



以預測市場匯聚資訊支援決策之方法： 從機率預測到類別判定

林鴻文^a 童振源^{b,*} 葉家興^c

^a 中山大學(大陸)南方學院經濟學與商務管理系

^b 政治大學國家發展研究所

^c 香港中文大學金融學系

摘要

本文成功建構預測市場的最適價格門檻，作為判定預測事件是否發生的依據，將事件發生的機率預測 (probabilistic forecasting) 轉換成事件發生與否的類別預測 (categorical forecasting)，作為公共政策與企業決策的重要依據。以 2006-2011 年的未來事件交易所 7,881 個預測事件樣本分析，本文透過計算所有價格的準確率尋找最適價格門檻，並且比較價格 50 門檻與最高價門檻準則的準確率。本文發現：全部預測事件樣本的最適價門檻為價格 60，其判定預測事件的準確率比價格 50 門檻與最高價門檻準則都要高，具有統計上的顯著性，並且通過樣本外測試。

關鍵詞：預測市場、機率預測、類別預測、最適價格門檻、支援決策

A Method to Support Decision-making Based upon the Information Aggregation of Prediction Markets: from Probabilistic Forecasting to Categorical Forecasting

Hung-Wen Lin^a Chen-Yuan Tung^b Jason Yeh^c

^aDepartment of Economics and Business Management, Nanfang College of Sun Yat-Sen University

^bGraduate Institute of Development Studies, National Chengchi University

^cDepartment of Finance, The Chinese University of Hong Kong

Abstract

We successfully construct a methodology to find optimal price thresholds for prediction markets as the basis of categorizing the occurrence of the prediction events. This method enables us to transform the probabilistic forecasting of event occurrence

* 通訊作者 電子郵件：ctung@nccu.edu.tw

DOI: 10.6188/JEB.2014.16(2).01



to the categorical forecasting, thereby creating an important reference for public policy and corporate decision-making. Based upon 7,881 samples of prediction events from the Exchange of Future Events between 2006 and 2011, we determine optimal price thresholds and compare the accuracy rates with those obtained by two benchmark rules. This paper finds that the price of 60 is the optimal price threshold for the all sample studied, which yields an accuracy rate higher than those obtained by the common threshold of 50 and the highest-price threshold. The superiority is statistically significant and is also robust in out-of-sample tests.

Keywords: prediction markets, probabilistic forecasting, categorical forecasting, optimal price threshold, decision-making support.

1. 緒論

「預測市場」，又稱之為「資訊市場」，其運作類似一般的「期貨市場」，讓參與者買賣「未來事件合約」的期貨市場，透過這個網路平台機制來彙整各方面的資訊，預測未來事件發生的結果，或者匯聚眾人的智慧，判斷未來事件的發展趨勢與眾人偏好。透過預測市場的資訊匯流所達成的預測與創新功能比起傳統的民意調查、焦點團體或訪談調查更具有優勢，能夠領先、即時、連續、長期提供決策與避險的重要參考依據。

最近 25 年來，「預測市場」(prediction markets) 的研究方法已經普遍被運用到預測各類事件或領域，包括政治、經濟、運動、娛樂、產品開發、政策研發、商業活動及企業管理等等。目前較為活躍的公開預測市場包括：愛荷華電子市場 (Iowa Electronic Markets)、好萊塢證券交易所 (Hollywood Stock Exchange)、Intrade、iPredict、遠見交易所 (Foresight Exchange)、及未來事件交易所 (Exchange of Future Events)。¹

2007 年 5 月，包括 Kenneth Arrow, Daniel Kahneman, Thomas Schelling 與 Vernon Smith 四位諾貝爾經濟學獎得主在內的二十五位知名經濟學家連署發表「預測市場宣言」(Statement on Prediction Markets)。他們共同認為預測市場可以顯著提昇公部門和私部門的決策品質 (Arrow et al., 2007)。隔年 5 月，Arrow 再度領銜二十二位經濟學家，在著名的《科學》(Science) 雜誌發表「預測市場的許諾」

¹ 網址分別為：www.biz.uiowa.edu/iem; www.hsx.com; www.intrade.com; www.ipredict.co.nz; www.ideosphere.com; xfuture.org.



(The Promise of Prediction Markets)，再次強調群眾智慧的預測能力 (Arrow et al., 2008)。

已有不少國際大企業利用預測市場的網路平台開發公司員工的集體智慧或匯聚眾人資訊協助企業進行決策，至少包括惠普 (HP)、雅虎 (Yahoo)、微軟 (Microsoft)、英特爾 (Intel)、奇異 (General Electric)、百思買 (Best Buy)、良策 (Rite-Solution)、西門子 (Siemens) 及谷歌 (Google)。這些大企業預測標的包括產品研發、銷售預測、財務狀況、行銷策略等公司重要發展指標 (LaComb et al., 2007; Spears et al., 2009; Lavoie, 2009; Soukhoroukova et al., 2012)。

說明預測市場的準確性高於其他預測方法的文獻相當豐富，顯示預測市場作為資訊彙整的平台已經相當成熟。比對預測市場與蓋洛普、哈里斯、紐約時報等大型媒體共 964 場民調，在 1988~2004 年間預測美國總統大選的表現，結果顯示預測市場的準確率比傳統民調的次數高出 74% (Berg et al., 2008)。其他文獻至少包括：(Ortner, 1998; Pennock et al., 2001; Wolfers and Leigh, 2002; Brüggelambert, 2004; Servan-Schreiber et al., 2004; Wolfers and Zitzewitz, 2004; Gürkaynak and Wolfers, 2005; Leigh and Wolfers, 2006; Tung et al., 2011)。

「預測市場」的預測標的可以分成兩類預測事件：「是否發生型」事件與「落點預測型」事件。「是否發生型」事件是預測某事件是否發生，例如某候選人是否當選或台股指數是否上萬點；「落點預測型」事件是預測該事件的落點，例如候選人得票率或經濟成長率。絕大部分的上述文獻都是針對「落點預測型」事件的預測結果進行實證分析，比較與傳統預測方法的預測準確度。

針對「是否發生型」事件的預測結果，已有研究都顯示 (Manski, 2006; Gjerstad, 2005; Wolfers and Zitzewitz, 2006; Christiansen, 2007; Ottaviani and Sorensen, 2007)，預測市場的價格可以視為預測事件發生的機率。近期研究也顯示 (童振源等人, 2011a, 2011b)，預測市場的預測事件價格可以作為候選人當選之機率。然而，「是否發生型」事件的結果只有該事件是否發生，關切重點應該是預測市場對該預測事件結果的預測是否會發生，而不是該預測事件結果發生的機率，因為很多決策或避險是建立在該事件是否發生的基礎上。

因此，本文將利用未來事件交易所的交易資料，建構預測市場的最適價格門檻，作為判定預測事件是否發生的依據，將事件發生的機率預測 (probabilistic forecasting) 轉換成事件發生與否的類別預測 (categorical forecasting)。因應政府與企業等預測單位之需要，本文透過計算所有價格的準確率尋找最適價格門檻，將預測市場客觀產生的最終交易價格 (反映事件發生機率)，轉換成事件發生與否的判定基準及公共政策與企業決策的重要依據。

文獻上使用「門檻」的觀念經常出現在天氣預報、醫療診斷、圖形判讀等研究。



例如，天氣預報必須設下最佳門檻規則，以判別「雨天」、「陰天」、「多雲」等天氣型態，門檻設定攸關預報的準確度與效益（Murphy, 1977; Murphy and Brown, 1984）。在醫療行為上，每項醫療診斷需要一個標準門檻，作為是否積極治療的依據，使得發生誤診的機率達最低（Pauker and Kassirer, 1980）。其他門檻研究：以區別分析方法建構門檻辨識灰色圖像（Otsu, 1979）；使用最小誤差法應用在影像分類軟體，其最小誤差法亦具備極大化分組準確率的基礎（Kittler and Illingworth, 1986）。

童振源等人（2011a）及童振源等人（2011b）針對 2006-2009 年的台灣選舉預測事件研究，以預測事件的最後一天加權平均價格超過某特定價格門檻，則預測市場預測該事件會發生；如果預測事件的最後一天加權平均價格低於某特定價格門檻，則預測市場預測該事件不會發生。他們以價格 100 除以有效候選人數（number of effective candidate, NEC）作為價格門檻。² 在選舉研究當中，「有效候選人數」是衡量選舉競爭激烈程度，其概念上，同一選區兩位候選人的得票率相當，且該選區其他候選人的得票率接近零，則 NEC 會趨近 2；同理，三位候選人的得票率不相上下，則 NEC 會接近 3；因此「有效候選人數」愈大，則選舉競爭愈激烈，則每位候選人當選機率愈低（Cox, 1997）。因此，預測事件的最後一天加權平均價格超過此價格門檻，則表示該預測事件應該會發生。他們也以價格 50（事件發生機率為 50%）及最高價（在同一個選區的候選人當中，最高價的預測事件發生機率最高）作為選舉預測事件的價格門檻，進行穩健性測試。

但是，前述的研究並沒有比較以上三種價格門檻對於選舉預測準確度的影響，只是將價格 50 與最高價作為穩健性測試的門檻。在三種門檻的命中率當中，價格 100 除以有效候選人數作為價格門檻的表現似乎較好，但是缺乏全面而有系統的比較分析，而且僅侷限在選舉預測事件。McHugh and Jackson（2012）分析 357 個是否預測型事件，發現價格門檻從價格 55 到 80 變動都不會影響到市場預測的準確率。可惜樣本太少，而且沒有詳細的分析資料。

價格是預測市場的最重要訊息，是各種資訊彙整的均衡結果。然而，每個預測事件的特徵及交易過程都不一樣，難以一概而論每個預測事件是否有一致的價格門檻。不過，為了方便判定預測事件的預測結果是否發生，我們確實需要一個適當的價格門檻判定該預測事件是否發生，而不是將價格當作該預測事件的發生機率。因此，本文的目的即是尋找最適當的價格門檻，極大化預測市場的預測準確率，以便預先判定預測事件的預測結果是否會發生，讓我們適當管控該事件發生與不發生的相關風險，或

² 有效候選人數的定義為： $NEC = \frac{1}{\sum_i p_i^2}$ ，其中 p_i 為第 i 個候選人的得票率， n 為候選人數。「有效合約數量」、「事件合約數量」、「合約交易筆數」、「合約交易人數」、「合約可交易天數」與「合約交易口數」等變數對於最適門檻和準確率詳細討論請參見附錄。



進行該事件相關的決策。

除了傳統上的價格 50 門檻及最高價門檻之外，本文將建構最適價格門檻作為判定預測事件是否發生的準則。此外，童振源等人（2011a）與童振源等人（2011b）均分析價格 100 除以有效候選人數的價格門檻，但是本文並不只針對選舉，所以希望建構最適價門檻，以便一體適用到其他類型預測事件。本文章節安排如下：第二節為研究方法與資料敘述，第三節為全部預測事件的最適價門檻準則之實證分析，第四節為各類型預測事件的最適價門檻準則之實證分析，第五節為三種價格門檻準則的樣本外測試與比較，最後則是結論。

2. 研究方法與資料

本文的資料來源為未來事件交易所 2006 年 7 月至 2011 年 8 月所有「是否發生型」預測事件，³ 但不包括個別預測事件，一個預測事件組至少有兩個預測事件，才能進行最高價門檻的準確率計算與比較。減去交易量為零的預測事件，全部的預測事件數量共計 7,881 個，政治類有 867 個、選舉類（一定是政治類）有 650 個、經濟類有 2,761 個、社會類有 259 個、兩岸類有 411 個、國際類有 1,216 個、運動類有 2,437 個、娛樂類有 1,730 個，所有預測事件的統計量見表 1。

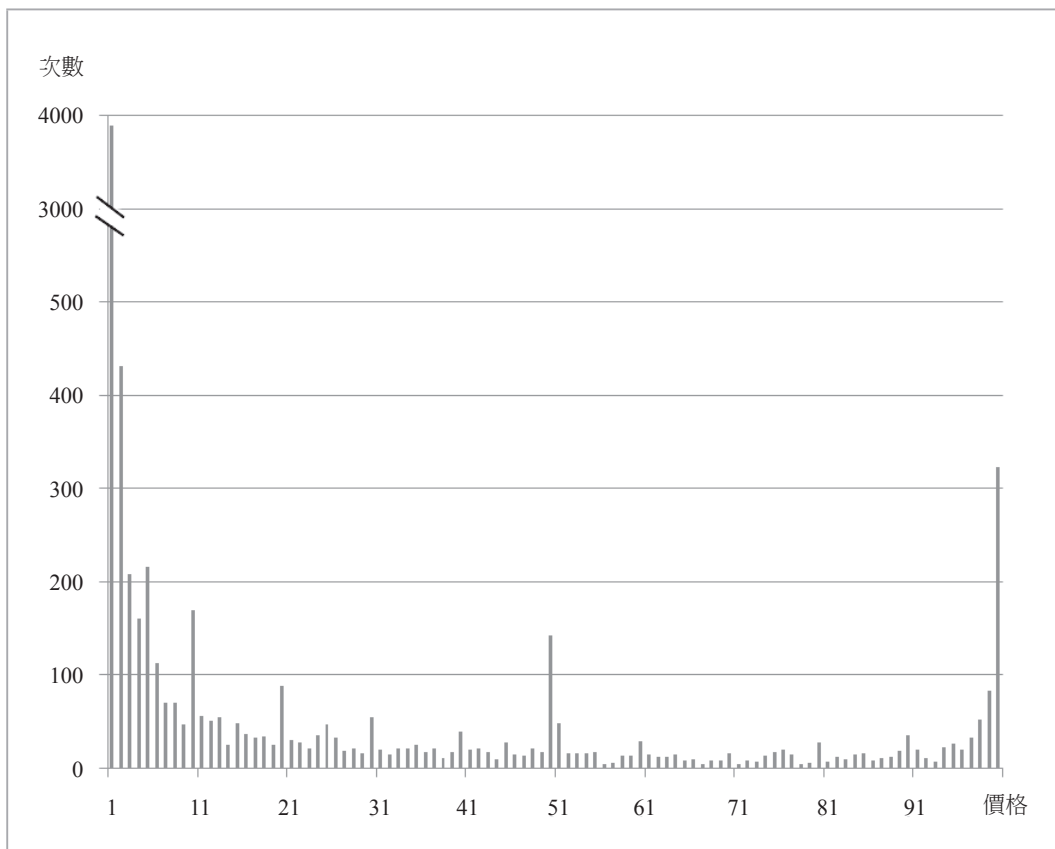
▼ 表 1 未來事件交易所預測事件的全部與分類敘述統計量：2006-2011

合約類別	事件數量	平均成交筆數	平均交易人數	平均交易天數	平均成交口數
全部	7881	108.04	33.44	46.41	27994.19
政治	867	457.17	110.96	115.26	130414.77
選舉	650	559.55	128.43	108.76	166862.54
經濟	2761	54.89	22.4	39.52	9176.59
社會	259	70.24	31.15	72.98	8783.93
兩岸	411	101.21	39.16	117.73	15383.45
國際	1216	53.3	20.33	36.69	9237.47
運動	2437	106.6	32.21	46.22	31297.64
娛樂	1730	27.19	15.83	25.23	3569.27

³ 未來事件交易所網址為 <http://xfuture.org>，為 2006 年 7 月由國立政治大學預測市場研究中心所創立，目前由未來事件交易所股份有限公司經營。



圖 1 按照價格的所有預測事件的次數分布，絕大部分的預測事件分布在價格低與價格高的兩個區域。當然，價格高的預測事件之發生機率高，價格低的預測事件之發生機率低；這兩個區域的預測事件應該是呈現預測市場的強烈共識。不過，即便發生機率高或低，都不代表該預測事件必然會發生或不發生；預測事件的價格確實隱含諸多資訊，但是我們無法從預測事件的價格直接判斷該事件被預測會發生或不會發生。



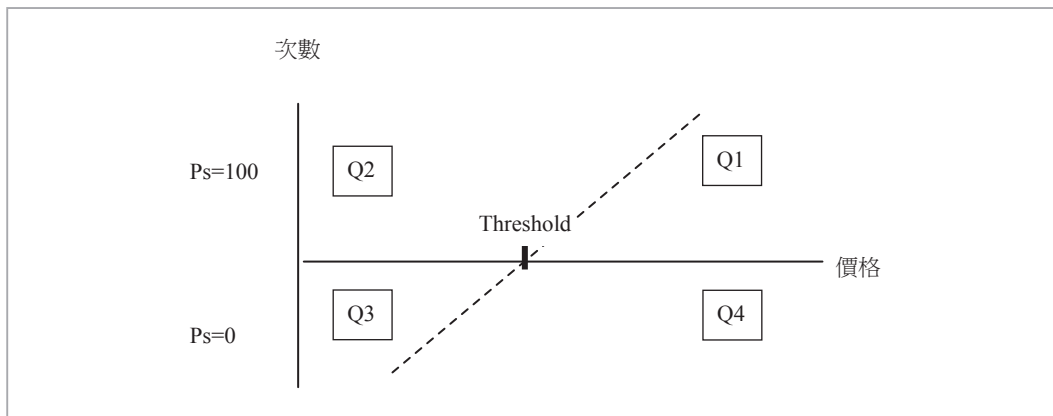
▲ 圖 1 未來事件交易所預測事件分布圖：按照價格

本文希望找到一個最適當的價格門檻，判定預測事件是否會發生。如果某預測事件價格在此價格門檻以上，便判定該預測事件為「預測會發生」；某預測事件價格在此價格門檻以下，便判定該預測事件為「預測不會發生」。然而，預測與事實結果很難完美，所以可能某預測事件價格在此價格門檻以上而判定該預測事件為「預測會發生」，但卻「沒有發生」；或者，某預測事件價格在此價格門檻以下而判定該預測事件為「預測不會發生」，但卻「真的發生」。

圖 2 為預測事件的價格與實際結果的分布示意圖。假設價格門檻為「Threshold」，所有預測事件在此價格門檻以上（預測該事件會發生），而且事實



結果為發生的預測事件分布在 Q1（預測正確組），事實結果為沒有發生的預測事件分布在 Q4（預測錯誤組）。相對的，所有預測事件在此價格門檻以下（預測該事件不會發生），但事實結果為發生的預測事件分布在 Q2（預測錯誤組），事實結果為沒有發生的預測事件分布在 Q3（預測正確組）。「預測正確組」由 Q1 與 Q3 組成，「預測錯誤組」由 Q2 與 Q4 組成。若將價格門檻設得太高，將縮小 Q1、擴大 Q2；若將價格門檻降得過低，將縮小 Q3、擴大 Q4。因此，這個價格門檻要使 $Q1 + Q3$ 極大化，使 $Q2 + Q4$ 極小化，才能作為日後準確判定某預測事件是否會發生的準則。



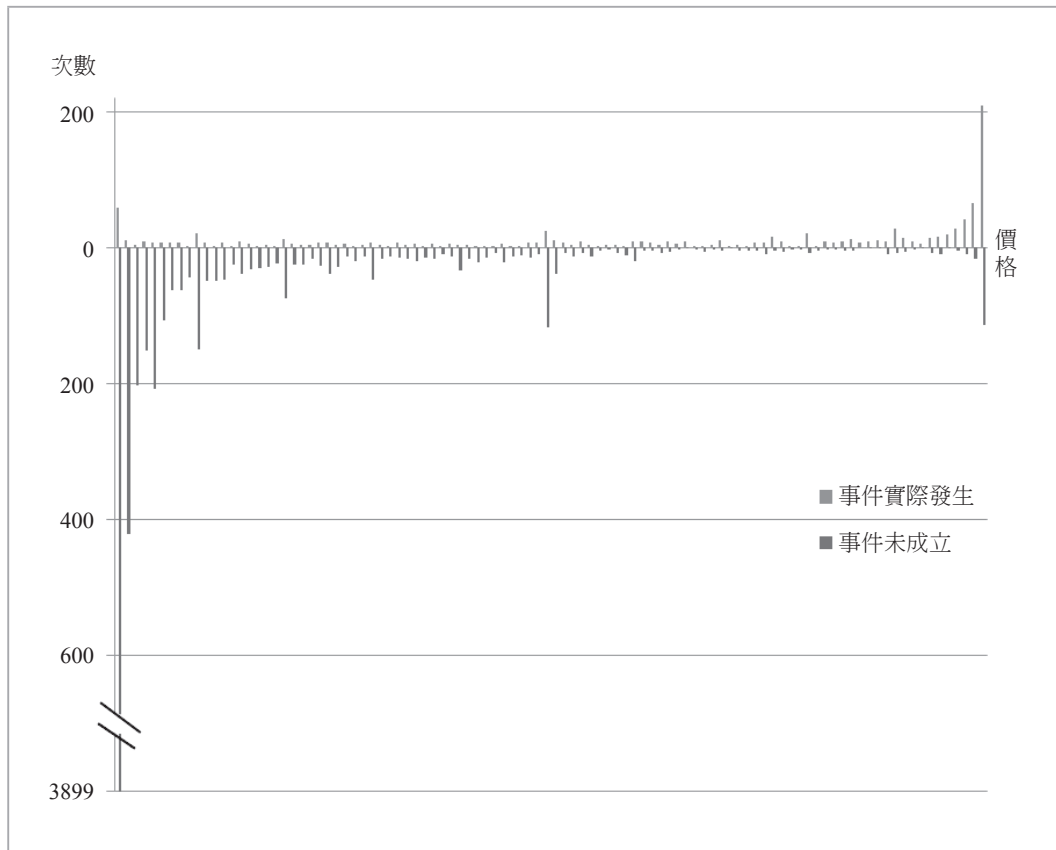
▲ 圖 2 預測事件分布示意圖：按照價格與實際結果

各個價格門檻準則的準確率計算方式如下：

$$\text{價格門檻準則的準確率} = \frac{Q1 + Q3}{\sum Qi} \quad (1)$$

方程式 (1) 當中的 Q1 為該合約價格高於門檻，且事後確實發生的合約數量；而 Q3 為合約價格小於門檻，且事後事件未成立的合約數量；方程式 (1) 的分母，即為圖 2：Q1、Q2、Q3 與 Q4 所有合約的總數。圖 3 為按照預測事件價格與實際結果區分之未來事件交易所的全部預測事件分布圖。很明顯的，絕大部分價格低的預測事件都是沒有發生，絕大部分價格高的預測事件都是發生。然而，從圖 3 來看，幾乎無法找到一個價格門檻，使得預測事件價格在價格門檻以上的預測事件都是發生，同時使預測事件價格在價格門檻以下的預測事件都沒有發生。

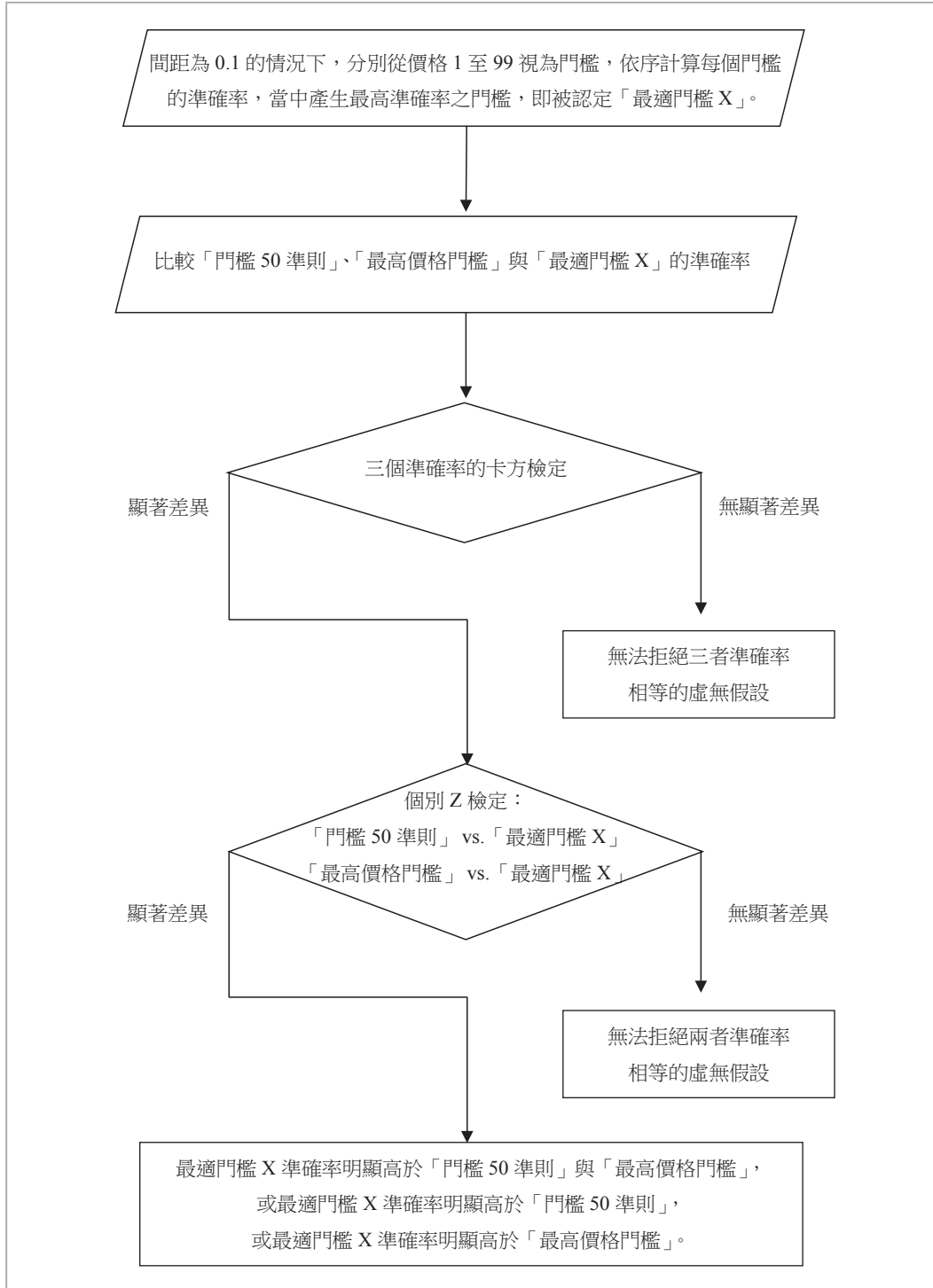
本文將從價格 1 至 99 逐一計算個別價格門檻的準確率，將準確率最高的價格門檻當做「最適門檻」。此外，本文還計算 50 為價格門檻的準確率及最高價門檻的準確率。最適價格門檻與價格 50 門檻準則都是固定價格門檻，但是最高價門檻準則為變動的價格門檻，視預測事件的群組之最高價而定，最高價可能是 95，也可能是



▲ 圖 3 未來事件交易所之預測事件分布圖：按照價格與實際結果

35。例如，在台北市長選舉時有五位候選人，最後只會有一個候選人當選，因此這五位候選人的事件期貨價格最高者應該被認定為最可能當選之候選人。不過，不同的選舉態勢會有不同的事件期貨價格分布。例如，如果一位候選人獨霸，則該候選人的事件期貨價格可能高達 95，其他候選人的事件期貨價格可能都是個位數；如果三強鼎立，則三位候選人的事件期貨價格可能都趨近 30，另外兩位候選人的事件期貨價格可能是個位數。

在計算完各項價格門檻的準確率之後，本文以卡方檢定判定三個準確度的差異性是否存在。其次，本文再以 Z 檢定，測試任意二價格門檻的準確率在統計上的顯著差異，比較最適價格門檻、價格 50 門檻及最高價門檻的準確率。以下便先分析全部預測事件的最適價格門檻與比較三種價格門檻準則的準確度，其整個研究流程圖如圖 4。

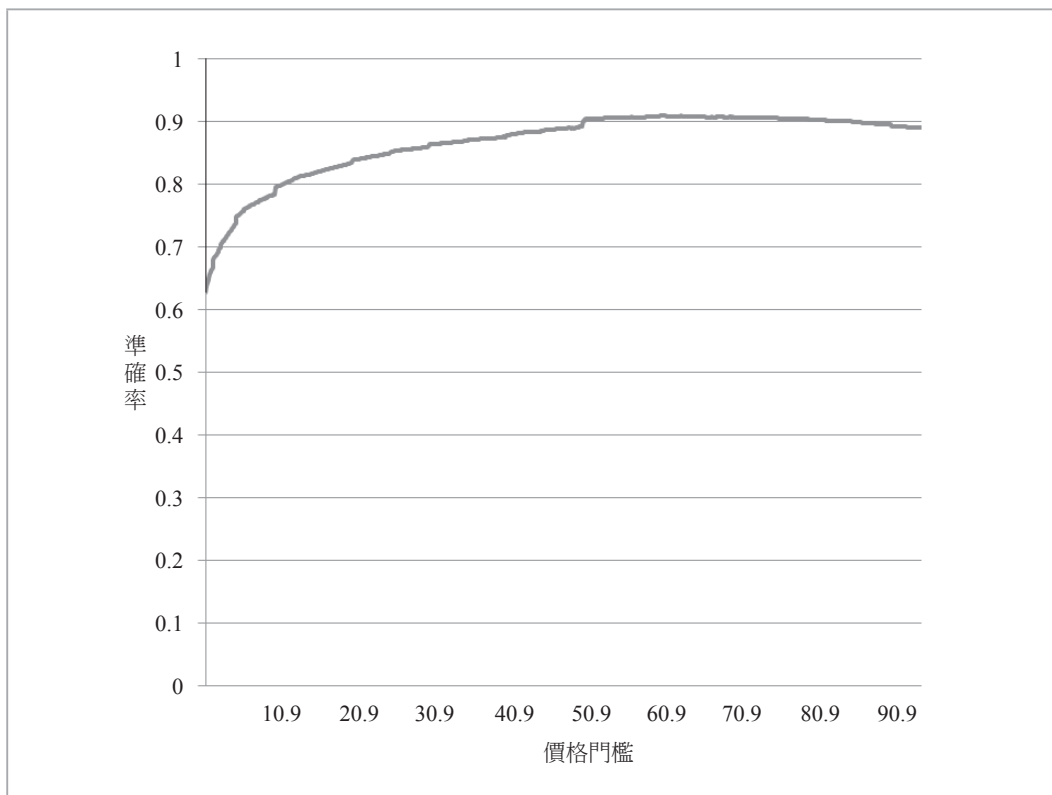


▲ 圖 4 研究流程圖



3. 全部預測事件的最適價門檻準則之實證分析

從價格門檻 1 到價格門檻 99，圖 5 為全部預測事件之不同價格門檻的準確率。很明顯的，價格 50 作為價格門檻的準確率並非最高，也就是 50 非最適價格門檻。根據價格 50 以後的價格門檻計算，全部預測事件的準確率有顯著提高，準確率的最高點出現在價格 60。本文測試所有價格的準確率，理論或客觀上可找到一個最高準確率的門檻，此價格即為「最適門檻」。



▲ 圖 5 整體預測事件之不同價格門檻的準確率

此外，為確認這些最適門檻的穩定性，本文進行 $K=10$ 的 K -fold 交叉驗證，將原先樣本隨機區分成 10 等份，每次使用其中九等份作為訓練樣本，計算出各自的最適門檻價格，總共十個門檻值的平均價格為：60.1625，訓練樣本的平均準確率為：90.89%，樣本外十分之一等份的測試準確率為 90.88%。為更精確估計，本文再重新執行 20 次 $K=10$ 的 K -fold，其結果： K -fold 最適門檻的平均數為 60.21、變異數為 0.0201，測試準確率的平均數為 90.86、變異數為 0.0008，此交叉驗證的變異很小，且平均值與圖 5 的最適價趨於一致，顯示此最適價格門檻具有穩定性。



除了價格 50 的價格門檻與最適價格門檻 60 之外，本文再以最高價門檻計算全部 7881 個預測事件的準確率，結果在表 2。以價格 50 為價格門檻的全部預測事件準確率為 89.47%，以最高價為價格門檻的全部預測事件準確率為 88.45%，以最適價為價格門檻的準確率為 90.79%。最適價似乎是全部預測事件的最好價格門檻，準確率至少可以提昇 1-2 個百分點。

▼ 表 2 不同價格門檻之全部預測事件的準確率：2006-2011（單位：%）

樣本數	價格 50	最高價	最適價（60）
7881	89.47	88.45	90.79

根據表 2 可知門檻的準確率排名，最適價預測準則最好，價格 50 為次佳與最高價準則最差，但這三者準確率的差距，是否在統計上顯著差異，則我們可用卡方檢定來檢驗。底下表 3 為各準則下的實際預測結果。

▼ 表 3 全部 7881 合約分別在三個門檻準則之觀測個數

	價格 50	最高價	最適價（60）	總合
實際正確預測	7051	6971	7155	21177
實際錯誤預測	830	910	726	2466
總合	7881	7881	7881	23643

觀察上方表 3，最適價門檻準則的正確預測個數最多，最高價門檻準則的正確預測個數最少，為凸顯實際預測結果的差距，正確預測個數可和樣本期望個數作比較，而樣本期望值可利用表 3 計算得到：觀測值的兩個邊際比例相乘，之後再乘以全部樣本的總合，其結果如表 4。

▼ 表 4 全部 7881 合約分別在三個門檻準則之期望個數

	價格 50	最高價	最適價（60）
期望正確預測	7059	7059	7059
期望錯誤預測	822	822	822

說明：1. $7059 = (7881/23643) \times (21177/23643) \times 23643$ 。

2. $822 = (7881/23643) \times (2466/23643) \times 23643$ 。

利用上述兩個表格數據和卡方檢定值公式，可算出卡方檢定值為 23.22，此數值大於自由度為 2 且顯著水準為 5% 的卡方臨界值 5.99，故在統計上，三個預測準則準



確率（正確預測合約數量）存在顯著差異。計算過程如方程式 (2)：

$$\begin{aligned}\chi^2 &= \sum_{i=1}^3 \sum_{j=1}^2 \frac{(f_{ij} - e_{ij})^2}{e_{ij}}, \text{ 其中 } f_{ij} \text{ 為觀測個數和 } e_{ij} \text{ 為理論期望值} & (2) \\ &= \frac{(7051 - 7059)^2}{7059} + \frac{(830 - 822)^2}{822} + \frac{(6971 - 7059)^2}{7059} + \frac{(910 - 822)^2}{822} \\ &\quad + \frac{(7155 - 7059)^2}{7059} + \frac{(726 - 822)^2}{822} \\ &= 23.22 > \chi_{0.05}^2(2) = 5.99\end{aligned}$$

本文總共使用 7,881 個樣本合約，因此我們以大樣本的兩個母體比例 z 統計量，其檢定的樣本分配為：

$$\widehat{p}_i - \widehat{p}_j = N \left(p_i - p_j, \frac{p_i(1-p_i)}{n_i} + \frac{p_j(1-p_j)}{n_j} \right) \quad (3)$$

與卡方檢定使用相同表 3 的資料， \widehat{p}_i 、 \widehat{p}_j 即為任兩個準則的準確率，引入方程式 (3) 得到 z 值。

▼ 表 5 全部預測事件之三種價格門檻準確率的 Z 檢定

虛無假設	$H_0: P_2 \geq P_1$	$H_0: P_3 \geq P_1$	$H_0: P_3 \geq P_2$
Z 值	-2.0434**	2.7788***	4.8196***

說明：1. P_1 為價格門檻 50 準則的準確率， P_2 為最高價門檻準則的準確率， P_3 為最適價門檻準則的準確率。

2. *** 表示 1% 的顯著水準，** 表示 5% 的顯著水準，以及 * 表示 10% 的顯著水準。

儘管最適價在全部預測事件作為價格門檻的準確率顯著比其他兩項價格門檻準則的準確率要高，但是每類預測事件的特性及交易狀況都不一樣，因此接下來將針對各類預測事件的最適價門檻準則進行實證分析。

4. 各類預測事件的最適價門檻準則之實證分析

表 6 呈現各類預測事件的最適價門檻及各類預測事件之三種價格門檻準則的準確率。大致上，各類預測事件的最適價門檻都在價格 60 附近，但是社會類為價格 50，娛樂類為價格 78。其次，除了政治類（包括選舉類）與社會類以最高價門檻準則判定預測事件的準確率最高之外，⁴ 其他類別的預測事件都呈現最適價門檻準則是三種價格門檻準則中的準確率最高。第三，雖然政治類與社會類的最高價準則判定預測事

⁴ 政治類別的合約當中，存在 74.97% 的選舉類別合約（650/867）。



件的準確率高於其他兩項價格門檻準則，但是雙方的差距都相當小，不到一個百分點。

▼ 表 6 各類預測事件之三種價格門檻的準確率

合約類別	樣本數	門檻 50	最高價	最適價	最適價門檻
政治類 *	867	94.69	95.27**	94.93	62
經濟類 *	2761	91.13	87.87	91.49**	60
社會類	259	94.21	94.98**	94.21	50
兩岸類 *	411	91.24	89.29	94.16**	60
國際類	1216	82.24	82.89	85.94**	60
運動類 *	2437	87.44	87.94	89.62**	60
選舉類 *	650	94.15	94.62**	94.46	62
娛樂類 *	1730	87.17	86.76	89.42**	78

說明：打 * 號的各類預測事件存在多個相同準確率之最適價門檻，為精簡篇幅目的，本文僅列出一個，最接近全部預測事件之最適價格門檻 60 的價格數值。打 ** 號的數值為三項價格門檻準確率較高者。

表 7 為卡方檢定結果，除了政治類（包括選舉類）與社會類之外，不同類別預測事件之三種價格門檻準確率差異性都在統計上顯著。

▼ 表 7 三種價格門檻的正確預測個數之卡方檢定

合約類別	p 值
政治類	0.8584
經濟類	0.0000***
社會類	0.9062
兩岸類	0.0411**
國際類	0.0310**
運動類	0.0462**
選舉類	0.9342
娛樂類	0.0365**

說明：1. 卡方檢定的虛無假設 H_0 ：三個門檻準則的準確率無差異。

2. *** 表示 1% 的顯著水準，** 表示 5% 的顯著水準，以及 * 表示 10% 的顯著水準。

表 8 為不同類別預測事件之三種價格門檻準確率的 Z 檢定，以分別比較任兩個價格門檻準則的準確率。在經濟類的預測事件，最高價準則與最適價格門檻的準確率，顯著高於價格 50 準則之準確率。在兩岸類、國際類、運動類、娛樂類的預測



事件，最適價格門檻的準確率，都顯著高於價格 50 門檻與最高價門檻準則。整體而言，最適價格門檻在各類預測事件的預測準確率呈現相當不錯表現。

▼ 表 8 不同類別預測事件之三種價格門檻準確率的 Z 檢定

	$H_0: P_2 \geq P_1$	$H_0: P_3 \geq P_1$	$H_0: P_3 \geq P_2$	備註
政治類	0.55	0.23	-0.33	不顯著
經濟類	-3.96***	0.47	4.43***	P_2 最差
社會類	0.39	0.00	-0.39	不顯著
兩岸類	-0.94	1.61*	2.54***	P_3 最好
國際類	0.42	2.5***	2.08**	P_3 最好
運動類	0.53	2.39***	1.86**	P_3 最好
選舉類	0.37	0.24	-0.13	不顯著
娛樂類	-0.36	2.06**	2.42***	P_3 最好

說明：1. P_1 為價格門檻 50 準則的準確率， P_2 為最高價門檻準則的準確率， P_3 為最適價門檻準則的準確率。

2. *** 表示 1% 的顯著水準，** 表示 5% 的顯著水準，以及 * 表示 10% 的顯著水準。

在政治類（選舉）與社會類預測事件，最高價準則的準確率稍微高於最適價門檻準確率，但統計上尚未達到顯著的檢定結果。也就是說，以統計學角度而言，最適價門檻的準確率仍不亞於最高價準則準確率。其次，價格 50 門檻準則顯然都不是最適當判定預測事件是否發生的價格門檻。

5. 三種價格門檻準則的樣本外測試與比較

本文前半部分使用全部預測事件，驗證最適價格門檻準則較價格 50 門檻與最高價門檻準則具有優勢。本節以 2006 至 2009 年的預測事件作為訓練樣本，再針對 2010 至 2011 年的預測事件進行樣本外測試，⁵ 計算最適價門檻準則、最高價門檻準則和價格 50 門檻準則的預測準確率。進行樣本外測試時，本文將樣本分成全部事件及政治、經濟、社會、兩岸、國際、運動、選舉與娛樂等八大類預測事件進行分析。

首先，本文針對 2006 至 2009 年的樣本內樣本尋找最適價門檻，全部事件的最適價門檻為 60，除了娛樂類預測事件的最適價門檻為 78 較高之外，其他類預測事件的最適價門檻分布在 50-62 之間。再者，以三種不同價格門檻的準則計算樣本外樣本的預測準確率。在全部預測事件方面，最適價門檻準則的預測準確率領先另外二個準則

⁵ 本文收集資料至 2011 年 8 月，因 2011 年僅有 8 個月的合約交易記錄，故使用 2010 年以後的合約，作為樣本外測試對象。



2-3 個百分點。在政治類（包括選舉類）預測事件方面，最高價門檻準則領先其他兩項價格門檻準則；在社會類與兩岸類預測事件方面，三項價格門檻準則平分秋色。除此之外，最適價門檻準則的表現都優於其他兩項準則。（見表 9）

▼ 表 9 2010-2011 年樣本外測試

	2010 年前最適門檻	訓練樣本數	2010-2011 年測試樣本數	門檻 50 準則之準確率	最高價準則之準確率	最適門檻準則之準確率
全部事件	60	4803	3078	88.66	87.91	90.35*
政治類	62	715	152	96.71	97.37*	94.74
經濟類	51	2078	683	93.12	85.36	93.27*
社會類	50	223	36	100	100	100
兩岸類	56	313	98	95.92	95.92	95.92
國際類	56	300	916	80.13	82.31	83.52*
運動類	62	1201	1236	85.19	87.46	87.94*
選舉類	62	517	133	96.24	96.99*	93.98
娛樂類	78	785	945	88.04	88.25	90.26*

說明：* 為準確率最高的價格門檻準則。

在卡方檢定結果，三項價格門檻準則的差異只有在全部樣本、經濟類、運動類是顯著的，其他類事件可能因為樣本數過少而造成統計上的不顯著。（見表 10）

▼ 表 10 不同類別預測事件之三種價格門檻準確率的卡方檢定：樣本外測試

$H_0: P_1=P_2=P_3$	p 值
全部事件	0.0000***
政治類	0.4519
經濟類	0.0000***
國際類	0.1611
運動類	0.0968*
選舉類	0.4499
娛樂類	0.2380

說明：*** 表示 1% 的顯著水準，** 表示 5% 的顯著水準，以及 * 表示 10% 的顯著水準。

在個別準確率的差異 Z 檢定，在全部事件與娛樂類，最適價門檻準則表現顯著



優於其他兩項價格門檻準則；⁶ 在國際類，最適價門檻準則表現顯著優於價格 50 門檻準則；在經濟類，最高價門檻準則表現顯著遜於其他兩項準則；在運動類，價格 50 門檻準則顯著表現最差；在政治類（選舉類），三項價格門檻準則的兩兩比較準確率都不顯著。⁷（見表 11）

表 11 不同類別預測事件之三項價格門檻準確率的 Z 檢定：樣本外測試

Z 值	$H_0: P_2 \geq P_1$	$H_0: P_3 \geq P_1$	$H_0: P_3 \geq P_2$	備註
全部事件	-0.91	2.16**	3.08***	P ₃ 最好
政治類	0.34	-0.85	-1.18	不顯著
經濟類	-4.66***	0.11	4.77***	P ₂ 最差
國際類	1.20	1.88**	0.69	P ₃ 好過 P ₁
運動類	1.64**	2.01**	0.36	P ₁ 最差
選舉類	0.34	-0.86	-1.19	不顯著
娛樂類	0.14	1.55*	1.41*	P ₃ 最好

說明：*** 表示 1% 的顯著水準，** 表示 5% 的顯著水準，以及 * 表示 10% 的顯著水準。

根據表 9 的 2006 至 2009 年樣本，政治類事件的最高價門檻準則表現較好，社會與兩岸類的三項價格門檻準則表現平手。為釐清這些特殊狀況的可能原因，本文彙整各類預測事件交易價格之統計資料。相較於其他類別事件，政治類預測事件的交易價格平均普遍較高，而且交易價格的標準差也比較大，最高價的平均值較大，最高價的標準差較小，最高價與次高價差額之平均數也顯著高於其他類別。這顯示，政治類預測事件的最高價普遍較高（85 以上），也就是市場共識相當強，或者選舉的態勢強弱懸殊，所以市場的判斷相當明確，也較容易判定預測事件的結果。此外，社會與兩岸類的最高價平均值及最高價與次高價差額之平均數也都相當高，展現市場的共識相當強烈或預測事件的結果相當容易判斷，但是交易價格平均並不高、交易價格的標準差比較小，這可能是社會與兩岸類的三項價格門檻準則之準確率難分軒輊的原因。（見表 12）

⁶ 由於社會類與兩岸類的樣本外預測事件之三項價格門檻的準確率完全相同，所以本文不再針對這兩類樣本外事件進行卡方檢定。

⁷ 由於社會類與兩岸類的樣本外預測事件之三項價格門檻的準確率完全相同，所以本文不再針對這兩類樣本外事件進行 Z 檢定。



▼ 表 12 各類預測事件交易價格之統計數據

變數	政治	選舉	社會	經濟	兩岸	國際	運動	娛樂
交易價格 (Pi) 的平均數	23.17	25.25	17.99	15.61	17.52	21.76	18.79	17.65
交易價格的標準差	36.70	37.31	33.35	29.49	32.70	31.24	30.28	28.27
Pi-50 之平均數	43.86	42.88	44.97	43.51	44.32	39.33	41.16	40.52
Pi-50 之標準差	11.89	12.80	10.55	12.61	12.58	15.04	13.99	14.27
最高價的平均值	87.01	85.11	87.31	71.12	80.72	61.80	70.81	71.72
最高價的標準差	17.46	17.68	20.30	29.43	29.32	30.43	28.01	26.40
最高價與次高價差額之平均數	71.50	67.74	71.98	48.22	67.84	24.65	40.61	48.97
最高價與次高價差額之標準差	31.45	32.58	27.95	34.43	34.50	29.46	36.04	33.78
交易價格與最適價絕對差額之平均數	50.73	49.25	44.97	51.02	51.47	46.65	48.62	62.75
交易價格與最適價絕對差額之標準差	16.74	17.75	10.55	15.40	14.91	16.18	15.83	22.44

說明：符號 || 為取絕對值。

6. 結論

預測市場的資訊匯流提供領先、即時、連續、長期的重要資訊，作為決策與避險的重要參考依據。然而，透過網路匯聚的集體智慧資訊，乃是以最終收盤價格呈現，反映參與者對事件發生的機率預測。如何將此機率預測轉換成類別預測，以佐助公共政策與企業決策，則是研究者必須處理的研究課題。為了方便判定預測市場上的預測事件之預測結果是否發生，本文成功利用未來事件交易所的交易資料，建構最適價門檻準則，極大化預測市場的準確率，以便預先判定預測事件的預測結果是否會發生，讓我們適當管控該事件發生與不發生的相關風險。

以 2006 至 2011 年的未來事件交易所 7,881 個預測事件樣本分析，全部預測事件的最適價門檻為價格 60，其判定預測事件的準確率比價格 50 門檻與最高價門檻準則都要高，而且具有統計上的顯著性。以分類事件而言，除了政治類與社會類之外，最適價門檻準則普遍比其他兩種價格門檻準則要好，而且大部分具有統計上的顯著性。在政治類與社會類的預測事件方面，雖然最高價門檻準則的準確率比最適價門檻要好，但並不具統計上的顯著性。

在樣本外測試與比較方面，以 2006 至 2009 年預測事件為樣本內樣本，2010 至 2011 年預測事件為樣本外樣本，全部預測事件的最適價門檻為價格 60，其判定預測事件的準確率比價格 50 門檻與最高價門檻準則都要高，而且具有統計上的顯著性。除了政治類之外，大部分類別的預測事件，皆呈現最適價門檻準則較其他兩項價格



門檻準則要好；不過，部分類別可能因為樣本數太少（例如，社會類與兩岸類，未來事件交易所在 2010 至 2011 年分別發行 36、98 個合約，但是運動類卻有 1,236 個合約），其差異性分析都不顯著。在政治類預測事件方面，雖然最高價門檻準則的準確率比最適價門檻要好，但並不具統計上的顯著性。

總體而言，以價格 60 為最適價門檻之準確率，普遍高於價格 50 門檻或最高價門檻的準則，其可作為判定某預測事件是否會發生的價格門檻。政治類預測事件仍可以沿用最高價門檻準則，其他類別預測事件則以最適價門檻準則較具優勢。透過預測事件價格的分析與比較，我們可以更有信心利用最適價格門檻判定預測事件的結果，以便我們作決策與進行避險的重要參考。

本研究結論條列如下：

- (1) 以極大化預測市場準確率為目的，本文建構最適價門檻準則，其價格門檻可使市場從機率預測轉成類別判定，讓我們適當管控該事件發生與不發生的風險；
- (2) 針對全部樣本，「最適價門檻準則」比「門檻 50 準則」、「最高價門檻準則」具備更高的準確率，且大部分具有統計上的顯著性，亦通過樣本外測試；
- (3) 針對分類樣本，除了政治類與社會類的預測事件，分類的預測事件也都呈現「最適價門檻準則」比「門檻 50 準則」、「最高價門檻準則」具備更高的準確率，而且大部分具有統計上的顯著性。

未來，我們可以給以預測事件結果不同權重，以建構新的最適價格門檻。本文將發生與不發生的正確預測視為同樣的效益，但兩者的影響可能不一樣，兩種正確預測對政府與企業決策的價值也未必相同。例如，發生車禍與否的後果完全是生離死別，正確預測「車禍不發生」的效益與正確預測「車禍發生」的效益不同。因此，新的權重最適價格門檻，應將不同權重的效益考量進去，以便建構類似天氣預報與醫療診斷判定的門檻。

此外，我們應該可以從更多的交易資訊與預測市場型態，例如交易人數、時間、數量、交易者特質、預測事件特質等等，進一步建構模型，在事前有更多資訊判定預測事件發生的結果。這將讓我們更從容利用預測市場的資訊，支援企業、社會及政府的決策。最後，本論文研究未來事件交易所的交易資料，是一個個案研究，而非通例結論，此模型參數未必適用到其他預測市場，需要以此模型引入至其他預測市場，進行相同步驟的分析，才能確認其最適門檻的價格。

參考文獻

童振源、周子全、林繼文、林馨怡（2011a）。2009 年台灣縣市長選舉預測分析。《選舉研究》，18（1），63-94。



- 童振源、周子全、林繼文、林馨怡 (2011b)。選舉結果機率之分析：以 2006 年與 2008 年台灣選舉為例。《臺灣民主季刊》，8 (3)，135-159。
- Arrow, K. J., Forsythe, R., Gorham, M., et al. (2008). The promise of prediction markets. *Science*, 320(5878), 877-878.
- Arrow, K. J., Sunder, S., Forsythe, R., et al. (2007). *Statement on prediction markets* (AEI-Brookings Joint Center Related Publication No. 07-11). Retrieved May 6, 2014, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=984584
- Berg, J. E., Nelson, F. D., & Rietz, T. A. (2008). Prediction market accuracy in the long run. *International Journal of Forecasting*, 24(2), 285-300.
- Brüggelambert, G. (2004). Information and efficiency in political stock markets: Using computerized markets to predict election results. *Applied Economics*, 36(7), 753-768.
- Christiansen, J. D. (2007). Prediction markets: Practical experiments in small markets and behaviours observed. *Journal of Prediction Markets*, 1(1), 17-41.
- Cox, G. W. (1997). *Making votes count: Strategic coordination in the world's electoral systems*. Cambridge: Cambridge University Press.
- Gjerstad, S. (2005). *Risk aversion, beliefs, and prediction market equilibrium* (Mimeo). Retrieved May 6, 2014, from http://www.aeaweb.org/assa/2006/0106_1015_0701.pdf
- Gürkaynak, R., & Wolfers, J. (2005). *Macroeconomic derivatives: An initial analysis of market-based macro forecasts, uncertainty, and risk* (Working Paper 2005-26). Retrieved May 6, 2014, from <http://www.frbsf.org/publications/economics/papers/2005/wp05-26bk.pdf>
- Kittler, J., & Illingworth, J. (1986). Minimum error thresholding. *Pattern Recognition*, 19(1), 41-47.
- LaComb, C. A., Barnett, J. A., & Pan, Q. (2007). The imagination market. *Information Systems Frontier*, 9(2-3), 245-256.
- Lavoie, J. (2009). The innovation engine at rite-solutions: Lessons from the CEO. *Journal of Prediction Markets*, 3(1), 1-11.
- Leigh, A., & Wolfers, J. (2006). Competing approaches to forecasting elections: Economic models, opinion polling and prediction markets. *Economic Record*, 82(258), 325-340.
- Manski, C. F. (2006). Interpreting the Predictions of Prediction Markets. *Economics Letters*, 91(3), 425-429.
- McHugh, P., & Jackson, A. L. (2012). Prediction market accuracy: The impact of size, incentives, context and interpretation. *Journal of Prediction Markets*, 6(2), 22-46.
- Murphy, A. H. (1977). The value of climatological, categorical and probabilistic forecasts



- in the cost-loss ratio situation. *Monthly Weather Review*, 105(7), 803-816.
- Murphy, A. H., & Brown, B. G. (1984). A comparative evaluation of objective and subjective weather forecasts in the United States. *Journal of Forecasting*, 3(4), 369-393.
- Ortner, G. (1998). *Forecasting markets-An industrial application* (Mimeo). German: Technical University of Vienna.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62-66.
- Ottaviani, M., & Sorensen, P. N. (2007). *Aggregation of information and beliefs in predictionmarkets* (Working Paper). Retrieved May 13, 2014, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.320.7486&rep=rep1&type=pdf>
- Pauker, S. G., & Kassirer, J. P. (1980). The threshold approach to clinical decision making. *The New England Journal of Medicine*, 302(20), 1109-1117.
- Pennock, D. M., Lawrence, S., Nielsen, F. A., & Giles, C. L. (2001). Extracting collective probabilistic forecasts from web games. *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, USA.
- Servan-Schreiber, E., Wolfers, J., Pennock, D. M., & Galebach, B. (2004). Prediction markets: Does money matter? *Electronic Markets*, 14(3), 243-251.
- Soukhoroukova, A., Spann, M., & Skiera, B. (2012). Sourcing, filtering, and evaluating new product ideas: An empirical exploration of the performance of idea markets. *Journal of Product Innovation Management*, 29(1), 100-112.
- Spears, B., LaComb, C., Interrante, J., Barnett, J., & Senturk-Dogonaksoy, D. (2009). Examining trader behavior in idea markets: An implementation of GE's imagination markets. *Journal of Prediction Markets*, 3(1), 17-39.
- Tung, C. Y., Chou, T. C., Lin, J. W., & Lin, H. Y. (2011). Comparing the forecasting accuracy of prediction markets and polls for Taiwan's presidential and mayoral elections. *Journal of Prediction Markets*, 5(3), 1-26.
- Wolfers, J., & Leigh, A. (2002). Three tools for forecasting federal elections lessons from 2001. *Australian Journal of Political Science*, 37(2), 223-240.
- Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, 18(2), 107-126.
- Wolfers, J., & Zitzewitz, E. (2006). *Interpreting prediction market prices as probabilities* (Working Paper No. 12200). Retrieved May 6, 2014, from <http://www.nber.org/papers/w12200>



附錄

根據第一節對有效候選人數 (NEC) 的定義，NEC 的計算過程需要真實的得票率，因此 NEC 為事後變數。本文以事前預測為研究目的，且研究對象不只包含選舉類合約，故我們必須以事件發生前的交易價格代替得票率，此變數稱為「有效合約數量」，以交易價格替代得票率，計算方法與 NEC 相同：

$\frac{1}{\sum_i^n V_i^2}$ ， V_i 為第 i 個候選人的「交易價格」/100， n 為候選人數，其表示預測事件的預測困難度。

以下討論「有效合約數量」、「事件合約數量」、「合約交易筆數」、「合約交易人數」、「合約可交易天數」與「合約交易口數」等變數，對於最適門檻和準確率的影響。我們先利用各個變數的中位數作分界，將本文的研究樣本區分成兩組，該變數在第一組的平均數將小於 (或等於) 中位數，相對此變數在第二組的平均數大於中位數，故我們列出下表比較這兩組的最適門檻與準確率：

▼ 表 六項變數對最適門檻與準確度之影響

	中位數	最適門檻	準確率
有效合約數量	1.26	(81.00,60.10)	(92.37,89.21)
事件合約數量	9	(59.90,60.20)	(90.64,90.57)
合約交易筆數	14	(78.40,59.90)	(89.39,91.83)
合約交易人數	13	(78.40,59.90)	(89.62,91.68)
合約可交易天數	14	(60.20,52.40)	(87.63,93.71)
合約交易口數	1235	(60.20,59.90)	(88.85,92.30)

說明：1. 括號內第一個數值，代表在小於或等於中位數樣本下所產生的數據，而括號內第二個數值，代表在大於中位數樣本下所產生的數據。

2. 此外，部分選區交易冷落，可能產生多數合約的價格接近零，根據有效合約數量定義：1 除以「(交易價格/100)平方值的加總」，將使有效合約數量發生不合理極端值，此處有效樣本為有效合約數量在 10 以下的合約，總共 5,972 個事件合約。

觀察上表，發現的確有部分的因素將使最適門檻異於 60，包括「有效合約數量」、「合約交易比數」與「合約交易人數」。以「有效合約數量」為例，小於或等於中位數的樣本，其試算的最適門檻為 81.00 (括號第一個數據)，對於大於中位數的樣本而言，其最適門檻為 60.10，兩門檻存在明顯差距，就「選舉合約」的有效合約數量低於 1.26 的樣本來說，該組當中的每一選區之有效合約皆接近 1 個，表示同一選區候選人的價格差距拉開，預測事件的預測困難程度相當低，所以價格門檻自然比較高。(根據上述定義，若一選區有 N 個候選人，選情激烈選區的每人價格為 N 分之一，則有效合約數量為 N)



以「合約交易筆數」與「合約交易人數」而言，市場交易冷落代表匯集資訊較弱的合約，可能需要較高的最適門檻，以反應市場的強烈共識。但針對這些變數，全部樣本無法像類別分組具備明確定義來劃定分組範圍，中位數分組僅為其中一種區分樣本的方式。未來的研究可討論更多種分組方法，作全面性和更詳盡的分析最適門檻值。