

使用 Google 搜尋工具改善臺灣消費模型推估能力*

劉欣姿**

- | | |
|-------------|-------------|
| 壹、研究緣起與目的 | 肆、實證結果 |
| 貳、文獻回顧 | 伍、研究結論與政策建議 |
| 參、研究方法與研究過程 | |

摘 要

國外研究指出網路行為與現實世界的經濟指標之間有很強的關聯，且每人平均所得較高國家之網路使用者愈可能用網路搜尋未來的資料，亦有許多國外文獻發現增加 Google 搜尋指標，有助於民間消費的推估。本研究運用 Google 搜尋趨勢的資料建構指標，來測試是否能對臺灣消費的推估有助益。為更即掌握消費動向，本文選擇商業營業額作為消費的高頻率指標，進行實證研究，結果如下：

1. Google 搜尋指標對商業營業額的推估有助益，但程度有限，可能與受限於資料長度、及選擇搜尋關鍵字有關。
2. 商業營業額與前期顯著正相關，表示商業營業額存在平滑性。
3. 股價指數、長短期利差與商業營業額顯著相關，且這 2 項指標有助於商業營業額推估。
4. 消費者信心指數、工業及服務業經常性平均薪資、失業率等指標未領先商業營業額。
5. 工業及服務業經常性平均薪資與商業營業額的關係較不顯著。

因應網路資訊時代，指標的建構方式與資料來源可更多元，建議未來可運用 Google 搜尋工具，建構反映不同總體資訊的 Google 搜尋指標，作為重要總體指標推估之參考。

* 本文參加國發會 103 年度研究發展作品評選，榮獲經濟及財金政策類佳作獎。

** 作者為經濟發展處科員。

Use Google Searches for Nowcasting Taiwan's Consumption

Hsin-Tzu Liu

Officer

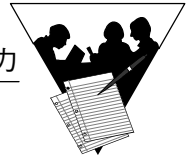
Economic Development Department, NDC

Abstract

Many studies indicate that there is a strong link between network behavior and economic indicators in the real world. The country's higher per capita GDP users may search more future data on the Internet. This study uses Google Trend's data to construct Google indicators and tests to see if it will be effective for Taiwan's consumption estimation. This article chooses sales of trade and food services as consumption's high frequency indicator. The empirical results are as follows:

1. Google indicators are useful for sales of trade and food services estimation, but its forecasting performance may be limited by the data length and the search keywords.
2. Trade and food services have a high positive autocorrelation.
3. TAIEX's average closing price and interest rate spread have a significant relationship with trade and food services.
4. Consumer confidence, average regular earnings in industry and services, and the unemployment rate don't have lead than trade and food services.
5. The connection between average regular earnings in industry and services and trade and food services is not significant.

In response to the Internet time, constructing indicators methods and data sources can lead to more diversity. This study recommends using the Google search tool to construct Google indicators reflecting the macroeconomic situation in the future.



壹、研究緣起與目的

一、研究緣起

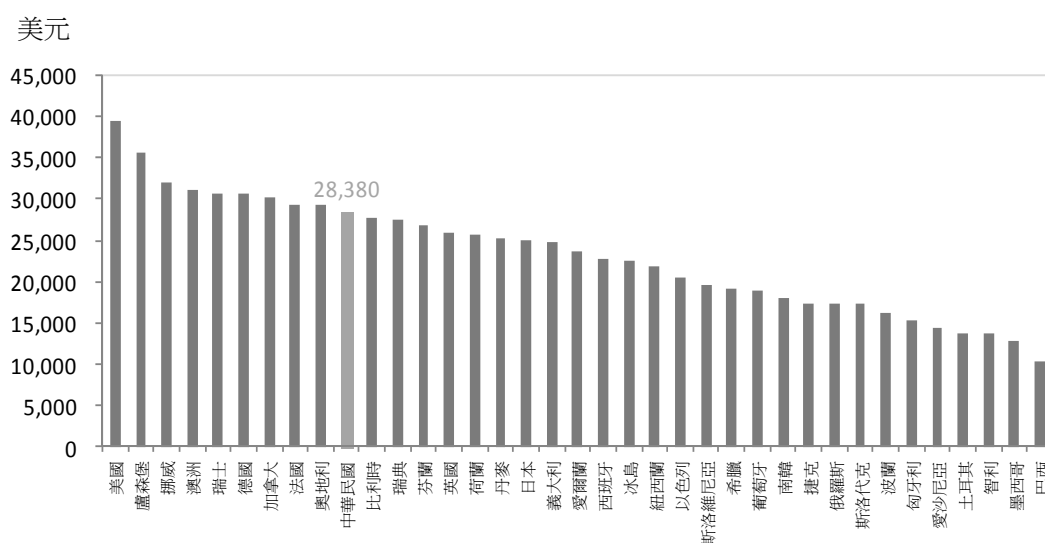
Preis, Moat, Stanley and Bishop (2012) 為了解網路搜尋條件與真實世界的關聯性，並進行跨國比較分析，研究發現相較於搜尋過去資料，每人平均所得較高國家之網路使用者愈可能用網路搜尋未來的資料¹，報告中亦證實網路行為與現實世界的經濟指標之間有很強的關聯。而根據行政院主計總處 2014 年國民幸福指數統計經家庭消費 PPP 調整後之 2011 年每人可支配所得² 為 2.8 萬美元 (PPP)，與 OECD 國家比較，37 國家中我國排名為第 10 名，優於調查亞洲國家 (日本及南韓)；參見圖 1。

其實在國際上已經廣泛運用 Google 搜尋工具在研究中，如 Choi and Varian (2009) 主張網路搜尋不僅是衡量面臨流感時就醫行為的可靠指標，亦可在官方數據出現前，提供預測經濟活動的有用資訊。而 Askitas and Zimmermann (2009) 則是運用在失業率的預測上，他們證實搜尋關鍵字與失業率有高度相關。D'Amuri and Marcucci (2010) 則是用 Google 搜尋工具建立網路工作尋找指標 (Internet job-search indicator) 作為預測美國失業率的領先指標。Kholodilin, Podstawski and Siliverstovs (2010) 研究發現加上 Google 搜尋指標的一階自我迴歸模型 (Autoregressive

¹ 未來的資料係觀察使用者每週搜尋未來一年的資料，例如給定 2010 年，觀察使用者搜尋 2011 年資料的筆數。

² 依據國民所得帳定義，係指家庭部門及對家庭服務之民間非營利機構 (NPISH) 之所得毛額 (薪資、產業主所得、財產所得及移轉收入淨額) 加上政府對家庭提供的實物社會移轉，減所得及財產稅、社會安全捐及折舊，以家庭消費購買力平價 (PPP) 換算，並折算為平均每人金額。

Model, AR) 有助於改善美國民間消費推估。而 Toth and Hajdu (2012) 則是運用這項工具推估匈牙利的零售銷售、民間消費及汽車銷售，其結果亦優於傳統一階的 AR 模型。

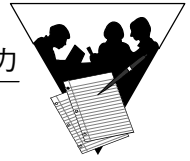


資料來源：行政院主計總處

圖 1 OECD 美好幸福指數指標—每人可支配所得 (PPP) 2011 年資料

二、研究目的

由於國外研究指出網路行為與現實世界的經濟指標之間有很強的關聯，每人平均所得較高國家之網路使用者愈可能用網路搜尋未來的資料，加上許多國外文獻發現增加 Google 搜尋指標，有助於民間消費的推估。而我國每人可支配所得相對較高，且目前國內傳統消費模型多使用官方發布的過去資料。鑑此，本研究目的，期望透過網路資料建構 Google 搜尋指標，來增加消費模型推估能力。



貳、文獻回顧

一、經濟變數之選取

彭素玲、周濟（2001）探討如何用高頻率資料（月或日資料）來改進總體經濟季預測，由於臺灣 GDP 支出面的高頻率較生產面來得多，因此，研究從支出面高頻指標，透過主成分分析法，萃取出相當的資訊，做為 CQM 模型預測實質 GDP、名目 GDP 與 GDP 平減指數的參考值。其中關於消費的部分，選用每月商業營業額作為民間消費（名目）之高頻率綜合性指標。

林建甫、周麗芳、何金巡（2005）為建構臺灣總體經濟計量模型，以模擬油價變動對國內景氣及政府財政之影響效果，模型之建構是根據總體經濟和計量經濟理論，估計變數間的經濟行為，並透過國民會計制度結合而成的聯立方程式體系。模型之總體經濟部門包括：商品市場、貨幣市場、外匯市場、勞動市場。其中民間消費由家庭可支配所得、貨幣存量、股價指數、消費習性、利率、失業率、所得分配不均度及人口數量共同決定。

黃舜卿（2008）研究發現領先 1 季的消費者信心指數對當季的實質消費支出成長率具有解釋能力，若加上參考國外文獻選擇實質民間消費支出、實質可支配所得、就業人口、股價指數、中央銀行定期存單利率（初級市場 92 天至 182 天）等總體經濟變數建構的模型，差異不大，但當時我國消費者信心指數相較於一般的總體經濟變數，能提供額外有限的訊息。當時消費者信心指數資料期間自 2001 年第 1 季至 2006 年第 4 季，推論受限於時間數列不長，影響消費者信心指數預測未來消費的能力。

彭素玲、郭迺鋒、周濟、方文秀 (2009) 同時考量少子化與老齡化、M 型所得分配趨勢、與產業結構改變等因素對臺灣消費之影響。為了解民間消費結構之變化趨勢，觀察 1994 年至 2005 年家計消費調查資料的 12 類消費支出的變化。

Kholodilin, Podstawski and Siliverstovs (2010) 比較每月實質消費之自我相關模型，或加上消費者信心指數³，或加上金融變數⁴，或加上 Google 搜尋指標 (Google Indicators)，資料期間 2005 年 1 月至 2009 年 12 月，研究結果發現，比起僅用 1 階 AR 模型，加上 Google 搜尋指標的模型較精確。

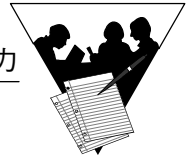
二、Google 搜尋指標之建立

D'Amuri and Marcucci (2010) 使用 Google 搜尋透視 (Google Insight) 的週資料轉換成月資料後再做季調處理，總共選擇 220 筆時間數列資料，轉換成年成長率，資料期間為 2005 年 1 月至 2009 年 12 月。透過 Stock and Watson (1990、2002) 的主成分分析法找出這 220 筆資料之主要因子數列，排除對主要因子貢獻較小的數列，選擇貢獻較大的前 10 條，利用這 10 條數列之主要因子數列作為預測消費的 Google 搜尋指標 (Google Indicators)。

Toth and Hajdu (2012) 則是以家庭消費為方向，決定 Google 搜尋的類別，包含食物飲料、酒、家用配料、居家修繕、融資、能源用電、健康、汽車零件、網路電信、款待、電影、電

³ 美國 Conference Board 消費者信心指數及密西根大學消費者信心指數。

⁴ 短期利率、長期利率、長短期利差、標準普爾 500 指數、或前述指標每年的差額。



動遊戲、書、藝術、教育、托嬰等，然後篩選與民間消費有顯著相關的類別數列，以主成分分析法，萃取主要因子數列。

黃月盈（2012）參考 Stock and Watson（1998、2002）利用主成分分析法建立可預測景氣變化的擴散指標，若比較主成分分析法與綜合指數法在歷次景氣循環的領先期數，兩者在各次循環的領先期數均相同，波動幅度亦大致相仿。

參、研究方法與研究過程

一、資料選取與處理

（一）資料選取

根據張志揚（2013）研究曾指出，若我們能充分掌握月資料的走勢，將對即期季模型-GDP 的短期預測有較大的助益，但對長期的預測則幫助不大。為即時掌握消費短期變化，本研究參考文獻選擇商業營業額（月資料）作為消費的高頻率指標。因主計總處的名目民間消費為季資料，為確認名目民間消費與商業營業額之關係，必須先將零售及餐飲業營業額月資料轉換成季資料，再進行比較，計算 2001 年第 1 季至 2014 年第 1 季兩者相關係數高達 0.99，表示商業營業額可作為消費的高頻指標。

由於 Google 搜尋趨勢的資料最早自 2004 年 1 月起，因此，本文資料期間由 2004 年 1 月至 2014 年 5 月，參考國內外文獻、102 年度行政院主計總處家庭消費支出按消費型態分，選取的變數資料如表 1。

表 1 變數資料名稱及來源

類別	變數	資料來源
消費 C_t	商業營業額	經濟部
信心調查 X_t	消費者信心指數	中央大學台灣經濟發展研究中心
勞動市場 Y_{it}	工業及服務業經常性平均薪資	行政院主計總處
	失業率	行政院主計總處
金融 Z_{jt}	股價指數	台灣證券交易所
	長短期利差 (10 年期公債殖利率-31 到 90 天商業本票利率)	中央銀行
Google 搜尋指標 G_t	食品、飲料、菸 ⁵ 、服裝 ⁶ 、鞋子、襪、房地產 ⁷ 、水電、瓦斯、燃料、家具 ⁸ 、清潔 ⁹ 、醫療、保健、運輸、交通、通訊、休閒、文化、教育、餐廳、旅館	Google 搜尋趨勢 ¹⁰

(二) 資料處理

無論是自我迴歸模型或主成分分析法，均要求變數資料必須為定態數列，因此，本研究採用二階段 Hodrick-Prescott (HP) 濾波法將資料剔除長期趨勢後，並平滑化與標準化，以符合定態的條件。有關 HP 濾波法之設定，本研究參酌徐志宏、周大森 (2010) 進行第一、二階段參數之設定。

⁵ 因菸草搜尋結果不理想，研究過程中以菸作為搜尋選擇字詞。

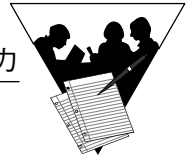
⁶ 因衣著搜尋結果不理想，研究過程中以服裝作為搜尋選擇字詞。

⁷ 因住宅服務搜尋不到資料，研究過程中以房地產作為搜尋選擇字詞。

⁸ 因家具設備搜尋結果不理想，研究過程中以家具作為搜尋選擇字詞。

⁹ 因家務維護搜尋不到資料，研究過程中以清潔作為搜尋選擇字詞。

¹⁰ 搜尋選擇字詞在臺灣，每段期間熱門度變化。



本研究參考 102 年度行政院主計總處家庭消費支出按消費型態分，選取 Google 搜尋指標的構成項目，但未剔除與零售及餐飲業營業額較不顯著相關的變數。而在估算 Google 搜尋指標，係用 Stock and Watson (2002) 主成分分析法，參照黃裕烈 (2012) 採用最小平方法估算出主要因子數列，經指數化處理，得到 Google 搜尋指標。

二、模型的建置

本研究之目的為證實 Google 搜尋指標對消費推估有無幫助，故並沒有使用總體計量一般均衡模型，而是採用一階 AR 模型¹¹，模型設定如下：

$$\text{模型 1 : } C_t = a + b_1 C_{t-1} + \varepsilon_t$$

$$\text{模型 2 : } C_t = a + b_1 C_{t-1} + b_2 X_t + \varepsilon_t$$

$$\text{模型 3 : } C_t = a + b_1 C_{t-1} + b_2 X_t + b_{3i} Y_{it} + \varepsilon_t$$

$$\text{模型 4 : } C_t = a + b_1 C_{t-1} + b_2 X_t + b_{3i} Y_{it} + b_{4j} Z_{jt} + \varepsilon_t$$

$$\text{模型 5 : } C_t = a + b_1 C_{t-1} + b_2 X_t + b_{3i} Y_{it} + b_{4j} Z_{jt} + b_5 G_t + \varepsilon_t$$

而模型中各變數的最適期數，依照與商業營業額最相關的領先期數決定。

三、推估能力評估

為了評估模型的推估能力，本研究採用文獻上常用的衡量指標均方根誤 (root mean square error, RMSE)、絕對誤差平均值 (mean absolute error, MAE)。

¹¹ 參考 Kholodilin, Podstawski and Siliverstovs (2010)、Toth and Hajdu (2012) 的研究方法。

估計誤差 $e_{it} = y_{it} - \hat{y}_{it}$, $i = 1, 2$, $t = 1, \dots, n$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n e_{it}^2} \quad MAE = \frac{1}{n} \sum_{i=1}^n |e_{it}|$$

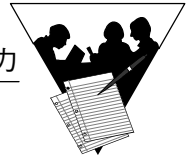
肆、實證結果

一、最適期數的決定

本文是透模型各變數與商業營業額最相關的領先期數，決定模型中各變數的最適期數。觀察各變數與商業營業額領先 0 至 12 期的相關係數，發現消費者信心指數、工業及服務業經常性平均薪資、失業率、最相關的領先期數為當期，而商業營業額及股價指數領先 1 個月，長短期利差領先 6 個月，Google 搜尋指標的領先期數為 12 個月，參見表 2。

表 2 各變數與商業營業額最相關的領先期數及相關係數

變數	領先期數	相關係數
商業營業額	1	0.9738
消費者信心指數	0	0.8062
工業及服務業經常性平均薪資	0	0.8207
失業率	0	-0.6400
股價指數	1	0.9420
長短期利差	6	0.2310
Google 搜尋指標	12	-0.0935



二、模型的估計結果

接著使用表 2 各變數的最適期數及 2004 年 1 月至 2014 年 5 月的數列資料，進行模型的估計。由表 3 可知，不論使用那個模型，商業營業額均顯著受前期影響。至於其他變數方面，在多數模型中消費者信心指數、股價指數、長短期利差具顯著性，而工業及服務業經常性平均薪資則較不顯著。從模型 4 與模型 5 可看出，模型 5 的 R^2 及 Adjust R^2 較高，而其 AIC 與 SC 的數值均較模型 4 來得小，顯示 Google 搜尋指標對商業營業額的推估可能有幫助。

表 3 模型的估計結果-整段期間

係數	模型 1	模型 2	模型 3	模型 4	模型 5
a	2.622	0.490	-37.310	-29.183*	-23.440*
b_1	0.974*	0.886*	0.926*	0.859*	0.874*
b_2		0.109*	0.125*	0.084*	0.145*
b_{31}			0.130*	0.027	-0.023
b_{32}			0.191*	0.152*	0.174*
b_{41}				0.291*	0.301*
b_{42}				-0.122*	-0.233*
b_5					-0.003*
R^2	0.949	0.954	0.972	0.978	0.984
Adjust R^2	0.948	0.953	0.971	0.977	0.982
AIC	0.551	0.468	0.001	-0.174	-0.392
SC	0.598	0.536	0.115	-0.010	-0.198

註：*為在 1% 下拒絕虛無假設。

三、模型的推估能力

因為與商業營業額最相關領先期數為 0 之變數，對於推估目標而言屬於同期的資訊，不適合用在預測或推估上，因此在評估模型推估能力之前，應去除掉最相關領先期數為當期之變數，如消費者信心指數、工業及服務業經常性平均薪資、失業率。然後，將 2004 年 1 月至 2012 年 12 月資料當成樣本內資料估計模型係數，用 EViews 動態推估¹² 2013 年 1 月至 2014 年 5 月的商業營業額。由表 4 可知，模型 5 的推估績效未優於模型 4，顯示 Google 搜尋指標對商業營業額的推估助益程度有限，可能與受限於資料長度、及選擇搜尋關鍵字有關。

表 4 推估績效之比較

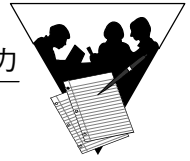
	模型 1	模型 2	模型 3	模型 4	模型 5
RMSE	0.4793	0.4793	0.4793	0.3150	0.4186
MAE	0.3670	0.3670	0.3670	0.2297	0.3260

伍、研究結論與政策建議

一、研究結論

本文選擇商業營業額作為消費的高頻率指標，使用商業營業額、消費者信心指數、工業及服務業經常性平均薪資、失業率、股價指數、長短期利差等重要經濟指標，及 Google 搜尋指標進行商業營業額推估。本次研究發現有

¹² EViews 動態預測是將推估資料重覆代入預測。以模型 1 為例，帶入 2013 年 1 月推估值，推估 2 月資料，然後帶入模型，可得到 3 月資料，以此類推，重覆代入。



- (一) Google 搜尋指標對商業營業額的推估有助益，但程度有限，可能與受限於資料長度、及選擇搜尋關鍵字有關。
- (二) 商業營業額與前期顯著正相關，表示商業營業額存在平滑性。
- (三) 長短期利差、股價指數等指標與商業營業額顯著相關，且這 2 項指標有助於商業營業額推估。
- (四) 消費者信心指數、工業及服務業經常性平均薪資、失業率等指標未領先商業營業額。
- (五) 工業及服務業經常性平均薪資與商業營業額的關係較不顯著。

二、政策建議

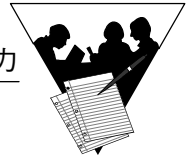
隨著網路普及化，我國的經濟活動與一般社會大眾的行為與網路密切連結，此時，我們在經濟指標的資料來源，不應侷限於傳統的統計資料。本文不僅試著透過 Google 搜尋工具，建構新的指標，同時檢視傳統經濟指標與商業營業額相關性，作為政策之參考，本文提出二點政策建議如下：

(一) 增加 Google 搜尋指標之研究

國際上已經廣泛運用 Google 搜尋工具，在研究面臨流感時就醫行為，以及消費、失業率、零售銷售額、住宅出售、遊客人數等預測。但國內很少有人探討、研究 Google 搜尋指標。建議未來增加 Google 搜尋指標之研究，有助於在官方數據出現前，提供政府預測經濟活動之參考。

(二) 研究季消費推估的參考指標

由於消費者信心指數、工業及服務業經常性平均薪資、股價指數等指標有助於商業營業額推估。而商業營業額作為消費的高頻率指標，建議未來可研究前述指標是否能改善每季消費推估。



參考文獻

1. 行政院主計總處 (2014), 國民幸福指數年報。
2. 林建甫、周麗芳、何金巡 (2005), 「建構台灣總體計量模型:油價、景氣與政府財政的總體經濟計量分析」, 總體經濟計量模型研討會, 高雄大學。
3. 徐志宏、周大森 (2010), 「近期台灣景氣循環峰谷之認定」, 經濟研究, 第 10 期, 行政院經濟建設委員會經濟研究處。
4. 張志揚 (2013), 「台灣總體經濟即期季模型之建立—運用月資料改善國民所得預測」, 中央銀行季刊, 第三十五卷第三期, 37-60。
5. 黃舜卿 (2008), 「我國消費者信心指數與民間消費支出之研究」, 經濟研究, 第 8 期, 行政院經濟建設委員會經濟研究處。
6. 黃裕烈 (2012), 「臺灣景氣基準循環指數之檢討與改進」, 行政院經濟建設委員會委託研究報告。
7. 黃月盈 (2012), 「建構景氣指標方法之研析」, 經濟研究, 第 12 期, 行政院經濟建設委員會經濟研究處。
8. 彭素玲、周濟 (2001) 「台灣總體經濟即期季模型之建立與應用」, 臺灣經濟預測與政策, 32:1, 頁 1-64
9. 彭素玲、郭迺鋒、周濟、方文秀 (2009), 「人口年齡結構、所得分配與產業結構轉型對臺灣民間消費與總體產出之影響」, 臺灣經濟預測與政策, 39:2, 51-101。
10. Askatas, H. and K.F. Zimmermann (2009), “Google Econometrics and Unemployment Forecasting.” IZA Discussion Papers No. 4201.
11. Choi, H. and H. Varian (2009), “Predicting the Present with Google Trends.” Technical report, Google Inc.
12. D’Amuri, F., and J. Marcucci (2009), “Google it! Forecasting the US unemployment rate with a Google Job search index.” FEEM Working Paper No. 31. 2010.
13. Good Trend, <http://www.google.com/trends/>
14. Kholodilin, K. A., M. Podstawski, and B. Siliverstovs (2010), “Do Google Searches Help in Nowcasting Private Consumption?: A Real-Time Evidence for the US,” KOF Working Paper No. 256.
15. Preis, T., H. S. Moat, H. E. Stanley, and S. R. Bishop (2012), “Quantifying the Advantage of Looking Forward”. Scientific Reports 2: Published 05 April 2012.
16. Stock, J. and M. Watson (1999), “Forecasting inflation”. Journal of Monetary

Economics No 44, 293-335.

17. Stock, J. and M. Watson (2002) , “Macroeconomic Forecasting using diffusion indexes”. *Journal of Business & Economic*, 20 (2) , 147-162.
18. Toth, I. J. and M. Hajdu (2013) , “Google as a tool for nowcasting household consumption: estimations on Hungarian data.” In 31th CIRET Conference, Vienna. https://www.ciret.org/conferences/vienna_2012/papers/upload/p_218-934263.pdf. Accessed (Vol. 7) .