

## 《信報》專欄〈一名經人〉

### 民調建指數模型 希拉里五盤四勝

預測美國總統選舉結果，除了以計量經濟模型（但大多由政治學者建立）推算外，亦有另一大類稱為指數模型（index models）可測結果。兩類有何不同？模型要測結果，方程的左邊肯定是某黨的支持或得票率，至於右邊則為一堆因素，但關鍵是孰輕孰重。計量模型的精髓是以歷史數據計出各因素的比重，但指數模型則以模型外的資料（prior knowledge）判斷，譬如候選人的生平、過往的領導能力等，但非吹水，一樣有數計。

#### 一、傳記模型

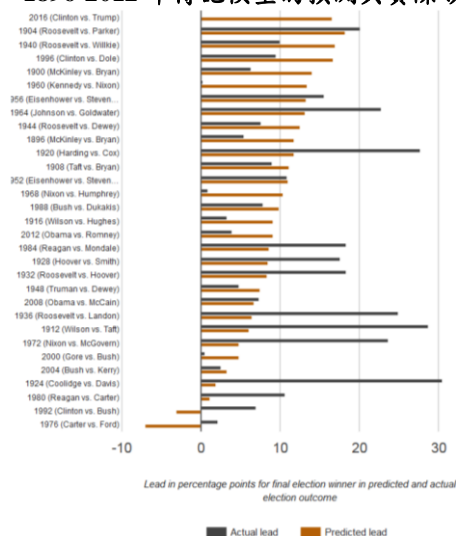
模型由市場系教授 J. Scott Armstrong 和管理系教授 Andreas Graefe 共同提出，基於 58 個可能影響選舉結果的傳記線索。研究員以心理學手法問及領袖特質，以薈萃分析（meta-analyses）找出所需具備的條件如智慧、身高、有否離婚等，其實都頗常識的。在 58 個條件中，8 個關於家庭，9 個關於教育，14 個關於政治生活，7 個關於怎抗逆，另外 20 個關於其他；而這 58 個條件中，47 個為是非題，11 個為比較題，皆 1 或 0。這 58 條 1 或 0 分的題目權重相同，將之加起來，便得兩候選人的傳記分數（ $D, R$ ）。

至於實際上執政黨候選人的得票率（ $V$ ）則以下列公式推算：

$$V = 19.9 + 60.5D / (D + R)$$

希拉里與特朗普的得分分別為 19 和 11，將值塞進右方即得希拉里得票率 58.3%。那麼模型的往績表現又如何？從 1896 至 2012 年卅次選舉中，只失手過兩次【圖一】。

圖一：1896-2012 年傳記模型的預測與實際領先度



## 二、事件指數模型

這也由上述二人提出，基於選民對候選人處理事件的能力。方法是列出大堆議題，然後基於 Roper Center for Public Opinion Research 的調查數據得分將各項填上 1 或 0，再將所有項的分數加總除以項的總數得出分數比例 ( $I$ )，執政黨的得票率公式如下：

$$V = 41.8 + 18.0I$$

從今年 7 月 10 日的調查得分計算，希拉里得票率將為 56.2%。那模型有幾準呢？據樣本外預測所見，無論以兩個月或選舉前夕預測，皆是歷來僅出錯一次【表一】。

表一：事件指數模型的預測與實際的平均絕對誤差 (MAE)

Election	Actual	63 days prior to Election Day		Election Eve	
		Predicted	AE	Predicted	AE
1972	61.8	53.4	8.4	54.1	7.7
1976	48.9	47.0	2.0	44.6	4.4
1980	44.7	52.9*	8.2	53.5*	8.8
1984	59.2	53.4	5.7	55.2	3.9
1988	53.9	51.5	2.4	52.4	1.5
1992	46.5	46.7	0.1	45.3	1.2
1996	54.7	57.7	3.0	58.1	3.4
2000	50.3	51.4	1.1	53.1	2.8
2004	51.2	48.8*	2.4	51.5	0.3
2008	46.3	48.1	1.7	48.0	1.7
MAE			3.5		3.6

\*wrong prediction of the popular vote winner

## 三、事件與領袖模型

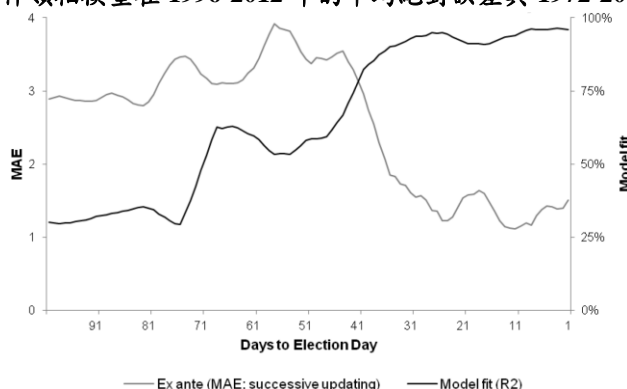
這同樣由 Andreas Graefe 提出，包括候選人處理事件與領導能力。在處理事件上，資料來自兩個調查：一是候選人面對當前最大的問題，二是其處理這問題的能力 ( $I$ )；至於領導能力 ( $L$ )，同樣取自調查。調查分別在選前 100 日與 1 日進行，方程如下：

$$V_{100} = 26.5 + 27.6I_{100} + 23.2L_{100}$$

$$V_1 = 9.6 + 49.6I_1 + 30.7L_1$$

選前 100 日的三個系數皆不統計顯著，但到選前 1 日則三個都很顯著。難道真要到 1 日前才知結果？以過往數據計算模型的誤差與  $R^2$ ，可見踏入選前一個月，誤差跌定而  $R^2$  升定，故約一個月前的調查已可靠【圖二】。目前的數據是希拉里得票率為 52.1%。

圖二：事件領袖模型在 1996-2012 年的平均絕對誤差與 1972-2012 年的  $R^2$



#### 四、大事模型

這同樣由傳記模型的二人提出，方法跟上述的相近，但僅聚焦於大事之上。首先，從調查中認出選民最關注的一件大事，之後找出選民於此大事上支持執政黨的比例（*BI*）。以1972至2012年選前11日的調查數據推算，執政黨候選人的得票率如下：

$$V = 25.9 + 0.5BI$$

最新計算顯示，希拉里得票率為51.7%。研究員將此模型與上次的計量模型比較了2000至08年三次大選表現，可見平均絕對誤差雖然比上不足但亦比下有餘【表二】。

**表二：大事模型的預測與實際的平均絕對誤差**

Model	Approx. date of forecast	2000	2004	2008	MAE
BI-M	Late August / early September	6.8	0.8	3.1	3.6
<b>Econometric models</b>					
Norpoth	January	4.7	3.5	3.6	3.9
Lockerbie	May / June	10.0	6.4	4.5	7.0
Fair	Late July	0.5	6.3	2.2	3.0
Abramowitz	Late July / August	2.9	2.5	0.6	2.0
Holbrook	Late August / early September	10.0	3.3	2.0	5.1
Lewis-Beck and Tien	Late August	5.1	1.3*	3.6	3.3
Wlezien and Erikson	Late August	4.9	0.5	1.5	2.3
Campbell	Early September	2.5	2.6	6.4*	3.8
MAE (error of the typical econometric model)		5.1	3.6	2.6	3.8

\* predicted the wrong winner

#### 五、白宮鑰匙模型

這由歷史學教授 Allan Lichtman 與數學地球物理學兼地震學教授 Vladimir Keilis-Borok 於1981年首次提出，列出十餘條入主白宮的條件【表三】。若答「是 (True, *T*)」的不多於5項，執政黨勝；若佔6項或以上，則執政黨敗。執政黨得票率方程如下：

$$V = 37.2 + 1.8T$$

今回答「是」的有7項，即希拉里得票率為49.7%，是五個模型中唯一輸的一個。

表三：白宮鑰匙模型所列的入主白宮「鑰匙」

1. Has the incumbent party been in office more than a single term?
2. Did the incumbent party gain more than 50% of the vote cast in the previous election?\*
3. Was there major third party activity during the election year?
4. Was there a serious contest for the nomination of the incumbent party candidate?
5. Was the incumbent party candidate the sitting president?
6. Was the election year a time of recession or depression?
7. Was the yearly mean per capita rate of growth in real gross national product during the incumbent administration equal to or greater than the mean rate in the previous 8 years and equal to or greater than 1%?†
8. Did the incumbent president initiate major changes in national policy?
9. Was there major social unrest in the nation during the incumbent administration?
10. Was the incumbent administration tainted by major scandal?
11. Is the incumbent party candidate charismatic or a national hero?
12. Is the challenging party candidate charismatic or a national hero?

與計量模型不同，基於調查的指數模型，五個中有四個料希拉里勝。

羅家聰  
環球金融市場部