

國立政治大學經濟學系碩士班

碩士學位論文

COVID-19 疫苗施打與網路輿情聲量關係：以
Moderna 疫苗為例

How the internet opinion and sentiment influence the
willingness of getting vaccinated? Case of the Moderna covid-
19 vaccine

指導教授：王信實 博士

研究生：曾偉恩 撰

中華民國 111 年 6 月

謝辭

本論文最後在碩二暑假完成，首先先感謝我爸媽無私鼓勵我讓從機械系到研究所轉為經濟所，在生活不需要擔心，可以自由暢遊在學術。

接著要感謝我兄弟楊士逸同學，如果沒有你與我無數次衝撞學術思想，這篇論文完備程度不會有如此，再來就是感謝王信實老師在實證上透過經濟直覺與模型結果相互驗證，以及孫懋嘉老師在社群資料以自身經驗給予在經濟上不同多維度思維，讓計量經濟模型應用在社群資料提高到更深的層次，最後要感謝廖仁哲老師，老師幫助我在使用合成控制法上運用自如，還有許許多多萍水相逢的網友以及業界前輩提點資料與模型使用。

再次感謝這兩年來幫助過我的老師、同學、家人與朋友，在此送上最誠摯感謝。



摘要

近年來 COVID-19 大肆傳染，政府積極傳遞疫苗的相關資訊來防止疫情擴散，由於民眾接收 COVID-19 資訊不只是來自政府，還有一部分來自網路輿情，而網路上存在許多真假難辨的資訊，造就民眾產生施打疫苗的疑慮，使疫苗施打量無法達到政府預期，因此若政府釐清網路資訊與施打量的因果關係，或許能提高疫苗施打量。在此透過合成控制法 (Synthetic Control Method)，使用 OpView 資料庫的聲量資料，以及衛生福利部疾病管制署提供 AstraZeneca、BioNTech、Moderna、Medigen 四種疫苗在施打量，發現網路輿情與施打量之間存在相關性後，並嘗試找出其因果關係。

關鍵詞：網路輿情、聲量、疫苗施打、因果關係、合成控制法



ABSTRACT

During the COVID-19 pandemic, the governments around the world actively disseminated the vaccine information and promoted the vaccination to prevent the epidemic. The mis- and dis-information about the vaccination on the internet usually makes people worried and thus decreases the willingness of vaccination. By using the Synthetic Control Method and the OpView data, as well as the AstraZeneca, BioNTech, Moderna, and Medigen vaccines provided by Taiwan Centers for Disease Control, this study investigates the negative causal relationship between the internet public opinion and vaccination. It is helpful to increase the number of people vaccinated by clarifying the causal relationship between internet information and vaccination.

Keywords : Internet Public Opinion 、 Volume 、 Vaccine Injection 、 Causality 、 Synthetic Control Method

目次

第一章 緒論.....	1
第二章 文獻回顧.....	6
第一節 COVID-19 衝擊與信念形成.....	6
第二節 網路輿情與疫苗施打意願.....	6
第三節 合成控制模型架構.....	7
第四節 合成控制模型演進.....	8
第三章 資料介紹.....	10
第一節 資料庫來源.....	10
第二節 資料處理.....	11
第四章 研究方法.....	17
第一節 合成控制法.....	17
第二節 模型設定.....	18
第五章 模型結果.....	19
第一節 合成結果.....	19
第二節 安慰劑檢定.....	23
第六章 結論.....	25
第一節 主要研究結果與貢獻.....	25
第二節 研究限制與未來研究方向.....	25
參考文獻.....	27
附錄.....	1

表次

表 1 衛福部與自訂施打期數.....	12
表 2 SYNTHETIC CONTROL METHOD 合成結果.....	20
表 3 2 種疫苗的權重 w	20
表 4 解釋變數權重.....	21
表 5 OpVIEW 關鍵字搜尋.....	1

圖次

圖 1 MODERNA 疫苗施打量 12 期時間趨勢.....	12
圖 2 負聲量建構 BOLLINGER BANDS.....	15
圖 3 正聲量建構 BOLLINGER BANDS.....	15
圖 4 負聲量 1 - 10 期之間相關性 (ACF).....	16
圖 5 負聲量 1 - 10 期之間偏相關性 (PACF).....	16
圖 6 施打量時間趨勢：MODERNA VS. SYNTHETIC MODERNA.....	21
圖 7 MODERNA VS. SYNTHETIC MODERNA 之間差距.....	23
圖 8 合成 MODERNA 施打量差距和安慰劑在不同控制組下.....	23

第一章 緒論

近年來網路快速發展，民眾在網路上自由地表達看法。相反地這讓網路上帶風向以及不實醫療資訊層出不窮，衍伸出政府無法控管的外部性。例如 2021 年 YouTube 上流傳肝膽排石法，宣稱透過喝橄欖油排膽結石，經由網路傳播可能讓民眾誤食造成不必要醫療浪費。各國在近年來也發生類似網路不當傳播，導致許多醫療成本，政府開始重視網路影響不單單是民眾自由發表個人意見的平台，更可能會透過網路改變民眾做決策。各國陸續也做出相對應的措施，來解決這世代網路蓬勃發展所造成影響。

2021 年 COVID-19 席捲全球，起初還沒研發出疫苗，為了讓重症人數與死亡人數下降，各國開始封城，民眾對於 COVID-19 資訊很多自於網路上資訊，本研究針對網路輿情進行分析。隨著時間推移，疫苗陸續分發到各國家，但由於網路輿情影響有一方民眾認為疫苗副作用死亡機率遠大於得 COVID-19 死亡機率，因此疫苗施打量在期初一直都不高，可能是網路輿情影響民眾對於疫苗看法。

在國外很多國家一開始就讓民眾自由選擇要打哪一種品牌的疫苗，像是美國食品及藥物管理局已核疫苗廠牌包含 BioNTech、Moderna、Johnson，以及世界衛生組織核准緊急使用疫苗包含 AstraZeneca、Convishield、Janssen，一開始施打量在各國持續上升，直到覆蓋率到達一定後，覆蓋率就無法再持續上升。像是在美國矽谷地區的施打率很高，而在某些地區施打率就不高，可能是因為能接受到資訊並不多，對於疫苗會抱持懷疑，也有可能是因為生活圈高度重疊，導致在此地區民眾能夠接受到的資訊過於單一，大部分接受到訊息來自相同生活圈之間在網路上互相分享文章或是貼文，在 Loomba & Figueiredo et al. (2021)提到在美國相信家人與朋友與受到錯誤訊息影響有關，因此施打量可能是因為朋友與家人之間

共同生活圈導致，即便政府會公布正確疫苗知識，依舊相信身旁朋友提供資訊，會演變成只要有人接受到網路上錯誤資訊，將會快速傳播，接受到錯誤資訊民眾之後會難以去更改對於疫苗的看法，產生對疫苗疑慮讓施打量無法持續提高。

本研究嘗試找到網路輿情與民眾施打量之間是否有影響，在此針對 Moderna 疫苗，Moderna 疫苗是台灣全面開放施打的第二種疫苗，為了確保研究可靠度，不適合使用太晚施打疫苗，會讓影響關係誤判導致研究不具可靠性，而選擇太早施打疫苗會面臨到其他疫苗還沒開始施打問題，此時期間選擇就為重要，基於以上理由，本研究採用 Moderna 施打量作為研究主題。先從找到網路輿情與疫苗施打量之間相關性，而再研究網路輿情與民眾施打量之間是否具有因果關係，之後政府可以透過網路輿情與 Moderna 施打量之間因果關係來讓民眾增加施打意願，解決疫苗施打疑慮，提高民眾施打量，盡快讓每個城市之間覆蓋率提供，達到全國人民群體免疫。

本論文透過 OpView 資料庫，針對 AstraZeneca、BioNTech、Moderna、Medigen 疫苗相關關鍵字查詢，每篇網路貼文或留言透過背後演算法計算出每一篇貼文或是留言的正向、中立、負向情緒聲量，其值皆為 0-100，在看正面、負面、中立聲量哪個高，此貼文或是留言標記情緒為該情緒聲量最高者之情緒。透過三種正、負、中立的 Moderna 情緒聲量時間趨勢圖，發現在第 5 期 (9/15-9/23) 變化最大，使用正、負面情緒聲量所建構 Bollinger Bands 驗證變化巨大的區間為第 5 期，並將此期當作介入效果。Bollinger Bands 上下通道在傳統上採用 Simple Moving Average(簡單移動平均)，傳統 Bollinger Bands 是將過去歷史事件給予相同權重，但民眾在形成對疫苗信念，會依據越近期發生事件給予較高權重，由於此特性採取與傳統方法特性不同，考慮使用 Exponential Moving Average(指數移動平均)能夠將近期所發生事件給予更大權重，參考 Vergura (2020)。透過三種疫苗包含

AstraZeneca、BioNTech、Medigen 的施打量使用 Abadie, Diamond & Hainmueller (2010)提出合成控制法 (Synthetic Control Method)合成出控制組施打量來確認與真實 Moderna 施打量之間介入效果的因果關係。

在後續第二章文獻回顧中主要分為四部分，第一部最近幾年文獻對於 COVID-19 帶來衝擊與民眾決策之關係，第二部分則是網路輿情與疫苗施打意願之間關係在行為經濟學上看法。第三部分介紹合成控制法 (Synthetic Control Method)模型架構以及此模型執行上遇到困難，第四部分介紹在合成控制法 (Synthetic Control Method)之後模型是如何針對模型限制改良。第三章共有二部份，第一部分為資料庫來源，本研究使用 OpView 資料庫提供關鍵字搜索，可以查詢貼文被標記是哪一種情緒聲量，以及衛生福利部疾病管制署提供全台四種疫苗在各期之間的施打量，第二部分為情緒聲量與衛生福利部疾病管制署提供四種疫苗施打量在資料上進行選取、分割、分析，其中選取資料期間為 2021/08/23 - 2021/11/09，在分割上考慮到在台灣施打疫苗不像國外能夠隨時都可以去施打，由於台灣特殊性是必須由衛生福利部疾病管制署公布預約時間，依序編號為幾期，而預約完成後再去預約時間施打，大部分公布可以施打時間通常落在一個禮拜內，在此特殊性下將施打量與關鍵字搜索得出情緒聲量日資料加總期資料，接著把四種疫苗關鍵字搜索之情緒聲量與四種疫苗施打量劃分成 12 期。透過三種正、負、中立的 Moderna 情緒聲量趨勢圖，發現在第 5 期 (9/15 - 9/23)變化最大，經查證可能是美國食藥署與歐盟藥品主管機關發布報告為施打 BioNTech 與 Moderna 疫苗跟心肌炎、心包膜炎之間有可能有關係，最後透過正、負情緒聲量計算 Bollinger Bands 後，驗證突破上通道為第 5 期加以驗證，而第 5 期為突破上通道作為 Moderna 施打量介入效果。¹接著為了確認影響下一期施打量變化全部都來自第 5

¹ 建構 Bollinger Bands 是使用日資料的 Moderna 負面情緒聲量。

期的介入效果，分別將 Moderna 負面情緒聲量進行 Autocorrelation Function (稱 ACF)與 Partial Autocorrelation Function (稱 PACF)分析，以確保效果是全部或是大部分都來自第 5 期 (9/15 – 9/23)介入效果。²

第四章的研究方法分別有二部分，第一部分介紹合成控制法 (Synthetic Control Method)如何透過控制組與實驗組之間介入前結果 (pre-intervention outcome)極小化差距，合成出控制組之演算法，第二部分為模型設定由合成控制法 (Synthetic Control Method)在此研究下定義變數，此研究 outcome variable (結果變數)為 Moderna 施打量。

第五章為二部分，第一部分呈現合成控制法 (Synthetic Control Method)合成 Moderna 施打量與真實 Moderna 施打量，第二部分為呈現介入效果透過安慰劑檢定 (placebo test)是否具統計顯著，以及合成結果與分析施打量變化。

第六章總結第一部份為本論文的結果與貢獻，根據 Loomba & Figueiredo et al. (2021)在網路錯綜複雜框架下，民眾受到網路輿情因而改變決策或是改變對於事件信念，是很難真正找出原因，本研究針對台灣在 COVID-19 期間，民眾接受 COVID-19 資訊很多都來自網路上資訊，就大大降低其他因素可能會去影響民眾對於有關 COVID-19 各種事件影響。³透過這項特殊性讓民眾對於網路影響導致疫苗意願改變，使用 Moderna 負面情緒聲量並將第 5 期當成介入效果，此效果影響與 Moderna 施打量可能為因果關係，第二部分說明目前在此議題上研究疫苗施打因果關係遭遇到限制。除了有資料上限制，合成控制法 (Synthetic Control Method)上模型本身限制，提及到未來提供其他研究者可能將此議題延伸主題，像是在其他國家是否也能將本研究結果具穩健性 (robustness)，或是透過 Valero

² ACF 與 PACF 分析使用期資料的 Moderna 負面情緒聲量。

³ 民眾接收 COVID-19 資訊可以從新聞媒體、報章雜誌、網路平台，由於網路普及，不論是哪一種接收管道，網路平台很多都會從新聞媒體、報章雜誌再次傳播。

(2015)提出 Demeaned Synthetic Control、Chen (2020)提出 Model-based Synthetic Control、Ben-Michael, Feller & Rothstein (2021)提出 Augmented Synthetic Control 解決模型上限制作為研究的方向。



第二章 文獻回顧

本章先介紹這幾年 COVID-19 對於民眾帶來哪些衝擊以及衝擊對於民眾決策是如何進行傳遞，第二節將介紹網路在 COVID-19、疫苗、民眾行為決策之間扮演的角色以行為經濟學的角度切入，第三節介紹本研究使用合成控制法 (Synthetic Control Method) 之模型架構與執行上的困難，最後一節是介紹近期合成控制法 (Synthetic Control Method) 模型上改良與演進。

第一節 COVID-19 衝擊與信念形成

Fetzer, Hensel, Hermle & Roth (2020) 研究 COVID-19 對於民眾經濟焦慮如何形成以及之間因果關係，而經濟焦慮透過 Google 搜尋四種關鍵字為 Recession、Stock Market Crach、Conspiracy、Survivalism 之搜尋強度來衡量。民眾對於死亡率與傳染率的信念與經濟焦慮狀況有顯著關係，在經過實驗設計過後，發現受試者對高死亡率在統計上擔憂明顯增加，對於給定有關傳染性的資訊，民眾明顯對於經濟狀況擔憂降低，以上這兩點死亡率與傳染性是否有訊息不對稱在形成經濟焦慮扮演了重要因素。陸續針對民眾信念預期不同研究與討論，Malmendier and Nagel (2011) 發現個人在更新自己信念時，個人對於自己近期發生的事件會給予過的重視，在 Tversky and Kahneman (1973) 特別針對重大更是如此。

第二節 網路輿情與疫苗施打意願

最近網路快速發展，假訊息大量充斥網路上，民眾接受到假訊息而去改變疫苗施打意願，將會造成接種疫苗產生外部成本，在 Loomba & Figueiredo et al. (2021) 研究表明 COVID-19 假訊息傳播可能會對疫苗接種意願產生影響，但在該研究並

未複製現實世界的網路平台。解決假訊息在網路平台上傳播是 COVID-19 疫苗接種重要因素，在網路上在訊息傳遞是快速，即便是假訊息很短暫傳播後，在長期下民眾對於接種疫苗意願將難以依此改變。針對疫苗意願研究在 Saleska & Choi (2021)從行為經濟學討論有可能因素，本研究在美國發現民眾對於疫苗開發異常速度和疫苗本身副作用相關報導可能會是影響施打意願的原因，分別從行為經濟學角度提出認知偏誤、負面偏誤、樂觀偏誤這三種偏誤，這些偏誤都圍繞在一開始認為疫苗施打意願如何，這些意願就會像魔鬼氈一樣難以消除，民眾傾向去找出與自己心中認為網路資訊，可能忽略真正資訊，即使政府在站公共衛生角度大力推廣施打疫苗的利益遠大於確診 COVID-19。

第三節 合成控制模型架構

Abadie & Gardeazabal (2003) 和 Abadie, Diamond & Hainmueller (2010)提出合成控制法 (Synthetic Control Method)，此方法透過控制組做一系列加權平均當作未受介入效果下的實驗組合成出合適控制組，在透過合成出來控制組與實驗組差距可以得知介入效果。Abadie, Diamond & Hainmueller (2010)設定為控制組群體透過凸組合 (convex combination)來進行加權平均，來達到介入效果前與實驗組擬合，以下為合成控制目標函數與限制式：

$$\begin{aligned} & (\mathbf{X}_1 - \mathbf{X}_0\mathbf{W})'\mathbf{V}(\mathbf{X}_1 - \mathbf{X}_0) & (1) \\ & \text{subject to } w_j \geq 0 (j = 1, 2, \dots, J) \text{ 且 } w_1 + \dots + w_j = 1 \end{aligned}$$

選取特定的 \mathbf{V} 、 \mathbf{W} 使得目標(1)極小，透過最適 \mathbf{V} 、 \mathbf{W} 採用凸組合去加權平均，若控制組資料採用凸組合後與介入前的實驗組相去甚遠，找不到合適 \mathbf{W} 可以用凸組合合成出合適控制組。

第四節 合成控制模型演進

Abadie & Gardeazabal (2003)是第一篇提出合成控制法 (Synthetic Control Method)並應用，該研究主要討論恐怖攻擊是否對於西班牙巴斯克地區(Basque country)造成經濟影響，而 Abadie, Diamond & Hainmueller (2010)這篇研究是將合成控制數學性質加以討論，該研究討論美國加州實施菸草管制政策 (Proposition 99)對加州菸草銷售是否具有因果關係，本篇討論介入效果下，提出安慰劑檢定 (placeo test)，第一次提出合成控制介入效果的統計檢定，後續 Abadie, Diamond & Hainmueller (2015)在本論文中為了再介入前更加擬合成控制組與實驗組之間，提出將介入前期間劃分成訓練期 (training period)與驗證期 (validation period)，在論文最後有做穩健性檢定 (robustness test)。

以上三篇合成控制是藉由凸組合 (convex combination) 進行合成，後續會提及相對於這三篇傳統合成控制法 (Synthetic Control Method)不同的改良模型，而在傳統合成控制法 (Synthetic Control Method)過程中有可能會產生內插偏誤 (interpolation bias)，此時 Abadie & L'Hour (2021)提供可能可以使用 Penalized Synthetic Control，此模型不同於傳統合成控制法 (Synthetic Control Method)裡面加入懲罰項來解決內插偏誤的問題。

而傳統合成控制法 (Synthetic Control Method)的另一個問題為差擬合 (poor matching)，為了解決此問題，包含 Ferman & Pinto (2021)的 Demeaned Synthetic Control，透過介入前控制組與實驗組在個體差異透過去除平均方式，Chen (2020)的 Model-based Synthetic Control 是透過 Model Average 方式來解決差擬合，Ben-Michael, Feller & Rothstein (2021)的 Augmented Synthetic Control，透過放寬凸組合達到讓權重允許可以是負的來解決。而 Chen (2022)的 Regularization Synthetic Control 一文中，計算各種合成控制方法的偏誤 (bias)性質及預測均方誤差

(MSPE)，以上整理了到現在 2022 以前合成控制方法的改良與演進。



第三章 資料介紹

在本章分別有二節，第一節主要是介紹本研究使用的資料來源，其中有 OpView 資料庫以及衛生福利部疾病管制署所提供 COVID-19 四種疫苗施打量，第二節為資料處理，針對在第一節提到資料分別進行資料選取、分割、分析三步驟。

第一節 資料庫來源

根據意藍資訊的「OpView」資料庫提供台灣有關各式各樣輿情分析相關資料，資料主要來自討論區：Ptt，社群網站：Facebook、Instagram 等平台，部落格、新聞，這些來源都包含了主文與回文，該資料庫可以透過鎖定關鍵字搜尋來取的主文與回文，會透過背後演算法得出每篇主文與回文正、負、中立情緒值，其值介於 0 - 100 之間，在正、負、中情緒哪個高將情緒標記於此。礙於資料庫是企業營利組織產品，不公開背後演算法是如何計算出情緒值，之後研究者使用此資料庫需要注意。本研究針對四種疫苗進行關鍵字搜尋如附錄，時間為 2021/01/01 - 2021/11/09。

衛生福利部疾病管制署提供針對 COVID-19 疫苗施打量，分別是 AstraZeneca、BioNTech、Moderna、Medigen 疫苗施打量，時間為 2021/3/22 - 2021/11/09，其中 BioNTech 施打日期為 9/22 後施打，之前零星數量為國外回台補登記。⁴Medigen 疫苗施打日期為 8/23 後施打，之前零星數量為參與 Medigen 臨床試驗施打量。在施打日期資料上，在第三劑以前是採預約制，民眾選擇合適地點與日期去施打，並非隨到隨打，而少部分民眾施打殘劑，政府無法進行統計。

⁴ 9/22 之前 BioNTech 施打量都補為 0，為了避免控制組被污染 (contaminate)，讓介入效果不夠純 (pure)。

第二節 資料處理

假設網路影響在當期對民眾影響大部分都在下一期會反應，而民眾接受到有關疫情的資訊很多是透過網路上資訊所獲得，根據政府提供疫苗施打期數，認為探討本期在網路輿情發生所有事件，若民眾在當期更改施打意願，但卻只能在下一期對疫苗施打量做出改變。^{5,6}資料採用從 2021/08/23 - 2021/11/09，將期數透過合適此研究進行自訂期數如表 1。⁷

在表 1 中第二行為衛生福利部疾病管制署公布施打疫苗期數，在第三行為本研究所自訂期數，由於政府公布期數並不完全，為了讓資料連續在每一期前後去補齊日期而去訂出此研究期數，第四行裡面天數是為了讓資料連續，導致不同期數之間日期數是不相同，未來研究者可能針對此問題劃分平衡期數，最後在施打量採取日期是 2021/08/23 - 2021/11/09，而不是在 11/12 是為了配合情緒聲量資料，總共期數為 12 期。

本研究資料將四種疫苗施打量資料劃分期數與 OpView 資料庫提供正、負、中立聲量也將其劃分成 12 期，彙總後需要 AstraZeneca 疫苗為 12 期的 OpView 正、負、中立聲量，12 期的施打量，BioNTech 疫苗為 12 期的 OpView 正、負、中立聲量，12 期的施打量，Moderna 疫苗為 12 期的 OpView 正、負、中立聲量，12 期的施打量，Medigen 疫苗為 12 期的 OpView 正、負、中立聲量，12 期的施打量，以上資料期間為 2021/08/23 - 2021/11/09。

OpView 聲量對於每則文章都會正、負、中立聲量，無法確定是哪一種聲量主導在網路輿情上討論，以下透過三種聲量時間趨勢圖確定是哪種聲量來主導如圖 1。

⁵ 這裡的當期是指本研究自行定義的期數，由於每一期的日資料數為不平衡，會有些微影響。

⁶ 在此需假設網路輿情影響民眾改變決策大部分會在下一期反應，即使有些人反應不在下一期，由於本研究探討是當期網路輿情對於民眾下一期決策改變之間因果關係，後續會在結果強調。

⁷ 採用時間從 8/23 開始，在台灣 8/23 是 Medigen 開始全面施打，因此才有足夠的控制組。

表 1 衛福部與自訂施打期數

	衛福部明定期數	施打日期	天數
第 1 期	06 期(含加開)	8/23 - 8/29	7
第 2 期	-	8/30 - 9/2	4
第 3 期	07 期(含加開)	9/3 - 9/10	8
第 4 期	-	9/11 - 9/14	4
第 5 期	08 期	9/15 - 9/23	9
第 6 期	09 期	9/24 - 10/2	9
第 7 期	-	10/3 - 10/5	3
第 8 期	10 期	10/6 - 10/14	9
第 9 期	11 期	10/15 - 10/21	7
第 10 期	-	10/22 - 10/28	7
第 11 期	12 期	10/29 - 11/3	6
第 12 期	13 期	11/4 - 11/12	9

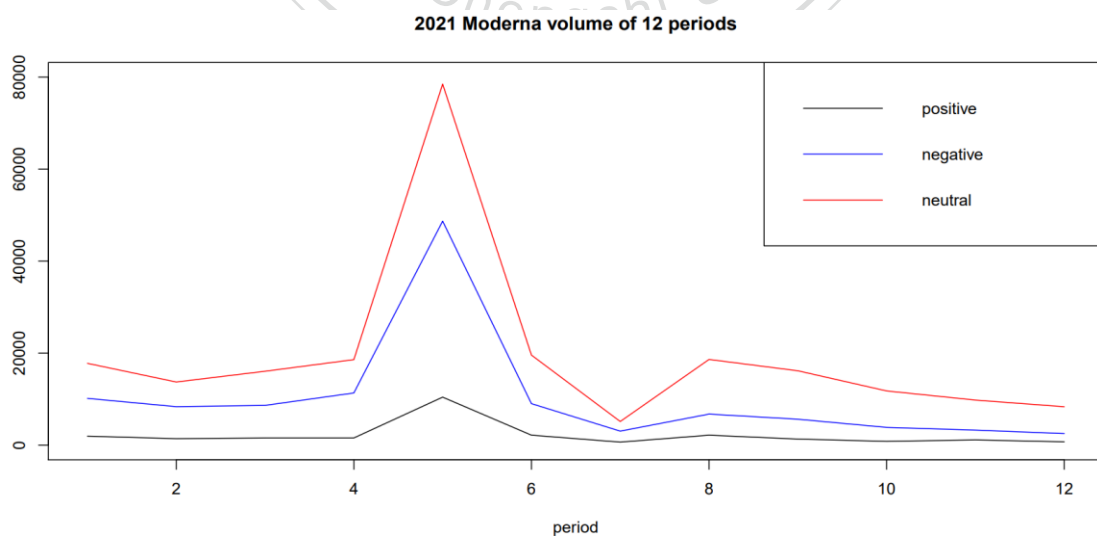


圖 1 Moderna 疫苗施打量 12 期時間趨勢

圖 1 裡正面聲量波動相較於中立、負聲量還來得小，正、負、中立聲量有共同趨勢，意味整個影響 Moderna 施打意願透過負聲量所支配在網路輿情聲量。⁸ 在第 5 期發現三種都一致向上，經查證可能為美國食藥署與歐盟藥品主管機關發布報告為施打 BioNTech 與 Moderna 疫苗跟心肌炎、心包膜炎之間有可能有關係，接下來針對負、正面聲量來建構 Bollinger Bands，透過突破上通道來確認。

使用日資料聲量計算出 Bollinger Bands，在傳統情況下會搭配 Simple Moving Average (簡單移動平均)，本研究為了捕捉網路輿情發生事件後當天波動最劇烈，隨時間經過後會遞減效果。⁹如同先前在 Malmendier and Nagel (2011)提到民眾會對近期發生事件給予更大權重來形成信念相呼應，選擇使用 Exponential Moving Average (指數移動平均)來挑選波動劇烈的期間，其透過突破上通道來挑選介入效果。

$$EMA_t = \alpha x_t + (1 - \alpha)EMA_{t-1} \quad (2)$$

其中 $\alpha = \frac{2}{n+1}$ 、 $EMA_0 = x_0$

在透過(2)可以建構出 Bollinger Bands 上下通道為(3)、(4)。

$$UB_7 = EMA_7 + 1.96\sigma_7 \quad (3)$$

$$LB_7 = EMA_7 - 1.96\sigma_7 \quad (4)$$

為了與研究所自訂期數來對應，最接近就是設 7 天，移動平均設為 7 天，上

⁸ 假設中立聲量在整個 12 期的時間內對民眾施打意願是沒有影響，雖然第 5 期正面聲相較於負面也有較小波動，後續會透過合成控制法 (Synthetic Control Method)來說明，介入效果是由負面聲量所主導。

⁹ 使用日資料的 Moderna 疫苗聲量來建構，若採用加總過後的期資料，無法捕捉到重大事件。

下通道 1.96 標準差，即常態假設下均數為 7 標準差為 1.96，包含了 95%信賴區間。¹⁰

首先負聲量建構 Bollinger Bands，通過上通道認為該期波動較高，透過民眾在網路輿情討論劇烈，即關鍵字搜索 Moderna 疫苗得到負面情緒聲量，在此區間負面聲量內網路輿情上有重大事件所導致，由圖 1 在第 5 期有重大事件，以下負面情緒聲量加以驗證。

圖 2 裡下面呈現出突破上通道，由於負面聲量越大對於民眾施打意願影響才具有改變可能性，可看出第 5 期 (9/15 - 9/23)與第 9 期 (10/15 - 10/21)突破上通道，選擇第 5 期當作介入效果，拋棄第 9 期是因為控制組可能受到除了介入效果以外因素導致控制組被汙染 (contaminate)，第 5 期 (9/15 - 9/23)前有 4 期後有 7 期，選擇此期原因來自於介入前有較長區間可以去做合成控制法 (Synthetic Control Method)合成，而在介入效果第 5 期(9/15 - 9/23)控制組 BioNTech 與 Medigen 剛施打不久，降低被汙染 (contaminate)的疑慮。

接下來透過正面情緒聲量來建構 Bollinger Bands 加以確認，第 5 期是否一樣會突破上通道。圖 3 裡下面可以發現第 5 期一樣會突破上通道，加以驗證之前提到重大事件在這第 5 期波動劇烈，足以作為介入效果，後續分析介入效果針對負面情緒聲量來進行研究。

¹⁰ 由於 $n = 7$ ， $\alpha = 0.25$ ，由於資料負聲量的特性具有一周為循環，以一個禮拜作為基準。

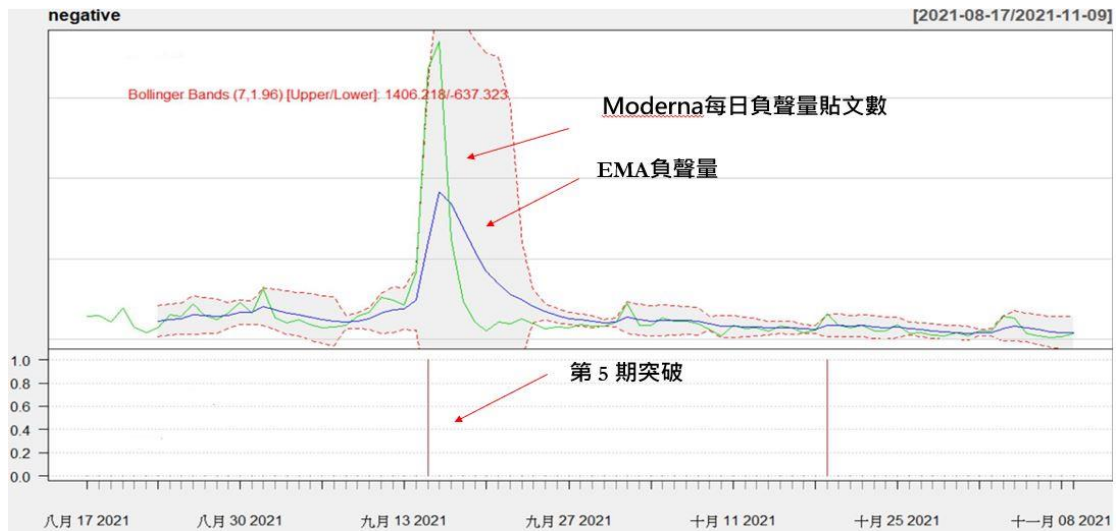


圖 2 負聲量建構 Bollinger Bands

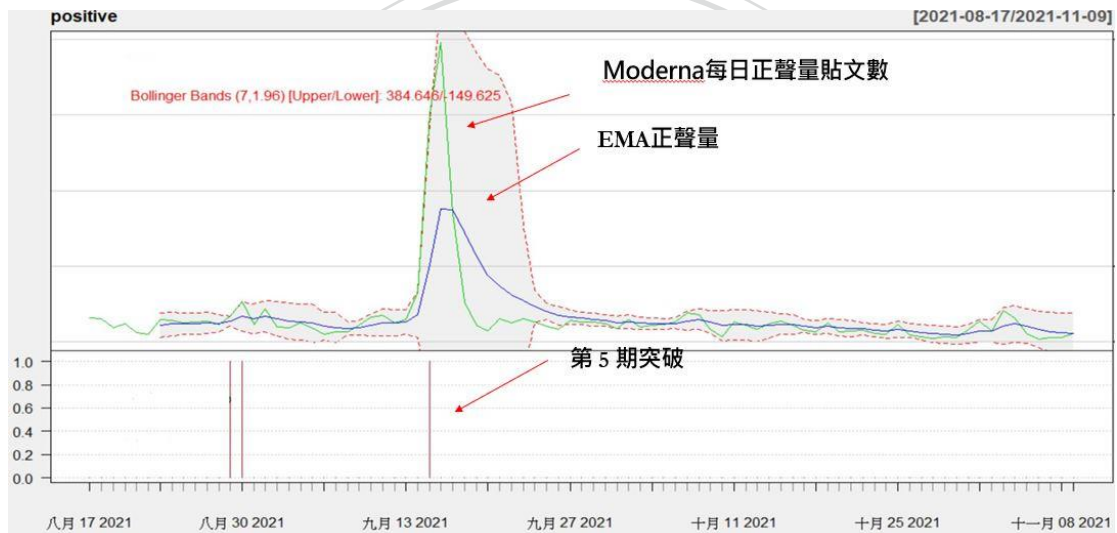


圖 3 正聲量建構 Bollinger Bands

確定介入效果後由先前提到效果後民眾施打疫苗意願需要下一期才會做出改變，下一期施打量影響也有可能來自於更久之負面聲量前期數所影響，接下來需要確認影響下一期施打量是完全來自於上一期介入效果，沒有更早之前效果。

需要透過 Moderna 負面情緒聲量資料跨期之間的相關性來確認，採用 Auto Correlation Function (稱 ACF)、Partial Auto Correlation Function (稱 PACF)。^{11,12}

¹¹ 少部分人並未下一期反應施打量，在第 6 期 (9/24 - 10/2) 施打量影響可能殘有之前不是受到第 5 期 (9/15 - 9/23) 影響，並在下一期未去反應施打量，而是到第 6 期 (9/24 - 10/2) 才去施打疫苗，這施打量影響在後面合成控制法 (Synthetic Control Method) 合成不會影響結果，只會影響介入效果大小。

¹² ACF 與 PACF 是用來確定負面聲量遞延效果導致影響後續的民眾施打意願，與當期負面聲量影響民眾下一期是否施打無關。

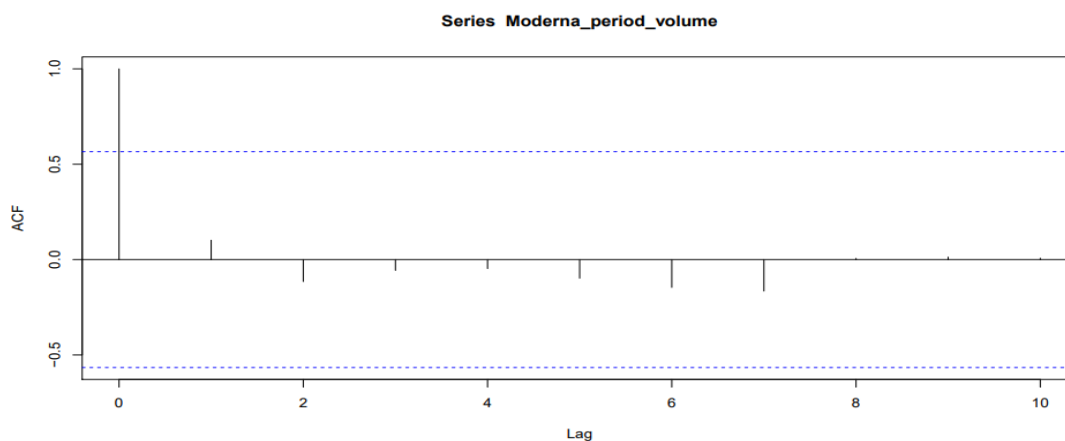


圖 4 負聲量 1 - 10 期之間相關性 (ACF)

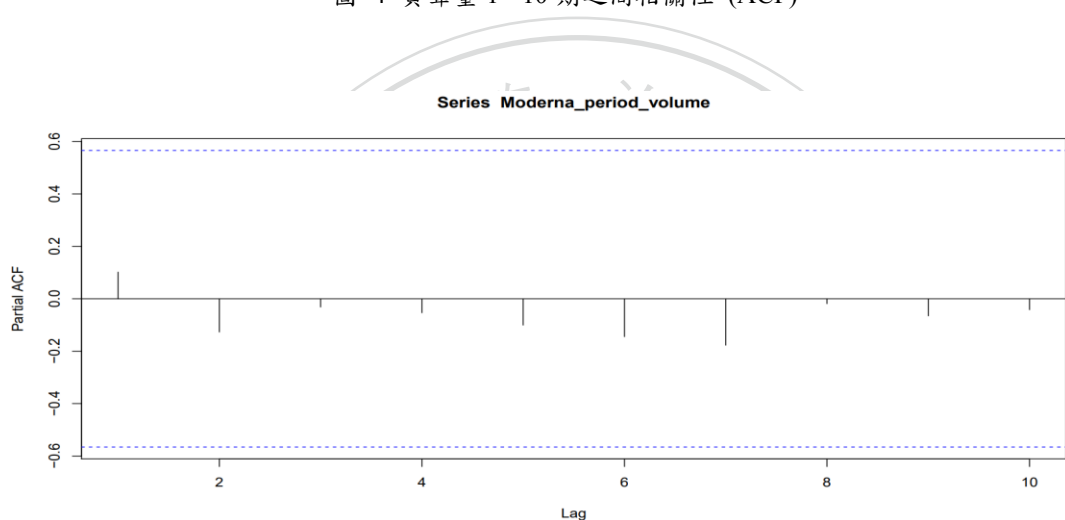


圖 5 負聲量 1 - 10 期之間偏相關性 (PACF)

圖 4 與圖 5 分別為 Moderna 負面情緒聲量的 ACF 與 PACF，可看出皆不超過 95%信賴區間，確認在第 5 期 (9/15 - 9/23)介入效果後得下一期影響 Moderna 施打量完全由第 5 期介入效果所導致，沒有來自比第 5 期更早之前的影響。¹³

¹³ 採用 Moderna 疫苗負面情緒聲量之期資料進行 ACF 與 PACF 分析，因為民眾改變 Moderna 施打意願並做出改變是透過期資料來進行。

第四章 研究方法

本章將介紹二部分，第一節為合成控制法 (Synthetic Control Method)模型架構，第二節為本研究在合成控制法 (Synthetic Control Method)下是如何設定變數。

第一節 合成控制法

根據 Abadie, Diamond & Hainmueller (2010)架構：

$$\arg \min \| \mathbf{X}_1 - \mathbf{X}_0 \mathbf{W} \|_V = \sqrt{(\mathbf{X}_1 - \mathbf{X}_0 \mathbf{W})' \mathbf{V} (\mathbf{X}_1 - \mathbf{X}_0 \mathbf{W})} \quad (5)$$

\mathbf{X}_1 是 $(k \times 1)$ 向量由 $k - 1$ 個解釋變數加上 1 個共變項(實驗組未受到介入效果)， \mathbf{X}_0 是 $(k \times J)$ 由 $k - 1$ 個解釋變數加上 1 個共變項(各自控制組未受到介入效果)， J 個控制組數所組成的矩陣， $\mathbf{W} = (w_1, \dots, w_J)'$ 是 $(J \times 1)$ 非負向量且加總為 1， \mathbf{V} 是 $k \times k$ 的對稱且半正定矩陣，用來當作控制組解釋變數權重加權後反映跟實驗組之間差距。

$$\mathbf{X}_1^* = \mathbf{X}_0 \mathbf{W}^* \quad (6)$$

將以上步驟挑選不同 \mathbf{V} 可得不同 \mathbf{W} ，挑選合適 \mathbf{V} 透過 \mathbf{W}^* 跟 \mathbf{X}_0 做凸組合 (convex combination) (6)後得到 \mathbf{X}_1^* 使(5)最小，即求得 $\| \mathbf{X}_1 - \mathbf{X}_0 \mathbf{W}^* \|$ 。¹⁴

¹⁴ 反事實 (counterfactual)與介入前結果 (pre-intervention outcome)之間為非線性關係，採用凸組合內插誤差。而當凸組合在控制組之間找不到權重，使得介入前就合成得不合，會發生差擬合 (poor matching)。

第二節 模型設定

解釋變數 X_1 與 X_0 為平均負聲量 (1-5 期內將各疫苗負面情緒聲量平均)、平均正聲量 (1-5 期內將各疫苗正面情緒聲量平均)、平均中立聲量 (1-5 期內將各疫苗中立情緒聲量平均)、落後 1 期施打量 (各疫苗在第 5 期前 1 期施打量)、落後 2 期施打量 (各疫苗在第 5 期前 2 期施打量)、落後 3 期施打量 (各疫苗在第 5 期前 3 期施打量)、落後 4 期施打量 (各疫苗在第 5 期前 4 期施打量)、落後 5 期施打量 (各疫苗在第 5 期前 5 期施打量)。控制組為 BioNTech、Moderna，實驗組為 Moderna，結果變數 (outcome variable) 為 Moderna 施打量，時間為 1 - 12 期，介入效果為第 5 期。¹⁵

以上使用解釋變數挑選原則是根據 Fetzer, Hensel, Hermle & Roth (2020) 透過 Google Trends 使用關鍵字搜尋建立於 0 - 100 數值來衡量搜尋強度，而在此研究裡認為關鍵字對於經濟焦慮是有關係。而本論文採用此概念，透過 OpView 資料庫提供關鍵字搜尋，使用情緒聲量來當作解釋變數加以控制除了介入效果以外的因素。¹⁶ 在落後各期的疫苗施打量是根據 Abadie, Diamond & Hainmueller (2010) 在此研究放入落後期數的每人菸草銷售量，本論文參考此方式放入施打量的落後期數，而正、負、中立聲量採取 1 - 5 期平均同樣根據於此。

¹⁵ 在 8/23 之前 AstraZeneca 就已經開始施打，在介入效果後 AstraZeneca 會受到影響，不能將 AstraZeneca 放入控制組。

¹⁶ 台灣與其他國家疫苗爆發時間點不同因此形成相對於國外的平行世界，在本研究使用時間為 2021/08/23 - 2021/11/09 期間，國外獲取疫情有關的資訊像是 Twitter, Bloomberg 等相較於國內網路輿情對民眾施打意願影響較小，另外 OpView 資料庫幾乎包含台灣用戶會使用網路平台，捕捉了大部分傳遞 COVID-19 資訊。

第五章 模型結果

本章一共有二節，第一節部分為合成控制法 (Synthetic Control Method) 合成出來結果呈現，第二節為安慰劑檢定 (placebo test)，檢定合成控制組 Moderna 疫苗施打量與真實 Moderna 疫苗施打量之間介入效果是否統計顯著。

以下呈現期數與表 1 總期數 12 期不同，因為網路輿情是透過聲量來衡量，而本研究之前假設網路對民眾影響大部分會在下一期反應，則當介入效果為第 5 期 (9/15 - 9/23)，關注於下一期第 6 期 (9/24 - 10/2) 變化。

第一節 合成結果

透過合成出來結果得到合成數值以及所使用的控制組使權重為多少，先是由表 2 合成出來數值各為多少，第一行為合成控制法 (Synthetic Control Method) 放入解釋變數作為控制其他可能會影響介入效果的變數，第二行為真實 Moderna 疫苗施打量第三行為透過控制組為 BioNTech、Medigen 施打量合成出來作為 Moderna 疫苗施打量控制組，第四行為將控制組 BioNTech、Medigen 合成出控制組，其權重為相同係數去做合成，即用簡單平均方法去合成出來。

接著呈現出來是在合成控制，控制組之間使用多少權重去合成出 Moderna 疫苗施打量的控制組。表 3 可以發現 BioNTech 權重最大，由 Saleska & Choi (2021) 認為民眾會對疫苗副作用跟政府態度來形成信念，在台灣 BioNTech 跟 Moderna 兩種疫苗沒有太多負面消息，雖然這兩種疫苗都會機率有心肌炎副作用，而相較於 AstraZeneca 與 Medigen 來說，民眾對這兩種疫苗疑慮似乎沒有這麼大，認為可能在台灣，民眾對 BioNTech 跟 Moderna 來說是持有正面信念，後續合成 Moderna 疫苗施打量與真實 Moderna 疫苗施打量之間形況能夠交叉驗證。

表 2 Synthetic Control Method 合成結果

變數	Moderna	合成 Moderna	控制組平均疫苗
平均負聲量	17433.0	21389.241	28287.4
平均正聲量	3371.0	4728.898	6513.7
平均中立聲量	28922.6	37289.446	45288.9
落後 1 期施打量	54028.0	62662.449	326915.0
落後 2 期施打量	34398.0	236.647	1234.0
落後 3 期施打量	27258.0	5457.303	28521.0
落後 4 期施打量	15249.0	803.829	4197.5
落後 5 期施打量	80101.0	114270.050	64641.0

表 3 2 種疫苗的權重w

疫苗	BioNTech	Medigen
權重	90.4%	9.6%

接著是由合成 Moderna 施打量使用解釋變數的比例呈現。表 4 第一行為合成控制放入的解釋變數，第二行為合成所使用解釋變數比例，發現正聲量比例比負聲量比例來的高，而與表 3 的權重所表示 BioNTech 佔合成 Moderna 施打量權重很高，可能來自於有相同正面信念提供資訊是相近，代表在此是由聲量來加以驗證。

表 4 解釋變數權重

變數	權重
負聲量	14.7%
正聲量	18.5%
中立聲量	23%
落後 1 期施打量	7.4%
落後 2 期施打量	13.8%
落後 3 期施打量	1.7%
落後 4 期施打量	14.7%
落後 5 期施打量	6.2%

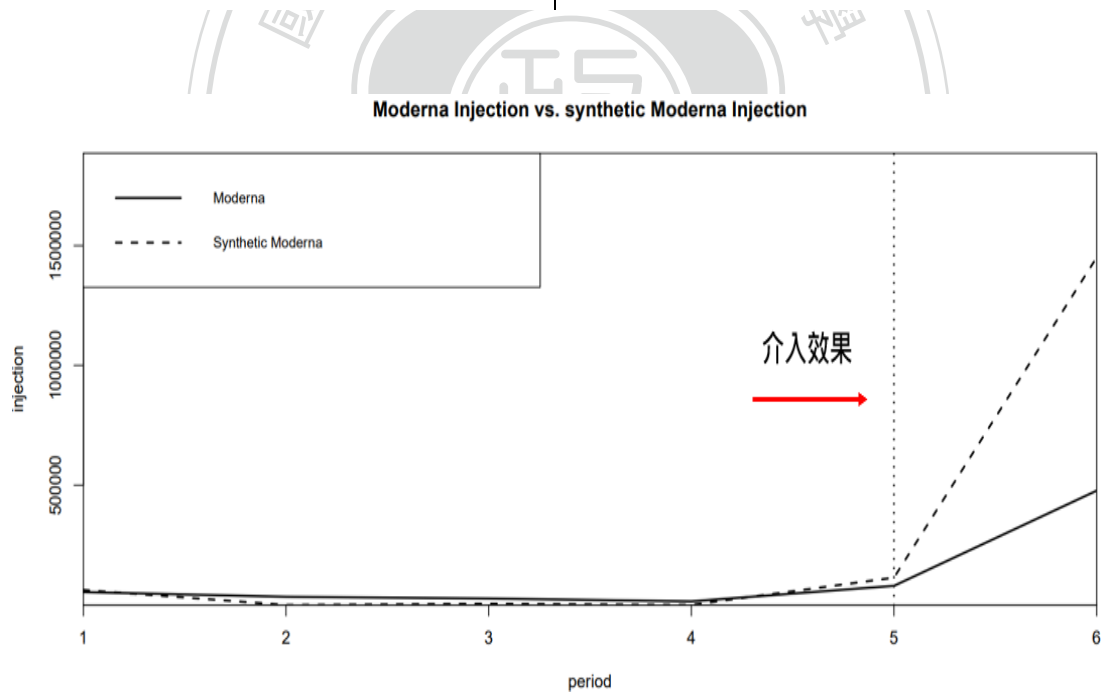


圖 6 施打量時間趨勢： Moderna vs. Synthetic Moderna

接著是將合成 Moderna 施打量與真實施打量作圖在一起，並將介入效果呈現。圖 6 可以驗證民眾對 Moderna 疫苗信念是正面，因為當介入效果在第 5 期

發生後，不論是 Moderna vs. Synthetic Moderna 都還是上升，意味著介入效果透過負面聲量所挑選出來下，而施打量在介入效果影響下，下一期還是會上升，但 Moderna 施打量上升沒有 Synthetic Moderna 還多，因為圖 1 提及介入效果由負面聲量支配所導致。

之後更進一步看合成出 Moderna 施打量與真實 Moderna 施打量之間差距，若在第 5 期 (9/15 - 9/23) 介入效果之前差距越小，即控制住除了介入效果以外的因素影響施打量，則介入效果後的下一期差距具有統計顯著，有證據證明介入效果與 Moderna 施打量之間有因果關係。

之前提到當民眾 Moderna 疫苗施打是正面，隨著時間增加會持續增加，當負面情緒聲量引起的介入效果使第 5 期 (9/15 - 9/23) 的下一期第 6 期 (9/24 - 10/2) 真實 Moderna 施打量與合成 Moderna 施打量下降，即便有可能在第 5 期受到正面聲量影響施打量，由圖 1 正、負聲量在第 5 期都有較大波動，因為民眾對於 Moderna 信念是正面影響會包含正面聲量所引起的影響，最後若介入效果能夠具統計顯著就足以說明有因果關係。

接下來將合成 Moderna 施打量與真實 Moderna 施打量差距作圖與介入效果呈現。圖 7 是真實 Moderna 施打量與合成 Moderna 施打量之間差距，在第 5 期介入前幾乎與原本 Moderna 施打量沒有區別，在介入效果後差距變成負，即真實 Moderna 施打量小於合成 Moderna 施打量，接下來可以進一步去做安慰劑檢定 (placebo test) 去驗證介入效果後的下一期第 6 期 (9/24 - 10/2)，由合成出 Moderna 施打量與真實 Moderna 施打量在第 6 期差距夠大，有證據支持之間的因果相關。

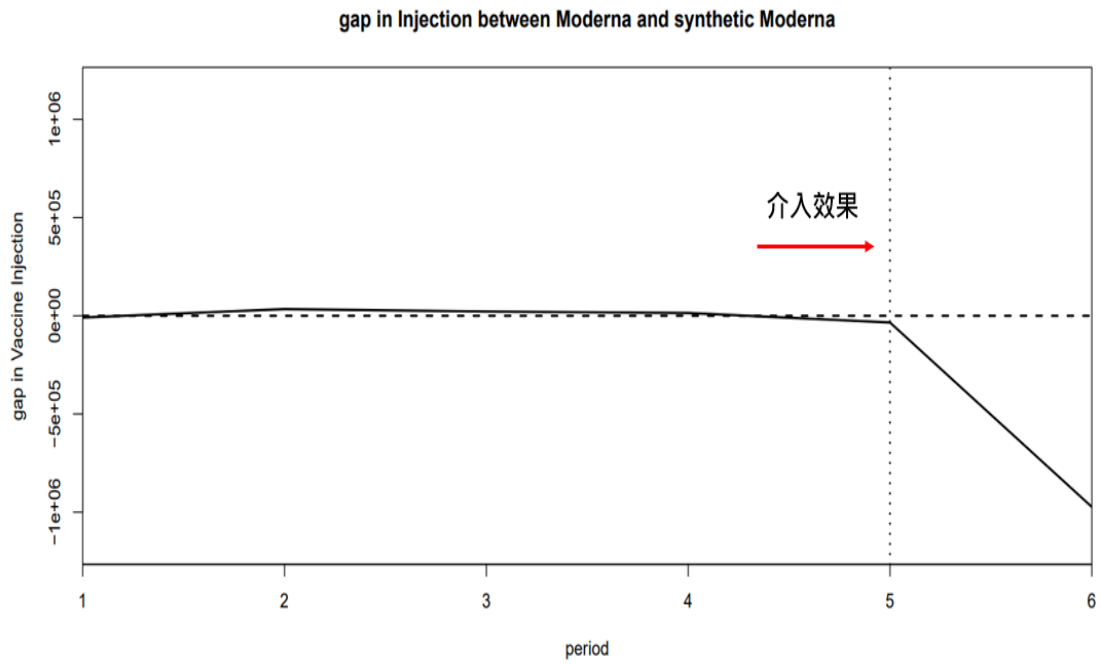


圖 7 Moderna vs. Synthetic Moderna 之間差距

第二節 安慰劑檢定

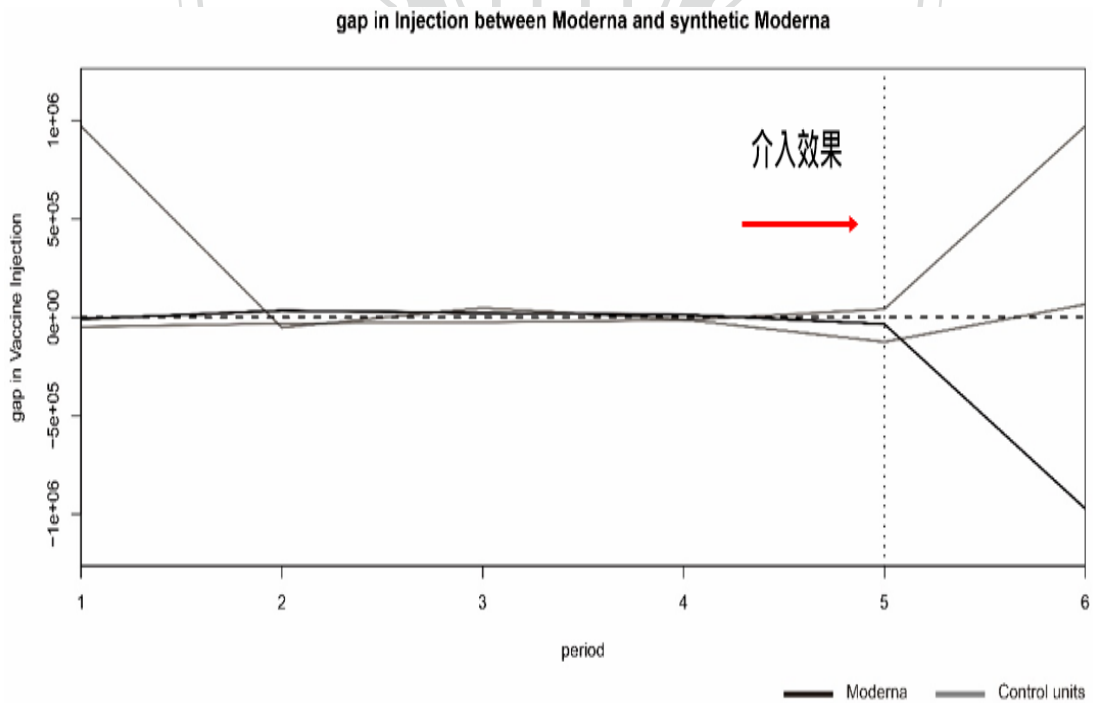


圖 8 合成 Moderna 施打量差距和安慰劑在不同控制組下

安慰劑檢定 (placebo test)原理是在介入效果固定為第 5 期 (9/15 - 9/23)下，透過隨機從控制組取一個來當新的實驗組，而剩下的當作控制組，重新使用合成控制法合成，再將真實施打量與合成施打量之間差距作圖，而當介入前若組合成控制法所得到合成施打量之差距都接近零，代表每次合成都控制除了介入效果以外的因素。當介入效果後，唯有 Moderna 當成實驗組所合成出來的合成施打量與真實 Moderna 施打量差距為負，而負效果是另外兩條合成施打量與真實施打量差距之最下側，而介入效果是在此架構下最顯著，圖 8 為不同疫苗當成實驗組所得到真實施打量與合成施打量差距，在最上面那一條是實驗組為 BioNTech，中間那條是實驗組為 Medigen，看出在介入效果是第 5 期後下一期 Moderna 施打量與其他不是 Moderna 實驗組比較下差距都是負最大，意味著機率 1/3 出現負最大，概念為 p-value 檢定法一樣，在此架構下是最顯著。

第六章 結論

本章討論有二節，第一節討論本研究在網路輿情與疫苗施打量上的結果與貢獻，第二節是在該研究上限制，資料上以及模型上的限制，提供未來研究者後續研究方向。

第一節 主要研究結果與貢獻

本研究針對這次 COVID-19 席捲全世界，民眾對於政府不信任跟疫苗製作流程過於快速，試著找出在民眾施打疫苗意願與網路輿情之間因果關係。分析結果後發現網路輿情負面聲量與 Moderna 疫苗施打量下降之間，有證據支持有因果關係，但未去探討其他疫苗是否也依舊可成立。根據 Loomba & Figueiredo et al. (2021)網路平台上關係複雜，難以建構出系統性模型去捕捉網路輿情與施打疫苗之間因果關係。在資料處理上透過台灣疫苗施打是需要分期預約制特殊性，加上台灣在疫情期間與其他國家處於平行世界，全球疫情爆發時台灣只有零星境外移入，可以有效排除台灣社交平台上雜訊，以及網路輿情傳播對民眾對於疫苗施打形成信念反應在施打量，基於以上特性，本研究透過合成控制法 (Synthetic Control Method)驗證在第 5 期 (9/15 - 9/23)作為介入效果，真實 Moderna 施打量與介入效果之間有證據支持具有因果關係。

第二節 研究限制與未來研究方向

本研究未來針對其他疫苗與網路輿情之間是否也具有因果關係以及在其他國家下是否也具有因果關係，另外可以針對民眾在網路輿情上對於其他事件是否會因事件性質不同形成信念有所不一樣。

一、 關於資料限制

本研究在搜尋關鍵字，並未排除只有 Moderna 有關的疫苗資訊，即沒有排除查詢該疫苗只有該疫苗的資訊，會導致挑選介入效果不夠純 (pure)，會使得在判斷因果之間關係產生疑慮。

本研究遇到控制組會受到汙染問題，必須減少一組對照組，在台灣只開放四種疫苗讓民眾去施打，控制組能夠選擇數量被限制下，會讓合成控制法 (Synthetic Control Method) 在合成時候難以合成相當好。之後在探討相關因果議題需要特別注意控制組是否足夠，以免發生資料上限制。

二、 關於模型限制

根據 Abadie, Diamond & Hainmueller (2010) 架構下是會遇到內插偏誤 (interpolation bias)，當反事實 (counterfactual) 與介入前結果 (pre-intervention outcome) 之間為非線性關係。Abadie & L'Hour (2021) 提出 Penalized Synthetic Control 透過處罰項來處理此偏誤。

而當凸組合 (convex combination) 在控制組之間找不到權重，使得介入前就合成得不合，會發生差擬合 (poor matching)，Valero (2015), Doudchenko and Imbens (2017) 提出 Penalized Regressions 解除凸組合限制，使用最小平方法來解決但可能會產生外置 (extra position) 問題，還有其他 Valero (2015) 提出 Demeaned Synthetic Control，在介入效果前先扣除平均後再去合成，Chen (2020) 提出 Model-based Synthetic Control 與 Ben-Michael, Feller & Rothstein (2021) 的 Augmented Synthetic Control 都是解決差擬合 (poor matching) 問題。

參考文獻

一、中文文獻：

陳宜廷 (2019), “臺灣與南韓之經濟成長比較-合成控制法下的反事實分析”, 臺灣經濟預測與政策(中央研究院經濟研究所), 50(1), 1-410.

二、英文文獻：

Abadie, A. & J. Gardeazabal (2003), “The Economic Costs of Conflict: A Case Study of the Basque Country”, *The American Economic Review*, 93, 112 – 132.

Abadie, A., A. Diamond & J. Hainmueller (2010), “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program”, *Journal of the American Statistical Association*, 105:490, 493–505.

Abadie, A., A. Diamond & J. Hainmueller (2015), “Comparative Politics and the Synthetic Control Method”, *American Journal of Political Science*, 59, 495–510.

Abadie, A. & J. L’Hour (2021), “A Penalized Synthetic Control Estimator for Disaggregated Data”, *Journal of the American Statistical Association*, 116:536, 1817-1834.

Abadie, A. (2021), “Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects”, *Journal of Economic Literature*, 59(2), 391-425.

Ben-Michael, E., A. Feller & J. Rothstein (2021), “The Augmented Synthetic Control Method”, *Journal of the American Statistical Association*, 116:536, 1789-1803.

Chen, Y.-T. (2020), “A distributional synthetic control method for policy evaluation”, *Journal of Applied Econometrics*, 35, 505-525.

- Chen, Y.-T. (2022), “Regularization of Synthetic Controls for Policy Evaluation”,
Department of Finance National Taiwan University.
- Doudchenko, N. & G. W. Imbens (2016), “Balancing, Regression, Difference-in-difference and synthetic control methods: A synthesis”, *NBER Working Paper*.
- Ferman, B. & C. Pinto (2021), “Synthetic controls with imperfect pretreatment fit”,
Quantitative Economics, 12, 1197-1221.
- Fetzer, T., L. Hensel, J. Hermle & C. Roth (2020), “Coronavirus Perception and Economic Anxiety”, *Review of Economics and Statistics*, 2021, 103 (5), 968-978.
- Loomba, S. & A. D. Figueiredo et al. (2021), “Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA”, *Nature Human Behaviour*, 5, 337-348.
- Malmendier, U. & S. Nagel (2011), “Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?”, *The Quarterly Journal of Economics*, 126 (1), 373-416.
- Saleska, J. & L. & Choi, K., R. (2021), “A behavioral economics. perspective on the COVID-19 vaccine amid public mistrust”, *TBM*, 11, 821-825.
- Tversky, A. & D. Kahneman (1973), “Availability: A Heuristic for Judging Frequency and Probability”, *Cognitive Psychology*, 5 (2), 207-232.
- Valero, R. (2015), “Synthetic Control Method versus Standard Statistical Techniques: a Comparison for Labor Market Reforms”, *Working paper*, University of Alincante.
- Vergura, S. (2020), “Bollinger Bands Based on Exponential Moving Average for Statistical Monitoring of Multi-Array Photovoltaic Systems”, *Energies*, 13, 3992.

三、中文書籍

伊藤公一朗(王美娟譯)(2018)，《數據分析的力量》，台灣東販。

四、英文書籍

Cunning, S. (2021), “Causal inference: The Mixtape”, Yale University.

Angrist, J. D. & Pischke, J. S. (2009), “Most Harmless Econometrics: An Empiricist’s Companion”, Princeton University.



附錄

表 5 OpView 關鍵字搜尋

主題	疫苗 (AstraZeneca)	疫苗 (Medigen)	疫苗 (Moderna)	疫苗 (BioNTech)
關鍵字	(AZ&死亡) (AZ&血栓) (AZ&副作用) (AZ&長者) (AZ&反應) (AZ&施打率) (AZ&年輕) (AZ&症狀) (AZ&第二劑) (AZ&第一劑) (AZ&保護力) (AZ&發燒) (AZ&混打) (AZ&預約) (AZ&我 OK*你先打)	(高端&EUA) (高端&三期) (高端&股價) (高端&登記) (高端&我 OK*你先打) (高端&副作用) (高端&死亡) (高端&免疫橋接) (高端&第一劑) (高端&第二劑) (高端&不良反應) (高端&出國) (支持國產疫苗) (高端&保護力) (高端&解盲) (高端&球員兼裁判) (高端&反應)	(莫德納&死亡) (莫德納&副作用) (莫德納&長者) (莫德納&反應) (莫德納&施打率) (莫德納&年輕) (莫德納&症狀) (莫德納&第二劑) (莫德納&第一劑) (莫德納&保護力) (莫德納&發燒) (莫德納&混打) (莫德納&預約) (莫德納&我 OK*你先打) (莫德納&心肌炎)	(BNT&死亡) (BNT&副作用) (BNT&長者) (BNT&反應) (BNT&施打率) (BNT&年輕) (BNT&症狀) (BNT&第三劑) (BNT&第二劑) (BNT&第一劑) (BNT&保護力) (BNT&發燒) (BNT&混打) (BNT&預約) (BNT&心肌炎)
來源	全選	全選	全選	全選
時間	2021/01/01- 2021/11/09	2021/01/01- 2021/11/09	2021/01/01- 2021/11/09	2021/01/01- 2021/11/09
資料數	約 150 萬	200 萬	約 90 萬	約 80 萬

表 5 設定 2021/1/1 - 2021/11/09 期間搜尋以上關鍵字，並利用每篇文章標記情緒計算出正、負、中立聲量，即計算正、負、中立文章數，而四種疫苗關鍵字搜尋為 AZ、高端、莫德納、BNT。

