

葉同學	回饋建議
書審老師A	同學的研究邏輯十分清晰，也能從成果中看出對於求知慾的熱忱，以及探究歷程中的推敲也都相當嚴謹。相當優秀的一個作品
書審老師B	這個作品結合政治學、統計學。加上學生具有優異的科技能力，為一優秀的跨領域的自主學習。作品各個部件得呈現十分完整，從中可以看見學生再思辯、邏輯、統計、資訊等能力的學習和進步，最後能將自主學習成果撰寫成小論文，讓整個學習過程更顯充實完整。
書審老師C	小論文與自主學習屬性不一，從多元角度來看，自主學習較偏向於學科以外的自我體悟與實踐，這篇文章雖然就小論文來講或許不錯，但是就自主學習的角度來講，我希望看到更多自己生活中的觀察、提問，學習計畫規劃、試錯與省思。

自主學習： 候選人學歷 對得票率的影響

南山高中
葉同學

普通科
指導老師：古展倫

百字簡述：
我因關注時事而反思，研究候選人學歷與得票率的關聯。結合英文與程式語言，從學術資料庫中統整文獻，學習統計設計多元迴歸分析。成果不僅成功量化學歷對得票率的影響，也培養以跨領域技能與思考解決問題的能力。

摘要

本歷程著重於我撰寫有關候選人學歷小論文的經歷。在過程中，我結合自身所學，並多元攫取知識，透過自身固有的優勢能力精進自我，最終以「多元迴歸分析」為研究工具，量化學歷對於得票率的具體影響。本文成功地量化了學歷對於得票率的具體影響，而我也看到自己對於發掘與解決不同知識領域及問題的熱情。

我期許自己能將這段學習的歷程轉化為啟發與動力，帶領我在踏入更高的殿堂、學習更深更廣的知識後，探詢與解決更多制度上的困惑與問題，協助創造更好的制度體系。



較近者為正在與同學討論撰寫內容的我

目錄

動機與 研究問題

對時事的熱情

1

簡述我的學習動機，並說明研究的流程與架構

學習與 研究過程

困難的跨領域學習

2

描述選題的障礙，以及如何結合自身優勢學習程式語言與撰寫小論文

研究說明

3

透過發表會簡報說明本研究的研究過程與結論

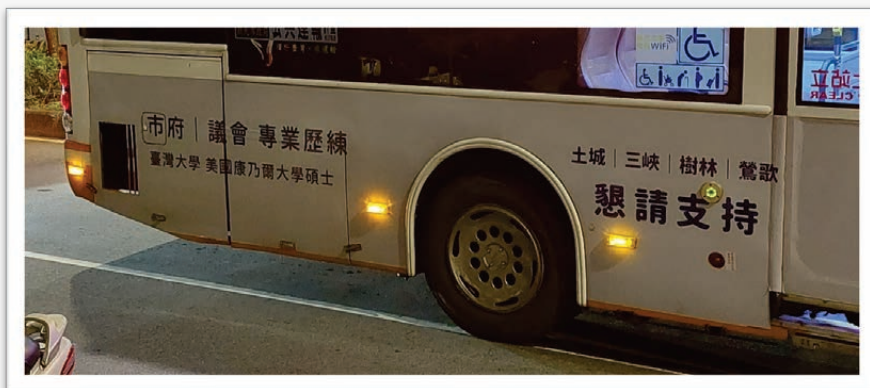
心得與展望

4

針對小論文的不足之處說明，同時思考未來能夠學習與精進的能力

點選即可前往各章節

動機： 對時事議題的關注



放學等公車時看到的選舉廣告。以大版面標示學歷的情形吸引我的注意

2022 年的地方選舉中，興起了一股討論候選人學歷的風潮，作為時常關注時事的人，這非常吸引我的目光。然而，大多數人的討論多半集中在攻擊候選人的學歷真實與否，以及相關的論文爭議。這激起了我的興趣：**到底學歷對於候選人的得票率是否有幫助？**

具體來說，假設學歷對於候選人的得票是無效的，那為何候選人又時常至高階學府進修，並大幅度的在廣告上曝光學歷呢？或者問，選民真的有那麼容易受到候選人的學歷而影響到投票抉擇呢？回答「**候選人學歷對得票的影響**」，事實上可以部分接近台灣選民的政治能力、政治人物的進修疑慮以及社會上的其他問題，而這種探索答案的過程，讓我決定進行這樣的研究。

研究問題： 主觀與客觀的界定

這個問題實際上有兩種不同的回答取向：

- 以候選人為核心的「學歷對候選人得票率的影響」
- 從選民為出發點的「候選人學歷是否會影響到選民的投票抉擇」

從方法論而言，這兩種研究取向會導向不同的方法。就「學歷對候選人得票率的影響」而言，我們應當使用選舉資料庫的總體候選人資料進行分析，方能有足夠的數據量化具體影響。而以「候選人學歷是否會影響到選民的投票抉擇」的疑問而言，我們應當針對選民發放問卷調查，以了解會影響投票取向的特質有哪些。本文較為接近前者。

需要說明的是，雖然本文採用了與人（選民）較為無關的方法，但為了區分出選民對不同學歷的評價高低，我們不使用傳統的「高中」、「大學」與「研究所」等類型化資料作為變數控制，而是使用邀請選民進行評價的方式執行（方式詳後）。這使得原本較為客觀取向的觀察研究滲入了人為的成分。這種創新的研究方法一方面為兩種不同的研究取向尋求了統整的可能性。然而另一方面而言，也可能會偏移我們原始的研究問題或形成干擾，而這也是我們需要注意的。

學習過程： 結合自身優勢與資源

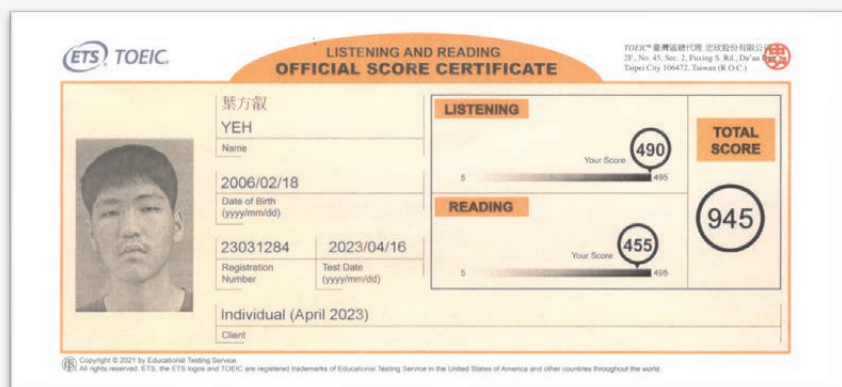
我首先開始在網路上查找相關資料，以檢驗是否有人做過類似的研究。經過初步的查找，並沒有能夠直接回應我的研究問題的研究。因此，我開始從幾個開放取用的政治科學期刊上閱覽我有興趣的標題的文章。經過進一步的閱讀，我發現許多選舉研究多半採用統計科學測量特定現象的正確與否。對於社會組的我而言，這是非常新奇的發現。

我一直以來都認為社會科學，特別是政治學，與理工方面的數學領域不會有太大的交集，沒想到在政治科學中是恰恰相反的。我對於新事物與跨領域的結合一直都有相當的熱忱，因此在持續查找文獻外，我也立刻投入統計學的學習。

作為一個高中學生，我學習相關領域的時間與能力受到很大的限制。幸好，我透過結合自身在高中兩年中所培養的能力，如英語、資料查找與基礎 C++ 語言能力，加速了我的學習。以下我將分項說明：

外語能力

我在高二即在多益考試獲得 945 分的成績。準備考試的過程中，我的英文閱讀與聽力能力獲得了顯著的成長。



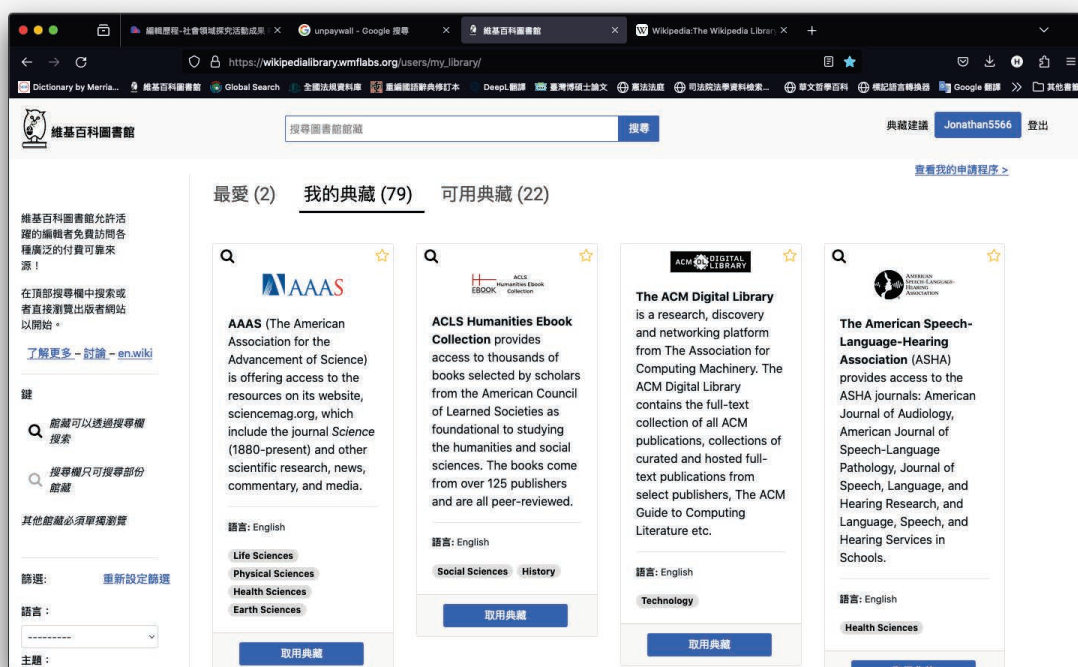
而在我的自學過程中，外語能力成為我突破瓶頸的利器。英文讓我深入瞭解國內外最新的研究成果與方法，這些寶貴的資源不僅豐富我的文獻探討段落，更充實了我的研究方法。

我曾陷入認為自己已經蒐集了相當多的相關文獻，卻未能發現關聯程度更高研究的困境。然而，我運用英文關鍵字進行搜尋後，驚喜地發現國外已有許多研究探討學歷之於選舉的相關議題，為我的研究提供了巨大助力，也讓我深刻的明白，作為一個研究者應該要先行閱覽許多不同語言與面向的資料，否則到頭來只是做了一個和別人相當雷同的研究、重複研究問題，不一定能促進科學進步。

結合自身優勢

多元的資料檢索能力

此外，查找文獻時，時常有許多文獻是需要付費才能使用的。在沒有學術機構的網路支援下，我們往往無法負擔要求的費用，導致漏讀許多重要的文獻。我主要使用國立公共資訊圖書館數位資源所提供的華藝論文資料庫功能，以及我自己在參與維基百科編輯時所獲得的維基百科圖書館 (The Wikipedia Library) 閱覽權，充實自己的文獻閱讀庫。



維基百科圖書館的截圖。近 80 種的資源成為我研究上的重要助力

學習 R 語言

在統計語言 R 的學習上，考量學習指標、資料整理等邊緣功能耗時費力，因此我直接向大學教授索取其研究的 R 語法檔，透過實際研究運作來認識其編碼方式。

R 語言與一般的程式語言不同，並非由單一的可執行檔所組成，比較接近一種與電腦的即時溝通媒介。透過與電腦對話，逐步引導其產出我們所需要的統計結果。

我的學習方法主要憑藉自己過往學習 C++ 語言時所習得的程式碼閱讀理解能力，猜測原始碼的功能，並直接在軟體中實作驗證猜想。必要時，則透過網路文獻來了解如何調整繪圖的細節。這是教授提供原始碼中的一段繪圖程式：

```
ggplot(full_result[complete.cases(full_result),], aes(x=noc, y=co)) +  
  geom_line(aes(y=0), linetype="dashed", colour=gray, alpha=0.4) + geom_point(size=3) +  
  geom_errorbar(aes(ymin=co-1.96*se, ymax=co+1.96*se))
```

這段原始碼看來複雜，其實仔細分析起來相當單純。首先，我們透過 `ggplot` 呼叫繪圖程式，並將圖命名為 `full_result`，並指定資料從資料集 `complete.cases(full_result)` 中取得。`aes(x=noc, y=co)` 則表達資料的橫軸與縱軸分別為資料集中的 `noc` 與 `co` 兩行中的資料。其後的程式碼也可以用類似的方法推得大概的意思。

對我而言，比較困難的情況是某段程式碼採用了我先前未使用過的統計表達（如上方的 `geom_errorbar` 即是在圖上繪製出 ± 1.96 的標準差），這時候我就需要透過手邊的統計學書籍去了解大概的意思。

有趣的是，我一開始雖然想透過直接研究程式碼，結合高中數學課上所習得的基礎統計學概念快速上手這個程式語言，但由於所學實在有限，我最後還是慢慢的在學習 R 語言的過程中把我購買的統計學科普書籍閱讀完畢，算是另一個粗淺但深刻的學習插曲。

困境： 多元的問題解決方案

我主要使用 R studio 進行 R 語言的撰寫。由於這是一個我從未使用過的程式，因此我在處理困境時，採用幾種不一樣的解決問題策略，主要可分為：

1. 上網尋找答案
2. 詢問 ChatGPT

透過網路查詢

R 語言作為一個成熟的程式語言，網路上的開源資源非常豐富與精實。值得說明的是，大部分的資源都是以英文的形式提供，這使我再一次感謝自己過往在語言上的投資。

詢問 ChatGPT

當時，AI 的興起引起大家一陣討論，而我也搭上這個熱潮，使用 ChatGPT 來解答我的許多疑惑。然而，AI 不免可能會提供錯誤的答案，因此我在詢問完 AI 後更會主動上網查看該函式庫的原始說明，從而盡量避免被誤導的可能。

使用 AI 解答疑惑是否是個正確的方法其實相當有爭議，但我認為在技術、軟體工程方面是可行的。首先，許多專業人士使用 AI 協助撰寫統計程式碼，再者，透過一定程度的資料搜尋，我們可以避免被 AI 誤導。最後，我們可以透過交互比對程式碼確認（將程式碼重新要求 ChatGPT 解釋，並重新要求其輸出另一種程式碼並比對其相似性）。如此的反覆與交互確認，很大程度上能保證我們獲得的程式碼是可用且正確的。



Tse-min Lin

17小時 · 🌐

#chatgpt與統計學習
#紅樓夢前80回與後40回書寫風格的差異

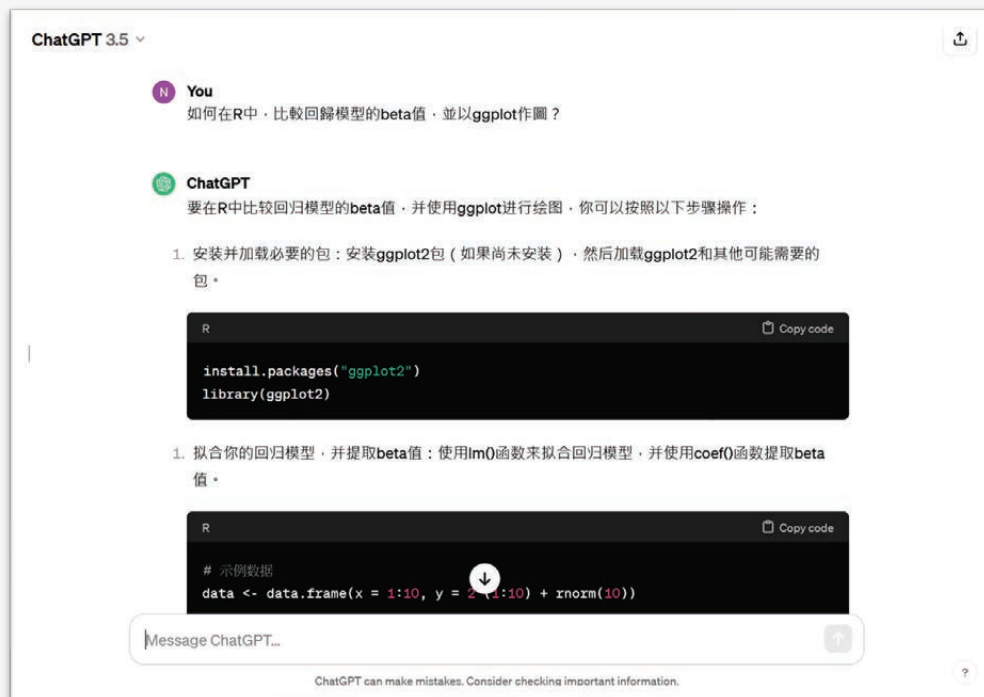
《紅樓夢》後40回作者與前80回作者是否同一人已經有很多人研究過了，用現代資料科學方法研究者亦不乏其人。以下我無意對這問題重新做一全面的分析，而只是以前80回與後40回書寫風格是否有所差異為例，展示如何用ChatGPT所產生的R程式碼很容易地來做這樣的研究。

我以下聚焦在「句子長度」這一點來比較前80回與後40回書寫風格的差異。這裡首先有幾點說明：

(一) 《紅樓夢》各版本原來是沒有標點的。現在我們看到的標點都是後人所加。既然這樣，研究句子的長短究竟有何意義？我個人覺得：即使在沒有標點的時代，不同作者書寫的語氣和文體仍然可能長短有所不同。後人將原著加上標點後，句子的長短在相當程度之內仍然可以反映原作者書寫風格的特異性。以下的分析係在這樣的假設之下而做。

(二) 如果分析的結果顯示前80回與後40回句子的長短的確有所差異，是否便可據以判斷前80回與後40回的作者不是同一人？這個問題的答案顯然是否定的：且不說不同的作者寫出的句子可能長短差不多，《紅樓夢》的寫作，所謂「十年辛苦不尋常」，十年之間，同一作者的寫作風格也可能有所改變。因為這樣，單單句子長短不同頂多

使用 ChatGPT 教學統計的教授。這讓我了解到這類科技工具在較為嚴謹的自我查證下也可成為自身學習時的有力幫手。



使用 ChatGPT 詢問統計問題的截圖。需要注意的是，一旦使用這類工具，自己就應該進行更多的查證工作，才能避免受到誤導。

研究過程： 漫長的挑戰與突破

本文的研究過程為期一個學期，但實際上以本文的研究手法，加上固有的課業壓力，這樣的時間是遠遠不夠的。不過，我們在有限的時間內依然完成了一個相對精緻的研究，特別是在量化具體影響部分。最主要消耗時間的部分，除了上文所提及的 R 語言的學習與撰寫，還有將中央選舉委員會的選舉資料轉換成適合我們進行統計的樣態，以及我們邀請評審針對學歷進行評價的過程。

由於總候選人將近 700 名，這部份的作業也需要針對每一名候選人手動輸入，因此耗費了不少時間。以下我們將針對兩種作業的過程與困境進行說明。

候選人資料編碼

中選會在資料平台上有提供基本的選舉資料檔，但其中僅列出候選人出生年、號次、性別與黨籍，至於我們研究關鍵的學歷、經歷等，則僅出現在選舉公報上。

因此，我們首先從選舉公報上將每位候選人的學歷與經歷複製到 Excel 表格之中。在這個過程中，我們曾經想透過資料爬蟲結合文字識別的方式，將選舉公報上的資料自動串接到表格中，但由於選舉公報的排版定位邏輯以我們的能力很難透過程式找到文字的位置，我們最後還是只能以人工方式一一剪貼。

事實上，這是在只有 700 名分析對象的情況下才能使用的權宜之計。如果我的研究客體更多，這種人工剪貼的方式顯然是不切實際的，最終仍要透過網路資料爬蟲才能充分解決這個問題，如何具體掌握這項技能將會是我未來努力的方向。



S	T	U	V	W	X	Y
最高學歷	DP	HT	un	ma	phd	tp
輔仁大學應用心理學系碩士	0.6599999999999999	岩山里里長	1	1	0	0
美國加州專業心理大學博士	0.82	國大代表	1	1	1	1
美國田納西州立大學企管碩士	0.71	無	1	1	0	0
崇右企業管理專科學校	0.37	無	0	0	0	0
國立臺灣海洋大學文學士	0.53	民主進步黨台北市黨部執行委員	1	0	0	0
清華大學社會所碩士	0.77	外交部歐洲司科長	1	1	0	1
淡江大學	0.45999999999999996	立法委員	1	0	0	0
美國東密西根大學文學碩士	0.82	第九-十二屆臺北市議員	1	1	0	1
高雄應用科技大學外文系	0.5599999999999999	第十二屆台北市議員	1	0	0	0
台大國發所兩岸組碩士畢業	0.79	現任台北市議員	1	1	0	1

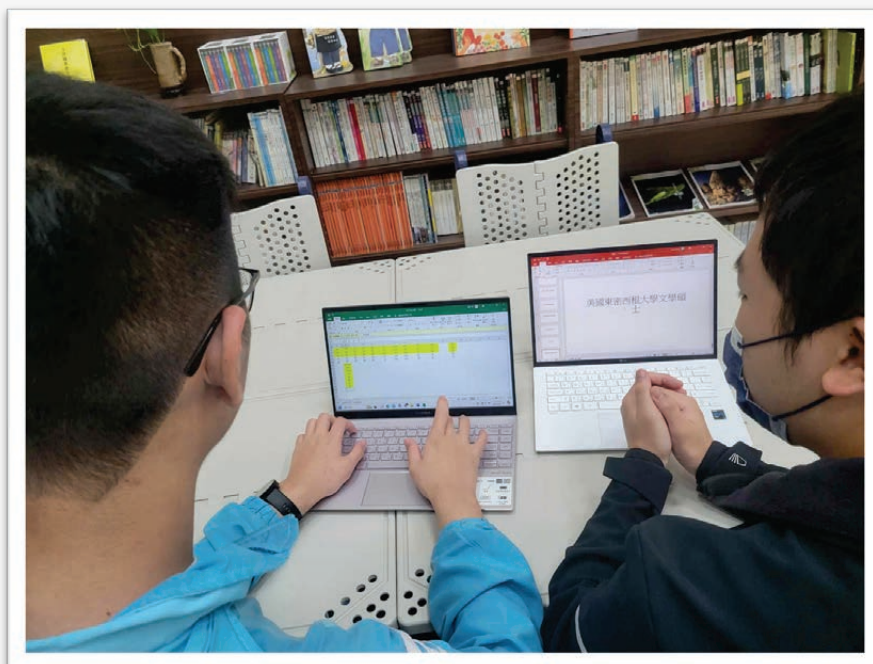
最後產出的資料成果，點選圖片即可查看

處理學歷評分

誠如前文提及，為了綜合主觀與客觀兩種研究取向，我們邀請 3 位評審針對學歷進行評分。我們將整理好的學歷資料一一複製到簡報上，並透過撥放簡報的方式請評審提供學歷評分，而我則在旁邊紀錄評審的評分。這樣的方式有兩個優點：

- 每位評審都是獨立評分，排除各自受到影響的可能性
- 評審看不到候選人除了學歷以外的任何資料，從而使評審的評分能夠不受偏移，專注在學歷上

但缺點也十分明顯，就是我需要重複一樣的漫長過程三次。當然，為了確保研究品質，這樣的犧牲也是在所難免的。



右方為正在進行評分的評審 C。左方則是正在輸入評分的我本人。可以注意到右方的電腦顯示僅有候選人學歷的簡報檔案

研究說明

本校在期末時舉辦了小論文發表會，讓各組上台分享在學期中進行的各項探究成果。我們的小論文也有參與該次的發表。因此，接下來我將透過該次報告的簡報檔，結合部分文字稿的內容，說明我的小論文內容。

需要說明的是，雖然在該次簡報中我的口語表達並無太大的問題，但我在肢體語言上仍會出現晃動等不自覺表現，這或許會影響到聽眾的吸收，而這也是我未來需要努力改進的方向。



手持剪報棒者為正在進行簡報的我本人
點選圖片即可查看該次報告的側拍段落

候選人學歷對得票率的影響

以2018年直轄市議員選舉為例

葉方敬 歐鴻勳 趙柏任 2023/01/03

投影片的封面頁



我們首先透過兩則選舉文宣，引發與會者對於「學歷」在選戰中的作用的共鳴與回憶。在報告時，我們向聽眾提示兩則選舉文宣都大幅度的宣揚了候選人的學歷背景，並向觀眾提出「是否有看過類似的宣傳方式」的問題，進一步引導觀眾了解我們的研究問題。

研究目的

25%



隨後，我們希望透過具體的舉例，說明學歷對於選舉的可能影響。這是某一年苗栗縣立法委員選舉的法院判決案例。在這個真實案例中，兩個候選人的得票率差距為 25%。然而，得票率較高的候選人卻有偽造學歷的嫌疑，這引起了得票率較低的候選人的不滿，認為另名候選人透過偽造的學歷獲得了更高的得票因而當選，因此提出選舉無效的訴訟。

研究目的

「以兩個候選人25%的得票率差距而言，這顯然不是學歷造成的影響。」

我們摘引判決書原文，說明法院認為兩個候選人 25% 的得票差距顯然並非學歷所造成的。

研究目的

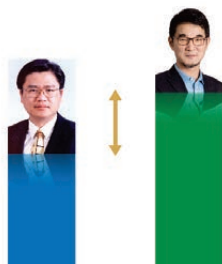
「以兩個候選人25%的得票率差距而言，
這顯然不是學歷造成的影響。」

無文獻支持

我們進一步說明，法院所謂「顯然不是」其實是依據經驗法則的判斷，並沒有文獻的支撐。

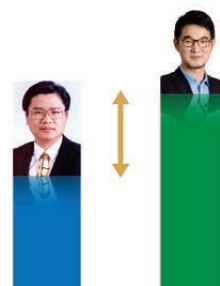
研究目的

10%



研究目的

15%



我們邀請觀眾思考，假設今天的得票率差距並非 25%，而是更低的 10% 或 15%，法院的這個推理依然能夠被適用嗎？又或者當差距只有 5% 時，法院能夠果斷排除這樣的影響可能來自於學歷與否？綜合以上，我們發現司法實務上確實存在這樣的量化學歷對得票率影響的需求。

研究目的

- 實證選舉情況
- 為司法實務提供**量化指標**

至此，我們歸納出本研究「實證選舉研究」與「為司法實務提供具體量化指標」的目的。

文獻回顧：學歷影響的理論層面解釋

	選民認同感	作為能力評斷指標
高學歷選民	易選擇高學歷候選人	易選擇高學歷候選人
低學歷選民	易選擇低學歷候選人	易選擇高學歷候選人

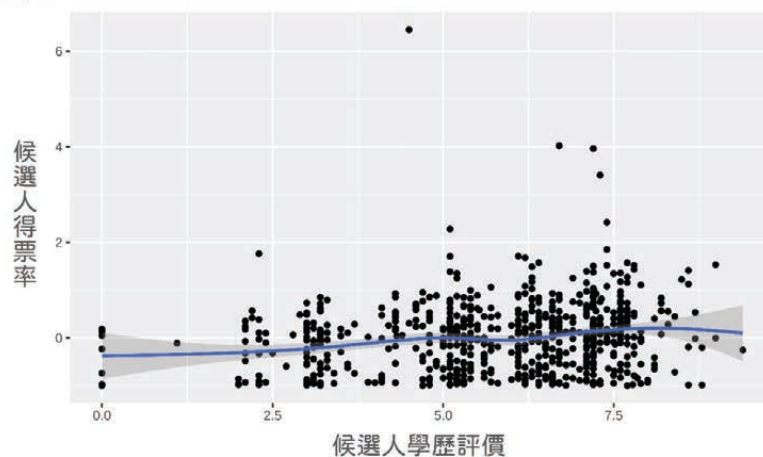
我們接著說明文獻上對於選民投票抉擇與學歷關係的分析。透過表格我們可以分析出「選民認同感」與「能力評斷指標」兩種說法，分別會導致不同出身的選民選擇不一樣的候選人。較為一致的是「能力評斷指標」的說法，不論選民的背景均會選擇學歷較高的候選人。而若以「認同感」作為立論基礎，則高學歷選民易選擇高學歷候選人，低學歷選民易選擇相近學歷背景的候選人。

研究假設

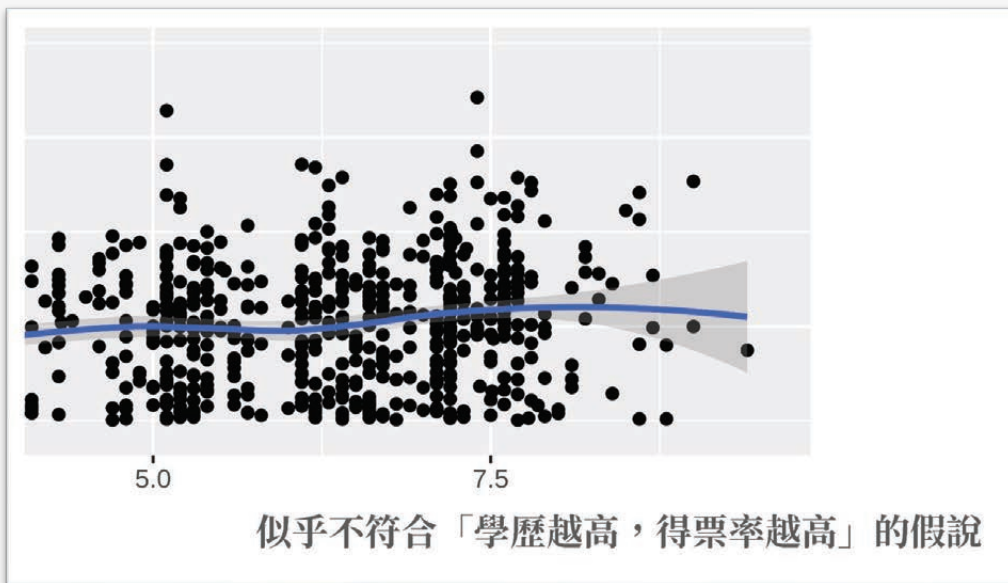
- H₁：學歷越高的候選人，得票率越高
- H₂：教育程度較低的選區，較容易受到學歷因素影響
- H₃：教育程度較高的選民，較喜歡高教育程度的候選人。而低教育程度選區無此現象。
- 詳請見書面，頁4

綜合以上的討論，我們提出三種不同的假設，並說明將在接下來依據我們的研究成果來檢驗這些假設。

資料視覺化



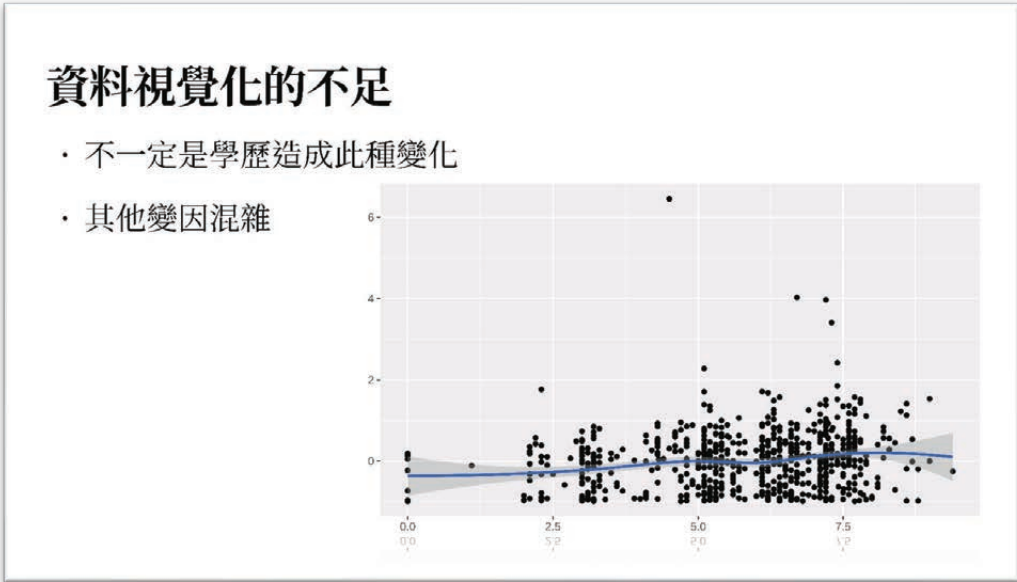
我們首先透過單回歸繪製成圖，檢驗「學歷越高的候選人得票率越高的假說」。藍線即可反映該現象的存在與否，而我們可以發現該條線並非單純向上的直線，而是有高有低，因此該假說無法被檢驗為真。甚至若我們放大高學歷的區間來看（請見下一頁），極高學歷的區間其實有得票率向下的趨勢，而這其實由於極高學歷區間的候選人數量極少，而容易受到極端值影響所導致的現象。



實際的情況卻是有高有低，顯見「學歷越高、得票率越高」的假設無法透過單純的繪圖推知。

資料視覺化的不足

- 不一定是學歷造成此種變化
- 其他變因混雜



然而從數據層面觀察，沒有出現預期的現象，可能是因為其他因子（如政黨、年齡等）干涉到了學歷對得票率的影響，以至於從圖形上來看並不顯著。因此，我們嘗試採用多元迴歸分析的方式，透過電腦推算每個因子的具體影響程度。

多元迴歸分析

- 分離其他變因的影響
- 需要控制的變因：性別、黨籍、年齡、連任與否、政治參與經歷
- 透過控制文獻中，最主要影響得票率的變因，就可以獲得**真實的學歷對得票率的影響**

我們接著說明在回歸模型中會控制的參數，包含性別、黨籍、年齡等。這些變數是文獻指出最主要影響得票率的變因。**透過多元回歸模型分離這些變因**，我們即可得到學歷對於得票率的真實影響。

多元回歸分析

書面頁6 表2 候選人得票率之多元迴歸分析

- 括號內為t鑑定數值，代表數值是否顯著異於0。如果不顯著異於0，代表該變因對得票率**沒有影響**。
- 數值可以直接理解為對得票率的影響程度。由於我們有先對候選人得票率進行處理，該數值可以直接換算成對得票率的影響
- 數值 $\times 10 \times$ 預期得票率

	(1)	(2)	(3)	(4)	(5)	(6)
學歷水準	0.08262 (5.154)	0.0802 (5.028)	0.0869 (5.296)	0.04754 (3.151)	0.02995 (2.290)	0.0319 (2.465)
男性		-0.1683 (-2.920)	-0.1789 (0.002)	-0.1011 (-1.932)	-0.0636 (1.406)	-0.0638 (-1.420)
年齡			0.0043 (0.0893)	-0.0083 (-3.343)	-0.0074 (-3.419)	-0.0086 (-3.825)
中國國民黨					0.8962 (13.372)	0.844 (22.255)
民主進步黨					0.4285 (7.506)	0.3765 (5.153)
現任				0.7022 (12.816)	0.3764 (6.952)	0.2143 (3.190)
經歷						0.1030 (3.990)
常數項	-0.4690 (-4.872)	-0.3461 (-3.300)	-0.5903 (-3.326)	0.0866 (-0.527)	-0.2771 (-1.921)	0.3081 (-2.110)
R ²	0.03563	0.0460	0.04866	0.2309	0.4263	0.4378

這頁是我們多元迴歸分析的總表。由於數據在解讀上較為繁雜，我們請聽者直接觀察紅框處，這就是在加入所有測量項目後學歷對於得票率的具體影響。而其他統計上的細節與解讀受限於篇幅，請見本文隨附的小論文本文。

多元回歸分析（續）

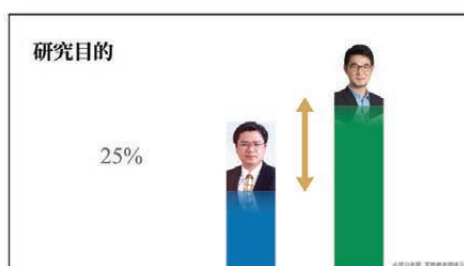
$$\text{學歷影響得票率程度} = \frac{100\% \times 10 \times 0.0319}{\text{同選區候選人數目}}$$

候選人數量	2	3	7	11	15
影響程度	15%	10.6%	4%	2%	2%

由於迴歸分析的結果較難感覺，且與選區候選人數量也有關（請見論文本關於 ADV 項目設計的說明），因此我們將具體的對得票率的影響按照選區人數製作成表格加以說明。

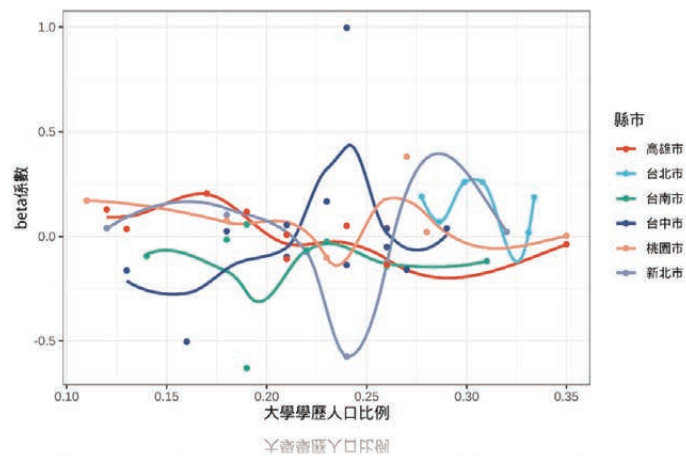
多元回歸分析（續）

候選人數量	2	3	7	11	15
影響程度	15%	10.6%	4%	2%	2%



我們結合最一開始提到的當選無效案例，說明在該案中法官的推理是符合本文的實證結論，因為在兩人競選的前提下，學歷最多只能影響 15% 的得票，遠遠少於兩人 25% 的得票差距。

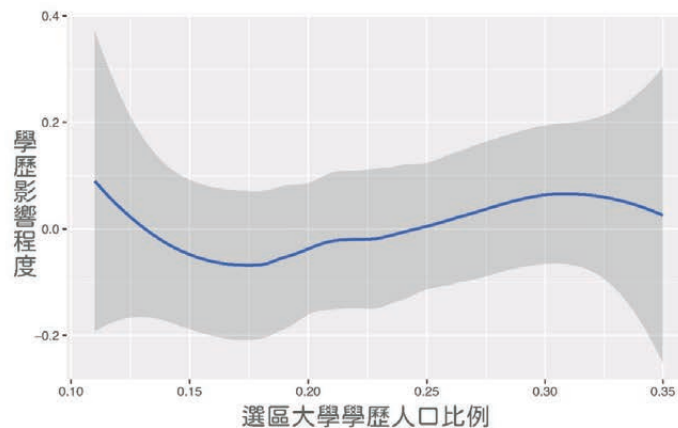
地區差異



我們嘗試透過分類縣市的方式來觀察不同縣市的影響程度差異，但無法透過圖形得到一致的解釋。

地區差異

- 與假說 H_2 及 H_3 均不符
- 文獻上無對應解釋



而為了驗證最一開始的假說 2 與假說 3，我們將選區的大學學歷人口數與學歷影響程度（迴歸係數）繪製成 X Y 分布圖。圖形有高有低，且文獻上也缺乏一致的解釋，故兩個與選民背景有關的假說都被否定。

需要說明的是，以上兩張投影片的內容與詮釋方法都存在問題，請見我在後文的討論（點選即可前往）

研究結論

1. 候選人學歷與得票率為顯著正相關
2. 在控制黨籍、年齡、性別、政治參與經驗等因素後，學歷仍會對得票率造成相當大的影響
3. 我國選舉實證上，對於學歷的重視程度，不存在一致的地區差異，相關理論需要謹慎挪用

本文的研究結論可以簡單摘要如下：

「學歷對於得票率的影響，透過迴歸模型控制了其他解釋變數後，可發現候選人的學歷水準最多會影響預期得票率（100%除以同選區候選人數）31.9%的得票。進一步將兩大黨與小黨的數據分別進行迴歸分析，可以發現候選人學歷評分對兩大黨候選人得票率的影響高於小黨或無黨籍候選人。我們認為這是政黨造成候選人宣傳量能的差異所造成的影響。」（詳見小論文本文頁9）

我們重申了研究分析的各項結論，與此同時，也一併指出了本文的研究限制。首先，本研究的評審組成均為高學歷的年輕知識份子，在評價候選人上，雖大致具備內在一致性，惟仍可能與普遍選民的認知存在差異。因此我們建議可強化評審的多元性。其次，我們的研究乃從數據面相出發，對於「造成這些現象」的原因僅能透過推論猜測結合數據佐證，並不一定代表真正產生這些結果的原因。因此，我們建議能夠有更多結合社會心理學領域的研究參與其中，才能尋得更貼近真實的答案。

心得與省思： 跨學科視野下的交匯點

這次的學習對我而言是個擴展自己視野的途徑。一方面而言，我看到了各種不同學科、領域結合的可能。另一方面，我透過自己的努力鑽進了一個我不曾接觸的領域。在過程中我心中一直有種莫名的興奮與幹勁——我想，我對於自己探詢知識的解答這件事情是有熱情的，這股熱忱更驅使數學不好的我一頭埋進統計學之中，十分令我震驚。

然而，在學習之中，我還是有些地方值得繼續精進。首先，我主要透過閱讀期刊論文的方式了解研究方法與理論基礎。這樣的方式可以讓我快速了解相關議題前人的量化方法與技巧，以及他們所使用的工具。這也是為什麼我決定使用 R 語言而非 Excel 等較為簡易的統計軟體。然而，這樣的方法也會對研究產生一定的負面影響。由於我並沒有接受過完整的政治科學訓練，我閱讀的論文很大程度牽制了我的行文模式與脈絡。也因此我需要更深入地檢視論文結構，以確保它是否最有效地對應我的研究問題及邏輯。

再者，學習統計語言 R 時，考量學習指標、資料整理等功能耗時費力，因此我直接向大學教授索要其研究的 R 語法檔，透過實際案例來了解其編碼方式。這樣的學習捷徑具有省時的特點，然而終非正途。我未來期待透過多元的校內與校際選課持續提升自我，特別是在程式語言上的基礎功。

回答研究問題與否

此外，本文在回答研究問題與否這個基礎上就產生了動搖。若就本文標題的「候選人學歷對得票率的影響」而言，本文外觀上看似是一客觀的觀察候選人學歷與得票率之間關連的研究，但若觀察研究方法，可發現我們採用了評審評分的方式類型化候選人學歷，而非單純的類別化參數。因此，本文所謂的「學歷」是否是學歷本身，又或者只是一個「學歷觀感」，其實存在爭議。

當然，若在「提供司法實務實證」這個研究目的上，本文相當程度地回答了研究問題。在資料類型化等細節上，雖然存在一些誤差與可議之處，但仍是相對完整的研究與分析。

研究方法 · 選區差異

需要特別指出的是，我們應避免單純依賴圖形（第 24 頁的兩張簡報）來解釋地區間的差異。理想的做法為，先根據理論假設建立一個模型，如假設學歷與得票率之間存在線性關係（即學歷越高，得票率越高），或者假設選民對候選人年齡的偏好呈現拋物線關係（即選民傾向於支持不老年輕的中年候選人），並使用實際的數據來檢驗這些假設是否成立。這樣可以避免僅憑圖形觀察而導果為因的謬誤。相較於我們小論文目前採用的模式，此方法更能驗證理論並確保研究結果可靠性，從而提升研究品質。

再者，我們目前僅根據選區平均教育程度分類的方法也不完整。理想的地區分類應建立在堅實的理論基礎之上，而非先行隨意分類，隨後根據圖形結果胡亂猜測可能的原因。此外，單純地以整體學歷高低區分，由於每個截面上的選區都很少，容易受到極端值影響，較難反映整體趨勢。以上兩者都可能導致我們得到的圖形只是湊巧出現，與選民的真實意向無關。

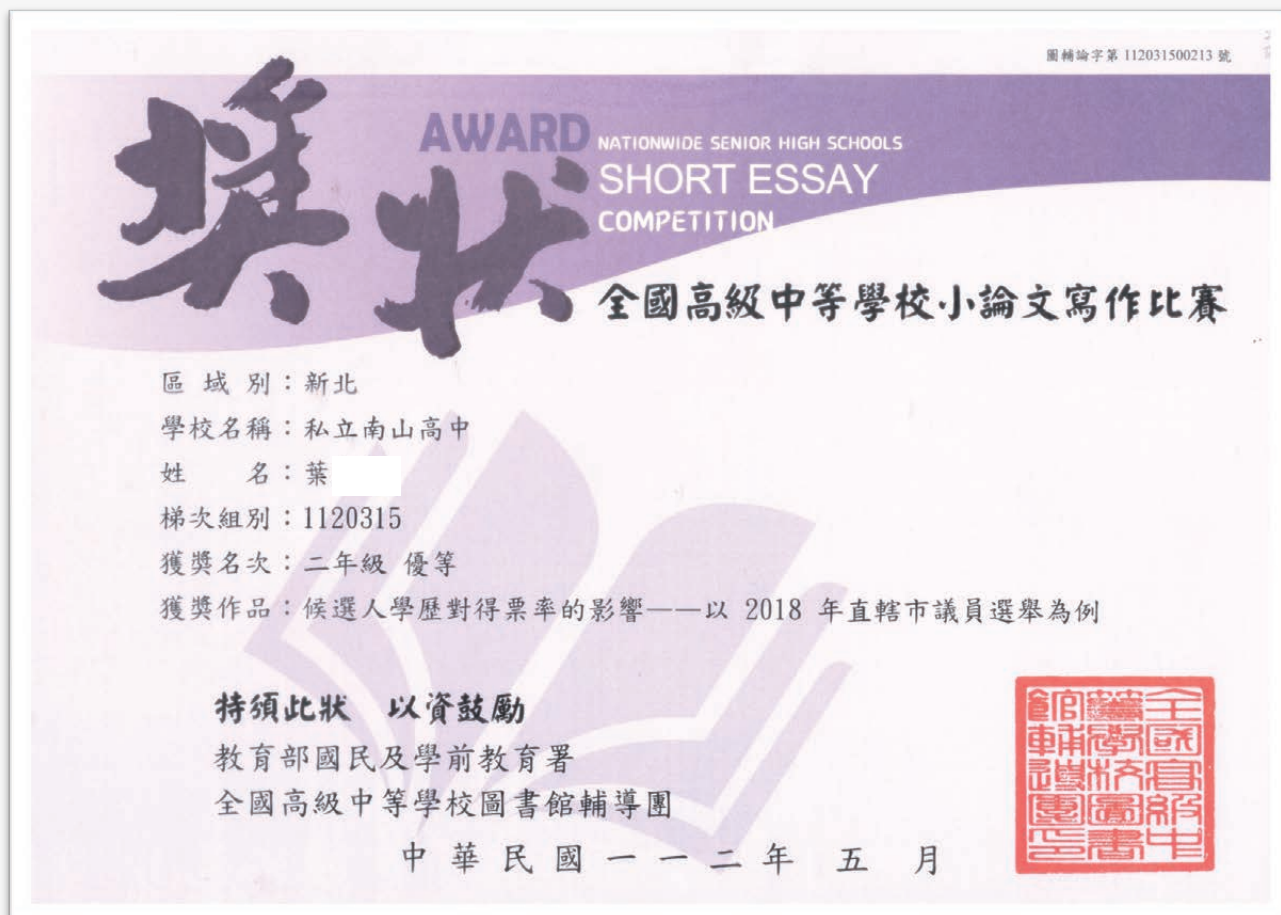
實際上，區域的意義是選區上的選民所賦予的，而非區域本身。因此，一個更合適的分類方法可能是先按照直轄市、產業結構、人口流動等具體的人口因素進行分組，然後依據統計圖形評估是否存在明確的趨勢。這樣的方法可能更有助於避免研究中出現無法找到特定趨勢，甚至得到錯誤的狀況。

結論與展望

綜合而言，本次小論文我成功踏入了一個我未曾想像過的領域。統計學與政治科學的結合十分引人入勝，而作為初步的學習，我也完成了階段性的挖掘與領悟。我將帶著這些知識在未來探索更多不同學科專業的議題。舉凡而言，刑事案件的不同特質，是否會導致不同的量刑結果？又或者，不同犯罪者的再犯率與刑度的關聯。這些也許都是我們可以透過不同工具解決的制度問題，我也期許自己有一天能夠結合所學，創造一個更理想與完備的制度體系。

附件

一、 本小論文獲中學生網站法政類 小論文優等獎之獎狀



二、 小論文本文（請接下頁）

壹、前言

在 2022 年九合一大選中，關於候選人學歷、論文等議題受到相當程度關注。而此種關注，是基於純粹的議題炒作、還是候選人的學歷真正會影響到其得票率，深值我們關心。我們可以從許多政治人物選舉文宣上觀察到，候選人傾向將自身學歷作為宣傳重點，政治人物繼續就讀在職專班的情況亦十分常見。從此觀點而言，政治人物在選戰策略上，似乎認為學歷是相當重要的因素。然而，這種選擇偏誤是否存在，仍然需要實證研究予以分析。我們也觀察到司法實務上，有將學歷影響程度量化的需求存在。因此，本文從 2018 年直轄市議員選舉出發，以多元迴歸分析為主要工具，試圖量化相關指標。

貳、文獻探討

一般而言，所謂候選人評價（candidate evaluation）是選民評價候選人能力的一種模式。如候選人外貌、經驗、學歷、政治能力等，都可被認為是候選人評價的一環。在 Osborne & Slivinski（2006）建構的民主選舉框架中，候選人評價是三大影響投票的因素——選民能力、政黨認同、候選人能力之中，對候選人能力的一種評斷樣態。傳統上會認為候選人評價可能是一浮動指標，對選舉觀察而言並不具有意義。不過歷來研究表明，實際上選民對候選人的評價，具有其理性判斷基礎（林佳龍，2000）。

Yahoo 奇摩新聞（2022）針對學歷關注現象所發佈的網路民調中，僅 1.8%的作答者認為擔任地方首長，學歷是最重要的特質。而值得注意的是，有 48.6%的作答者認為，「施政能力」是擔任地方首長最重要的特質。然而，事實上關於候選人的「施政能力」具體為何，對於沒有相關職務經驗的候選人而言，選民難以衡量。我們認為，在其他候選人評價變數缺乏的情況下，選舉人會以候選人的學歷作為評斷施政能力的標準之一。況且，尚有 41.8%的作答者認為候選人的道德標準是其考量的最重要因素。誠如 Besley & Reynal-Querol（2011）指出，教育程度實際上會影響政治參與者的道德操守。因此我們會發現，在這調查中，作答者認為重要的因素，實際上都與學歷無法脫勾。這提醒了我們，在事實上無法抽離學歷與其他候選人評價的前提下，並不能直接以問卷實驗法處理研究問題。

在選制差異部分，Besley & Reynal-Querol（2011）透過比較近百餘年來民主及獨裁政體領導人的學歷背景，認為民主制度相較於獨裁體制更容易產生高教育程度的領導人。他們認為，之所以比較結果如此，本質上是由於獨裁制度領導者的選拔，更傾向於關注繼任者的效忠程度，而非民主選舉制度所關注之候選人評價。

我們觀察到在司法實務部分，臺灣高等法院臺南分院（2012）曾審理一因雲林立委當選人偽造學歷而引起的當選無效之訴。在判決之中，法院在針對「偽造學歷是否會影響得票」的論述中，認為候選人學歷並非影響得票率的重要因素。並且法院認為，以原告及被告（偽造學歷者）之間約 25%之得票差距，顯然並非學歷影響所造成之。惟法院缺乏相關文獻支持，主要是依據經驗法則推論。也因此，在國內缺乏相關文獻證據的情況下，本文不僅為實證選舉活動，更具有為相關實務界提供具體量化指標的作用。

關於選民結構所造成的差異影響，林智堅（2017）提出「學歷較高的選民傾向投給學歷較高的候選人，而學歷較低者無此影響」的假說，惟缺乏相關文獻支持。與此對應地，van Noord et al.（2022）則透過問卷實驗，得出「較低教育程度的投票權人，較容易受到學歷因素的影響而做出選擇」的推論。事實上，這類問題也有許多以社會心理學的角度切入的理論，可大致歸因為「認同感」、「將其視為能力評斷指標」兩種面向。從第一種面向出發，我們可以推導出「較低教育程度的選民，會偏好低教育程度的候選人」及「高教育程度的選民，偏好高教育程度的候選人」兩種論點。但有趣的是，相關研究中，沒有證據表明高教育程度選民受學歷影響程度較大。至於「將其視為能力評斷指標」，則在實證研究上，高教育

程度及低教育程度選民均有偏好高教育程度候選人的實證（Spruyt & Kuppens, 2014），而關於我國地區差異情況，並沒有一致性的量化整理，也因此相關現象值得我們觀察。

參、研究方法

本文擬以觀察型研究（observational study）為主，從數據面以迴歸分析為出發點，探討候選人學歷對於得票率之關聯。以下就共變數選擇、評分方法、統計技術等要點說明。

一、研究數據選用

本文主要以 2018 年直轄市議員選舉之資料，驗證在我國，候選人學歷對其得票率之相關影響。至於為何使用 2018 年之資料，主要是本文研究目的上，希望得到一在沒有重大意外情況下的實證。本文認為，2022 年地方選舉對於學歷議題的炒作，可能已超出通常情況下，台灣選民對學歷的重視程度，因此，我們採用 2018 年的數據，以期歸納出較符合我國民情常態之數據。

（一）、數據範圍

本文之所以選用直轄市議員選舉作為資料搜集之範圍，乃基於以下二重要因素：

1、關注度受限於市長選舉

如同王宏恩（2014）所述，縣/市議員選舉，實際上受限於縣/市長選舉，關注度較低。這創造了一個有利我們驗證學歷影響的實驗環境——受關注程度較低，對於任一特定候選人評價較不明確。換言之，投票者對於要投票給誰，並不清晰（傅明穎，1998）。在這種情況下，我們更可以假定，選民會利用相對單純的選舉公報/候選人廣告上，清晰、不需要過多思考即可評價之學歷項目，來選擇候選人。當然，此種架設，需要建立在候選人有相當程度地針對自身的學歷進行宣傳。本文針對此的假定是，在多數地區，候選人均會某種程度的彰顯自己的學歷供選民參考，不論是在選舉公報上，或者是路邊廣告之中。不表明學歷者僅佔相當少之比例，不至影響研究結果。因此，我們可以假定這類影響，為一平均為 0 之隨機誤差。

2、減少遺漏變數

同時，我們認為，就縣市議員選舉而言，分析直轄市議員選舉之遺漏變數較少，較能精確地從數據面反應我們的研究問題。本研究中可能的遺漏變數，乃選民有關情感取向的投票行為。本文參考傅明穎（1998）的實驗結果，假定在直轄市中，基於選民整體教育水準高於縣的差異，這類人情/情感對投票的影響較為少見，選民會回歸到自身的候選人評價進行抉擇。另一種遺漏變數可能著重在異常競選方法，如以樁腳的方式非常態選舉策略。依據吳親恩（2008）之研究，此類異常競選，在都市化程度較高的直轄市較少見。因此，本文以直轄市作為觀察客體，可以消弭此類遺漏變數造成的影響。換言之，在減少這類遺漏變數後，我們所統計的變數，是可以完整反映選舉本身的。

（二）、實驗變數設計說明

本文所選用變數的代表意義說明如下。有關候選人客觀資訊的資料來源為中央選舉委員會選舉資料庫（中央選舉委員會，無日期）。筆者整理後的原始資料請見葉方叡等（2022）。

1、學歷以外的因子

筆者整理的原始資料中，性別之變數代號為 MALE，是為一 0 代表女性，1 代表男性之虛擬變數。並且資料中以多項虛擬變數處理候選人推薦政黨，均以 1 代表符合該描述，0 代表不符合該描述，包含 KMT（候選人為中國國民黨推薦）、DPP（候選人為民主進步黨推薦）、OTHER（候選人為其他政黨推薦）及 NP（候選人聲明為無黨籍）。選用此變數，源於在資料上我們可以觀察到在許多地區，獲政黨提名的候選人，其得票率顯著高於無黨籍或小黨的候選人，故有加以比較之必要。

同時，我們考量到在跨選區的比較上，不同的選區，意味著期望得票率不同，為了直接比較整體數據，本文參考王宏恩（2014），將得票率與預期得票率的比例（原始資料中註為 ADV 項）設定為得票率減期望得票率（100%除以同選區候選人數目），再除以期望得票率。

此外，資料中我們將政治參與經驗分為三類，分別是縣市級民意代表或更高層次的政治經驗，評為 3；除前項以外之民選公職或兩大黨地方/中央黨部重要幹部（如主委、中央決策委員），評為 2；其他政治相關工作，如議員/立法委員之辦公室助理，則評斷為 1。加入此變數，係源自 Atkeson & Hamel（2020）針對候選人經歷對選舉有所影響的結論，實有詳加比較之必要。

2、候選人學歷

本文針對候選人的學歷部分，本文採用的評審評分的方式評價候選人學歷。在社會科學中，缺乏一對於「學歷」之評分基準——特別是在其「觀感」部分。較多的作法是控制教育程度及時長，如 Besley & Reynal-Querol（2011）針對領導人教育程度與國家經濟狀況關聯的觀察。這主要是因為在行為社會學門的研究中，較為注重的是學歷所帶來的「能力」影響，而非本文所關注的觀感層次。

若再進一步思考，我們或許可直接以大學入學成績作為變數控制，然而這種方法無法比較出研究所及國外學歷與國內大學之間的影響差異。研究所學歷及海外學歷對選民投票抉擇影響程度不明的前提下，此方法並不能呼應本文的研究問題。

在缺乏評價候選人學歷的相關可靠標準的情況下，本文認為如李承達、駱明慶（2008）透過邀請數名評審，直接針對候選人的美貌程度評分，作為評斷候選人美貌媒介之方法最為恰當。以美貌作為變數，實際上與本文欲處理的學歷項目，有相當程度之相似——這類變數，在研究上都不易有客觀之標準，而當我們將目光望向本文所關注之投票者時，若我們直接將相關問題由評審評分，有助於我們比較候選人的學歷變數，而非受限於客觀指標。本文共邀請 3 位獨立評審，針對候選人學歷項進行評分。在評分時，請評審以 0~10 分進行評分，並允許輸入一位小數。而評審除了學歷外，無法看到任何其他能影響其判斷的候選人資訊。評審之特質描述請見附表。

誠然，理想的評審產生方式，是直接從各個選舉區投票權人中隨機抽樣。受限於資源，本文所邀請的評審，對於選民結構而言，並不具有代表性。然而，這也是此類探討選民觀感的研究所需要面對的問題——在不可能了解所有選民對於候選人特質評價的情況下，我們要如何確保評審之評分是可反映整體選民觀感的？本文認為，這類「觀感」，其實本質上是一個排序的過程——我們以 0~10 分為評分範圍，有人可能會認為博士學歷可以加 3 分，有人可能會

認為只能加 1 分。然而，這種評審之間存在的差異，實際上在個別評審中，依然可以形成對一群候選人學歷的排序。因此，此評分方式，實際上具有其可信基礎。再者，我們可能需要考慮地區差異的問題。易言之，各地選民對於各學歷之間的評分尺度，可能會有所不同。然而，如上文所述，這種差異，最終仍然會形成對候選人學歷指標的排序。也因此，在我們以 ADV 項控制各個選區所造成的預期得票率差異後，應不生改變研究結果的影響。

二、統計方法

本文之目的在於量化候選人學歷對於得票率的影響。在本文不可能重現所有選舉環境變數的情況下，本文以多元迴歸模型控制除了學歷以外，在文獻中被對於候選人得票率有顯著影響的因子，對於得票率的影響。換言之，本文透過多元迴歸分析抽離其他主要因子對於得票率的影響，推算候選人學歷對得票率的真正改變。

誠然，所有這類型試圖量化影響力的研究，都會面臨到遺漏變數的問題。換言之，就是對於「模型所包含的變數，是否能完整描述選舉」的疑慮。本文在已經盡量掌握相關資料的情形下，仍然有相關的問題，但不影響研究本身，詳細會在相關疑慮產生處討論。

此外，本文亦採用視覺化分析對資料進行簡單的整體觀察。繪圖部分採用 R 語言的 ggplot2 函式庫完成。

三、研究假設

本文在參酌如前文文獻探討所述之研究問題後，提出假設如下：

H₁：候選人學歷評價與得票率之間呈現正相關。

H₂：整體教育程度較低的選舉區，候選人的得票率較容易受到學歷因素的影響（van Noord et al., 2022）。

H₃：整體教育程度較高的選民傾向投給學歷較高的候選人，而學歷較低的選區無此現象（林智堅，2017）。

四、研究流程

本文統計部分之研究流程，首先探討各變數對學歷評分的影響，以了解不同背景候選人的學歷差異，並針對全體數據以視覺化工具，先行觀察總體學歷與得票率之關聯趨勢，其後以多元迴歸模型作為工具，試著抽離其他變因對得票率的影響，計算出學歷對得票率影響的具體數值。其後，本文將不同選區的情形分開進行迴歸分析，並透過視覺化工具，檢視不同背景的選區對得票率的重視程度差距。最後，總結上開統計分析，提供本文之觀察與結論。

肆、研究分析與結果

一、影響學歷評價之因素

由於評審並不可能從學歷去對應到個別候選人。因此，當我們分析學歷評價是如何改變時，可以忽略評審對於特定候選人的喜惡所造成的干擾，可直接以評分平均值檢視，不需進行個別分析。

我們觀察各特質候選人的學歷評分平均，可以發現，兩大黨候選人的學歷評價，是顯著高於平均值的（國民黨平均：5.89；民進黨平均：5.986；全體候選人總平均：

5.754。國民黨與民進黨候選人彼此差異並不顯著）。本文認為，這可能是基於兩大黨在初選的過程中，就已經將候選人學歷變數納入考慮的結果，特別是在如今初選多採用民調方式進行，此種設想更為可能，但此種假說仍需要進一步驗證。同時，男性與女性候選人，有相當大的差異（男性候選人平均：5.69；女性候選人平均：5.872）。我們認為，此種差距是來自政黨因素——女性候選人之中，有 60.9% 的候選人屬於兩大黨，而男性僅有 54% 為兩大黨推薦之候選人。再考量兩大黨候選人學歷水準較平均值為高的影響下，可以推論此種差距並非來自於性別本身。此外，年齡因素部分，年齡越小的候選人，其學歷平均越高。而上述這些特質的影響，需要透過多元迴歸分析去釐清其具體改變程度。

表 1 候選人學歷評價之多元迴歸分析

	常數	性別	連任	民進黨	國民黨	年齡	經歷	學士	碩士	博士	海外學歷
beta 係數	4.180	-0.024	0.140	0.061	-0.007	-0.014	-0.067	2.243	1.227	0.923	0.800
t-value	4.18	-0.34	1.32	-0.65	-0.08	-4.22	-1.65	23.96	15.59	4.77	6.83
R ²	0.7422										

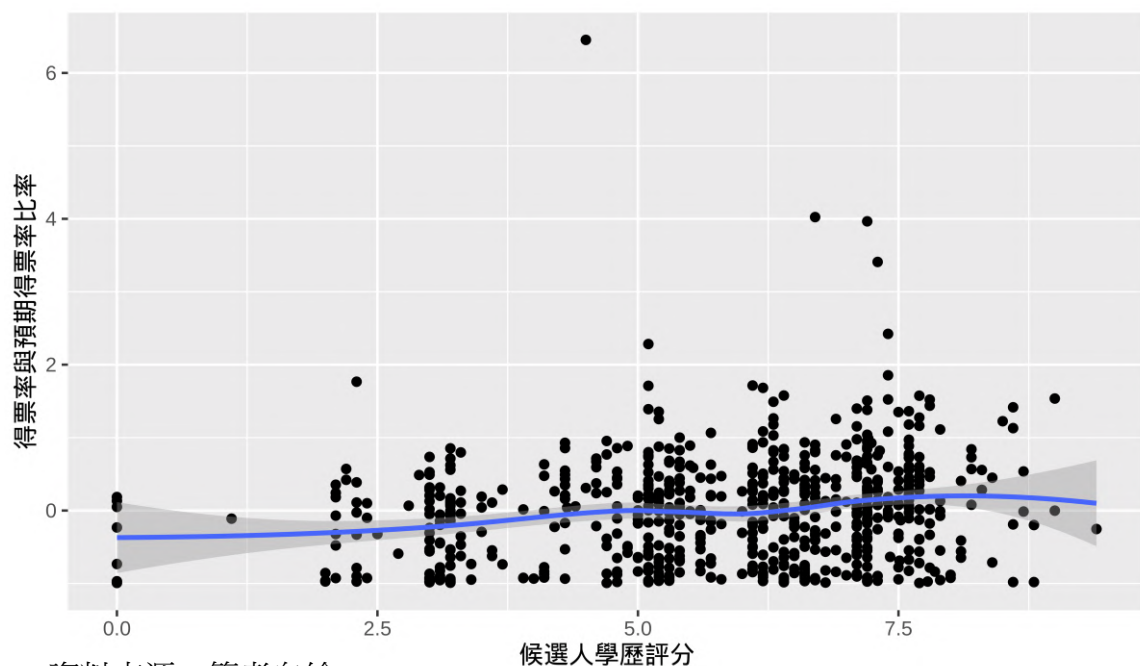
*應變數為學歷評價

表 1 中，我們以多元迴歸分析，探討影響學歷評價之因素。其中，性別、黨派因子都是不顯著的，顯示前文關於平均值的推論並不精確。值得關注的是，連任因子的係數顯著異於 0，我們後文將針對此點進一步討論。

影響最大的還是學歷本身。若候選人具有大學、碩士、博士學歷，則分別可在學歷評分上多獲得 2.24、1.22 及 0.92 分，且不論是 *p* 值或 *t* 鑑定均相當顯著。而候選人若具有海外學歷，不論學位，在控制其他變因的情況下，對於學歷評分有 0.8 分之正向效果，可能顯示對當今社會而言，海外學歷仍然相當吸睛。

二、學歷影響？視覺化分析

圖 1 候選人 ADV 與候選人學歷評分之分佈圖



資料來源：筆者自繪

表 2 候選人得票率之多元迴歸分析

	(1)	(2)	(3)	(4)	(5)	(6)
學歷水準*	0.08262 (5.154)	0.0802 (5.028)	0.0869 (5.296)	0.04754 (3.151)	0.02995 (2.290)	0.0319 (2.465)
性別		-0.1683 (-2.920)	-0.1789 (0.002)	-0.1011 (-1.932)	-0.0636 (1.406)	-0.0638 (-1.420)
年齡			0.0043 (0.0893)	-0.0083 (-3.343)	-0.0074 (-3.419)	-0.0086 (-3.825)
中國國民黨					0.8962 (13.372)	0.844 (22.255)
民主進步黨					0.4285 (7.506)	0.3765 (5.153)
現任				0.7022 (12.816)	0.3764 (6.952)	0.2143 (3.190)
經歷						0.1030 (3.990)
常數項	-0.4690 (-4.872)	-0.3461 (-3.300)	-0.5903 (-3.326)	0.0866 (-0.527)	-0.2771 (-1.921)	0.3081 (-2.110)
R²	0.03563	0.0460	0.04866	0.2309	0.4263	0.4978

註：括號中為 t 值，反映回歸係數是否顯著異於 0，簡單而言，若其大於 1.96 或小於 -1.96，可以推論該回歸係數有意義。表中所有模型的應變數均為 ADV 項。

*：考慮到不同數量級的變數不可一同分析。因此我們在此階段操作時，先將評審的原始評分統一除以 10 後進行分析。

首先，我們對整體資料進行視覺化檢視。針對整體選舉的分析，可視為多次實驗，可以有效平均極端案個案的影響力，較能反應整體趨勢。我們以 ADV 項控制選區對於得票率之影響後將每位候選人的學歷評價（X 軸）與 ADV 項（Y 軸）的分佈繪製成圖 1。藍線為 ggplot2 中 smooth 函式所推算的平滑曲線，其繪圖方式為在每個 X 截面上進行線性迴歸（王宏恩，2014），而灰色區塊是 95% 信賴區間。假設假說 H_1 為真，那我們理應觀察到藍線的斜率為相當顯著的正值，但圖形中並無此一趨勢。並且在相對低學歷的選舉人中，似乎存在學歷的負向影響。至於相對高學歷候選人區，則沒有特別的一致趨勢。但值得注意的是，在學歷極高（>8）的區域，似乎存在某種得票率隨學歷上升而降低的趨勢。極端案例發生在新北市第二選舉區的林麗容及黃偉榮，兩者學歷評分分別為 9.1 及 8.7，而其 ADV 項則分別為 -0.98 及 -0.97，在人數極少之極端高學歷區，此二極端案例所造成的曲線變化，不宜過度推論為整體趨勢。此外，雖然迴歸直線大致顯示學歷評分越高，ADV 項越高，兩者之 Pearson 相關係數為 0.1061，但 $R^2 = 0.00297$ ， $p = 0.2463$ ，並沒有特定統計上之意義，同樣無法解釋相關現象。

三、學歷影響幅度？迴歸模型分析

第二種方法，我們採用多元迴歸分析來處理研究問題。多元迴歸分析相較於單純的視覺化檢視，其優勢在於透過數理化的模型建構，我們可透過量化方式觀察特定變數對於應變數的影響幅度。並且，迴歸模型在資料上，也可以形成對於變數的控制，避免因單一數據推論，而忽略其他變數造成的影響。

(一) 整體數據分析

表 2 將多種變數——「性別」、「學歷水準」、「黨籍」、「爭取連任」、「經歷」等項目，逐步加入迴歸分析中討論。首先，在僅討論學歷水準本身的模型(1)中，beta 係數為 0.082。而模型(2)加入性別後，可以觀察到性別的 beta 係數為 -0.16，這代表女性在得票率上，似乎較有優勢。模型(3)加入年齡變因，年齡本身並不顯著，其他數值並沒有明顯的改變。模型(5)加入政黨因素後，現任及學歷水準係數，均出現顯著的降低。

模型(6)加入候選人經歷因素後，除了模型的解釋力 (R^2) 顯著提升，我們可以觀察到，特別是兩大黨，候選人的政黨是影響得票率最主要的變因。這點也可以從模型(5)加入政黨因素後，其他變數的係數顯著下降推論而得。從原始資料而言，國民黨籍候選人的平均 ADV 值為 0.6215，而民進黨候選人為 0.1695。相比之下，非屬於兩大黨的候選人，其 ADV 平均值為 -0.4627，有鮮明差距。而現任者之 beta 值為 0.21。換言之，作為現任者，在控制其它變數後，現任者這一特質本身，平均可替候選人增加預期得票率 20% 的選票，相當顯著。與此相比，我們關注的學歷水準 beta 值下降到 0.0319，但依然在 1% 的水準下顯著異於零，虛無假設被拒絕。而這代表著，在控制其他變數後，學歷水準每增加一個標準差，候選人會多獲得預期得票率 5.4% 的選票。而在最極端情況下，候選人的學歷可以為候選人增加預期得票率約 30% 的得票。

關於性別差異部分，在黨籍的影響加入控制後，其顯著性有所降低。這也可以呼應前文對於此的假設。在學歷評價上，性別應該是主要受到推薦政黨的影響，而非其本身。

我們也將有考慮(6)及未考慮學歷水準因素(7)的模型所預測之 beta 值互相進行 t 鑑定。如前文所述，假設學歷因子只是包含了某種參數的「代理變數」，那在加入考慮及未考慮學歷變因的模型，各項變數的 beta 係數前後，不應該產生顯著的改變，且學歷項目之 beta 值應也不顯著(李承達、駱明慶，2008)。將變數在考慮學歷影響與未考慮學歷影響的 beta 係數進行 t 鑑定。除了性別因子 ($t = 0.2619$)，其餘係數前後的差距在統計上均顯著不同。因此我們確信，學歷對於得票率的影響，並非其他因子的代理變數，具有其獨立意義。綜上，假說 H_1 是被接受的。

表 3 候選人票率之多元迴歸分析(續)

	(7)
學歷水準	
男性	-0.633 (-1.49)
年齡	-0.101 (-4.81)
中國國民黨	0.852 (14.35)
民主進步黨	0.380 (6.51)
現任	0.236 (3.53)
經歷	0.100 (3.90)
常數項	-0.059 (-0.58)
R 平方	0.4394

(二) 不同候選人特質的影響幅度？

1、小黨候選人

小黨的候選人，由於除去了影響最大的黨籍變數影響，這些候選人的得票率是受到何種因素影響，也有助於本文研究問題的理解。

無黨籍或小黨的候選人，影響得票率最多的變因是現任與否(係數為 0.846, $p < 0.001$)。在無黨及小黨候選人的情況下，學歷評分每上升一個標準差，可替候選人增加預期得票率 2.8% 的選票。相較於國民黨候選人的 8.9%，以及民進黨候選人的 5.8%，本文認為這種差異是基於政黨具有更多地方資

源，可以透過宣傳，使得縱然在選戰關注度受限的直轄市議員選舉中，部分選民仍能夠加強對政黨候選人學歷的印象。此部分涉及到候選人宣傳量能的遺漏變數問題，而這種差異會發生在政黨差異之中，故此種變數應已被作為代理變數，包含於政黨這項因子之中。

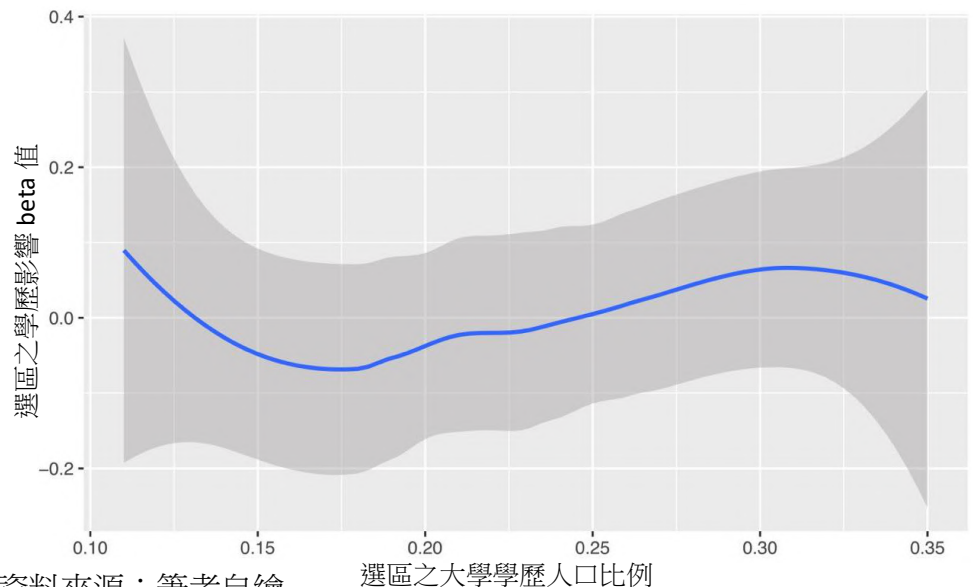
2、爭取連任的現任者

由於我們在前文觀察到，爭取連任的候選人這一變數，對於學歷評價會有所影響。因此，我們也以迴歸模型觀察影響其得票率的因素。單迴歸模型顯示，學歷水準項的 beta 係數為 0.04740 ($p = 0.0246$)，但在加入政黨、年齡等因素後，學歷水準係數的顯著性消失 ($p = 0.6240$)，且整個模型中，僅有國民黨這一項變數的 beta 係數在 10% 顯著水準下依然顯著。這可能顯示，爭取連任者主要是受到某種本文所搜集的資料中無法表現的變數，而影響其得票率。此部分仍有待後續研究評估，但至少在本文所掌握到的全部數據而言，學歷評價對於爭取連任的候選人得票率影響不大，可以推論是由於連任者參數本質上包含了許多政治能力的變數，進一步稀釋掉其他變因的影響。

四、地區差異

我們也嘗試以各個選區為單位，將該選區具有大學學歷的人口比例與各個選區進行多元迴歸分析。而由於部分小選區中，有候選人全部都屬於兩大黨，或者全部都是男性，抑或是全部都為爭取連任的候選人，使用迴歸模型推估時，會造成共線性問題。遇到此狀況時，本文優先選擇剔除性別變數。如果仍然出現共線性問題，由於刪除黨籍或現任者變數都會造成學歷項之 beta 值巨大偏差（請見上文模型(5)），因此我們直接刪除該選區的數據。

圖 2 各選區教育程度與學歷影響程度趨勢圖



資料來源：筆者自繪

圖 2 中我們將學歷評分項之 beta 值與各選區之大學學歷人口之分佈，透過 ggplot 之 smooth 函式繪製成趨勢圖。假設假說 H_2 或 H_3 為真，我們理應觀察到相對於中間區外，左右兩端的曲線應較為凸起，然而，我們沒有辦法透過圖 2 觀察到相關現象——選區學歷較高之處，有影響程度下降的趨勢存在；而整體教育程度較低的區域，beta 值也有高有低。此外，單迴歸模型顯示係數為 0.0244， t 值則為 0.6，拒絕對立假說。綜合以上，假說 H_2 與 H_3 的應予拒絕，換言之，我國選舉，從實證資料而言，無法歸納出特定教育程度的選區對於候選人學歷之特定投票傾向，與文獻中說法並不相符。

然而，從數據面觀查無此現象，可能是某種選民內心流動的結果，有待後續研究論證，特別是社會心理學層面。此外，若從政黨配票的視角出發。在大多數選區，主要的參選人多為兩大黨的候選人。而在兩大黨採用配票策略以最大化整體得票率的情況下

(王業立, 2006), 亦可推論是導致兩大黨候選人之間的差異均不顯著, 進而造成整體資料視覺化上, 無法觀察出各區域的差異。當然, 由於公開資訊缺乏關於兩大黨選戰策略在個別選區有採用配票與否的相關資料, 相關假說在驗證上有困難, 值得後續研究加以分析。

五、現實案例討論

前文我們已經確認了學歷評價確實會造成得票率的正向改變。然而, 這個數據如何量化? 我們從前文所提到的雲林縣立法委員選舉學歷紀載爭議為例, 在不考慮政黨加成影響、並假定在 2012 至 2018 年這段期間, 選民對於候選人學歷之重視程度沒有顯著改變的情況下, 以本文所得出的實證數據而言(採用模型(6)之係數), 就算考慮最極端的情況(偽造學歷的候選人全部掌握了學歷因素, 意即, 其學歷可以被量化為 10 分), 學歷的影響最多也僅能改變 $0.0319 \times 10 \times 0.5$ (模型(6)中, 學歷評分項之 beta 係數) $\times 10$ (假設其獲得了最高的學歷評分) $\times 0.5$ (預期得票率), 約為 7.9% 的得票。當然, 我們的模型是建立在關注度較低的議員選舉上, 與立委選舉情形存在一定程度的差異。但在立委選舉關注度更高, 選民依賴其他指標進行候選人評價的判斷比例會更高的情況下, 可以想見這種影響只會更低。綜合以上, 法院認定兩者 25% 的得票差距主要並非學歷造成的(臺灣高等法院臺南分院, 2012), 是合乎本文實證資料的論述。

伍、研究結論與建議

本文以 2018 直轄市議員選舉為例, 探討學歷對候選人得票率的影響, 以瞭解學歷在候選人政治生涯中所扮演的角色, 同時為相關司法實務提供具體量化指標。

在學歷評分部分, 我們觀察到主要兩大政黨的候選人, 其學歷評分顯著高於無黨籍或小黨的候選人。本文認為, 這種現象可推論候選人的學歷本身就會對政黨初選, 特別是兩大黨初選的結果產生影響。

而學歷對於得票率的影響, 透過迴歸模型控制了其他解釋變數後, 可發現候選人的學歷水準, 最多會影響預期得票率(100%除以同選區候選人數) 31.9% 的得票。進一步將兩大黨與小黨的數據分別進行迴歸分析, 可以發現候選人學歷評分對兩大黨候選人得票率的影響高於小黨或無黨籍候選人。我們認為這是政黨造成候選人宣傳量能的差異所造成的影響。綜合迴歸模型, 我們認為台灣選舉中, 確實存在好的學歷評價可以幫助候選人得票的實證。但回到社會層面, 從模型得出的數據而言, 可以發現學歷並不是最主要的影響因子, 換言之, 台灣選民對於政治人物學歷的態度, 並沒有如想像般重視。也因此, 政治人物對於學歷的盲目追求, 相關現象更值得我們反思。若能回歸到政治參與經歷、議題表述等層面, 對於政治人物自身的得票率, 乃至於社會整體, 才更有幫助。

此外, 本文發現, 從實證角度無法佐證在許多文獻中記載, 關於候選人學歷對不同教育程度選民影響差異的相關論述。由於此一議題, 較少研究者從如本文的觀察型研究視角出發, 因此本文建議相關研究可以加入更多相關縱橫比較, 進一步歸納相關現象。

研究方法層面, 基於欠缺客觀評分學歷方法的前提下, 本文採用評審評分的方式以評價候選人學歷。雖然相關的研究也有類似的評價方法, 且此方法最能呼應我們的研究問題, 但筆者也建議在後續研究中, 可考慮強化評審背景的多元性。

需要特別澄清者, 本文並無意圖針對造成「選民為何因候選人學歷而改變投票傾向」的命題研究, 主要原因乃是本文在研究方法選用上, 一開始由於研究目的的設計為數理具體化相關的影響, 而捨棄了問卷調查一途。因此, 本文雖然已經成功量化候選人學歷對得票率的

影響，但關於台灣選民內心中是如何流動，造成此種現象，仍值得更加關注個體的實踐研究，特別是社會心理學層面的觀察。

陸、參考文獻

- Atkeson, L.R., Hamel, B.T. (2020). Fit for the Job: Candidate Qualifications and Vote Choice in Low Information Elections. *Political Behavior*, 42, 59–82.
- BESLEY, T., & REYNAL-QUEROL, M. (2011). Do Democracies Select More Educated Leaders? *The American Political Science Review*, 105(3), 552–566.
- Osborne, Martin J. and Al Slivinski, (1996), A Model of Political Competition with Citizen Candidates, *Quarterly Journal of Economics*, 111(1), 65-96.
- Spruyt, B., & Kuppens, T. (2014). Warm, cold, competent or incompetent? An empirical assessment of public perceptions of the higher and less educated. *Current Sociology*, 63(7), 1058–1077.
- van Noord, J., Kuppens, T., Spruyt, B., & Spears, R. (2022). When and Why People Prefer Higher Educated Politicians: Ingroup Bias, Deference, and Resistance. *Personality and Social Psychology Bulletin*, 2, 23-41. <https://doi.org/10.1177/0146167222107779>
- Yahoo 奇摩新聞 (2022 年 07 月 07 日)。你是否會重視候選人的學歷？。
<https://perma.cc/FM48-KRU9>
- 中央選舉委員會 (無日期)。選舉及公投資料庫。<https://db.cec.gov.tw/>
- 王宏恩 (2014)。喜獲籤王？初探我國選舉中的排序效應：以 2002 縣市議員選舉為例。 *政治科學論叢*，62，195-220。
- 王業立 (2006)。 *比較選舉制度*。五南出版社。
- 李承達、駱明慶 (2008)。美貌對候選人得票率的影響——以 2004 年立委選舉為例。 *經濟論文叢刊*，36(1)，67-113。
- 林智堅 (2017)。 *三人競選之中槓桿者的政治社會基礎及其影響——以 2014 新竹市長選舉為例*。國立臺灣大學社會科學院國家發展研究所：碩士論文。<https://perma.cc/P8Q9-2DAN>
- 林佳龍 (2000)。台灣民主化與政黨體系的變遷：菁英與群眾的選舉連結。 *台灣政治學刊*，4，3-55。
- 吳親恩 (2008)。地方議會金權政治的變化：司法判決書的分析。 *台灣政治學刊*，12(2)，165-212。
- 傅明穎 (1998)。北市選民的候選人評價與投票決定。 *台灣政治學刊*，3，195-243。
- 葉方叡、歐鴻勳、趙柏任 (2022 年 12 月 15 日)。2018 直轄市議員選舉學歷水準評斷資料集。<https://doi.org/10.6084/m9.figshare.21731900>
- 臺灣高等法院臺南分院 (2012 年 11 月 27 日)。臺灣高等法院臺南分院 101 年度選上字第 9 號民事判決。<https://ppt.cc/fTJa5x>

附表 評審特質基本描述與評審間評分相關係數互聯表

評審	年齡	性別	居住地	教育程度	與 A 評審之 Pearson 相關係數	與 B 評審之 Pearson 相關係數
A	24	男	新北市	碩士	1	0.4224
B	36	女	高雄市	大學	0.4224	1
C	28	男	桃園市	博士	0.4261	0.5712