

國立政治大學經濟學系

碩士學位論文

主場優勢對淨勝分之影響：

以 P-League 為例

The Impact of Home-Court Advantage on Point
Differential: Evidence from P-League Official Data

指導教授：黃柏鈞 博士

研究生：吳竹孟 撰

中華民國 113 年 3 月

摘要

本研究採用 P-League 2021-2022 和 2022-2023 賽季之官方網站的比賽記錄和觀眾入場統計探討主場優勢與淨勝分及其他籃球數據的因果關係。本文主要使用迴歸分析以及 Stefan Wager and Susan Athey (2018) 提出之分析架構，透過迴歸分析觀察主場優勢對各項統計數據 (如淨勝分、勝率) 的影響，再使用因果森林方法處理高維度數據，檢驗異質性的處理效應。

首先，迴歸分析中，我們發現球隊在主場比賽，淨勝分平均會提高 3 分以上，勝率平均提高 8% 以上，交互作用的部分，在主場，勇士平均會減少 1.962 次犯規，並提高二分命中率 4.186%，領航猿則增加 1.882 個犯規、二分命中率下降 2.237%，鋼鐵人增加 1.403 個犯規，夢想家二分命中率下降 0.451%。雖然整體上主場會為球隊帶來優勢，但主場優勢會根據不同的球隊而改變，而非全體一致，甚至有些球隊在主場會表現更差。

接下來透過因果森林方法，我們發現觀眾、比賽間隔、球星數量和不同球隊都會影響主場優勢的大小。觀眾越多，主場優勢越大；比賽間隔在 2-5 天之間，主場優勢最大；球星越多，主場優勢越小。勇士和攻城獅主場優勢較大，國王和領航猿主場優勢較小。

綜合分析上述兩種方法的分析結果，我們確立的主場優勢的存在，以及其存在的原因和具體影響。雖然本文分析的是籃球上的主場優勢，但結果可以延伸到其他種球類運動，進而讓球團思考如何經營球隊文化，以及規劃球隊場館的建立，以最大化發揮自身球隊的主場優勢。

關鍵詞：運動經濟學、主場優勢、因果森林、決策樹

Abstract

This study uses the official website's game records and audience attendance statistics of the P-League for the 2021-2022 and 2022-2023 seasons to explore the causal relationship between home advantage and net points as well as other basketball statistics. This paper primarily employs regression analysis and the analytical framework proposed by Stefan Wager and Susan Athey (2018) to observe the impact of home advantage on various statistical data (such as net points, winning percentage) through regression analysis, and then uses the causal forest method to handle high-dimensional data to examine the heterogeneous treatment effects.

In the regression analysis, we found that teams playing at home on average increased their net points by more than 3 points and their winning percentage by more than 8%. In terms of interaction effects, at home, the Warriors on average committed 1.962 fewer fouls and increased their two-point shooting percentage by 4.186%, while the Pilots increased their fouls by 1.882 and decreased their two-point shooting percentage by 2.237%, and the Steelers increased their fouls by 1.403, with the Dreamers experiencing a decrease in their two-point shooting percentage by 0.451%. Although home advantage generally benefits teams, it varies among different teams and is not universally beneficial; indeed, some teams perform worse at home.

Next, using the causal forest method, we discovered that the number of spectators, the interval between games, the number of star players, and the specific teams all affect the magnitude of home advantage. The more//

spectators, the greater the home advantage; the greatest home advantage is observed when the interval between games is 2-5 days; the more star players a team has, the smaller the home advantage. The Warriors and Lionneers enjoy greater home advantage, while the Kings and Pilots have less.

By synthesizing the results of both methods, we confirm the existence of home advantage, its causes, and its specific impacts. Although this analysis focuses on basketball, the results can be extended to other types of team sports, thereby prompting teams to consider how to cultivate team culture and plan the establishment of their venues to maximize their home advantage.

Keywords : : Sport Economic, Home Advantage, Causal Forest, Decision Tree

目次

1	緒論	1
2	文獻回顧	4
3	主場優勢的可能來源	8
4	實證模型	10
4.1	邏輯迴歸-檢驗主場獨立性	10
4.2	利用迴歸分析來評估主場優勢	11
4.3	因果森林	14
4.3.1	利用無偏性估計處理效果	15
4.3.2	從迴歸樹到因果樹與森林	16
4.3.3	誠實樹和森林	17
4.3.4	SHAP 原理	20
5	實證資料	22
6	實證結果	24
6.1	分析主場獨立性	24
6.2	分析主場優勢對淨勝分的影響	25
6.3	分析主場優勢對其他因素的影響	26
6.3.1	分析主場優勢對獲勝機率的影響	26
6.3.2	分析主場優勢對罰球命中率的影響	27
6.3.3	分析主場優勢對犯規數以及罰球數的影響	28

6.3.4	分析主場優勢對失誤、三分球命中率以及二分球命中率的影響	29
6.4	異質性分析-因果森林	29
6.4.1	決策樹分析結果	30
6.4.2	透過 SHAP 觀察主場優勢之異質性	32
6.4.3	因果森林方法之結論	33
7	結論與建議	35
	參考文獻	54

圖次

1	球星數量對晉級冠軍賽或季後賽機率的影響	38
2	因果森林決策樹	38
3	各特徵之 SHAP value	39
4	觀眾人數之 SHAP value	40
5	比賽間格之 SHAP value	41

表 次

1	Variables List	42
2	descriptive statistics	43
3	主場獨立性迴歸結果	44
4	淨勝分迴歸結果	45
5	勝負迴歸結果	46
6	罰球命中率迴歸結果	47
7	犯規數迴歸結果	48
8	罰球數迴歸結果	49
9	失誤數迴歸結果	50
10	二分球命中率迴歸結果	51
11	三分球命中率迴歸結果	52
12	Predicted Conditional Average Treatment Effect-Home court . . .	53

1 緒論

籃球運動，作為全球廣泛受歡迎的體育項目之一，其比賽結果受到多種因素影響。在這些因素中，「主場優勢」被廣泛認為對比賽結果有顯著影響。這一現象不僅在籃球中存在，在各種體育項目中都有所體現。主場優勢涉及的範圍包括心理學、社會學、環境因素以及生理學等多個層面。

關於主場優勢的研究已有深入探討。研究通常集中在分析主場球隊的勝率、得分差距、球員表現等指標，並將其與客場比賽進行比較。這些研究揭示了主場優勢的存在，並對其可能的成因提出了多種解釋，包括觀眾支持、場地熟悉度、旅行因素和裁判偏差等。觀眾對主場球隊的支持被認為是主場優勢的重要來源。研究發現，大型室內運動場的密集觀眾群體可透過提供社會支持和增強噪音，從而提高主場球隊的表現。此外，球隊對自家球場的熟悉度也是一個重要因素。研究顯示，球隊在搬遷至新場館前的最後一個賽季往往會有更高的勝率，反映了場地熟悉度對主場優勢的貢獻。旅行因素同樣對主場優勢有顯著影響。遠程旅行可能對客場球隊造成不利影響，包括時差、疲勞和生物鐘的不穩定。研究發現，旅行距離超過 200 英里的球隊在客場比賽時勝率更低。而裁判的決策偏差也是影響主場優勢的一個關鍵因素，研究表明裁判在主場觀眾的壓力下更可能做出對主場球隊有利的判罰。

從實證研究的角度，透過對 NBA 數據的回歸分析，研究人員發現，觀眾人數的增加顯著提高了主場球隊的勝率。具體來說，觀眾人數每增加一個標準差，主場球隊的勝率提高了 2.7%。此外，裁判對主場球隊的偏好也在統計上得到了

證實，這種偏見可能源於對家庭觀眾反應的心理反應。

然而，關於主場優勢的研究並非僅限於 NBA 或其他國際知名聯賽。各地的本土聯賽，如台灣的 P-League，同樣提供了研究這一現象的絕佳案例。P-League 作為台灣的專業籃球聯賽，其比賽的主場優勢效應可能與 NBA 存在差異，考量到文化、觀眾參與度和地理因素的不同。因此，本研究將聚焦於探討 P-League 中的主場優勢如何影響球隊的淨勝分。透過對比主客場的比賽數據提供一個新的視角，用以理解主場優勢在不同文化和聯賽背景下的動態。在本研究中，我們探討了台灣 P-League 2021-2022 和 2022-2023 賽季的主場優勢對淨勝分的影響，運用 OLS 和邏輯迴歸模型檢驗主場獨立性以及主場優勢對各項數據的影響，並使用因果森林方法以及 SHAP 值進行異質性分析。

我們發現主場比賽可以使球隊淨勝分平均增加 3.5 分，提升勝率約 8%。主場優勢對二分之一球命中率以及犯規數有顯著影響，特別是對勇士隊而言。但對於失誤數、罰球數、罰球命中率和三分球命中率無顯著影響。此外，觀眾人數和比賽間隔對主場優勢的大小也有影響，顯示主場優勢的存在是明確的，但其影響因素多樣。總體而言，我們的研究結果支持了社會助長理論在解釋主場優勢現象中的應用，並揭示了主場優勢在不同條件下的異質性。

本研究共包含 7 個章節。第 1 節為緒論，說明研究之時空背景、研究動機與本文章節的架構。第 2 節為文獻回顧，回顧國外對於主場優勢議題研究成果。第 3 節為探討 Moskowitz and Wertheim. (2012) 所提及的主場優勢來源。第 4 節為研究方法與模型，先敘述線性迴歸與邏輯迴歸檢驗架構，接著闡述 Wager and Athey (2018) 提出的因果森林推論過程，說明如何將隨機森林轉換為因果森林。

第 5 節為介紹使用的資料類別與形式。第 6 節為實證結果，我們將透過單一決策樹圖呈現因果關係與分群結果。最後，將於第 7 節作出總結，並提供後續研究的方向與建議。

2 文獻回顧

主場優勢一直是體育領域中一個引人關注的主題，研究人員探討主場優勢如何影響球隊的比賽成績。本文旨在回顧有關主場優勢如何影響比賽淨勝分的相關文獻。透過深入分析既有研究，我們將評估主場優勢對球隊在比賽中的淨勝分方面所產生的影響。

在體育運動中，主場優勢指的是在主場比賽時成功的概率較在客場比賽時高，這一現象在體育文獻中受到了廣泛的關注。我們可以在不同的體育項目中使用不同種類的競爭體系來評估主場優勢的存在：1. 平衡比賽；比賽在平衡的主場和客場賽程下進行 (例如，籃球、足球或棒球例行聯賽)。2. 不平衡比賽；比賽在不平衡的主場和客場賽程下進行 (例如，大滿貫網球賽、大型高爾夫錦標賽或國際足協世界盃)。Courneya and Carron (1992) 將主場優勢定義為「在平衡的主客場賽程下，主場隊在體育比賽中贏得超過 50% 的比賽的一致性結果」。只要比賽是平衡的，隊伍在主場和客場比賽中進行相等次數的比賽，整個比賽的主場優勢可以通過在主場比賽中獲勝的比賽數量來評估，並以已決定比賽的比例表示。Stefani (2008) 發現主場優勢在高強度的比賽中像是季後賽或冠軍賽更為明顯，相較例行賽。事實上，在籃球賽場上，比賽中主客場球隊的各種數據平均上的明顯差異都能用來說明主場優勢的存在，例如：罰球數的差異、籃板上的差異、命中率上的差異。

許多 NBA 主場優勢研究通過研究投籃命中率來分析其效果。Kotecki (2014) 使用基於表現的統計數據報告了顯著的主場優勢，特別是通過比較主場與

客場比賽中的投籃命中率、罰球命中率和得分，他顯示這些數據顯著表明主場優勢幫助球隊表現更好。Cao et al. (2011) 研究了壓力對 NBA 表現的影響。使用罰球作為他們感興趣的衡量標準，他們測試了主場球迷是否能分散對手球員的注意力，使其在罰球時承受壓力。然而，他們發現主場地位對罰球失誤的影響不具有實質性的統計顯著性。Harris and Roebber (2019) 使用兩分球、三分球和罰球作為感興趣的衡量標準來研究主場優勢。他們發現兩分球是主場優勢的最強預測因子。他們建議主場球隊應該嘗試投更多的兩分球，並迫使對手嘗試更多的兩分球投籃。這一策略將最大化主場優勢的好處，並給他們最好的獲勝機會。

在 NBA 因疫情而沒有觀眾的“泡沫”賽季中，由於傳統的主場優勢因素 (如觀眾支持、場地熟悉度和裁判偏差) 被剝離，為研究提供了一個獨特的觀察窗口。在這種情境下，客隊的勝率、罰球命中率以及犯規次數均出現了顯著變化，從而凸顯了主場優勢的影響。根據 Price and Yan (2021) 客場球隊的平均得分確實顯著增加了。這與我們關於主場勝率下降的直覺和結論相一致。如果客場球隊得分顯著增加而主場球隊沒有，那麼我們預計客場球隊將贏得更多的比賽。一個有趣的發現是，主場和客場球隊的所有投籃和得分數據至少都有小幅增加。儘管這些增加並非都具有顯著性，但這些增加正是 Carron et al. (2005) 在解釋缺乏觀眾時證據表明球隊表現更好時所報道的。這很重要，因為它與我們的結論一致，即主場優勢主要通過對客場球隊產生負面影響來影響比賽。如果觀眾導致整體表現下降，那麼主場優勢必須來自客場表現下降大於主場表現下降。這就是為什麼客場球隊能夠在取消主場優勢的情況下縮小與主場球隊的差距。

將主場優勢的影響分為主場效應和客場效應，為一些有趣的新見解提供了

空間。有許多文獻討論造成主場優勢的原因，通常是來自賽程、觀眾、裁判方面，賽程方面，賽程安排可能會造成球隊有不同的表現，例如：連續背靠背讓球隊勝率變低，另外，與比賽相關的旅行可能會導致主場優勢，因為運動員會因疲勞而受到影響，並且會干擾正常的例行程序。已經分析了與旅行相關的因素，包括兩個比賽場地之間的距離 (Clarke and Norman, 1995; Snyder and Purdy, 1985)、跨越的時區數量 (Balmer et al., 2001; Pace and Carron, 1992) 以及時差的影響 (Jehue et al., 1993; Recht et al., 2003)。結果顯示，旅行時間和距離對運動表現的影響較小或無顯著影響，但在沒有旅行涉及的當地的比賽中，主場優勢會減少 (Carron et al., 2005; Nevill and Holder, 1999; Pace and Carron, 1992; Pollard, 2006)。無論如何，儘管運動員普遍相信旅行會對比賽產生影響，但只有一小部分主場優勢的差異可以通過與旅行相關的因素解釋，因此將長距離旅行納入比賽活動中對比賽的干擾只有輕微影響。

觀眾方面，可以分為觀眾對球員的影響以及觀眾對裁判的影響，球員方面，雖然普遍認為觀眾會影響表現，但對於觀眾規模對表現的因果效應知之甚少。結果取決於所做的假設，支持性的觀眾通常被認為會提高期望，從而增加表現壓力，也可能提供情感上的舒適，減輕壓力。Zajonc (1965) 提出社會助長 (social facilitation) 的概念，當個體在他人的存在下，對於簡單或已經熟練的任務表現得更好，但對於複雜或不熟悉的任務則表現得更差。Zajonc (1980) 在社會促進的背景下定義了「單純」這個詞。他認為，社會促進文獻可以更容易地用單純存在（驅動效應）來解釋，而不是用針對目標（如評估擔憂）的特定動機來解釋。但 Cottrell et al. (1968) 指出，所有社會促進的實驗都涉及到既可能與他人的存在

相關的正面和負面結果的任務。也許社會引起的 醒程度是由於某種評估焦慮，而不僅僅是單純存在。Sanna and Shotland (1990) 也證明在觀眾面前執行記憶任務的參與者比單獨執行的參與者表現更好，但只有當他們預期觀眾給予正面評價時。當他們預期負面評價時，情況剛好相反。而 Taylor et al. (2010) 顯示支持性的觀眾會增加生物壓力指標所測量的壓力。這與被廣泛描述的觀眾帶來“主場優勢”的情況 (Boudreaux et al., 2017) 形成對比。當面臨更大的表現壓力時，表現通常會下降 (Dohmen, 2008 ; Harb-Wu and Krumer, 2017)。這種意外的負面結果被稱為“壓力下的失敗”(choking under pressure) (Hill et al., 2010 ; Baumeister, 1984)。裁判方面，當主場裁判吹判主場球員犯規時，會受到主場球迷的噓聲；相反地，吹判客場球員犯規時，會得到球迷的鼓舞和支持，這種球迷造成的場邊噪音會增加比賽的張力和比賽的觀賞性，但同時也讓裁判在這樣的壓力下更容易做出偏誤的吹判，而間接地讓主場球隊受益。有許多人認為裁判偏袒主隊是一種陰謀論，但事實上並非如此，大多數的裁判都是正直、專業的人士，他們會盡全力確保比賽的公平性，但他們無法不受人類心理的影響，裁判也是人，觀眾帶來的社會影響力是不可忽視的，可以在當事人完全沒有意識時影響其行為和決定。裁判在比賽中面對觀眾給予的龐大壓力，自然會做出能減輕自身壓力的判決，特別是一些模稜兩可的回合，判決通常會偏袒主隊，這也是造成主場優勢的重要原因。證據表明觀眾的影響取決於運動的種類以及其他中介因素，例如在主觀評判運動 (如足球) 中的裁判決定。事實上，有幾項研究表明，裁判可能會受到大觀眾的影響，傾向於支持主場隊伍 (Downward and Jones, 2007 ; Nevill et al., 2002 ; Sutter and Kocher, 2004)。顯然，觀眾的大小、支持的強烈程度或接近比賽場地

的程度都能夠影響運動員、教練和裁判的情緒狀態，甚至影響他們的注意力水平，從而影響運動表現，部分解釋了主場優勢現象。

3 主場優勢的可能來源

在考慮主場優勢成因之前，有幾個前提條件：有可觀的經濟利益促成球隊盡量爭取在主場獲勝。若能在主場發揮的越好，會有更多的觀眾願意買票入場，購買球隊的周邊商品。也會有更多廠商贊助球隊、購買廣告。聯盟也為主隊獲勝提供了誘因，對大部分球隊而言，觀眾人數和收入會跟比賽的勝率同步上漲，跟主場勝率同步上漲幅度更大，可以說聯盟希望透過一些舉措，為主場帶來些許合法的優勢。Moskowitz and Wertheim (2012) 提出四種慣常的主場優勢來源，我們將一一說明，並在後面的實證分析中驗證前兩種假說是否為真。

第一個傳統的解釋是主場觀眾的支持，球員在主場比賽時，有歡呼、有掌聲，賽前介紹時甚至還會有煙火，球員發揮更出色似乎非常符合邏輯。另一方面，當球員在客場比賽，罰球時一群觀眾朝球員起鬨，球員似乎會表現更差。檢測觀眾支持的作用存在一個問題：球迷的所有舉動不只影響到球員，也影響到防守球員、防守球員的隊友以及裁判。我們後面會使用罰球來檢測觀眾對球員的影響，因為罰球是球員與球迷之間的單獨互動，可以隔離其他因素。

第二個傳統解釋是對主隊有利的賽程安排，每個賽季，NBA 球隊有 20 多場連續比賽（背靠背），客場平均要打 14 場。光是這樣就足以影響 NBA 的主場優勢了。根據統計，相較於非連續比賽的客場正常勝率，這 14 場比賽的勝率僅

為 36%。這意味著，因為這種連續的賽程安排，每個賽季客隊會多輸 1-2 場比賽。換句話說，光是因為 NBA 的連續賽程安排，主隊基本上占盡優勢。但經過統計，我們發現 P-League 主隊及客隊背靠背的次數總共是 78 次，而主場就佔了 77 次，所以賽程安排優勢是不存在的，甚至可以說是對客場有利，在後面的章節，我們也會用迴歸分析觀察背靠背對比賽的影響。

第三個傳統解釋是客隊要承受舟車勞頓之苦，在 NBA 或其他國外聯賽中，城市之間距離較遠，四處奔波很可能會導致球員體力不支，進而影響比賽。但值得注意的是 NBA 的主場優勢在過去 100 年間是始終一致的，不管是搭乘噴射機或是巴士，球隊表現都一樣，前者勝率並不會比較高。另外，台灣的幅員較小，城市與城市之間交通非常便捷，所以更不會有這方面的影響。

第四個傳統解釋是更適合主隊的天氣，常常可以聽到 NBA 解說員說，來自熱帶地區的球隊，比如邁阿密熱火，到了寒冷地區的球場就會腳步踉蹌，比如克利夫蘭，但研究指出，氣候條件與主場優勢沒什麼關連。如果 NBA 球隊可以利用自己的主場氣候優勢，我們應該能觀察到寒冷地區的球隊贏下大部分寒冷地區的比賽；也能一進到明顯不同的氣候，球隊的表現會變差；最後，我們應該看到在室內比賽的球隊對主場氣候不敏感，因為他們相當於在控制了天氣條件的狀況下比賽。最後一種情形可以是對照組，如果對照組在沒有任何改變的情況下出現了同樣的結果，就代表導致主場優勢的並不是天氣狀況，而是其他東西。

由於後兩項傳統因素對於台灣的 P-League 並不太適用，在接下來的章節中，我們會分析前兩項傳統的主場來源，觀察其影響，並找出其他可能的主場優勢來源，以完整解釋主場優勢的影響。

4 實證模型

本文採用統計學中的迴歸分析來推論主場對於淨勝分的影響，但在這之前，我們需要先檢驗主場獨立性也就是聯盟分配主客場時，不應該受到任何控制變數的影響。並在後續分析更多變數，如失誤、命中率等等，找出主場優勢體現的面向。

4.1 邏輯迴歸-檢驗主場獨立性

如果 P-League 的主場分配不公平，它可能對主場優勢對勝利分差的影響產生各種效應，例如，不公平的主場分配可能導致某些球隊在主場擁有更多場次的熱情球迷。主場球迷的存在可以提高球隊士氣，激勵球員，並為對手創造更具挑戰性的環境。這有助於擁有有利主場分配的球隊實現更高的勝利分差。球迷的熱情和氛圍可以對球隊表現產生積極影響，增強主場優勢的效果。

鑒於上述提到的原因，確定主場分配是否公平對我們非常重要，在實證研究中，我們經常遇到因變數是“是/否”或“事件發生/事件未發生”的情況。在這種情況下，因變數僅具有兩種可能的值，可以編碼為 0 或 1，稱為二元變數或虛擬變數。在這些情況下，使用線性迴歸模型來研究影響因變數的因素是不適當的。相反，需要使用概率模型，其中最常用的是邏輯迴歸模型，這也是本研究選擇的模型。

現在，如果我們將主場變數作為因變數 HOME，而將所有其他變數視為自

變數，我們可以得到要使用的邏輯迴歸模型：

$$P(\text{HOME} = 1|x_i) = \frac{1}{1 + \exp \left[-(\beta_0 + \sum_{k=1}^5 \beta_k D_{Tki} + \beta_6 D_{Bi} + \beta_7 X_{Si} + \beta_8 X_{Vi} + \sum_{k=9}^{33} \beta_k X_{Ri}) \right]} \quad (1)$$

D_{Bi} 代表是否為背靠背比賽的虛擬變數， X_{Si} 代表球星數量， X_{Vi} 代表觀眾人數， X_{Ri} 代表裁判。由於方程式過長，我們將式 1 中的五個球隊虛擬變數表示為 D_{T1} 到 D_{T5} ，邏輯迴歸的係數代表了在主場或客場比賽的概率。我們希望每個係數都不具有統計顯著性，這意味著聯盟對主場和客場比賽的分配是獨立和公平的，不受影響。

4.2 利用迴歸分析來評估主場優勢

我們將使用多元線性迴歸分析主場優勢對得分差的影響。線性關係依賴於兩個變數之間的聯繫，並運用線性方程的原則來估計這種聯繫的強度。迴歸分析是一種統計技術，利用變數之間的關係來進行解釋和預測。當假設存在線性關係時，迴歸分析使用線性方程進行統計決策和應用，通常被稱為線性迴歸。

在建立以得分差作為應變數的模型時，最重要的控制變數是 HOME，使我們能夠比較主場和客場比賽的淨勝分差異。接下來，在選擇要納入模型的其他控制變數時，我們應該選擇與非主場優勢影響的結果相關的因素，同時也會影響得分差，例如明星球員的數量、球隊實力、觀眾入場人數等等。

為了觀察不同控制變數對淨得分的影響，我們將分析多個模型。每個模型將

包括一個主場虛擬變數和其他控制變數，如明星球員的數量、觀眾入場人數等等。最後，在控制了之前的變數的情況下，將包括主場虛擬變數和球隊虛擬變數之間的交互項，以確定不同球隊之間的主場優勢是否存在差異。為了簡化模型，我們將除了 HOME 之外的控制變數表示為 X 。

$$\text{POINT}_i = \beta_0 + \beta_1 \text{HOME}_i + \beta_2 X_i + \epsilon_i \quad (2)$$

特別注意到模型 2。在 2021-2022 和 2022-2023 賽季的 P-League 中，共有 6 支球隊：勇士 (Braves)、國王 (Kings)、領航猿 (Pilots)、攻城獅 (Lioneers)、夢想家 (Dreamers) 和鋼鐵人 (Steelers)。通常在引入虛擬變數時，我們要求如果有“ m ”個定性變數，模型中應該包括“ $m-1$ ”個虛擬變數。否則，引入所有“ m ”個虛擬變數將導致解釋變數之間的完全多重共線性，通常稱為虛擬變數陷阱。因此，我們以 BRAVES 作為參考，並在模型中包括其他球隊 (KINGS、PILOTS、LIONEERS、DREAMERS、STEELERS) 來分析球隊的實力。我們將逐一控制不同因素來檢查得分的變化，最後我們將控制所有變數並使用球隊與主場的交互項來分析交互作用，觀察不同球隊的主場優勢。

我們也會使用其他因素作為應變數，為了簡化模型，我們將除了 HOME 之外的控制變數表示為 X 。首先，我們想了解主場優勢對勝率的影響，使用 WIN 作為被解釋變數是為了彌補淨勝分回歸的缺點，淨勝分迴歸的缺點是每單位分差對比賽的影響力不同。例如，假設主場能讓淨勝分提高 3 分，有兩場球賽，一場球賽的比數是 99:100(99 是客隊)，另一場的比數是 99:130(99 是客隊)，若現在

主客互換，第一場比賽的比數會變成 102:99，第二場比賽的比數變為 102:130，都是一樣 3 單位的變化，但第一場比賽勝負顛倒，第二場比賽則無關勝負。因此，這節我們將使用邏輯迴歸模型，來分析控制變數對勝負的影響。

因此，我們將建立一個以虛擬變量 WIN（1 代表贏球，0 代表輸球）作為應變數的邏輯迴歸模型：

$$P(\text{WIN}_i = 1|x_i) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 \text{HOME}_i + \beta_2 X_i)]} \quad (3)$$

正如前面提到的，觀眾規模可能會影響球員和裁判的表現。觀眾可以影響比賽的許多方面，例如提高主隊的士氣，使球員更加自信，鼓勵更好的防守。然而，我們無法將與這些方面相關的群體效應從同時存在的其他因素分開來進行測量。我們可以通過比賽中與所有其他因素分開的部分來孤立這些效應，比如罰球。罰球涉及球員與可能試圖分散或干擾他們的觀眾成員之間的個人互動，而且所有罰球都是標準化的，因為球場尺寸一致，罰球線到籃筐的距離在所有競技場中都相同。

所以，如果我們想知道透過觀眾和裁判，主場優勢會不會影響球員的罰球命中率，可以使用罰球命中率作為被解釋變數，建立一個多元迴歸模型。

$$\text{FT}\%_i = \beta_0 + \beta_1 \text{HOME}_i + \beta_2 X_i + \epsilon_i \quad (4)$$

同樣地，如果我們想要確定裁判是否偏袒主場球隊，我們可以使用被吹罰的犯規數量 FOULS 作為因變數：

$$\text{FOULS}_i = \beta_0 + \beta_1 \text{HOME}_i + \beta_2 X_i + \epsilon_i \quad (5)$$

同理，我們也會以失誤、罰球數、2PT%、3PT% 作為應變數，來觀察主場優勢造成的影響是否體現在這些面向，在第 6 節，我們會針對觀眾及裁判效應，結合社會心理學的概念作分析：

$$\text{TO}_i = \beta_0 + \beta_1 \text{HOME}_i + \beta_2 X_i + \epsilon_i \quad (6)$$

$$2\text{PT}\%_i = \beta_0 + \beta_1 \text{HOME}_i + \beta_2 X_i + \epsilon_i \quad (7)$$

$$3\text{PT}\%_i = \beta_0 + \beta_1 \text{HOME}_i + \beta_2 X_i + \epsilon_i \quad (8)$$

4.3 因果森林

因果森林是一種基於機器學習的因果推斷方法，它能夠處理高維度的數據，並估計異質性治療效果。在我們的研究中，主場優勢被視為一種“處理” (treatment)，而比賽的淨勝分是衡量效果的指標。我們的目標是估計不同球隊 (例如“勇士”、“國王”等)、觀眾人數和球星數量... 等等，對主場優勢對淨勝分的影響。前面章節所使用的迴歸分析呈現的大部分是平均效應，我們有加入球隊與主場的交互項來檢驗異質性的處理效應。我們會根據 Wager and Athey (2018) 提出之架構，應用因果森林方法對異質性做更完整的分析。

4.3.1 利用無偏性估計處理效果

“因果森林” (causal forest) 是一種用於估計異質性處理效應 (heterogeneous treatment effects) 的統計學方法。它結合了隨機森林 (random forests) 的非參數特性與因果推論的原理。下面我們將詳細說明因果森林的推導過程與邏輯：

對於每個個體，我們考慮在處理與未處理兩種情況下的潛在結果，分別記為 $Y_i^{(1)}$ 和 $Y_i^{(0)}$ 。式 (9) 定義了個體層面的處理效果 (treatment effect)，即對於特定特徵 x 的個體，其處理效果為：

$$\tau(x) = \mathbb{E}[Y_i^{(1)} - Y_i^{(0)} | X_i = x] \quad (9)$$

但估計函數 $\tau(x)$ 的主要困難在於我們只能觀察到兩個潛在結果 $Y_i^{(1)}$ 和 $Y_i^{(0)}$ 中的一個，因此作者提出一個關鍵的無偏性假設 (unconfoundedness)：

$$\{Y_i^{(0)}, Y_i^{(1)}\} \perp\!\!\!\perp W_i | X_i. \quad (10)$$

式 (10) 敘述了無偏性假設，它指出，在給定觀察特徵 X_i 的條件下，處理分配 W_i 與潛在結果 $Y_i^{(1)}$ 和 $Y_i^{(0)}$ 是獨立的。這個假設的背後動機是，假定特徵空間中鄰近的觀察點可以視為來自隨機實驗，從而使得基於鄰近觀察點的匹配和其他局部方法對於估計 $\tau(x)$ 是一致的。

式 (11) 是應用無偏性假設的結果，它提供了一種估計處理效果 $\tau(x)$ 的方法，這個表達式的核心思想是使用傾向得分來平衡處理組和對照組，以估計處理效果。

$$\mathbb{E} \left[Y_i \left(\frac{W_i}{e(x)} - \frac{1 - W_i}{1 - e(x)} \right) \mid X_i = x \right] = \tau(x), \quad \text{where } e(x) = \mathbb{E}[W_i \mid X_i = x] \quad (11)$$

4.3.2 從回歸樹到因果樹與森林

在高層次上，樹和森林可以被視為一種具有適應性鄰域度量的最近鄰方法。傳統的最近鄰方法 (例如 k-最近鄰) 會尋找與測試點 x 最接近的點，而樹基方法則在決策樹的結構下定義接近性，最接近 x 的點是落在相同葉子 (leaf) 中的點。

對於給定的測試點 x ，我們通過識別包含 x 的葉子 $L(x)$ 並設置 $\hat{\mu}(x)$ 來評估預測：

$$\hat{\mu}(x) = \frac{1}{|\{i : X_i \in L(x)\}|} \sum_{\{i : X_i \in L(x)\}} Y_i \quad (12)$$

從啟發性的角度來看，如果我們認為葉子 $L(x)$ 足夠小，以至於葉子內的反應 Y_i 大致相同分佈，那麼這種策略是有道理的。

在這樣的背景下，我們同樣希望將葉子視為足夠小，以至於與索引 i 對應的 (Y_i, W_i) 對，其中 $i \in L(x)$ ，就好像它們來自一個隨機實驗一樣。然後，自然而然地，我們可以估計在任何 $x \in L$ 中的處理效應，如下所示：

$$\hat{\tau}(x) = \frac{1}{|\{i : W_i = 1, X_i \in L\}|} \sum_{\{i : W_i = 1, X_i \in L\}} Y_i - \frac{1}{|\{i : W_i = 0, X_i \in L\}|} \sum_{\{i : W_i = 0, X_i \in L\}} Y_i \quad (13)$$

$\hat{\tau}(x)$ 是對於 x 的治療效果估計，計算方式是找到包含 x 的葉子 L ，然後計算該葉子中接受處理 ($W_i = 1$) 和未接受處理 ($W_i = 0$) 樣本的 Y_i 值差異的平均。

總結來說，本節介紹了如何從回歸樹過渡到因果樹，並說明了在因果樹中如何使用特定的方法來估計處理效果。這些方法不僅在技術上與回歸樹相似，而且在因果推論的框架下提供了有效的處理效果估計。

4.3.3 誠實樹和森林

這一節強調了因果森林結果的靈活性和適應性。對於各種因果森林，只要子樣本大小 s 以適當的速率增長，就可以達到一致性 (consistency) 和集中的漸近正態性 (centered asymptotic normality) (式 (14)、(15))。然而，這些結果要求個別樹滿足一個相當嚴格的條件，稱為「誠實」(honest)。一棵樹被認為是誠實的，如果對於每個訓練實例 i ，它只使用 Y_i 來估計葉子內的處理效果 τ (式 (13)) 或決定分割點的位置，但不是兩者都用，這種設計確保了估計的無偏性和一致性。以下我們將討論滿足這個條件的兩種因果森林算法。

$$\frac{\hat{\tau}(x) - \tau(x)}{\sqrt{\text{Var}[\hat{\tau}(x)]}} \Rightarrow N(0, 1) \quad (14)$$

$$\hat{V}_{IJ}(x) = \frac{n-1}{n-s} \left(\frac{n}{n-s} \right)^2 \sum_{i=1}^n \text{Cov}_*[\hat{\tau}_b^*(x), N_{ib}^*]^2 \quad (15)$$

第一種算法稱為雙樣本樹 (double-sample trees)，通過將訓練子樣本分為兩半 (I 和 J) 來實現誠實性。使用 J 樣本來決定分割點，同時保留 I 樣本進行葉子內的估計。這種設計允許樹在結構劃分和處理效果估計上保持獨立，從而減少偏差。

算法 1. Double-Sample Trees

雙樣本樹將可用的訓練數據分為兩部分：一半用於估計每個葉子內的期望反應，另一半用於放置分割點。

Input:

- n 個訓練實例，形式為 (X_i, Y_i) 用於回歸樹，或 (X_i, Y_i, W_i) 用於因果樹，其中 X_i 為特徵, Y_i 為反應, W_i 為處理分配。
- 葉子最小樣本數 k 。

Procedure:

1. 從 $\{1, \dots, n\}$ 中抽取一個大小為 s 的隨機子樣本 (抽出不放回)，並將其分成兩個大小分別為 $|I| = \lfloor s/2 \rfloor$ 和 $|J| = \lceil s/2 \rceil$ 的不相交集合。
2. 通過遞迴分割生長一棵樹：使用 J 樣本中的任何數據來選擇分割點，以及來自 I 樣本的 X - 或 W -觀測值，但不使用來自 I 樣本的 Y -觀測值。
3. 僅使用 I 樣本觀測值來估計葉子內的反應。

雙樣本回歸樹：使用式 (12) 對包含 x 的葉子進行預測，僅使用 I 樣本觀測值。分割標準是 CART 回歸樹的標準（最小化預測的均方誤差）。限制分割，使得樹的每個葉子必須包含 k 個或更多 I 樣本觀測值。

雙樣本因果樹：與雙樣本回歸樹的定義類似，但預測時我們使用式 (13) 在 I 樣本上估計 $\hat{\tau}(x)$ 。依據 Athey 和 Imbens (2016) 的研究，樹的分割點選擇是通過最大化 J 集中所有 i 的 $\hat{\tau}(x)$ 的變異數來實現的。此外，樹的每個葉子必須包含 k 個或更多每種處理類別的 I 樣本觀測值。

第二種算法稱為傾向樹 (propensity trees)，這些樹使用傾向得分 (propensity scores) 來平衡處理和對照組，從而減少由於特徵分佈不均勻引起的偏差。

算法 2. Propensity Trees

傾向樹僅使用處理分配指示器 W_i 來放置分割點，並保留反應 Y_i 用於估計。

Input:

- n 個訓練實例，形式為 (X_i, Y_i, W_i) ，其中 X_i 為特徵, Y_i 為反應, W_i 為處理分配。
- 葉子最小樣本數 k 。

Procedure:

1. 從 $\{1, \dots, n\}$ 中抽取一個大小為 $|I| = s$ (不放回) 的隨機子樣本 I 。
2. 使用 I 樣本訓練一棵分類樹，其中結果是處理分配，即對於 $i \in I$ 的 (X_i, W_i) 數對，樹的每個葉子必須包含 k 個或更多每種治療類別的觀測值。

3. 使用式 (13) 在包含 x 的葉子上估計 $\tau(x)$

在第 2 步中，分割點的選擇是通過優化某個標準來實現的，例如，CART 用於分類的基尼標準 *Breiman et al. 1984*。

總結來說，誠實樹隨機化 I 、 J 數據分割，可以在不浪費任何數據的情況下實現誠實性。此外，每個數據點都將參與一些樹的 I 和 J 樣本中，因此將被用於指定森林的結構和處理效果估計。這種方法實踐中可以改善均方誤差，相比於標準隨機森林表現更好。

雖然雙樣本樹 (double sample trees) 和傾向樹 (propensity trees) 都是基於樹的方法，但它們的應用和目的不同。雙樣本樹著重於估計異質性效應，而傾向樹著重於改進傾向分數的估計，以進行更準確的因果推斷。在本研究中，我們採用了雙樣本樹 (double-sample trees) 算法來分析數據，通過將數據分為兩部分，一部分用於決定分裂點，另一部分用於估計葉節點的效果，從而減少過擬合並提高估計的準確性。CausalForestDML 模組結合機器學習和因果森林的思想，提供了靈活的配置選項，適用於處理不同類型的變量和研究問題，使我們能夠更準確地估計處理效應，並識別出影響結果的關鍵特徵。

4.3.4 SHAP 原理

Shapley values 是已故博弈論大師 Lloyd Shapley 基於賽局理論 (cooperative game theory) 中解決合作對策問題而提出來解決方案，目的是來計算出當前的賽

局中不同玩家的貢獻從而分配收益。而 SHAP 正是由 Shapley 啓發而被發明的，如果將應用情境由博弈轉移到解釋模型的輸出上，會變成計算目前模型中所使用的特徵裡，不同特徵對於目標問題預測的輸出總共有多少的貢獻。

SHAP (Shapley additive explanations) 是一種解釋機器學習模型預測的方法。它主要用於解釋模型的每個特徵對特定預測的貢獻，使模型的預測過程變得透明且易於理解。SHAP 值尤其在複雜的非線性模型，如隨機森林、梯度提升機和深度神經網絡中，提供了洞察模型決策的重要途徑。

當模型是非線性的或輸入特徵不是獨立時，SHAP 值應該對所有可能的特徵排序計算加權平均值。SHAP 將這些條件期望與從博弈論的經典 Shapley 值組合到每個特徵的歸因值中，也就是根據式 (16) 進行計算

$$\phi_j = \sum_{S \subseteq \{x_1, \dots, x_p\} \setminus \{x_j\}} \frac{|S|!(p - |S| - 1)!}{p!} (f_x(S \cup \{x_j\}) - f_x(S)) \quad (16)$$

其中 $\{x_1, \dots, x_p\}$ 是所有輸入特徵的集合， p 為所有輸入特徵的數目， $\{x_1, \dots, x_p\} \setminus \{x_j\}$ 為不包括 $\{x_j\}$ 的所有輸入特徵可能的集合， $f_x(S)$ 為特徵子集 S 的條件期望值 $E[f(x)|x_S]$ ，而 $|S|!(p - |S| - 1)!/p!$ 是權重，為子集 S 的特徵組合占比，所有可能的子集 S 的特徵組合占比之和為 1。

5 實證資料

這些數據來自 P-League 官方網站的比賽記錄和觀眾入場統計。我們的樣本包括 2021-2022 和 2022-2023 賽季的六支球隊的比賽記錄。每支球隊在 2021-2022 賽季各參加了 30 場比賽，在 2022-2023 賽季各參加了 40 場比賽。我們將基於球隊為單位合併六個表格，每個表格包含 70 條數據項目，因此總共會有 418 條記錄。在這 418 條記錄的數據集中，有 209 場比賽是基於主場隊伍的數據，另外的 209 場比賽是基於客場隊伍的數據。合併表格中包含的變數如下：得分差異、球隊虛擬變量、週末虛擬變量、觀眾入場人數、主場/客場指標、淨勝分、明星球員數量、犯規、主場/客場指標和球隊之間的交互作用項，以及裁判。

表 1 列出了所有控制變數及其預期效應。在 P-League 中，總共有六支球隊。我們以勇士隊作為基準，使用五個虛擬變數來控制其他五支隊伍的實力。這五個虛擬變數的負值是預期的，因為勇士隊已經連續三年贏得冠軍，使他們成為當前 P-League 中最強大的球隊。因此，與勇士隊相比，其他五支隊伍的得分差異對他們來說是負面的。除了控制球隊實力外，我們還可以進一步分析球隊和主場之間的互動效應。

BACKTOBACK 用於控制由於連續比賽而導致的球員疲勞對比賽結果的影響。AUDIENCE 可能會影響球員和裁判的表現。傳統上，我們相信更多的觀眾有利於主場隊伍。STARS 是指每個賽季被選為年度最佳陣容第一隊的球員。每個賽季有五名球員入選。根據 Moskowitz and Wertheim (2012)，明星球員對比賽有顯著的影響，明顯增加了主場隊伍贏得比賽的概率。如圖 1 所示，儘管樣本

只有三年，但擁有更多的明星球員顯著有助於球隊進入季後賽和贏得冠軍。因此，在模型中控制明星球員是至關重要的。

在表 2 中，主場隊伍平均擁有比客場隊伍多 3.46 分的淨得分優勢、勝率高 8%。他們的二分之一球、三分球和罰球命中率也略高於客場隊伍。這符合預期，因為主場隊伍的球員可能更加熟悉球場和籃框，從而表現更佳並具有更高的投籃命中率。這也可能歸因於主場球迷的支持，這可能會提高球員投籃的信心。在主場比賽中，平均比客場比賽多出 0.16 次罰球的嘗試，且失誤次數稍微低於客場隊伍。這可能與裁判的決策有關，因為文獻表明，裁判在主場觀眾的壓力下更有可能做出支持主場隊伍的決策。

6 實證結果

第 4 節中，4-1 我們先使用邏輯迴歸模型檢驗主場獨立性，預期係數都是不顯著，才能支持我們認為聯盟分配主場是公平、獨立的論點。再來切入主題，4-2 使用多元線性回歸模型分析主場優勢對淨勝分的影響，並分別控制不同變數，也加入交乘項觀察各個球隊與其主場的交互作用。除了淨勝分之外，4-3 中，我們想知道主場優勢對勝率、FT%、FOULS、TO、FT、2PT% 以及 3PT% 的影響，使用 FT% 是爲了觀察觀眾影響球員的途徑，FOULS、FT 則是爲了觀察裁判影響比賽的途徑。邏輯迴歸的部分，係數原本是對數勝算比，但對數勝算比並不直觀，且不好解釋，故我們在表格中都已將係數調整爲邊際機率，以利於解釋意義。標準差也以穩健標準差呈現。

後續章節的安排如下。6-1 節將呈現主場獨立性之結果，6-2 節將呈現場優勢對淨勝分的影響之結果，6-3 節將呈現場優勢對勝率、FT%、FOULS、失誤數、FT、2PT% 以及 3PT% 的影響之結果。

6.1 分析主場獨立性

從表 3 可以看出除了 BACKTOBACK，其他控制變數皆不顯著，BACKTOBACK 在 1% 顯著水準下顯著，主要原因是因爲在 2021-2022 以及 2022-2023 賽季中，官方排定的賽程中，78 場背靠背比賽，有 77 場是主場球隊進行，這與我們在第四節傳統解釋中，對主場有利的賽程安排有所出入，顯然在 P-League 中，主隊無法享有這項優勢，可能是官方爲了讓球隊不用四處奔波，

所以刻意安排球隊連續在主場比賽。而其他本文所使用的控制變數並不會影響到主場的分配，符合我們的預期。

6.2 分析主場優勢對淨勝分的影響

表 4 是控制不同變數下，以淨勝分作為被解釋變數的多元迴歸模型得到的結果。模型 1 只控制 HOME，在顯著水準 5% 下，在主場比賽平均可以使球隊淨勝分增加 3.455 分。這樣的結果與我們預想中一致，主場優勢確實存在，我們常常看到競爭激烈的比賽中，兩支球隊緊咬住對手，幾分鐘內可以互換領先數次，可想而知，這是相當大的優勢。

模型 2 控制 HOME 以及五個球隊的虛擬變數，在顯著水準 5% 下，主場平均可以使淨勝分增加 3.461 分。

模型 3 控制 HOME 以及 BACKTOBACK，在顯著水準 5% 下，主場平均可以使球隊淨勝分增加 3.869 分。但 BACKTOBACK 係數並不顯著，代表連續比賽造成的疲勞，可能對 P-League 的球員影響不大，並不是會影響比賽勝負的關鍵。

模型 4 控制 HOME 以及 STARS，在顯著水準 5% 下，主場平均可以使球隊淨勝分增加 4.027 分。另外，在顯著水準 1% 下，STARS 的係數代表當球隊中每增加一位有入選年度第一隊的球員，就可以讓淨勝分多 4.856 分，有些球隊可能會有 2 至 3 位這種球員，可想而知，對比賽勝負會有壓倒性的影響。

模型 5 控制 HOME 以及 AUDIENCE，在顯著水準 5% 下，主場平均可以使球隊淨勝分增加 4.012 分。

模型 6 控制以上所有模型的變數，並加入球隊變數與 HOME 的交乘項以分析不同球隊與主場的交互作用。STAR 在顯著水準 1% 下顯著，係數為 3.523，代表球星在這個模型中依然對淨勝分有巨大影響。但主場及交互作用並不顯著。

綜合來看，我們可以得到幾個結論，HOME 幾乎在所有模型中都是顯著的，對每場淨勝分平均有 3.5 分以上的影響。BACKTOBACK 在所有模型中均不顯著，可能這些球員已經習慣這樣的賽程安排，緊湊比賽累積的疲勞也不會造成比賽中的劣勢。STARS 在模型 4 到模型 6 均在 1% 顯著水準下顯著，代表球星對淨勝分影響甚大，在比賽中扮演極為重要的角色。

6.3 分析主場優勢對其他因素的影響

5.2 節中，我們用多元線性迴歸模型探討主場優勢對淨勝分的影響，在這節中，我們將會分析主場優勢對獲勝機率、罰球命中率、犯規數、失誤數、罰球數、兩分球命中率及三分球命中率的影響，觀察主場優勢體現在哪些因素上面。

6.3.1 分析主場優勢對獲勝機率的影響

使用 WIN 作為被解釋變數是為了彌補淨勝分回歸的缺點，淨勝分迴歸的缺點是每單位分差對比賽的影響力不同。例如，有兩場球賽，一場球賽的比數是 99:100(99 是客隊)，另一場的比數是 99:130(99 是客隊)，若現在主客互換，由第四節的分析結果得知 HOME 對淨勝分的影響大約是 3 分，第一場比賽的比數會變成 102:99，第二場比賽的比數變為 102:130，都是一樣 3 單位的變化，但第

一場比賽勝負顛倒，第二場比賽則無關勝負。因此，這節我們將使用邏輯迴歸模型，來分析控制變數對勝負的影響。

我們可以從表 5 觀察到 HOME 係數在模型 1、模型 2 和模型 4 中，在至少 10% 顯著水準下顯著，係數的意義是若在主場比賽，獲勝機率可以提升 8.1% 左右。

STARS 在模型 4、模型 5 與模型 6 中，在 1% 顯著水準下顯著，且數值為 0.125、0.125 和 0.122，對勝負的影響跟 HOME 差不多，甚至更多，代表有球星的隊伍有著極大的優勢。另外，球隊與主場的交互作用同樣在邏輯迴歸模型中不顯著，總結來說，主場優勢對勝率的影響是顯著的，但個別球隊的主場優勢，並沒有顯著差異。

6.3.2 分析主場優勢對罰球命中率的影響

第 4 節的傳統主場優勢解釋中，有提到人群可能透過很多種途徑影響比賽，但我們無法將有關這些方面的群體效應與同時存在的其他因素分離出來測量，而罰球是一名球員與其他試圖干擾他、使他分心的觀眾之間的單獨互動，所以我們透過罰球來觀察觀眾對球員的影響。看過比賽的人都知道，客隊球員在罰球時，現場的觀眾總是無所不用其極的試圖干擾球員，大吼大叫、舉高自製的立牌還有無止盡的噓聲，因此通常我們認為球迷會導致球員的罰球命中率降低。

我們使用多元線性迴歸模型來分析主場對罰球命中率的影響。從表 6 中，我們發現 HOME 變數在所有模型中均不顯著。

結論是主場的觀眾吵雜，其實不太會影響球員的罰球，這很可能是因為身為職業球員，投籃已經成為肌肉記憶，像罰球這種固定程序已經不太會被外界干擾，但如果是業餘聯賽，情況可能就不一樣了，球員並沒有受過專業訓練，可能會輕易的被觀眾影響。

6.3.3 分析主場優勢對犯規數以及罰球數的影響

我們常常會聽到有人說抱怨裁判偏袒主隊，也就是所謂的”主場哨”，我們在這節將使用多元線性迴歸模型來分析主場對犯規數以及罰球數的影響，看犯規數及罰球數是否會因為主客隊而有所差別。

表 7 是主場對犯規數的迴歸，我們發現 HOME 變數在模型 1 到模型 5 中均不顯著。模型 6 中，HOME 在 5% 顯著水準下顯著，代表勇士隊在主場時，場均能減少 1.962 個犯規，交互作用項的部分，HOME_PILOTS、HOME_STEELERS 在至少 5% 顯著水準下顯著，分別為 3.844、3.365，意思是當領航猿在主場比賽時，場均會增加 1.882 個犯規；鋼鐵人在主場時，場均會增加 1.403 個犯規。

表 8 是主場對罰球數的迴歸，我們發現 HOME 在所有模型中均不顯著，代表沒有足夠證據支持在主場比賽能有較多的罰球數。

6.3.4 分析主場優勢對失誤、三分球命中率以及二分球命中率的影響

這節要探討的是更多主場優勢可能會體現的方面，我們會依序使用失誤、三分球命中率以及二分球命中率作為應變數的迴歸來分析。

表 9 是主場對失誤數的迴歸，我們發現 HOME 在所有模型中均不顯著，代表球員不太會因為在主場就減少或增加失誤。另外，交互作用皆不顯著。從這個模型中，我們得知主場對失誤數幾乎沒有影響。

表 10 是主場對兩分球命中率的迴歸，HOME 在模型 6 中，在 5% 顯著水準下顯著，代表勇士主場的兩分球命中率比客場平均高出 4.186%。交互作用的部分，模型 6 中，HOME_PILOTS、HOME_DREAMERS 分別在 5%、10% 顯著水準下顯著，代表這些球隊在主場時兩分球命中率降低了 2.238%、0.451%

表 11 是主場對三分球命中率的迴歸，HOME 在所有模型中均不顯著，代表球員不太會因為在主場就提升或降低三分球命中率。

結論是在這三個模型當中，主場對失誤及三分球命中率沒有顯著影響，只有二分球命中率會因主客場不同而有顯著變化，其中勇士隊二分命中率提高最多，領航猿、夢想家二分命中率稍降。

6.4 異質性分析-因果森林

我們將 P-League 2021-2022 以及 2022-2023 賽季共 418 筆資料拆分成 334 筆訓練資料及 84 筆測試資料，並使用 econml 中的 CausalForestDML 套件作分析。應變數為淨勝分，控制變數包含 5 個球隊虛擬變數、球星數量、觀眾人數以及比

賽間隔天數，這裡使用「比賽間隔天數」取代 BACKTOBACK，是因為間隔天數更能體現不同天數的間隔對主場優勢的影響，當間隔天數等於 0，就是背靠背比賽。

本章節主要將以因果森林圖來觀察各特徵對主場優勢造成的異質性，圖中同時呈現各分群之平均估計值 (CATE) 及樣本數，幫助我們了解影響的大小及範圍，並協助我們觀察較複雜的結構。

6.4.1 決策樹分析結果

主場的處理效果可能非常複雜，但我們可以利用決策樹圖形幫助我們建構一些簡單的分類規則，這些規則可以區分擁有哪些特徵的樣本會有強烈的主場優勢，哪些樣本保持中立，哪些樣本有主場劣勢。

圖 2 是基於因果森林模型產生的決策樹，用於說明主場優勢如何影響比賽的淨勝分，並顯示了不同特徵條件下這種影響的異質性。具體來說，這棵樹通過不同的觀眾人數和比賽間隔等特徵，將樣本劃分成多個子集，並為每個子集提供了主場優勢對淨勝分影響的 CATE，以及這些估計值的標準差和置信區間。

根節點提供了整個數據集的 CATE，為 2.595，標準差為 2.081，這表示在考慮所有特徵之前，主場優勢對淨勝分的平均影響。第一層是以觀眾人數做切割，可以看到左半邊顏色較淺 (淺綠)，代表樣本對主場的反應較輕微。右半邊顏色較深 (深綠)，代表樣本對主場的反應較強烈。對於觀眾人數小於或等於 5112.5 時，CATE 為 1.763，標準差為 1.313，顯示在這個範圍內，主場優勢對淨勝分

的影響相對較小。當觀眾人數大於 5112.5 時，CATE 上升至 4.763，標準差為 2.15，表明在觀眾人數較多的情況下，主場優勢對淨勝分的正面影響更為顯著。

第二層分割在觀眾人數較少的子集中，根據球星數量進一步劃分。例如，當球星數量小於或等於 0.5 時，CATE 進一步降低到 0.211，顯示當主隊有較少球星時，主場優勢對淨勝分的影響可能會降低。在觀眾人數較多的子集中，當比賽間隔小於或等於 2.5 天時，CATE 為 6.295，標準差為 0.739，這表示在短時間內連續主場比賽的情況下，主場優勢對淨勝分的影響顯著增加。在右側的一個葉節點中，對於觀眾人數大於 6587.5，且比賽間隔小於或等於 0.5 的子集，CATE 高達 5.953，標準差為 0.791，這可能表示在觀眾人數非常多且比賽間隔短的情況下，主場優勢對淨勝分的影響最為顯著。這其實稍微與普遍的認知不符，一般認為比賽間隔越長，球員就不會因疲勞而表現失常，也能有更多時間準備比賽。造成這樣結果的原因可能是因為比賽間隔太長的話，球員反而沒辦法維持比賽狀態。

這棵樹的解釋力量在於它展示了主場優勢對淨勝分的影響不是一成不變的，而是依賴於比賽的具體環境條件，以觀眾人數來說，當人數大於 6587.5，CATE 是 1.465，主場優勢極其微弱，有可能是當支持的觀眾過多時，反而帶給主隊球員過多壓力，導致表現失常 (Hill et al., 2010; Baumeister, 1984)。且觀眾人數被頻繁用作節點的分割條件，這通常表明在模型訓練過程中，觀眾人數是影響淨勝分的一個重要特徵，對主場優勢效應存在顯著的異質性影響。下一節我們會透過觀察 SHAP 值 (Shapley value) 來分析主場優勢的異質性，雖然與決策樹的構成邏輯不同，但有可能可以提供其他洞見，並相互解釋。

6.4.2 透過 SHAP 觀察主場優勢之異質性

圖 3 是一個 SHAP 值的散點圖，通常稱為蜂群圖 (beeswarm)，包含所有訓練資料，讓我們能了解模型是如何從訓練資料中學習，即模型在學習階段最重視哪些特徵。橫軸代表 SHAP 值，即每個特徵對模型預測淨勝分的貢獻程度，而縱軸列出了所有控制變數 (特徵)。特徵的排序是按照 SHAP 的平均絕對值，可以看做一種特徵重要性的排列圖，從這裡可以看到觀眾人數的重要性最高。圖中每個點代表一個觀測樣本，而點的顏色則根據特徵值的高低進行編碼，藍色代表低特徵值，紅色代表高特徵值。這些顏色的深淺代表特徵值在低到高的範圍內的變化。

可以看出特徵重要性依序是：觀眾人數、比賽間格、領航員、勇士、國王、工程師、夢想家、球星數量、鋼鐵人。觀眾人數欄位左半部的點呈現藍色，右半部的點呈現紅紫色，代表當樣本的觀眾人數數值小 (藍色) 時，SHAP 值也較小；當樣本的觀眾人數數值大 (紅紫色) 時，SHAP 值也較大。圖 4 是觀眾人數 SHAP 值之散點圖，基本上是呈現線性關係。比賽間格欄位最左邊的幾個點是藍色，中間則是紅藍混雜，可以看出比賽間格數值的大小與 SHAP 值可能是非線性的關係。圖 5 是比賽間格 SHAP 值之散點圖，看不出明顯趨勢，當比賽間格為 0 時，代表該場為背靠背比賽，48 分鐘裡來回跑動對體能有苛刻的要求，連續兩晚全場都拚盡全力打是極其困難的，比賽到第二個晚上，NBA 球隊勝率僅為 36% 左右 (Tobias Moskowitz, L. Jon Wertheim, 2012)。我們發現當比賽間格介於 2-5 天時，對淨勝分的正面影響最大。

6.4.3 因果森林方法之結論

由因果決策樹以及 SHAP 值可以得知觀眾人數及比賽間格是影響主場優勢最大的兩個因素。觀眾人數越多，主場優勢就更加明顯；比賽間格則是存在非線性的影響，推測是由於比賽間格過短會造成球員疲勞，而過長則會使球員無法維持良好的比賽狀態。

表 12 展示的是所有訓練資料的預測的條件平均處理效應。這個表格主要分析了不同特徵在中位數以下 (低水平) 與中位數以上 (高水平) 時的處理效應，並提供了標準差異和多重假設檢測 (MHT) 的 p 值。標準差異顯示了中位數以下和中位數以上之間的標準化差異，而多重假設檢測的 p 值用於檢測中位數以下和中位數以上的預測效應之間的差異是否統計顯著。前面的因果森林圖展示較複雜的結構，而表 12 只以中位數切割，提供較直觀的結果，幫助我們了解因果森林學習的過程以及觀察各特徵的數值大小對主場優勢的影響。所有球隊名都是虛擬變數，等於 1 代表那場是該球隊之比賽，等於 0 代表那場並非該球隊之比賽。

因此，我們發現勇士隊主場的平均淨勝分 (3.866 分) 相較其他球隊的主場平均淨勝分 (2.116 分) 高出許多，攻城獅亦然。而國王和領航員相較其他球隊，主場優勢都明顯較小，由圖 2 得知，當觀眾人數小於 5112.5 時，領航員在主場時，CATE 是 0.211，子葉顏色為白色，代表幾乎沒有主場優勢，提供我們部分因果森林結構的解釋。比賽間格不顯著，其原因為比賽間格對主場優勢是非線性的 (圖 5)，以中位數區分的話，無法體現其對主場優勢之影響。而觀眾人數在中位

數以上時，平均有 3.666 分的主場優勢，在中位數以下時，平均僅有 1.104 分的主場優勢，兩者有顯著差異。另外，球星數量在中位數以上時，平均有 2.131 分的主場優勢，在中位數以下時，平均有 2.611 分的主場優勢，我們認為是因為球星數量多的情況下，球隊實力較穩定和堅強，故主場帶來的優勢會比球星少的情況下還輕微。

7 結論與建議

本研究採用 P-League 2021-2022 和 2022-2023 賽季的官方比賽記錄和觀眾入場統計數據探討台灣籃球聯賽主場優勢對淨勝分之影響。本文主要使用 OLS 和邏輯迴歸模型對主場獨立性以及淨勝分、獲勝機率等進行分析，以及 Wager and Athey (2018) 提出之分析架構，透過因果森林方法進行高維度數據處理，分析不同特徵對主場優勢的異質性影響，再利用因果森林決策樹圖提供可解釋性。

主場獨立性分析中，我們發現與國外聯賽不同的是，P-League 比較傾向於安排主場球隊打背靠背比賽，可能是考慮球隊交通的便利性。除此之外，每支球隊均被公平的分配主場，無其他控制變數影響主客場的分配。

根據 OLS 以及邏輯迴歸分析，我們發現在主場比賽平均可以使球隊淨勝分增加 3.5 分左右，並提升勝率大約 8%，且球星數量也是影響淨勝分的重要因素之一。但不同球隊與主場的交互作用並不顯著。除了淨勝分和勝率，我們也將其他比賽中的變數當作應變數，觀察主場的影響。我們發現主場對罰球命中率、罰球數、失誤數以及三分球命中率均無影響，但對犯規數、兩分球命中率有影響，能使勇士隊場均減少 1.962 個犯規，但會使領航猿和鋼鐵人場均增加 1.882 個和 1.403 個犯規。同時使勇士的二分命中率平均提升 4.186%、領航猿及夢想家下降 2.237% 和 0.451%。

接下來，我們透過因果森林更深入探討主場優勢的異質性，由決策樹 (圖 2) 以及 SHAP 值 (圖 3) 顯示觀眾人數是相當重要的特徵，觀眾人數越多，主場優勢越大，可知觀眾對球員產生了正面的激勵效果，超過了潛在的壓力，從而幫助

主場球隊展現更出色的表現。比賽間隔也會影響到主場優勢，但其影響是非線性的，間隔 2-5 天是最適當的，過短或過長的時間皆會降低主場優勢。除此之外，我們也從表 12 發現勇士和攻城獅主場優勢較大，國王和領航猿主場優勢較微弱，以及球星數量越多，主場優勢越小。

綜合以上發現，我們可以檢驗 Moskowitz and Wertheim (2012) 提出的四種慣常的主場優勢來源。第一個傳統解釋是主場觀眾的支持，透過因果森林證明主場觀眾確實可以讓球隊在主場表現更佳，此一發現與社會助長 (social facilitation) 的概念相吻合，根據 Zajonc (1965) 的研究，個體在他人存在下，對於簡單或已經熟練的任務表現得更好，但對於學習或陌生的任務表現較差。在比賽中，主場觀眾的支持可能增加球隊成員的生理激活水平，從而提高他們在比賽中的表現，尤其是對於那些他們已經熟練掌握的技能 and 策略。因此，我們的研究結果進一步支持了社會助長理論在解釋主場優勢現象中的應用。第二個傳統解釋是對主場有利的賽程安排，在主場獨立性分析中，我們發現在 P-League 中，主場球隊更容易打背靠背比賽，可能是官方為求比賽地點的方便性而做此安排，所以 P-League 並不存在對主隊有利的賽程。第三個傳統解釋-客隊要承受舟車勞頓之苦和第四個傳統解釋-更適合主隊的天氣也同樣不適用於台灣籃球聯賽，因為台灣幅員較小，不需要長途跋涉。且各地區天氣差異相較於美國來說非常小，所以不同球隊所在地區的天氣差異並不足以解釋主場優勢。

綜合分析迴歸和因果森林結果，主場優勢的存在是無庸置疑的，但並非一致，不同的球隊、觀眾人數、比賽間隔和球星數量都會影響到主場優勢的大小，且優勢主要反映在二分之一球命中率上面，其中又以勇士隊提升最多。

最後，本文提供後續研究的兩大方向與建議：第一、因 P-League 僅成立不到 4 年，又因球隊數量變動，截至目前只有兩個賽季的資料可以使用，樣本數稍有不足，且因各球隊的主場球館也尚未完全定案，常常會有變動，因此球隊文化可能還需要更多時間來建立，這些因素也會影響主場優勢的形成。若隨著球隊數量穩定以及樣本數增加，能產生更穩健的研究成果。第二、本研究使用失誤及犯規數作為迴歸應變數來分析主場對裁判吹判的影響，但若使用誤判數作為應變數，也許更能夠捕捉到裁判吹判標準的變化。希望有更全面的數據讓我們探討更全面的主場優勢帶來之影響。

Figures

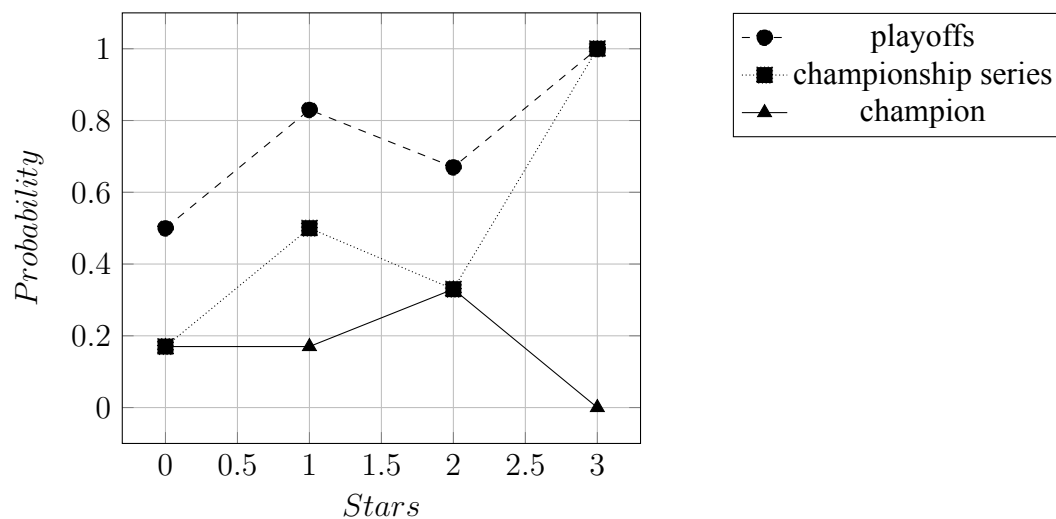


圖 1: 球星數量對晉級冠軍賽或季後賽機率的影響

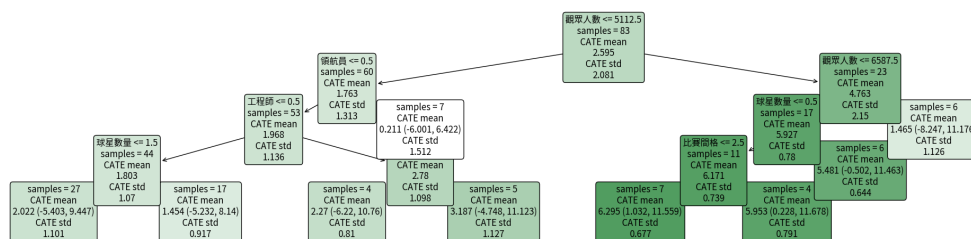


圖 2: 因果森林決策樹



圖 3: 各特徵之 SHAP value

說明: 橫軸代表 SHAP 值，縱軸為特徵，排序是按 shap 之平均絕對值。一個點代表一個樣本，顏色越紅表示特徵數值越大，顏色越藍表示特徵本身數值越小。

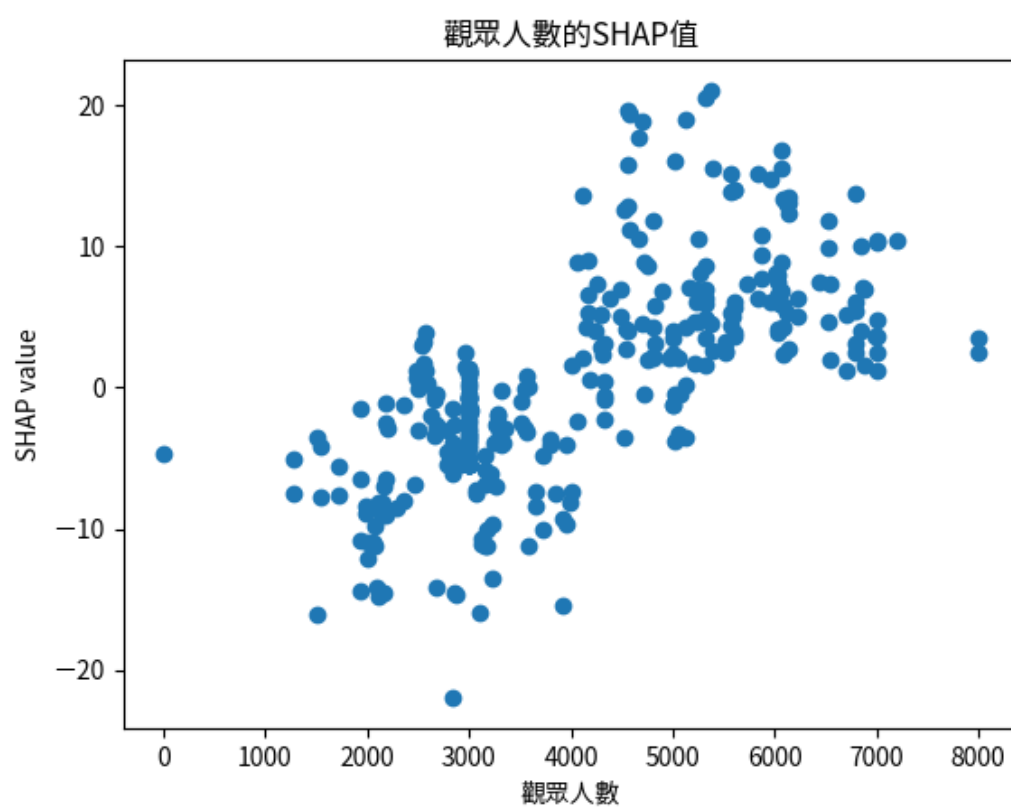


圖 4: 觀眾人數之 SHAP value

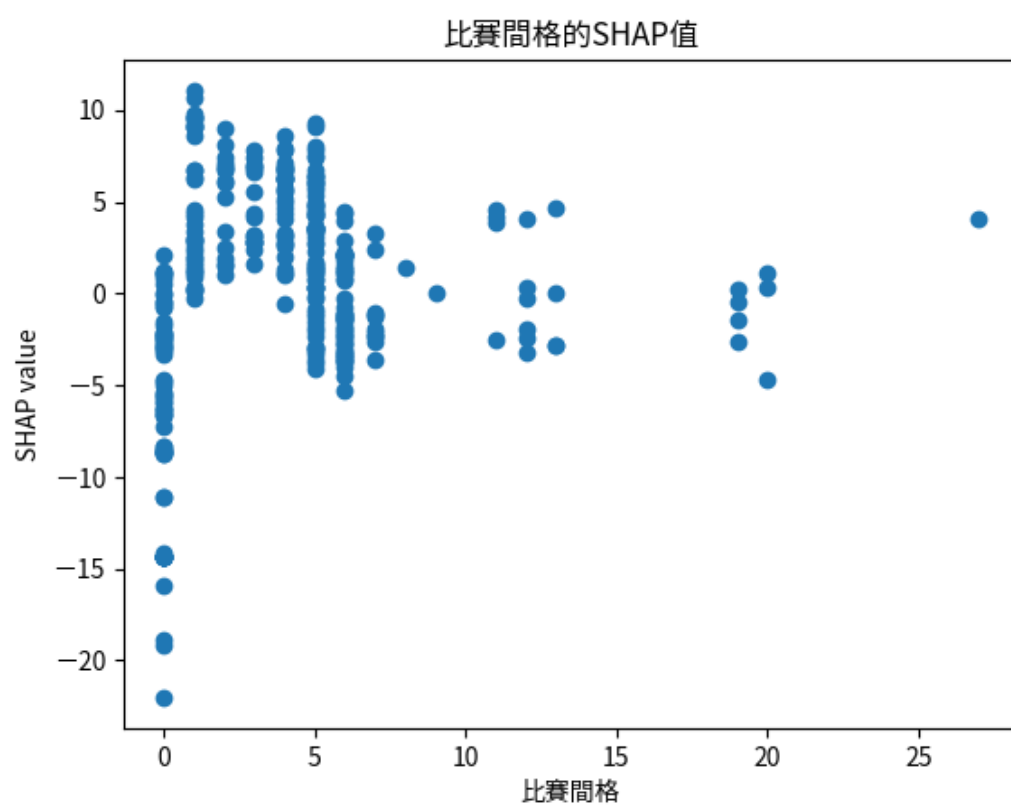


圖 5: 比賽間格之 SHAP value

Tables

表 1: Variables List

Variable	Description
POINT	表示特定球隊在比賽中的得分差的變數
WIN	表示球隊是否獲勝的虛擬變數
FT%	罰球命中率
FOULS	犯規數
FT	罰球數
TO	失誤數
2PT%	二分球命中率
3PT%	三分球命中率
HOME	代表比賽是否為主場的虛擬變數
KINGS	國王
PILOTS	領航猿
LIONEERS	攻城獅
DREAMERS	夢想家
STEELERS	鋼鐵人
BACKTOBACK	表示是否為背靠背比賽的虛擬變數
STARS	某個賽季被選為 P-League 年度第一隊的球員人數
AUDIENCE	比賽的觀眾出席人數

表 2: descriptive statistics			
	home	away	diff
POINT	1.73	-1.73	3.46** (1.46)
WIN	0.54	0.46	0.08* (0.05)
FT%	67.80	68.02	-0.22 (1.18)
FOULS	18.92	18.93	-0.01 (0.40)
FT	21.00	20.83	0.16 (0.72)
TO	16.04	16.15	-0.1 (0.42)
2PT%	47.88	46.83	1.05 (0.75)
3PT%	31.06	30.93	0.13 (0.83)
STARS	0.73	0.75	-0.02 (0.09)
AUDIENCE	4153.52	4153.52	0 (147.59)

Notes: *** p<0.01, ** p<0.05, * p<0.1

Notes: 括號內為穩健標準誤

表 3: 主場獨立性迴歸結果

	HOME
KINGS	0.048 (0.103)
PILOTS	0.003 (0.075)
LIONEERS	0.012 (0.075)
DREAMERS	-0.025 (0.079)
STEELERS	0.015 (0.075)
BACKTOBACK	0.939*** (0.185)
STARS	-0.019 (0.039)
AUDIENCE	0.000 (0.000)

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: 迴歸中也包含控制裁判，但由於裁判數量過多，為求簡潔而在表格中省略

Notes: 括號內為穩健標準誤

表 4: 淨勝分迴歸結果

POINT	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
HOME	3.455** (1.458)	3.461** (1.441)	3.869** (1.681)	4.027** (1.633)	4.012** (1.633)	3.523 (3.464)
BACKTOBACK			-1.122 (2.039)	-1.290 (2.021)	-1.251 (2.023)	-1.271 (2.032)
STARS				4.856*** (1.203)	4.862*** (1.208)	4.713*** (1.220)
AUDIENCE					0.000 (0.001)	-0.000 (0.001)
HOME_KINGS						-3.098 (4.710)
HOME_PILOTS						-2.669 (4.875)
HOME_LIONEERS						2.649 (4.717)
HOME_DREAMERS						3.788 (5.121)
HOME_STEELERS						2.269 (5.137)
控制球隊		✓	✓	✓	✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1

Notes: 球隊變數都以勇士隊為基準，僅讓其他五支球隊的虛擬變數進入模型

Notes: 若”控制球隊”欄位打勾，代表已控制所有球隊

Notes: 括號內為穩健標準誤

表 5: 勝負迴歸結果

WIN	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
HOME	0.081* (0.048)	0.081* (0.047)	0.068 (0.053)	0.073** (0.052)	0.072 (0.048)	0.101 (0.122)
BACKTOBACK			0.035 (0.070)	0.030 (0.069)	0.031 (0.069)	0.031 (0.070)
STARS				0.125*** (0.041)	0.125*** (0.041)	0.122*** (0.041)
AUDIENCE					0.000 (0.000)	0.000 (0.000)
HOME_KINGS						-0.113 (0.164)
HOME_PILOTS						-0.117 (0.170)
HOME_LIONEERS						-0.014 (0.159)
HOME_DREAMERS						0.089 (0.169)
HOME_STEELERS						-0.027 (0.177)
控制球隊		✓	✓	✓	✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1

Notes: 球隊變數都以勇士隊為基準，僅讓其他五支球隊的虛擬變數進入模型

Notes: 若”控制球隊”欄位打勾，代表已控制所有球隊

Notes: 括號內為穩健標準誤

表 6: 罰球命中率迴歸結果

FT%	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
HOME	-0.219 (1.177)	-0.213 (1.146)	-0.295 (1.304)	-0.296 (1.298)	-0.306 (1.296)	2.489 (3.022)
BACKTOBACK			0.225 (1.642)	0.227 (1.642)	0.252 (1.637)	0.173 (1.629)
STARS				-0.043 (1.075)	-0.394 (1.075)	-0.119 (1.093)
AUDIENCE					0.000 (0.000)	0.000 (0.000)
HOME_KINGS						-5.874 (3.943)
HOME_PILOTS						3.288 (4.322)
HOME_LIONEERS						-7.111* (3.803)
HOME_DREAMERS						-5.371 (4.120)
HOME_STEELERS						-1.431 (4.432)
控制球隊		✓	✓	✓	✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1

Notes: 球隊變數都以勇士隊為基準，僅讓其他五支球隊的虛擬變數進入模型

Notes: 若”控制球隊”欄位打勾，代表已控制所有球隊

Notes: 括號內代表穩健的標準誤

表 7: 犯規數迴歸結果

FOUL	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
HOME	-0.014 (0.403)	-0.012 (0.395)	-0.266 (0.435)	-0.247 (0.434)	-0.238 (0.446)	-1.962** (0.955)
BACKTOBACK			0.699 (0.598)	0.679 (0.595)	0.655 (0.573)	0.650 (0.597)
STARS				0.590* (0.346)	0.587* (0.347)	0.546 (0.345)
AUDIENCE					0.000 (0.000)	0.000** (0.000)
HOME_KINGS						1.296 (1.378)
HOME_PILOTS						3.844*** (1.293)
HOME_LIONEERS						0.559 (1.377)
HOME_DREAMERS						1.345 (1.404)
HOME_STEELERS						3.365** (1.320)
控制球隊		✓	✓	✓	✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1

Notes: 球隊變數都以勇士隊為基準，僅讓其他五支球隊的虛擬變數進入模型

Notes: 若”控制球隊”欄位打勾，代表已控制所有球隊

Notes: 括號內為穩健標準誤

表 8: 罰球數迴歸結果

FT	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
HOME	0.163 (0.716)	0.164 (0.699)	0.393 (0.784)	0.384 (0.786)	0.374 (0.787)	-2.373 (1.875)
BACKTOBACK			-0.631 (0.511)	-0.622 (0.963)	-0.595 (0.963)	-0.607 (0.964)
STARS				-0.263 (0.640)	-0.259 (0.642)	-0.418 (0.618)
AUDIENCE					0.000 (0.000)	-0.000 (0.000)
HOME_KINGS						-0.258 (2.578)
HOME_PILOTS						4.072* (2.435)
HOME_LIONEERS						2.899 (2.436)
HOME_DREAMERS						3.497 (2.488)
HOME_STEELERS						6.356** (2.601)
控制球隊		✓	✓	✓	✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1

Notes: 球隊變數都以勇士隊為基準，僅讓其他五支球隊的虛擬變數進入模型

Notes: 若”控制球隊”欄位打勾，代表已控制所有球隊

Notes: 括號內為穩健標準誤

表 9: 失誤數迴歸結果

TO	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
HOME	-0.110 (0.418)	-0.112 (0.416)	-0.154 (0.473)	-0.134 (0.471)	-0.134 (0.472)	0.843 (1.071)
BACKTOBACK			0.117 (0.648)	0.095 (0.642)	-0.096 (0.643)	0.078 (0.649)
STARS				0.626* (0.355)	0.626* (0.356)	0.603* (0.357)
AUDIENCE					0.000 (0.000)	0.000 (0.000)
HOME_KINGS						-1.135 (1.326)
HOME_PILOTS						-1.753 (1.471)
HOME_LIONEERS						-1.669 (1.599)
HOME_DREAMERS						-1.766 (1.524)
HOME_STEELERS						0.507 (1.503)
控制球隊		✓	✓	✓	✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1

Notes: 球隊變數都以勇士隊為基準，僅讓其他五支球隊的虛擬變數進入模型

Notes: 若”控制球隊”欄位打勾，代表已控制所有球隊

Notes: 括號內為穩健標準誤

表 10: 二分球命中率迴歸結果

2PT(%)	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
HOME	1.043 (0.751)	1.029 (0.742)	1.110 (0.824)	1.103 (0.825)	1.108 (0.827)	4.186** (1.642)
BACKTOBACK			-0.222 (1.138)	-0.214 (1.138)	-0.229 (1.139)	-0.259 (1.150)
STARS				-0.221 (0.657)	-0.224 (0.658)	-0.225 (0.663)
AUDIENCE					0.000 (0.000)	-0.000 (0.000)
HOME_KINGS						-2.471 (2.538)
HOME_PILOTS						-6.424** (2.553)
HOME_LIONEERS						-3.860 (2.405)
HOME_DREAMERS						-4.637* (2.619)
HOME_STEELERS						-1.031 (2.489)
控制球隊		✓	✓	✓	✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1

Notes: 球隊變數都以勇士隊為基準，僅讓其他五支球隊的虛擬變數進入模型

Notes: 若”控制球隊”欄位打勾，代表已控制所有球隊

Notes: 括號內為穩健標準誤

表 11: 三分球命中率迴歸結果

3PT(%)	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
HOME	0.133 (0.834)	0.125 (0.832)	0.148 (0.963)	0.169 (0.959)	0.175 (0.961)	2.832 (2.115)
BACKTOBACK			-0.064 (1.234)	-0.087 (1.239)	-0.103 (1.245)	-0.172 (1.250)
STARS				0.658 (0.718)	0.656 (0.716)	0.656 (0.713)
AUDIENCE					0.000 (0.000)	0.000 (0.000)
HOME_KINGS						-1.653 (2.718)
HOME_PILOTS						-6.390** (3.018)
HOME_LIONEERS						-5.991** (2.989)
HOME_DREAMERS						-2.108 (3.082)
HOME_STEELERS						0.340 (2.687)
控制球隊		✓	✓	✓	✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1

Notes: 球隊變數都以勇士隊為基準，僅讓其他五支球隊的虛擬變數進入模型

Notes: 若”控制球隊”欄位打勾，代表已控制所有球隊

Notes: 括號內為穩健標準誤

表 12: Predicted Conditional Average Treatment Effect-Home court

	(1)	(2)	(3)	(4)
	Predicted Treatment Effects		Std. diff.	MHT p-value
	Below median	Above median	(1) - (2)	(1) - (2)
勇士	2.116	3.866	-0.053	0.000***
國王	2.564	1.503	0.129	0.005***
領航猿	2.712	0.932	-0.043	0.000***
攻城獅	2.093	3.738	0.018	0.000***
夢想家	2.449	2.079	-0.047	0.329
鋼鐵人	2.384	2.388	-0.005	0.992
比賽間格	2.173	2.627	-0.390	0.113
觀眾人數	1.104	3.666	-0.042	0.000***
球星數量	2.611	2.131	0.090	0.093*

Notes: *** p<0.01, ** p<0.05, * p<0.1

參考文獻

- Balmer, N. J., Nevill, A. M., & Williams, A. M. (2001). "Home advantage in the Winter Olympics (1908-1998)." *Journal of Sports Sciences*, 19(2), 129-139.
- Baumeister, R. F. (1984). "Choking under pressure: Self-consciousness and paradoxical effects of incentives on skill-ful performance." *Journal of Personality and Social Psychology*, 46, 610-620.
- Boudreaux, Shane D. Sanders, and Bhavneet Walia., 2017. "A Natural Experiment to Determine the Crowd Effect Upon Home Court Advantage," *J. Sports Econ.*, 18 (7) (2017), pp. 737-749
- Cao, Z., J. Price, and D. F. Stone (2011). Performance under pressure in the NBA. *Journal of Sports Economics*, 12 (3), 231-252.
- Carron, A. V., Loughhead, T. M., & Bray, S. R. (2005). "The home advantage in sport competitions: Courneya and Carron's (1992) conceptual framework a decade later." *Journal of Sports Sciences*, 23(4), 395-407.
- Clarke, S. R., & Norman, J. M. (1995). "Home advantage of individual clubs in English soccer." *The Statistician*, 44(4), 509-521.
- Cottrell, N. B., Wack. D. L., Sekerak, G. J., & Rittle, R. H. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology*, 9, 245-250.
- Courneya, K. S., & Carron, A. V. (1992). "The home advantage in sport competi-

- tions: A literature review.” *Journal of Sport and Exercise Psychology*, 14(1), 13-27.
- Dohmen, T.J.(2008). “Do professionals choke under pressure?” *Journal of Economic Behaviour & Organization*, 65, 636-653.
- Downward, P., & Jones, M. (2007). “Effects of crowd size on referee decisions: Analysis of the FA Cup.” *Journal of Sports Sciences*, 25(14), 1541-1545.
- Stefan Wager & Susan Athey. (2018). “Estimation and Inference of Heterogeneous Treatment Effects using Random Forests,” *Journal of the American Statistical Association*, 113:523, 1228-1242.
- Harb-Wu, K., & Krumer, A. (2019). “Choking under pressure in front of a supportive audience: Evidence from professional biathlon.” *Journal of Economic Behaviour & Organization*, 166, 246–262.
- Harris, A. R. and P. J. Roebber (2019). NBA team home advantage: Identifying key factors using an artificial neural network. *PLOS ONE* 14 (7), 1–9.
- Hill, D., Hanton S., Matthews N., & Fleming S. (2010). “Choking in sport: A review.” *International Review of Sport and Exercise Psychology*, 3(1), 24–39.
- Jehue, R., Street, D., & Huizenga, R. (1993). “Effect of time zone and game time changes on team performance: National Football League.” *Medicine and Science in Sports and Exercise*, 25(1), 127-131.
- Kotecki, J. (2014). Estimating the effect of home court advantage on wins in the

- NBA. *The Park Place Economist*, 22 (1), 49–57.
- Michael Price, Jun Yan. (2021). The Effects of the NBA COVID Bubble on the NBA Playoffs: A Case Study for Home-Court Advantage. *American Journal of Undergraduate Research*, 18(4), 13-15.
- Nevill, A. M., & Holder, R. L. (1999). "Home advantage in sport: An overview of studies on the advantage of playing at home." *Sports Medicine*, 28(4), 221-236.
- Nevill, A. M., Balmer, N. J., & Williams, A. M. (2002). "The influence of crowd noise and experience upon refereeing decisions in football." *Psychology of Sport and Exercise*, 3(4), 261-272.
- Pace, A., & Carron, A. V. (1992). "Travel and the home advantage." *Canadian Journal of Sport Sciences*, 17(1), 60-64.
- Pollard, R. (2006). "Home advantage in soccer: Variations in its magnitude and a literature review of the inter-related factors associated with its existence." *Journal of Sport Behavior*, 29(2), 169-189.
- Recht, L. D., Lew, R. A., & Schwartz, W. J. (2003). "Baseball teams beaten by jet-lag." *Nature*, 377(6550), 583.
- Sanna, L. J., & Shotland, R. L. (1990). Valence of anticipated evaluation and social facilitation. *Journal of Experimental Social Psychology*, 26, 82-92.
- Snyder, E. E., & Purdy, D. A. (1985). "The home advantage in collegiate basketball." *Sociology of Sport Journal*, 2(4), 352-356.

- Stefani, R. (2008). "Measurement and interpretation of home advantage." *Statistical thinking in sports*, 203-216.
- Sutter, M., & Kocher, M. G. (2004). "Favoritism of agents - the case of referees' home bias." *Journal of Economic Psychology*, 25(4), 461-469.
- Taylor, S.E., Seeman, T.E., Eisenberger, N.I., Kozanian, T.A., Moore, A.N., Moons, W.G., 2010. "Effects of a supportive or an unsupportive audience on biological and psychological responses to stress." *J. Pers. Soc. Psychol.* 98 (1), 47.
- Tobias Moskowitz, L. Jon Wertheim. (2012). "Scorecasting: The Hidden Influences Behind How Sports Are Played and Games Are Won," 137-197.
- Robert B. Zajonc. (1965). "Social Facilitation." *Science*, 149(3681), 269-274.
- Zajonc, R. B. (1965). Social facilitation: A solution is suggested for an old social psychological problem. *Science*, 149, 269-274.
- Zajonc, R. B. (1980). Compresence. In P. B. Paulus (Ed.), *Psychology of group influence* (pp. 35-60). Hillsdale, NJ: Erlbaum.