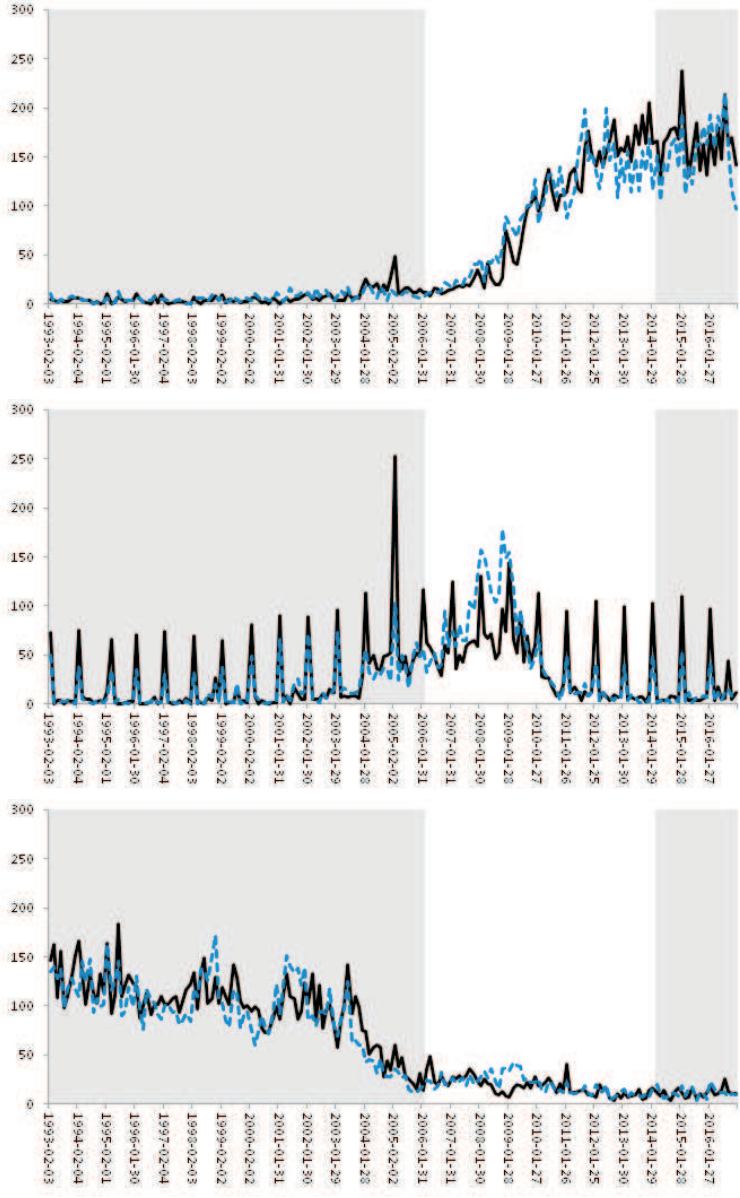


圖 3: 區分主題下正面 (實線) 與負面 (虛線) 字詞次數, 上圖爲「高度可持續就業」主題, 中圖爲「穩定物價」主題, 而下圖爲「適度的長期利率」主題, 圖中陰影用來區分不同 FED 主席的擔任期間

臺灣大學學術期刊論文著作權

圖 1 顯示的週期性字詞數高點是來自 FOMC 對穩定物價的看法, 而其出現的週期性表示「穩定物價」似已成爲 FOMC 每年初的例行性宣示。當然, 在特殊情況發生時 (如 2008 年這一段期間), FOMC 也還是會關注「穩定物價」這一主題。所以, 採用 MAP-PLSA 模型區分文件主題, 的確有助於我們發掘 Minutes 中更多訊息。

圖 4 爲 $PN_{i,k}$ (正/負面字詞數差) 數列的圖形, 其中陰影時段代表美國景氣循環的擴張期間 (同圖 2), 而圖中的上、中、下圖則分別描述「高度可持續就業」,「穩定物價」以及「適度的長期利率」三大主題的情況。我們首先觀察到, 在第一次衰退期, 較多的負面字詞數主要出現在「適度的長期利率」這個主題, 而在第二次衰退期 (金融海嘯時期), 三個主題都出現較多的負面字詞數, 但以在「穩定物價」這個主題上最多。此一結果與金融危機時期 FOMC 高度關注當時通貨膨脹率的態度相符合。其次, 在前兩次擴張期, FOMC 多數時候對「適度的長期利率」持較正面看法; 但在第三次擴張期, FOMC 對「高度可持續就業」則有較強的正面看法, 顯示了 FOMC 對就業市場的樂觀態度, 但對其他兩個主題, 則沒有明顯的正/負面態度。相形之下, 除了每年的例行性宣示以及第二次衰退期的負面看法外, FOMC 對於「穩定物價」似乎都沒有明顯的正/負面態度。這或許反映了 FOMC 在多數時期對於「穩定物價」這個使命成竹在胸, 所以其態度不受景氣影響。綜合而論, 相較於圖 1 以及圖 2 的結果, 在區分主題的做法下, 我們更容易理解 FOMC 對不同使命所持的態度。

4.3 迴歸分析

在這個小節中, 我們將檢視這些指標是否對台灣的經濟數據有解釋以及預測能力。本文依據 Sadique et al. (2013) 的做法, 估計以下的模型:

$$\text{模型 1:} y_{t+1} = \alpha + \sum_{i=0}^{m} \alpha_i y_{t-i} + \varepsilon_{t+1}, \tag{6}$$

$$\text{模型 2:} y_{t+1} = \alpha + \sum_{i=0}^{m} \alpha_i y_{t-i} + \beta_0 \times X_{t,0} + \varepsilon_{t+1}, \tag{7}$$

臺灣大學學術期刊資料庫

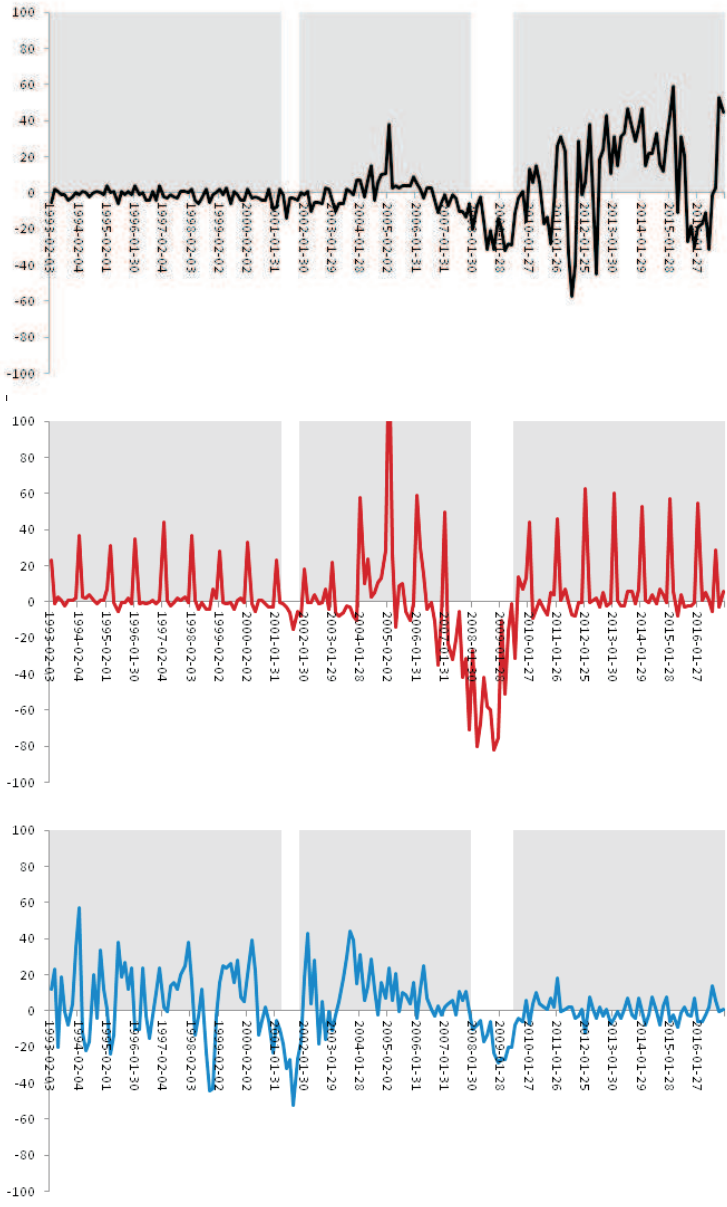


圖 4: 區分主題下正面減負面字詞次數, 其中上圖爲「高度可持續就業」主題, 中圖爲「穩定物價」主題, 而下圖爲「適度的長期利率」與其它主題, 圖中陰影部份爲美國景氣循環擴張期間

臺灣大學學術期刊資訊網

$$\text{模型 3:} y_{t+1} = \alpha + \sum_{i=0}^{m} \alpha_i y_{t-i} + \beta_1 \times X_{t,1} + \beta_2 \times X_{t,2} + \beta_3 \times X_{t,3} + \varepsilon_{t+1}, \tag{8}$$

其中 X 是 PN 或 Tone, $X_{t,k}$ 爲第 $k = \{1, 2, 3\}$ 個主題在 t 時間點 (依 FOMC 召開記者會的日期) 所編制的 PN 或是 Tone 指標, 而 $X_{t,0}$ 則是在不區分主題的情況下, 依據傳統文字探勘方法所編制的指標數列。由於 FOMC 召開記者會的日期並不固定 (例如1996年第一次會議是1月30日, 但1997年則是在2月4日), 所以在選擇對應的總體經濟數列時必需特別小心。在財經相關數據 y_{t+1} 的部份, 若是日資料, 則 y_{t+1} 爲 FOMC 召開記者會後下一個交易日; 若是週頻率的資料, 則 y_{t+1} 爲記者會後一個禮拜的交易日資料; 若是月頻率資料, 則 y_{t+1} 爲記者會召開月份下一個月的資料 (例如記者會是1月召開, 則 y_{t+1} 爲2月份資料)。式 (6) 至式 (8) 中還包括落後期的經濟數據 y_{t+1}, $i = \{0, 1, \cdots, m\}$, 主要是用來捕捉可能產生的序列相關性質, 其中 m 的選取則取決於 Schwarz information criterion (SIC), 而最大的落後期數爲4期; 換言之, 各模型的樣本個數爲 $192 - m$。此外, 由於「穩定物價」主題變數 $X_{t,2}$ 的季節成份明顯, 所以我們也在式 (8) 加入虛擬變數 (dummy variable), 以控制每年年初的例行性宣示效果。從式 (7) 至 (8) 中可知, 當 β_0 不顯著而 $\beta_k, k \in \{1, 2, 3\}$ 顯著時, 代表相較於傳統文字探勘所編制的全體指標 (即 $\text{Tone}_{t,0}$ 與 $\text{PN}_{t,0}$), 區分主題的指標數列能提供更顯著的解釋與預測能力。[6]

我們所考量的資料包括: 台灣加權股票指數 (簡稱台股指數), 美元兌換台幣匯率 (簡稱美元匯率) 與銀行業隔夜拆款利率 (簡稱隔夜拆款), 這些均爲日資料以及週資料。而領先指標不含趨勢指數 (簡稱領先指標), 同時指標不含趨勢指數 (簡稱同時指標), 工業生產指數 (簡稱工業生產), 海關出口值 (簡稱海關出口), 消費者物價指數 (簡稱 CPI), 躉售物價指標 (簡稱 WPI), 實質貨幣總計數 M1B (簡稱 M1B), 全體貨幣機構放款與投資 (簡稱貨幣機構放款與投資) 則爲月資料。資料來源爲台灣經濟新報, 主計總處統計資料庫以及我國的中央銀行, 時間則爲1993年至2006年。爲避免

[6]在進行假設檢定時, 我們採用 Newey and West (1987) 的標準差估計值 (Newey-West standard errors) 計算 t-statistic。

臺灣大學學術期刊資料庫

假性迴歸 (spurious regression) 的問題, 所有資料已轉換爲定態 (stationary) 的成長率 (或報酬率), 而 $PN_{t,k}$ 與 $Tone_{t,k}$ 指數本身就已是定態數列, 所以不做任何轉換。[7]

表 2 列出 $PN_{i,k}$ 的迴歸分析結果; 爲求簡潔, 我們列出模型 1 的 adjusted R^2 值 (但不列出估計係數), 以方便比較不同模型的配適結果。比較表中各模型的 adjusted R^2 值不難發現, 除了 CPI 成長率, 考量 $PN_{t,0}$ 或 $PN_{t,k}$ 指標的模型, 其配適結果都比模型 1 好, 而模型 3 的配適度也多數好過模型 2 (除了美元匯率報酬率和 CPI 成長率)。此外, 我們也觀察到,「適度的長期利率」指標的 $PN_{t,k=3}$ 對台灣的隔夜拆款成長率 (日以及週頻率) 有正向的解釋以及預測能力, 但當時間拉長到 1 個月後, $PN_{t,0}$ 與 $PN_{t,k}$, $k = 1, 2, 3$ 對隔夜拆款成長率已不具影響性 (因係數不顯著, 我們不列出此一估計結果)。

至於總體經濟月資料部份, 從表中可以發現, $PN_{t,0}$ 與 $PN_{t,k}, k = 1, 2, 3$ 指標對台灣的工業生產指數成長率, 海關出口值成長率, WPI 成長率, M1B 成長率與貨幣機構放款與投資成長率等變數都有解釋以及預測能力, 只是統計顯著性各有不同。從模型 3 的結果來看, 當 FOMC 對「高度可持續就業」主題持正面看法時, 表示未來美國經濟看好, 這對台灣的經濟有帶動作用, 所以是一個正面消息; 這會對台灣的海關出口值, 貨幣持有以及貨幣機構放款與投資等變數產生正向的影響。同樣地, 當 FOMC 對「穩定物價」主題持正面看法時, 表示未來美國物價穩定, 經濟前景看好, 所以對台灣的工業生產指數成長率, 海關出口值成長率, WPI 成長率等產生正面的影響效果。最後, 當 FOMC 對「適度的長期利率」主題持正面看法時, 表示美國的長期利率平穩, 這對美國經濟也是一個正面的消息, 所以也會對台灣的領先指標成長率, 工業生產指數與海關出口值產生正向的效果。

值得注意的一點是, 隔夜拆款成長率 (日資料) 和領先指標成長率都不受 $PN_{t,0}$ 的影響, 但卻受 $PN_{t,k=3}$ 的影響。這表示相較於傳統文字探勘所編制的全體指標, 區分主題的指標數列的確可提供不同的解釋與預測能力。表 3 爲 $Tone_{i,k}$ 的迴歸分析結果, 其結果跟表 2 中的係數符號大致相同,

[7] 爲節省篇幅, 本文就不列出單根檢定 (unit root test) 的結果; 有興趣的讀者可以向作者索取。

臺灣大學學術期刊資料庫

表 2: 指標數列 $PN_{i,k}$ 迴歸分析結果

頻率	模型 1	模型 2			模型 3						
	Adj R^2	β_0	t-stat	Adj R^2	β_1	t-stat	β_2	t-stat	β_3	t-stat	Adj R^2
日 隔夜拆款成長率	0.060	0.008	0.020	0.082	−0.005	−1.282	0.005	0.554	0.017	2.542**	0.089
美元匯率報酬率	0.026	0.002	1.339	0.058	0.002	0.888	0.002	1.323	0.000	0.159	0.053
台股指標報酬率	0.005	−0.005	−1.195	0.006	0.009	1.717*	−0.009	−1.520	0.009	1.300	0.022
週 隔夜拆款成長率	0.013	0.028	1.797*	0.045	0.034	1.225	0.010	0.354	0.060	2.821**	0.046
美元匯率報酬率	0.026	0.003	1.505	0.044	0.003	1.255	0.001	0.379	0.006	1.621	0.040
台股指標報酬率	0.005	−0.002	−0.250	0.009	−0.012	−0.541	0.003	0.342	−0.002	−0.096	0.018
月 領先指標成長率	0.910	−0.002	−1.084	0.915	−0.001	−0.627	0.000	0.090	−0.007	−2.980**	0.923
同時指標成長率	0.904	−0.001	−0.357	0.905	−0.001	−0.797	0.001	0.545	−0.004	−1.414	0.906
工業生產成長率	0.136	0.046	2.352**	0.185	−0.026	−0.701	0.079	2.120**	0.065	2.064**	0.197
海關出口成長率	0.161	0.064	2.983**	0.212	0.069	1.798*	0.125	2.874**	0.098	2.329**	0.244
CPI 成長率	0.045	−0.001	−0.831	0.043	−0.002	−0.681	−0.002	−0.863	0.002	0.600	0.038
WPI 成長率	0.284	0.004	1.909*	0.304	−0.001	−0.257	0.006	1.767*	0.006	1.239	0.305
M1B 成長率	0.093	−0.008	−2.699**	0.125	0.009	2.193**	−0.019	−2.958**	−0.007	−0.850	0.150
貨幣機構放款與投資成長率	0.080	0.002	2.053**	0.086	0.003	2.805**	0.000	−0.024	0.003	0.963	0.091

註: 符號 t-stat 爲 t-統計量, Adj R^2 爲 adjusted R^2; * 和 ** 表示在 10%和 5%的水準下顯著。各變數的落後期 m 選取準則是依據 SIC 選取, 因此模型的樣本數爲 $192-m$。

臺灣大學學術期刊資料庫

但對各變數的解釋以及預測能力則有些差異。例如在表2中日頻率的隔夜拆款成長率受 $PN_{t,k=3}$ 的影響，但表3中則顯示該變數受 $Tone_{t,k=2}$ 與 $Tone_{t,k=3}$ 所影響。此外，週頻率的美元匯率報酬率以及月頻率的領先指標成長率、海關出口成長、CPI、WPI 、M1B 成長率等變數也有類似的情況。我們認爲，這些差異主要是來自於 $PN_{t,k}$ 與 $Tone_{t,k}$ 指標編製方式不同而產生；但大多數結果均顯示，$Tone_{t,k}$ 對台灣經濟變數仍有影響效果。

考量「穩定物價」主題變數 $X_{t,2}$ 的季節因素，我們也在模型3加入虛擬變數 (dummy variable) 來控制每年年初的例行性宣示效果。從估計結果顯示，加入虛擬變數後係數的正負號大體不變，只有部分係數的顯著性受到影響；我們將這部分結果列於附錄一作爲對照。除此之外，我們也利用 General Inquirer 詞庫的分類方式，分成增加 (increase) 與減少 (decrease) 類別，再依據 Sadique et al. (2013) 的作法來建構另一新指標。即使這個詞庫的增加 (減少) 分類關鍵字數相當少，實證結果顯示這個指標對台灣財經變數仍有一些解釋能力；我們將這些結果列於附錄二以供參考。

4.4 預測結果

我們也嘗試利用模型1與模型3的估計結果進行樣本外的預測。由於2017年 FOMC Minutes 文件的公佈時點爲2月22日，4月5日，5月24日，7月5日，8月16日以及10月11日 (其餘2個時點因資料限制不納入預測考量)，爲配合這些時點，我們所收集的樣本外觀察時點 y_{T+i} 分別爲2017年3月、5月、6月、8月、9月以及11月等6期資料，其中隔夜拆款成長率、美元匯率報酬率、台股指標報酬率等三個變數是取每一個月的月底資料。我們的預測採遞迴方式 (recursive forecasts)，每次預測1期 (one-step ahead forecast)。表4列出各模型的預測均方誤差 (root mean square error forecast, RMSE)。從表4的結果來看，模型3的樣本外預測表現大致上都比傳統模型來的好 (RMSE 比較小)。例如，模型3之 $Tone_{i,k}$ 指標在工業生產成長率中的 RMSE 是模型1的1/2，在領先指標成長率的 RMSE 更只有模型1的1/3。換言之，加入相關指標會讓模型的樣本外預測表現更好。雖然模型3在這些變數的預測表現不錯，但仍有一些例外。例如隔夜拆款成長率，台股指標報酬率以及 WPI 成長率這三個變數，模型1的樣本外預測表現就

臺灣大學學術期刊資料庫

表 3: 指標數列 $\text{Tone}_{i,k}$ 迴歸分析結果

頻率	模型 1	模型 2			模型 3						
	Adj R^2	β_0	t-stat	Adj R^2	β_1	t-stat	β_2	t-stat	β_3	t-stat	Adj R^2
日 隔夜拆款成長率	0.060	1.757	1.743*	0.070	−0.295	−0.761	0.892	2.346**	1.202	1.713*	0.129
美元匯率報酬率	0.026	0.354	1.017	0.039	−0.053	−0.658	0.027	0.526	0.341	1.272	0.053
台股指標報酬率	0.005	1.298	−1.025	0.003	−0.393	−0.958	−0.070	−0.208	−0.408	−0.522	0.010
週 隔夜拆款成長率	0.013	8.619	2.048**	0.044	0.452	0.354	0.334	0.413	5.674	1.816*	0.055
美元匯率報酬率	0.026	0.741	1.417	0.037	0.326	1.137	−0.076	−0.764	0.666	1.805*	0.059
台股指標報酬率	0.005	−0.266	−0.091	0.010	0.377	0.503	0.506	0.598	−2.134	−1.134	0.007
月 領先指標成長率	0.910	−0.472	−1.061	0.914	0.035	0.438	−0.058	−1.327	−0.539	−1.277	0.919
同時指標成長率	0.904	−0.094	−0.154	0.904	0.031	0.353	0.018	0.387	−0.467	−0.996	0.908
工業生產成長率	0.136	13.139	2.643**	0.176	1.460	0.964	2.392	2.334**	3.129	1.853*	0.182
海關出口成長率	0.161	18.179	3.389**	0.203	0.968	0.556	2.388	1.331	9.324	2.366**	0.188
CPI 成長率	0.045	−0.337	−0.877	0.043	−0.251	−1.139	−0.195	−1.078	0.574	1.709*	0.063
WPI 成長率	0.284	1.388	2.017**	0.311	0.171	0.785	0.062	‘ 0.474	0.724	1.811*	0.326
M1B 成長率	0.093	−2.002	−2.115**	0.112	0.328	0.815	−0.627	−1.616	−0.465	−0.753	0.096
貨幣機構放款與投資成長率	0.080	0.437	1.647	0.084	0.461	2.714**	0.000	−0.002	0.322	1.582	0.108

註: 符號 t-stat 爲 t-統計量, Adj R^2 爲 adjusted R^2; * 和 ** 表示在 10%和 5%的水準下顯著。各變數的落後期 m 選取準則是依據 SIC 選取, 因此模型的樣本數爲 $192-m$。

臺灣大學學術期刊資料庫

表 4: 樣本外預測 RMSE 結果

變數	模型 1	模型 3 +PN$_{i,k}$	模型 3 +Tone$_{i,k}$
隔夜拆款成長率	4.082	4.206	4.115
美元匯率報酬率	0.546	0.481	0.506
台股指標報酬率	1.006	1.012	1.218
領先指標成長率	0.092	0.048	0.036
同時指標成長率	0.224	0.108	0.127
工業生產成長率	5.992	5.259	2.666
海關出口成長率	3.558	2.918	3.128
CPI 成長率	0.371	0.355	0.375
WPI 成長率	0.530	0.544	0.533
M1B 成長率	0.744	0.402	0.708
貨幣機構放款與投資成長率	0.422	0.179	0.422

註: RMSE 指 root mean square error forecast。

比較好。值得一提的是, 模型 3 因爲解釋變數都是變數落後項, 因此本身已包含了某種預測的結果, 所以即使樣本外的表現沒有比較好, 樣本內的結果仍有一定的參考價值。

5 結論

文字探勘的技術在國外已發展多年, 但其涉及的層面多屬資訊工程等相關的領域, 較少與財經議題相互結合。本文嘗試將文字探勘的技術運用在經濟預測的相關領域中, 希望能夠拋磚引玉, 讓國內學者對文字探勘的技術及其應用有更深一層的了解。特別是在資訊爆炸的時代, 許多傳統的經濟議題以及計量方法會因爲資訊工程日新月異的技術而改變。若能掌握此趨勢, 加入相關的領域知識, 就能對財經問題提出與傳統不同的見解。例如, 在傳統的計量分析中, 學者大多使用結構性的數據資料, 再透過計量方法來驗證學理或假說。研究者若能適當納入文字資訊 (非結構性的數據) 進行分析, 應可提供更豐富或更具說服力的實證結果。

相較於既有文獻, 本文的貢獻如下。首先, 雖然我們所用的模型是傳統

臺灣大學圖書館期刊資料庫

文字探勘的技術, 但我們加入了不少經濟相關領域的知識。例如, 我們考慮了 FOMC 對三大經濟使命的看法, 並將就業, 物價, 與利率等相關字詞來當成先驗資訊, 讓我們在使用文字探勘技術時更具合理性。其次, 相較於處理文字資訊的財經相關文獻, 我們考慮了複合字以及 FOMC 常用的專門術語, 利用 MAP-PLSA 模型來區分不同主題, 並以情緒分析技巧, 將 FOMC 對三大經濟使命的正/負面看法以數據方式呈現, 再利用迴歸分析, 探討 FOMC 對未來經濟所持之態度與台灣財經變數之間的關聯性。據我們所知, 這是國內、外文獻首次以這種分析方式進行實證研究。最後, 從實證結果中我們發現, 不同 FED 主席的擔任期間所關注的議題大不相同, 並且在每年第一次的 Minutes 文件中, 都會例行性地提到「穩定物價」這一個使命。此外, 我們也發現 FOMC 對三大經濟使命的正/負面看法可能會影響台灣的財經變數; 若沒有利用 MAP-PLSA 模型來區分不同主題, 我們就難以看到這些實證結果。

在後續的研究方面, 我們認爲中文文字探勘技術的精進會是一個很好的發展方向, 也是我們目前的研究重點。此外, 利用不同的文字探勘模型來區分主題 (如 Blei, Ng, and Jordan (2003) 年的 latent Dirichlet allocation 模型), 或是在文中處理一字多義 (polysemy), 同義字 (synonymy) 以及否定句 (negation), 還是利用不同的文字資料探討美國相關財經問題 (如 Huang and Kuan (2017)) 等, 都會是很好的研究方向。

附錄1

爲了讓本文內容更爲連貫, 我們將部份的實證結果列於此附錄中。首先, 由於「穩定物價」主題變數 $X_{t,2}$ 的季節成份明顯, 因此我們在式 (8) 中加入虛擬變數以進行迴歸分析:

$$y_{t+1} = \alpha + \sum_{i=0}^{m} \alpha_i y_{t-i} + \beta_1 \times X_{t,1} + \beta_2 \times X_{t,2} + \beta_3 \times X_{t,3} + \beta_4 \times \text{Dummy}_t + \varepsilon_{t+1},$$

其中 $\text{Dummy}_t = 1$ 爲每年年初第一次召開會議之時程, 而 $\text{Dummy}_t = 0$ 爲其它時間。我們嘗試透過此虛擬變數來補捉每年年初的例行性宣示效

臺灣人文與數位典藏資料庫

表 5: 指標數列 $PN_{i,k}$ 迴歸分析結果

頻率	模型 3 + 虛擬變數							
	β_1	t-stat	β_2	t-stat	β_3	t-stat	β_4	t-stat
日 隔夜拆款成長率	−0.003	−0.822	0.009	1.003	0.006	1.679*	0.909	1.502
美元匯率報酬率	0.002	0.753	0.003	1.335	0.000	0.061	−0.122	−1.166
台股指標報酬率	−0.008	−1.476	−0.014	−1.664*	0.010	1.514	0.622	1.631
週 隔夜拆款成長率	0.029	1.079	0.029	0.785	0.057	2.839**	−2.086	−1.056
美元匯率報酬率	0.002	0.874	0.005	1.302	0.005	1.581	−0.420	−1.628
台股指標報酬率	−0.010	−0.461	−0.004	−0.316	−0.001	−0.031	0.845	0.983
月 領先指標成長率	0.000	−0.624	0.000	−0.135	−0.007	−3.308**	0.070	0.304
同時指標成長率	−0.001	−0.995	0.000	0.137	−0.003	−1.424	0.048	0.207
工業生產成長率	−0.019	−0.520	0.043	1.761*	0.071	2.365**	4.528	1.179
海關出口成長率	0.063	1.654*	0.096	2.932**	0.103	2.581**	3.602	0.600
CPI 成長率	−0.003	−1.146	0.003	1.032	0.001	0.339	−0.670	−2.522**
WPI 成長率	0.000	−0.059	0.003	0.506	0.007	1.804	0.361	1.689*
M1B 成長率	0.006	1.612	−0.003	−0.794	−0.011	−1.522	−2.162	−3.578**
貨幣機構放款與投資成長率	0.003	1.702*	0.002	1.590	0.002	0.834	−0.279	−1.937*

註: 符號 t-stat 爲 t-統計量; *和 **表示在 10%和 5%的水準下顯著。

果。相關的迴歸分析結果列於表 5 與表 6。從表 5 與表 6 的結果中發現, 加入虛擬變數後各變數的係數正負符號與表 2, 表 3 的結果一致, 只是有一些變數的顯著性受到影響。例如表 5 $PN_{i,k}$ 指標數中的月頻率 WPI 與 M1B 成長率等變數, 原本這些變數會受「穩定物價」主題所影響, 但加入虛擬變數後, 其影響效果變不顯著。至於其它變數的部份, 是否加入虛擬變數對係數顯著性的影響差異不大。換言之, 這些指標對台灣的財經變數還是有解釋以及預測能力。

附錄 2

我們也利用 General Inquirer 詞庫的分類方式來建構指標, 並探討新建構的指標是否也對台灣的財經變數有解釋能力。依據 Sadique et al. (2013) 的解構指標方式, 從該詞庫中的增加 (increase) 與減少 (decrease) 分類找出 99(79) 個與增加 (減少) 相關的字詞, 並利用

臺灣大學學術期刊資料庫

表 6: 指標數列 $\text{Tone}_{i,k}$ 迴歸分析結果

頻率	模型 3 + 虛擬變數							
	β_1	t-stat	β_2	t-stat	β_3	t-stat	β_4	t-stat
日 隔夜拆款成長率	−0.343	−0.898	0.685	1.901*	1.435	1.825*	1.050	1.498
美元匯率報酬率	−0.055	−0.672	0.020	0.389	0.349	1.277	0.038	0.650
台股指標報酬率	−0.394	−0.970	−0.072	−0.205	−0.406	−0.504	0.013	0.064
週 隔夜拆款成長率	0.453	0.336	0.337	0.464	5.671	1.508	−0.015	−0.009
美元匯率報酬率	0.334	1.166	−0.046	−0.477	0.633	1.749*	−0.154	−0.634
台股指標報酬率	0.366	0.480	0.458	0.520	−2.082	−1.089	0.245	0.380
月 領先指標成長率	0.034	0.445	−0.061	−1.259	−0.535	−1.342	0.018	0.253
同時指標成長率	0.030	0.349	0.014	0.254	−0.461	−1.015	0.025	0.318
工業生產成長率	1.227	0.914	1.291	1.788*	4.138	1.109	6.044	1.592
海關出口成長率	−1.520	−0.938	1.097	0.751	10.976	2.751**	7.575	1.597
CPI 成長率	−0.224	−1.036	−0.111	−0.589	0.477	1.540	−0.447	−2.171**
WPI 成長率	0.141	0.628	−0.040	−0.309	0.831	2.238**	0.529	2.971**
M1B 成長率	0.413	1.269	−0.312	−0.905	−0.897	−1.493	−2.192	−3.894**
貨幣機構放款與投資成長率	0.469	2.743**	0.032	0.280	0.286	1.415	−0.165	−1.372

註: 符號 t-stat 爲 t-統計量; *和 **表示在 10%和 5%的水準下顯著。

$$\text{Change}_{i,k} = \frac{\text{增加字數}_{i,k} - \text{減少字數}_{i,k}}{\text{增加字數}_{i,k} + \text{減少字數}_{i,k}} \times 100,$$
$$k = 1, \cdots, K; \quad i = 1, \cdots, N,$$

公式建構指標, 進行相關的迴歸分析。建構不同指標之主要原因是文獻上在分析各主題正/負面看法時, 常會因爲解讀的不同而產生困擾。例如負面的「穩定物價」是指對通膨的隱憂, 還是來自於因通膨未達 2%而造成的影響; 本文也有類似的情況。類似這種語意分析 (semantic analysis) 的困擾是目前我們技術上無法克服的問題。當然, 一個解決之道就是以人工方式一篇篇進行判讀, 但這也失去了文字探勘的用意。另一種方式則是利用不同的分類方法再進行分析; 但這也只是從另一個面向來了解各主題的涵意, 還是無法百分之百解決語意分析的困擾。此外, 我們可以看到 General Inquirer 詞庫的增加 (減少) 分類關鍵字數相當少, 都不到 100 個字, 這樣的代表性或許不足, 因此在迴歸實證的解讀上可能要特別小心。但反觀本文

臺灣人文期刊資料庫

表 7: 指標數列 $\text{Change}_{i,k}$ 迴歸分析結果

頻率	模型 3					
	β_1	t-stat	β_2	t-stat	β_3	t-stat
日 隔夜拆款成長率	−0.319	−0.658	0.553	1.228	0.949	1.394
美元匯率報酬率	−0.084	−1.748*	0.038	0.832	0.142	0.590
台股指標報酬率	0.179	0.638	0.086	0.428	−0.881	−1.332
週 隔夜拆款成長率	3.096	2.882**	1.332	1.892*	3.976	1.808*
美元匯率報酬率	0.039	0.397	−0.037	−0.316	0.507	1.782*
台股指標報酬率	−0.948	−1.048	0.342	0.559	−2.540	−1.351
月 領先指標成長率	0.137	2.490**	−0.103	−2.445**	−0.285	−1.076
同時指標成長率	−0.073	−0.969	−0.035	−0.705	−0.283	−0.922
工業生產成長率	1.782	0.968	1.362	1.295	2.326	0.713
海關出口成長率	0.165	0.080	0.935	0.552	7.080	2.016**
CPI 成長率	0.109	0.622	−0.018	−0.121	0.411	1.653*
WPI 成長率	0.296	1.701*	−0.056	−0.350	0.291	1.036
M1B 成長率	0.029	0.073	−0.687	−2.253**	−0.863	−1.772*
貨幣機構放款與投資成長率	0.241	1.788*	0.023	0.203	0.193	0.935

註: 符號 t-stat 爲 t-統計量; *和 **表示在10%和5%的水準下顯著。

所採用的正/負面分類的關鍵字則有 2047/5468 個字, 相較之下有其一定的優勢。表 7 列出利用模型 3 的分析結果, 其中 X 是新建構的 Change 指標。從表中的結果也不難發現, 這指標對台灣的財經變數還是有一些解釋能力以及預測能力。換言之, 本文所採用的文字探勘技術以及 MAP-PLSA 模型確實可以幫助研究者探討台灣的財經變數走勢。

參考文獻

Banchs, Rafael (2012), *Text Mining with MATLAB*, New York: Springer Science and Business Media.

Berger, Helge, Jakob de Haan, and Jan-Egbert Sturm (2011), "Does Money Matter in the ECB Strategy? New Evidence Based on ECB Communication," *International Journal of Finance and Economics*, 16(1), 16–31.

臺灣大學學術期刊資料庫

Bernd, Hayoa and Matthias Neuenkirch (2013), "Do Federal Reserve Presidents Communicate with a Regional Bias?" *Journal of Macroeconomics*, 35(4), 62–72.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3(1), 993–1022.

Chien, Jen-Tzung and Meng-Sung Wu (2008), "Adaptive Bayesian Latent Semantic Analysis," *IEEE Transactions on Audio, Speech, and Language Processing*, 16(1), 198–207.

Doran, James S., David R. Peterson, and Price McKay (2012), "Earnings Conference Call Content and Stock Price: The Case of REITs," *Journal of Real Estate Finance and Economics*, 45(2), 402–434.

Hofmann, Thomas (1999), "Probabilistic Latent Semantic Indexing," *ACM SIGIR Forum*, 51(2), 211–218.

Huang, Yu-Lieh and Chung-Ming Kuan (2017), "Prediction with FOMC Minutes: An Application of Text Mining," Working Paper, National Tsing Hua University.

Jubinskia, Daniel and Marc Tomljanovich (2013), "Do FOMC Minutes Matter to Markets? An Intraday Analysis of FOMC Minutes Releases on Individual Equity Volatility and Returns," *Review of Financial Economics*, 22(3), 86–97.

Loughran, Tim and Bill McDonald (2011), "When Is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *Journal of Finance*, 66(1), 35–65.

——— (2014), "Measuring Readability in Financial Disclosures," *Journal of Finance*, 69(1), 1643–1671.

——— (2016), "Textual Analysis in Accounting and Finance: A Survey," *Journal of Accounting Research*, 54(1), 1187–1230.

Lucca, David O. and Francesco Trebbi (2009), "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements," NBER Working Paper, No. 15367.

Murphy, Kevin P. (2012), *Machine Learning: A Probabilistic Perspective*, Cambridge: MIT press.

Newey, Whitney K. and Kenneth D. West (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–708.

Rosa, Carlo (2013), "The Financial Market Effect of FOMC Minutes," *FRBNY Economic Policy Review*, 19(2), 67–81.

臺灣人壽保險研究資料庫

Sadique, Shibley, Francis In, Madhu Veeraraghavan, and Paul Wachtel (2013), "Soft Information and Economic Activity: Evidence from the Beige Book," *Journal of Macroeconomics*, 37(3), 81–92.

投稿日期: 2017年7月6日, 接受日期: 2018年4月23日

Text Mining of the FOMC Minutes and Forecasts of Taiwan Economic Variables

Yu-Lieh Huang

Department of Quantitative Finance, National Tsing-Hua University

Chung-Ming Kuan

Department of Finance, National Taiwan University, Center for Research in Econometric Theory and Applications (CRETA), National Taiwan University

In this paper we extract useful information from the minutes of the Federal Open Market Committee (FOMC) and examine how such information can help predict economic/financial variables. Based on the minutes during 1993–2016, we conduct sentiment analysis to determine the FOMC's attitude towards different topics, i.e., the mandates of the FED. Our approach is different from related studies in the following respects. First, we identify compound words which carry more specific meaning than do single words. Second, we adopt the topic model, MAP-PLSA, for estimating the conditional probabilities of these words/terms, which in turn can be used to classify sentences in the minutes under different topics. Third, the attitude towards each topic is determined by the "tone" of its sentences. We then proceed to evaluate whether the FOMC's attitude towards different topics can be used to improve economic forecasts.

Keywords: federal Open Market Committee, sentiment analysis, text mining, topic models

JEL classification: C10, E30, O11

臺灣大學學術期刊資料庫

臺灣大學學術期刊資料庫