國 立 政 治 大 學

統 計 學 系

碩 士 學 位 論 文

賽局理論與學習模型的實證研究

An Empirical Study of Game Theory and Learning Model

指導教授 ： 余清祥 博士

研　究　生 ： 陳冠儒 撰

中華民國一百年六月

謝辭

能夠完成這篇論文最感謝的當然是我的指導教授余清祥老師，剛開始接觸賽局理論時什麼都不懂，但是老師總是很有耐心的給予我許多指導，真的很感謝老師。此外也很感謝中研院社科中心的陽春雷老師，也給予我的論文許多的指導與方向。當然也非常感謝我的口試老師陳麗霞老師、陳怡如老師和溫在弘老師，在口試過程給予我許多寶貴的意見。

在碩班兩年的生活，陪伴我的家人和朋友們，與他們的相處都讓我非常開心，我的爸爸媽媽平常對我的關心，以及大哥二哥在我的人生旅途上，跟我說許多他們經歷過的經驗，非常感謝他們對我的關心，以及我家的狗-娜娜，每次回家看到她心情都會變得很好，超可愛的!還有在碩班的研究室一起度過的朋友們，金碩、心維、丞庭、卓卓、盈方、婉婷、大衛、小蘋果、詠翔、歐家、柯姐、慈慧、慧甄、雨慈、妮妮、建佑，還有佩茹，和你們相處真的很開心，謝謝你們！還有球隊的朋友，傳傑、跳哥、哲宇王、品豪、耀震、水哥、安卓、死魚、小豪、晉維、匯捷、挺誼，和你們一起練球吃宵夜也很開心，之後也希望可以再回來找你們打球！最後要感謝的是李宜真，謝謝你陪伴我，也常常給我鼓勵，在我心情不好的時候也總是能讓我勇敢往前，謝謝你，謝謝大家。

摘要

賽局理論(Game Theory)大多假設理性決策，單一回合賽局通常可由理論證明均衡(Equilibrium)或是最佳決策，然而如果賽局重複進行，不見得只存在單一均衡，光從理論推導可能無法找到所有均衡。以囚犯困境(Prisoner Dilemma)為例，理論均衡為不合作，若重複的賽局中存有互利關係，不合作可能不是最佳選擇。近年來，經濟學家藉由和統計實驗設計類似的賽局實驗(Game Experiment)，探討賽局在理論與實際間的差異，並以學習模型(Learning Model)描述參賽者的決策及行為，但學習模型的優劣大多依賴誤差大小判定，但誤差分析結果可能與資料有關(Data Dependent)。有鑑於學習模型在模型選取上的不足，本文引進統計分析的模型選取及殘差檢定，以實證資料、配合電腦模擬評估學習模型。

本文使用的實證資料，屬於囚犯困境的重複賽局(Repeated Game)，包括四種不同的實驗設定，參加賽局實驗者（或是「玩家」）為政治大學大學部學生；比較學習模型有四種：增強學習模型(Reinforcement Learning model)、延伸的增強學習模型(Extend Reinforcement Learning Model)、信念學習模型(Belief Learning Model)、加權經驗吸引模型(Experience-Weighted Attraction Model)。實證及模擬分析發現，增強學習模型較適合用於描述囚犯困境資料，無論是較小的誤差或是適合度分析，增強學習模型都有較佳的結果；另外，也發現玩家在不同實驗設定中的反應並不一致，將玩家分類後會有較佳的結果。

關鍵詞：囚犯兩難、重複賽局、學習模型、蒙地卡羅模擬、適合度檢定

Abstract

In game theory, the optimal strategy (or equilibrium) of one-shot games usually can be solved theoretically. But, the optimal strategies of repeated games are likely not unique and are more difficult to find. For example, the defection is the optimal decision for the one-shot Prisoner Dilemma (PD) game. But for the repeated PD game, if the players can benefit from cooperation between rounds then the defection won't be the only optimal rule. In recent years, economists design game experiments to explore the behavior in repeated games and use the learning models to evaluate the player's choices. Most of the evaluation criteria are based on the estimation and prediction errors, but the results are likely to be data dependent. In this study, we adapt the model selection process in regression analysis and apply the idea to evaluate learning models. We use empirical data, together with Monte Carlo simulation, to demonstrate the evaluation process.

The empirical data used are repeated PD game, including four different experimental settings, and the players of the game are from National Chengchi University in Taiwan. Also, we consider four learning models: Reinforcement learning (RL) model, Extend Reinforcement learning (ERL) model, Belief Learning (BL) model, and Experience-weighted attraction (EWA) model. We found that the RL model is more appropriate to describe the PD data. In addition, the behaviors of players in a group can be quite different and separating the players into different sets can reduce the estimation errors.

Key Words: prisoner dilemma, repeated game, learning model, Monte Carlo simulation, goodness-of-fit

目錄

表目錄

圖目錄

第一章　前言

第一節　　研究動機

賽局理論(Game Theory)是經濟學的一個分支，用於研究人類的決策及行為，有助於了解人們如何制訂策略，在日常生活中有許多賽局理論的應用。賽局理論主要在探討的問題是雙向互動，我的計算必須考慮你的計算，而你的計算也考慮了我的計算，為一門研究 「多人決策」（包括兩人）之間的問題。賽局中的每一個人的決策，會受到賽局中其他人的影響。個人的報酬不只根據自己的選擇，也取決對手的決定，因此使得賽局理論更加複雜。例如：市場上只有甲乙兩家公司，且生產的產品同質，若整個市場對產品的需求量固定，甲乙公司分別可以選擇製造高產量或是低產量，下表列出兩家公司不同策略所能得到的報酬。

表 1.1、市場報酬矩陣

乙公司 / 甲公司	高產量	低產量
高產量	(50，50)	(120，20)
低產量	(20，120)	(80，80)

註：(A，B)分別代表(甲公司，乙公司)的報酬　　單位：萬元

根據表 1.1 的報酬表，如果雙方選擇的是高產量，此時供給大於需求會導致價格變低，產量變高卻只能各獲利 50 萬元。如果甲公司選擇高產量，乙公司選擇低產量，生產高產量的甲公司可以獲得龐大利益 120 萬元，選擇低產量的乙公司只能獲得 20 萬元；反過來說，若乙公司選擇高產量，甲公司選擇低產量，乙公司的獲利較高；如果雙方都選擇低產量，因為供給小於需求使得價格上漲，提高邊際利潤，因此雙方都可以 獲利 80 萬元。此時，若你是甲(乙)公司的決策者，

你會選擇什麼策略？

當然實際上一定不如上述問題那麼簡單，但是多少可藉由簡化問題獲得有用的想法。由上例來看，如果乙公司選擇的是高產量的決策，甲公司該如何應對呢？如果甲公司選擇高產量，公司就會獲利 50 萬元，如果選擇低產量，公司僅能獲利 20 萬，相較之下理性的總經理應當會選擇高產量的決策；如果乙公司選擇的是低產量的決策，甲公司若選擇低產量，可以獲利 80 萬元，若選擇高產量則可以獲利 120 萬元，此時甲公司的總經理應當還是會選擇較高報酬的高產量。綜觀以上分析，不管乙公司選擇什麼策略，甲公司都會選擇高產量，對乙公司而言也是如此，因此最後雙方都會選擇高產量的決策，這也是著名的奈許均衡(Nash Equilibrium)。奈許均衡主要應用在非合作博弈，假設人是理性且自利，因此只選擇對自己最有利的策略，不將對手的報酬列入考量。以上例來看，不管對手的公司選擇為何，選擇高產量可獲得較高報酬，因此「高產量」就是均衡或是最佳決策。

表 1.2、囚犯困境償付矩陣

犯人乙 / 犯人甲	坦承	沉默
坦承	(-8，-8)	(0，-10)
沉默	(-10，0)	(-1，-1)

註：(A，B)分別代表(犯人甲，犯人乙)的報酬　　單位：年

上述兩家公司的範例，屬於眾所皆知的囚犯困境(Prisoner Dilemma)賽局。常見的敘述為有兩個囚犯，在一次案件中同時被逮捕，警方分別將兩人分開偵訊，如表 1.2 所示，若兩人都坦承，分別會被判八年刑期；若一人坦承一人沉默，坦承的人因轉為證人而無罪釋放，而沉默的人必須負擔十年刑期；如果兩人同時保

持沉默，則因罪證不足僅能以較輕微的罪名起訴，兩人只會被判一年刑期。從奈許均衡理論推測，若雙方為理性且自利的，最終都會選擇坦承。

從單一回合的囚犯困境，可以從理論基礎找到平衡點，但重複的囚犯困境就又不同了，以上述生產的例子來說，假設兩家公司的決策僅會影響一季的生產量，未來長久的日子裡，兩間公司每一季都會遇到相同的問題，雙方也都知道長期選擇低產量是對雙方最好的策略，但是每一次都陷入該選擇高產量還是低產量的決策問題，因為你不知道另一個搭擋到底是競爭對手，還是合作夥伴，此時若你是公司的總經理，又該如何做決策呢？

像這樣重複的囚犯兩難問題，可能存在兩個以上的最佳決策，有別於奈許均衡的單一平衡點，雙方有可能會為了長期合作而選擇合作。實證上，經濟學家以學習模型(Learning Model)解釋人在重複賽局的學習行為，但過去研究選擇模型時，通常是依據配適結果或交叉驗證的標準，缺乏統計分析比較有系統的模型選擇，因此本文希望引進迴歸分析的操作模式選取最佳學習模型。

第二節　研究目的

如果挑選學習模型以較小變異數或最大概似估計量為優劣準則，可能會和迴歸分析遭遇類似的問題，變數愈多、誤差愈小（或是 R^2 愈大），但 R^2 最大不見得是最適合的模型，還必須檢查殘差是否符合迴歸模型假設，如：服從常態分佈、資料獨立、變異數同質等。但是重複賽局的資料有時間先後關係，因此資料不服從獨立假設，增加分析的複雜性，因此建議以電腦模擬解決。

理論 ➡ 實驗設計 ➡ 電腦模擬

賽局一開始探討理論較多，從簡單的零和賽局證明「壞中取小」定理，到之後非合作博弈賽局證明奈許均衡，都是賽局理論之一。賽局理論主要探討的是單

回合的賽局，單回合的賽局通常存在均衡點，而重複賽局僅以理論較難證明存在單一均衡點。又或許可以推論有單一均衡點，但實際上人的決策卻未必會與理論結果一致，因此經濟學家透過實驗，記錄人在重複賽局中的決策行為，觀察實際上人的決策變化，並以學習模型解釋人的決策行為，此為實驗經濟。而除了實驗數據，本文也將電腦模擬應用在賽局上，探討理論上學習模型的特性。

本文實證的資料，從實驗經濟出發，利用實證資料帶入模型，與以往實驗不同的是多人重複的囚犯困境賽局，以及實證資料的遊戲規則。本文希望可以提出其他輔助選取模型的判斷標準，因此除了傳統的配適結果判斷模型的優劣以外，還希望可以透過蒙地卡羅模擬，檢驗資料是否符合模型假設，從符合假設的模型中，選取較佳的模型。其次，透過敏感度分析，希望可以找到表現最穩健的模型，如此即使判斷失誤，資料並非來自判斷的模型，也不會與實際結果差異太大。

本文第二章將介紹文獻探討，以及本文使用的學習模型、蒙地卡羅模擬法。第三章會介紹本文使用的實證資料以及研究方法。第四章是資料分析，第一節介紹學習模型的必要性，之後分成四個情境資料，探討不同模型在不同情境下的表現，以及探討學習模型的敏感度分析。第五章則是結論及建議的部分。

第二章　文獻探討

本文希望探討修正估計學習模型以及驗證模型假設，本章會先介紹賽局的發展，接著第一節介紹學習模型，包括學習模型的發展過程和模型本身的意義，第二節介紹蒙地卡羅模擬法，以及如何應用蒙地卡羅驗證本文模型假設。

賽局理論的發展是從 1928 年馮諾曼(J. von Neumann)首先從證明「壞中取小」定理，此定理僅適用在零和(Zero-Sum)賽局中，此賽局一方的虧損，會恰好等於另一方的報酬，雙方總利益為 0。在 1944 年馮諾曼和摩根斯坦(O. Morgenstern)出版他們合著的《賽局理論與經濟行為》(Theory of Games and Economic Behavior)，才開始注意到賽局理論可以用來分析許多經濟問題。

賽局的發展，繼而由電影「美麗境界」主人翁的奈許(J. F. Nash Jr.)提出『Non-cooperative Games』博士論文，以研究「多人非合作」之賽局為論述，並提出著名的「奈許均衡」的概念，奈許均衡主要應用在典型的囚犯困境，在理想的情況下，個人理性追求自身的利益得到的理論結果，即達到奈許均衡(也稱為非合作均衡)。因此奈許於 1994 年與哈桑尼(J.C.Harsanyi)及賽爾登(R.Selton)等賽局理論研究者，共同獲得諾貝爾經濟學獎。

理論上，單一的囚犯困境會達到奈許均衡，但是在固定對手的重複賽局中，玩家擁有懲罰對手的能力(即下次選擇不合作)，雖然這樣的懲罰較消極，但如果雙方玩家都長期選擇不合作對彼此都不會達到最大利益，因此有些玩家會傾向選擇合作釋出善意，達成長期合作雙贏的目標。

由上可知重複賽局有可能存在兩個以上平衡點，因此經濟學家希望透過實驗，了解人在重複賽局中的決策變化，並利用學習模型解釋人的學習行為。在兩人或多人賽局，受試者對其他受試者所形成的信念或許會隨著時間而改變，此一過程被解釋成信念學習；受試者根據過去經驗也許會跳離導致糟糕報償的決策，朝向

高報償的決策方向，此種過程被稱為增強學習。

第一節　學習模型

有些學習模型，玩家會根據過去的經驗來決定未來的決策，如：增強學習模型(Reinforcement Learning Model)、延伸的增強學習模型(Extend Reinforcement Learning Model)。學習模型會預測對手的選擇，以決定自己下一步最佳選擇，如：信念學習模型(Belief Learning Model)。此外，也有兼顧以上兩種的學習模型，如：加權經驗吸引模型(Experience-Weighted Attraction Model)。

配適學習模型需要的主要資訊是玩家在不同回合下所得到的報酬，這些模型都認為，玩家會根據過去自己的選擇以及過去遇到的玩家，進而影響下一次的選擇，本文使用的資料裡是重複的囚犯兩難賽局，表 2.1 提供了兩位玩家在不同選擇配對下的報酬，其中 C 和 D 分別代表合作以及不合作。

表 2.1、報酬矩陣

玩家 2 ＼ 玩家 1	合作(C)	不合作(D)
合作(C)	(8，8)	(1，12)
不合作(D)	(12，1)	(3，3)

註：(A，B)分別代表(玩家 1，玩家 2)的報酬　　單位：NT$

為了方便描述模型，先定義一些符號。假設玩家共有 N 人，遊戲有 T 回合，分別以 $i=1,\ldots,N$ 代表各玩家以及 $t=1,\ldots,T$ 代表各回合，$k=1,2$ 則分別代表選擇 C,D，$s_i(t)$ 代表 i 玩家在第 t 回合的選擇，$s_{-i}(t)$ 則是代表 i 玩家的對手在第 t 回合的選擇，$\pi_i(s_i^k(t), s_{-i}^{k^*}(t))$ 則是玩家 i 在第 t 回合選擇 k 遇到對手選擇 k^* 所得到的報酬，舉例來說，玩家 i 在第 6 回合選擇不合作但遇到合作的對手，其報酬為

$\pi_i(s_i^D(6), s_{-i}^C(6)) = 12$ 。

1. 隨機反應均衡(Quantal Response Equilibrium; QRE) (Richard and Thomas, 1995)：

此模型裡面，玩家選擇的不會是最佳選擇(如奈許均衡)，取而代之的是選擇"較佳"的策略，玩家在做決策時可能會犯錯，但是越嚴重的錯誤(報酬越低)，犯錯的機率就越小，其考慮對手也有可能犯錯的事實，式子如下：

$$P\left(s_i^k\right) = \frac{\exp\left(\lambda \cdot \sum_{k'} P\left(s_{-i}^{k'}\right) \cdot \pi_i\left(s_i^k, s_{-i}^{k'}\right)\right)}{\sum_{k^*} \exp\left(\lambda \cdot \sum_{k'} P\left(s_{-i}^{k'}\right) \cdot \pi_i\left(s_i^{k^*}, s_{-i}^{k'}\right)\right)}$$

$P\left(s_i^k\right)$ 為玩家 i 選擇 k 策略的機率，其形式為利用羅吉斯函數(Logit Function)轉換成機率，$P\left(s_{-i}^{k'}\right)$ 為玩家 i 預測對手會選擇 k' 的機率，參數 λ 可視為玩家對報酬的敏感度。如果玩家選擇 k 策略時，遇到的對手選擇 k' 的報酬越低，也就是 $\pi_i\left(s_i^k, s_{-i}^{k'}\right)$ 越少的話，玩家選擇錯誤(報酬較低)的策略，機率就越低，當然也和玩家預測對手選擇的機率有關，不過大致上來說玩家較不會選擇預期報酬較低的策略。

在估計參數時，由於無法得知玩家對於對手預測選擇為何，因此本文將 $P\left(s_{-i}^k\right)$ 視為所有玩家選擇 k 策略的平均機率，如果所有玩家選擇 k 策略的平均機率為 p_k ，玩家遇到選擇 k 策略的對手機率也應為 p_k ，也將此機率視為玩家預測對手選擇的機率($P\left(s_{-i}^k\right) = p_k$)。用此想法配適模型，其配適結果會剛好使預測下次選擇策略 k 的機率，即為平均機率。

2. 增強學習模型(Reinforcement learning model; RL) (Roth and Erev, 1995)：

此模型最開始是從行為心理學發展出來的，行為學家認為人們的行為，可以

由過去的學習經驗解釋，第一篇增強學習在策略以及賽局上應用的研究，是在 1955 年，由布希(Bush)和莫斯特勒(Mosteller)提出的，他們簡單定義了增強學習的規則，並且將它們應用在決定策略上，之後克羅斯(Cross, 1983)，將增強學習應用在經濟決策，非常不幸的是，一直等到十年後才有人注意到他重要的貢獻，之後陸陸續續有學者將增強學習應用在賽局上如：McAllister(1991)、Mookerjhee and Sopher(1994,1997)、Roth and Erev(1995)、Sarin and Vahid(2001)等。而最常見增強學習模型的形式如下：

$$A_i^k(t)=\begin{cases}\phi\cdot A_i^k(t-1)+\pi_i\left(s_i^k,s_{-i}(t)\right) & if \quad s_i^k=s_i(t)\\ \phi\cdot A_i^k(t-1) & if \quad s_i^k\neq s_i(t)\end{cases}$$

如果將 $I\left(s_i^k,s_i(t)\right)$ 當成指標函數，當 $s_i^k=s_i(t)$ 則為 1，$s_i^k\neq s_i(t)$ 則為 0，上式可化簡為 $A_i^k=\phi\cdot A_i^k(t-1)+I\left(s_i^k,s_i(t)\right)\cdot\pi_i\left(s_i^k,s_{-i}(t)\right)$ 。

A_i^k 為策略 k 對於玩家 i 的吸引，此值越大，代表越傾向選擇 k 的選項，可以注意到上式，當選擇 k 的策略時，所得到的報酬越高，其吸引就會增加越多，其下次選擇 k 的機率就會越高，如果選擇 k 得到的報酬很低，對於下一次選擇 k 的吸引就增加的較少，其特點是完全根據過去的經驗來調整下一次選擇策略的機率，而且如果上一次的選擇 k，下一次選擇 k 的機率只會增加不會減少，增加幅度是上一回合的報酬而定。這裡的 ϕ 值，則可以視為過去經驗的累積程度，ϕ 值越大，代表玩家較容易記起過去的經驗，ϕ 值越小，代表過去的經驗較容易被遺忘。

3. 延伸增強學習模型(Extend Reinforcement learning model; ERL)(Lai Y.H., 2005)：

以增強學習模型為基礎做細項調整，此模型認為光是以過去的經驗，似乎不足描述本文的實證資料，其修改後的式子為：

$$A_i^k(t)=\begin{cases}\phi\cdot A_i^k(t-1)+\pi_i\left(s_i^k,s_{-i}(t)\right)+\gamma\cdot\sum_{\tau=1}^{t-1}I\left(s_i^k,s_{-i}(\tau)\right) & if \quad s_i^k=s_i(t)\\ \phi\cdot A_i^k(t-1) & if \quad s_i^k\neq s_i(t)\end{cases}$$

與原本增強學習模型的不同點在於，新的項目 $\gamma\cdot\sum_{\tau=1}^{t-1}I\left(s_i^k,s_{-i}(\tau)\right)$ ，其中 γ 代表玩家過去選擇經驗的乘數，可以將此項視為玩家過去選擇的特性，因為有些玩家可能不在乎報酬到底是多少，這些玩家可能有特殊的傾向，舉個例子來說，有些玩家從頭到尾可能會全部選擇合作，或者全部選擇不合作，而延伸增強學習模型新增的項目，可以幫助捕捉那些玩家的特性。

4. 信念學習模(Belief learning model; BL)(Belief-Based Model)：

最早提出信念學習必須回溯至庫爾諾(Cournot, 1960)，他提出玩家選擇最佳策略時，是根據下一回合玩家認為其對手會選擇的策略而做調整，其最主要的思維是，玩家會傾向選擇預期最高報酬的選項，換句話說玩家只在乎對手的選擇，以此調整選擇下一次預期最高報酬的選項。

關於信念有很多不同的形式如：虛擬對策(Fictitious Play)(Brown, 1951)、庫爾諾最佳反應(Cournot Best Response) 以及加權虛擬對策(Weighted Fictitious Play)，其中加權虛擬對策，其特殊形式則包含了庫爾諾最佳反應以及虛擬對策，在加權虛擬對策，對於對手過去決策的信念表示為，過去選擇策略的次數比 s_{-i}^k ，令此為 $N_{-i}^k(0)$ ，正確的說玩家 i 的信念決定於對手選擇策略 s_{-i}^k 的次數，記為 $N_{-i}^k(0)$ ，總和為 $N(t)=\sum_{k=1}^{k_{-i}}N_{-i}^k(t)$ ，起始的信念為 $B_{-i}^k(0)=\frac{N_{-i}^k(0)}{N(0)}$

其中 $N_{-i}^k(0)>0$ 且 $N(0)>0$ ，信念則是會根據 ρ 來更新過去的資訊，並且加上實際回合玩家的選擇的組合。

$$B_{-i}^{k}(t)=\frac{\rho\cdot N_{-i}^{k}(t-1)+I\left(s_{-i}^{k},s_{-i}(t)\right)}{\sum_{k=1}^{k_{-i}}\rho\cdot N_{-i}^{k}(t-1)+I\left(s_{-i}^{k},s_{-i}(t)\right)}$$

再將 $N(t-1)$ 表示成之前的形式

$$B_{-i}^{k}(t)=\frac{\rho\cdot B_{-i}^{k}(t-1)+\dfrac{I\left(s_{-i}^{k},s_{-i}(t)\right)}{N(t-1)}}{\rho+\dfrac{1}{N(t-1)}}=\frac{\rho\cdot N(t-1)\cdot B_{-i}^{k}(t-1)+I\left(s_{-i}^{k},s_{-i}(t)\right)}{\rho\cdot N(t-1)+1}$$

當 $\rho=1$ 時，此模型就會化簡成虛擬對策，所有的觀察值都是一樣的，如果 $\rho=0$ 則是會化簡成庫爾諾最佳反應，正常來說 $0\le\rho\le1$。

可以計算在 t 回合時期望報酬為

$$\begin{aligned}E_{i}^{j}(t)&=\sum_{k=1}^{k_{-i}}\pi_{i}\left(s_{i}^{j},s_{-i}^{k}\right)\cdot B_{-i}^{k}(t)\\&=\frac{\rho\cdot N(t-1)\cdot E_{i}^{j}(t-1)+\pi_{i}\left(s_{i}^{j},s_{-i}(t)\right)}{\rho\cdot N(t-1)+1}\end{aligned}$$

這裡可以將期望報酬視為在增強學習裡面的"吸引"，而且也會跟著遊戲回合調整機率。而實際上，信念學習模型主要精神是根據玩家對對手選擇的信念來決定自己的策略，但是囚犯兩難的困境不合作的選項相較之下有絕對的優勢，而且需要對手過去的選擇資訊，因此信念學習模型不太適用在本文的實證資料。

以至於此，本文提出 Beta Binomial (以下簡稱 BB)的模型來解釋玩家以對手的選擇改變自己的策略，最原先的想法如下：

$$\begin{aligned}&X\mid p\sim B(n,p)\\&P\sim Beta(a,b)\\&\Rightarrow P\mid x\sim Beta(a+x,b+n-x)\\&\Rightarrow E(P\mid x)=\frac{a+x}{a+b+n}\end{aligned}$$

假設 P 來自 Beta 分佈，調整過後的機率期望值即為 $\dfrac{a+x}{a+b+n}$，將其改變為

學習模型的符號：

$$P_i^k(t+1)=\frac{a+x_{it}^k}{a+b+t}$$

其中 x_{it}^k 表示玩家 i 在到第 t 回合時，所遇到的玩家選擇 k 策略的次數，可以預見，當遇到選擇合作的人越多的時候，玩家應該會越傾向選擇合作，而起始值 a 、b 代表起始機率，用此式子來表達玩家以對手的選擇，調整本身策略的機率變化，並觀察是否適合應用在囚犯困境裡。

5. 加權經驗吸引模型(Experience-weighted attraction model; EWA) (Cramer and Ho, 1999)：

此模型同時包含了信念學習模型以及增強學習模型的特性，此模型有兩個要件會隨時間更新，一個是吸引 $A_i^j(t)$，一個是經驗權重 $N(t)$，經驗權重的更新如下：

$$N(t)=\rho\cdot N(t-1)+1\ ,\ t\geq 1$$

吸引的更新如下：

$$A_i^k(t)=\frac{\phi\cdot N(t-1)\cdot A_i^k(t-1)+\left[\delta+(1-\delta)\cdot I\left(s_i^k,s_i(t)\right)\right]\cdot\pi_i\left(s_i^k,s_{-i}(t)\right)}{N(t)}$$

其中報酬權重 $\left[\delta+(1-\delta)\cdot I\left(s_i^k,s_i(t)\right)\right]$ 是個關鍵，因為此項不只考慮了玩家選擇的報酬，同時也考慮玩家沒有選到有可能會得到的報酬。

加權經驗吸引模型同時包含信念學習模型以及增強學習模型，當 $\delta=0,\ N(0)=1,\ \rho=0$ 時，就化簡為增強學習模型，當 $\delta=1,\ \rho=\phi$ 時，就化簡為信念學習模型。

以上介紹五種模型裡面，有四種是學習模型，而過去吸引的初始值($A^k(0)$)通常以平均報酬決定，但是本文提出修正估計參數，此部分會在下一章節詳談。

而學習模型裡的吸引轉換成機率，通常都是以羅吉斯函數(Logit Function)轉換，轉換公式如下：

$$P_i^k(t)=\frac{e^{\lambda\cdot A_i^k(t-1)}}{\sum_{k^*} e^{\lambda\cdot A_i^{k^*}(t-1)}}$$

λ可當作玩家對於吸引力的敏感係數，當λ越高代表玩家對於吸引力的敏感度越高。

第二節　　蒙地卡羅模擬

蒙地卡羅模擬法又稱為統計模擬法、隨機抽樣技術，是一種模擬方法，以機率和統計理論為基礎的計算方法，使用隨機亂數來解決許多計算問題的方法。蒙地卡羅的方法早在西元 1777 年，法國布馮(Buffon)提出投針實驗的方法求圓周率，被認為是蒙地卡羅模擬的起源。

20 世紀高速電腦的出現，使得數學上得以用電腦快速處理大量的模擬，舉例來說，考慮單位方形內一不規則的圖形面積，蒙地卡羅模擬是一種隨機化的方式，向正方形「隨機」投擲 N 個點，落在不規則圖形內的個數為 M，當 N 很大時，圖形面積會逼近$\frac{M}{N}$，過去沒有電腦在實行此實驗可能會花費許多時間許多人力，但現在高速電腦的運算下，此實驗可能僅需幾秒就能解決。

蒙地卡羅模擬法，是基於大數法則的實證方法，當實驗的次數越多，其平均值也就會越趨近於理論值。其可以對繁複的模型設定做運算，例如投資模型組合的各種風險因子，特別是難以估算的非線性投資組合，因此使用蒙地卡羅模擬可以根據模型假設，模擬多次以後便會接近期望平均值。

應用蒙地卡羅模擬，假設資料來自某模型，根據模型的設定生成亂數後，如果資料的確能夠用模型解釋，則可將實際資料視為亂數之一，其模擬的亂數信賴

區間也應該要包含實際值，利用此特性可對模型假設做一些驗證，雖然即使實際值符合信賴區間也無法說明模型是適合的，但是若不符合信賴區間，則較能夠說明模型是不正確的。

第三章 資料介紹與研究方法

上一章節提到希望利用模型假設，產生亂數並驗證亂數信賴區間是否包含實證資料，本章第一節將介紹本文使用的賽局資料，包括其實驗流程以及實驗設定，第二節將介紹研究方法，包括實證資料特性、判斷選取模型的標準、修正估計參數以及如何驗證模型假設。

第一節 資料介紹

一、報酬矩陣以及配對規則

本文使用的資料為囚犯兩難賽局，其報酬矩陣如下：

表 3.1、報酬矩陣

玩家 2 \ 玩家 1	合作(C)	不合作(D)
合作(C)	(8，8)	(1，12)
不合作(D)	(12，1)	(3，3)

註：(A，B)分別代表(玩家 1，玩家 2)的報酬　　(單位：NT$)

表 3.1 為囚犯兩難的報酬矩陣，若玩家 1 和玩家 2 都選擇合作，皆能得到 8 元的報酬；若都選擇不合作，則只能得到 3 元的報酬；若一個選擇合作一個選擇不合作，選擇合作者僅得到 1 元，選擇不合作者能得到 12 元的報酬。其顯示會影響玩家報酬不僅只有自己的決策，對手的選擇也是非常重要的因素，實證資料是多人重複賽局，而如何將玩家配對在一起，本文設定的配對方式有下列四種：

RM(Random)：配對對象為隨機，玩家之前的選擇不影響下一次配對。

WH(Weight-History)：配對對象根據玩家前五回合的記錄，轉換成分數後，分數由高至低配對。其分數計算方式為 $T(t)=5\cdot I(t-1)+3\cdot I(t-2)+2\cdot I(t-3)+1\cdot I(t-4)+1\cdot I(t-5)$，其中 $I(t)$ 為指標函數，如果玩家選擇合作 $I(t)=1$，如果選擇不合作則 $I(t)=0$，設定此遊戲規則為採用費伯納數列當做計算 T 分數的係數，用意是希望玩家用此計算權重，較容易達到收斂。

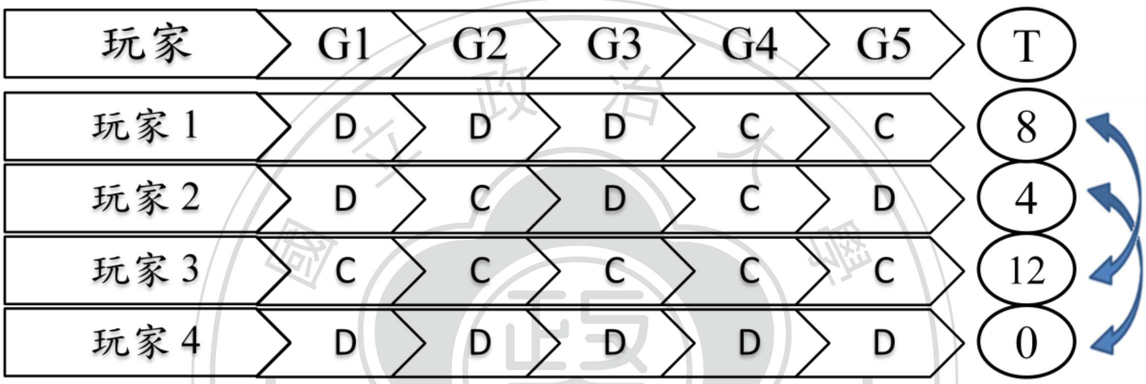


圖 3.1、WH 配對解說圖

從圖 3.1 可以看到 WH 的配對方式，以玩家 1 來說，在第 4 第 5 回合分別選擇合作，所以 T 分數是 8 分，玩家 2 分別在第 2、第 4 回合選擇合作，所以 T 分數是 4 分，玩家 3 全部選擇合作，T 分數是 12，玩家 4 全部選擇不合作，所以 T 分數是 0，計算出分數以後再以分數高低做配對，在下一回合的賽局，玩家 1 就會和玩家 3 配對，玩家 2 則是和玩家 4 做配對。

WH(c) (Weighted-History with attraction of Cooperate)：配對規則如同 WH，並告知玩家這組會有兩個固定選擇合作或不合作的玩家。最原始是由鞭子與胡蘿蔔(whip and carrot)的故事得到的想法，故事是若想使驢子往前走，當鞭子鞭策不動驢子時，便釣一根胡蘿蔔在驢子前面，誘使驢子往前走。如同此實驗設定，鞭子就像是安排兩個不合作的玩家，告訴玩家若選擇不合作必定會遇到不合作的人；胡蘿蔔就像是安排兩個合作的玩家，告訴玩家只要選擇合作，就必定會合作的對手與他們配對，以此兩種方法激勵玩家傾向選擇合作。

WH(p)(Weighted-History with Payoff Information)：配對規則如同 WH，並告知玩家現在報酬的排名。

二、實驗流程

上一小節介紹本文使用的賽局資料為多人重複賽局，分別會在四種不同的配對規則下進行，在每一個配對規則下，會有 5 組 14 人的玩家，每個配對規則總共有 70 個人進行遊戲，所以四種配對規則總共有 280 人，玩家是來自不同科系，國立政治大學的學生，玩家參加遊戲時就可得到 50 元的出席費，並且在結束時，給予玩家在賽局遊戲中得到的總報酬，玩家得到的報酬多寡，決定於在賽局遊戲中的策略為何。

實驗的進行，由玩家在電腦前進行操作，並由電腦計算後，進行下一次的配對，玩家則是會隨機分配到不同的遊戲規則。一開始會發一份說明，內容明確告知玩家此次實驗須了解的內容，包括選擇什麼樣的策略遇到什麼樣的對手，可得到多少報酬，以及不同的賽局規則，以確認玩家了解這些規則。但是該如何說明玩家的行為，是因為遊戲規則不同而改變呢？因此實驗時會先讓玩家在隨機配對的規則下進行遊戲，實驗流程如下表：

表 3.2、實驗流程

遊戲設定	階段 1	階段 2	階段 3
RM	RM-5	RM-25	RM-5
WH	RM-5	WH-25	RM-5
WH(c)	RM-5	WH(c)-25	RM-5
WH(p)	RM-5	WH(p)-25	RM-5

實驗流程如表 3.2 所示，玩家在不同遊戲規則下，會先進行 5 回合完全隨機配對的賽局，此時玩家皆不知道其分配的組別為何，接著會進行 25 回合不同規則的賽局，此時會告知玩家其所在何組，最後再進行 5 回合完全隨機配對賽局。階段 1 主要想觀察不同遊戲規則下的玩家，一開始是否是有差異的，階段 2 則是觀察不同遊戲規則下，玩家的行為是否有改變，如果階段 1 的遊戲結果顯示玩家

在各組之間是無差異的，且階段 2 玩家出現不同的行為，才能夠說明不同規則會影響玩家，階段3則是在玩家經過不同規則的遊戲之後，觀察其行為是否會改變。本文主要在探討階段 2 的資料，觀察不同遊戲規則下玩家選擇的變化。

表 3.3、階段 1 及階段 3 分組機率

	Trt2(RM)			Trt3(WH)	
group	p1(d)	p3(d)	group	p1(d)	p3(d)
6	0.714	0.786	11	0.629	0.743
7	0.557	0.686	12	0.743	0.914
8	0.657	0.900	13	0.686	0.814
9	0.629	0.914	14	0.529	0.743
10	0.643	0.771	15	0.571	0.829
avg.	0.640	0.811	avg.	0.631	0.809
std	0.057	0.096	std	0.095	0.075

表3.3取自Yang et al.(2007)，p1(D)和p3(D)分別表是階段1和階段3的不合作機率，該文章表示用Kruskal-Wallis rank sum test檢定，階段1的玩家是沒有差異的，如此便能說明，階段2玩家的行為若有改變，是因為規則設定的不同。

三、基本資料分析

上一小節提到賽局遊戲分成 3 階段進行，而本文主要分析的資料為階段 2 的資料，因為階段 2 的資料可以顯示不同遊戲規則下的影響，下表為不同情境資料的概要。

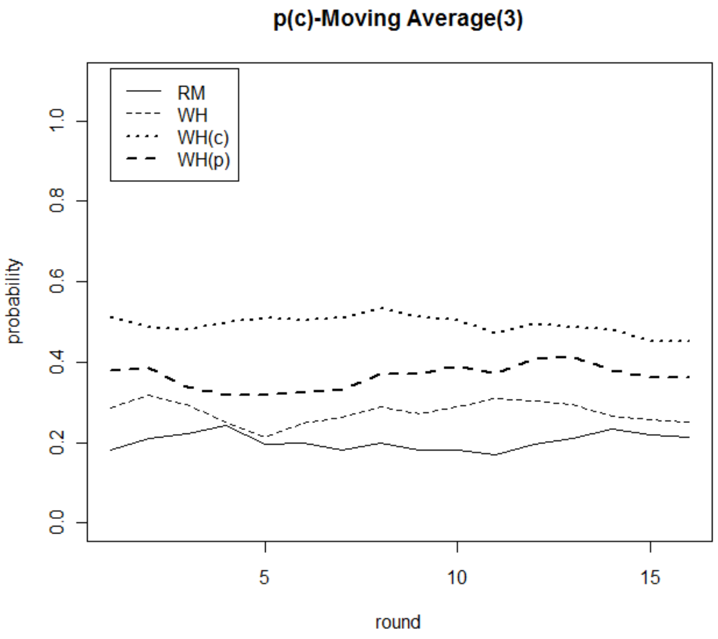
表 3.4、各情境資料概要

6~23 round	RM	WH	WH(c)	WH(p)
p(C)	0.202	0.273	0.494	0.367
p(C|C)	0.228	0.557	0.646	0.510
p(sw)	0.193	0.207	0.324	0.326

考慮玩家一開始玩家可能還在摸索配對，以及遊戲結束時會傾向選擇不合作，這裡僅將資料的 6~23 回合作分析。p(C)為平均合作機率，p(C|C)為玩家選擇合作的情況下，對手亦為合作的機率，p(sw)為轉換機率，意即玩家這次選擇與下次不同的機率。第二章提到在固定對手的重複賽局中，玩家擁有懲罰對手的能力(即下次選擇不合作)，但在完全隨機(RM)的配對下，因為玩家無固定的對手，其懲罰效果也幾乎不存在，即使是重複的囚犯困境的問題，預期玩家長期的決策也會漸漸傾向選擇不合作，從表 3.4 可以看到 RM 情境下的 p(C)是最小的，而 WH 開始到 WH(c)、WH(p)，整體的 p(C)都有提升，從這可以看出增加了配對的條件有助將玩家吸引至合作的選項，而 WH(c)與 WH(p)新增的條件也增加了玩家選擇合作的誘因，其 p(C)都比 WH 來的高。

RM 下的 p(C|C)為 0.228，與 RM 的 p(C)接近，在隨機配對下，這兩個值本應一樣，頗符合常理，而 WH(c)下的 p(C|C)卻比其他兩種情境還高出一成左右，其原因應是在 WH(c)的情境下，加入了固定選擇的玩家，如果固定選擇的玩家是合作的，p(C|C)比起其他兩個情境來的高也是符合常理的。

p(sw) 的部份 RM 跟 WH 差不多，但是在 WH(c)與 WH(p)裡 p(sw)則是增加了一成，可能是在 WH(c) 裡玩家已知會有兩個固定選項的玩家，如果是這兩個玩家是固定選擇合作的，當玩家已經到達頂端(分數最高)的時候，只要選擇不合作就有很大的機率可以與合作的對手配對，就可以賺進 12 元的酬勞，而自己的分數也不至於太低，只要在下一次繼續選擇合作，就能保持自己的分數；而如果固定玩家的選項是不合作，分數低的玩家知道選擇不合作，必然會與不合作的對手配對，就有動力選擇合作，使自己不要一直落於最底端，如此一來一往 p(sw)的機率就會提高。在 WH(c)下，則是認為當玩家看到自己的報酬不比別人高的時候就會開始轉換策略，以尋求較高的報酬策略，在轉換的過程中 p(sw)也會增加。



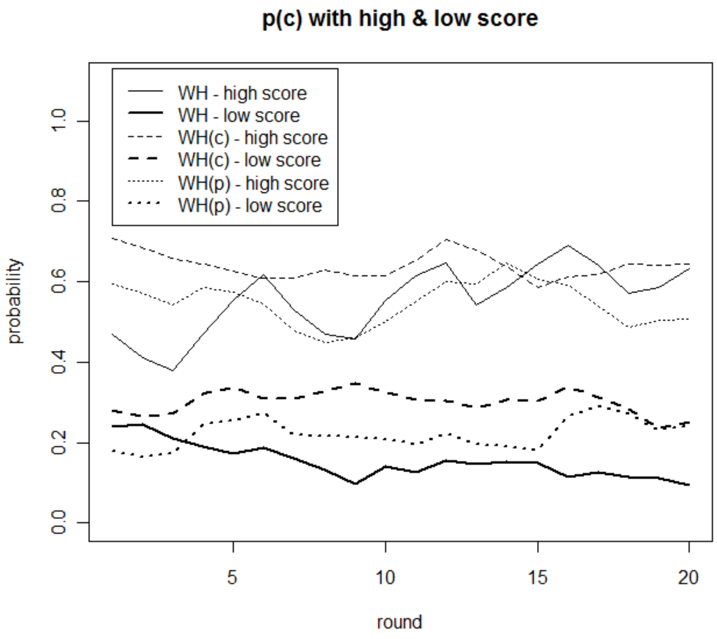
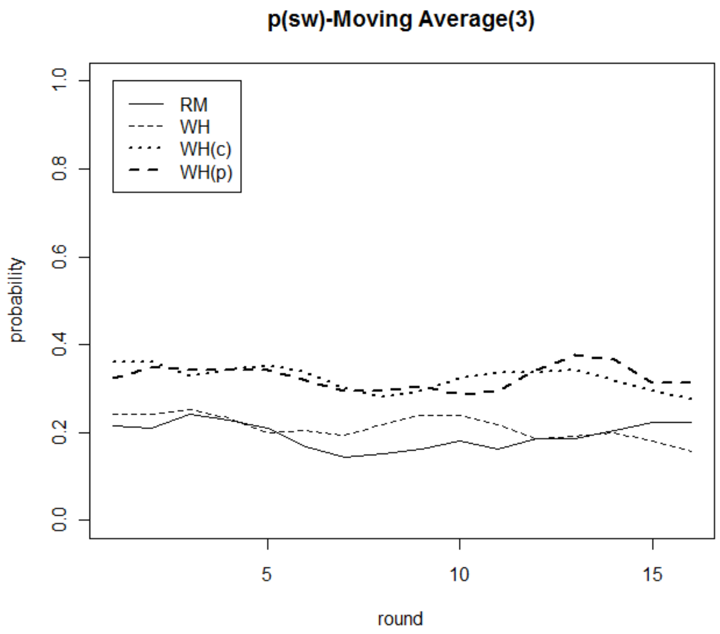


圖 3.2(a) 合作機率

圖 3.2(b) 高低分數的合作機率

從圖 3.2(a)與表 3.4 看到的一樣，以局部合作機率 p(C)來看，其在 4 種配對規則下的大小關係也會是 $WH(c) > WH(p) > WH > RM$，3.2(b)是 T 分數為 6 做界線畫出的圖，可以將 T 分數的高低視為合作與不合作的分水嶺，普遍可以看到分數較高的玩家會持續讓自己保持在頂端，所以選擇合作機率較高，分數較低的玩家，則是較難逃脫低分數的配對，所以普遍合作機率較低。



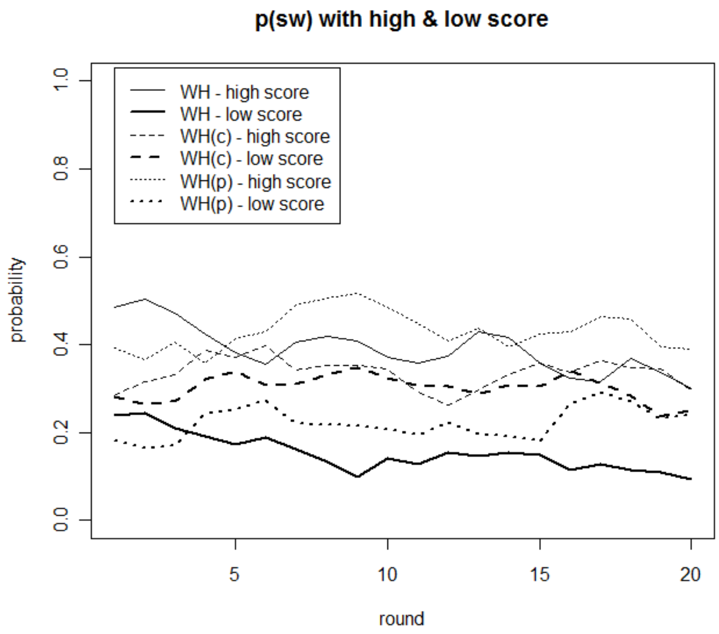


圖 3.3(a) 轉換機率

圖 3.3(b) 高低分數的轉換機率

從圖 3.3(a)可以看到 RM 與 WH 的 p(sw)是最低的，代表玩家在這兩種情境下的選擇比較不會改變，而在 WH 下可以看到 p(sw)有越來越低的趨勢，代表在這個情境裡的玩家會越來越傾向選擇某個選項，意即會越來越收斂，從圖 3.3 (b)也可以看到，WH 之下不管 T 分數是高是低，其 p(sw)都會漸漸遞減，也可以推測在這個情境的玩家一旦落入兩極(意即合作配對或不合作配對兩區)就會漸漸穩定下來，另一個方面在 WH 下，T 分數高的玩家會較分數低的玩家 p(sw)高，也可以推測有投機的玩家，當自己分數較高的時候就會選擇不合作，以賺取更多的酬勞。

從圖 3.3 (a)可以看到 WH(c)與 WH(p)的 p(sw)是差不多的，但是從圖 3.3 (a)看到 WH(c)不管 T 分數是高是低，其 p(sw)都差不多，表示不管在哪裡的玩家都會一直改變策略，即使在低分區的玩家也是如此，而 WH(p)的情況與 WH 較類似，分數高的玩家較容易轉換策略。

第二節　研究方法

過去通常以模型配適或是交叉驗證的結果選取模型，除了這些本文希望驗證模型的假設是否正確，在符合假設條件下選取較佳的模型，以及修正模型估計參數，希望修正後的模型對於資料會有較佳的配適結果，

一、判斷準則

過去判斷學習模型的標準通常有兩種，分別是：

1. 最小均方差、最大對數概似量或 AIC、BIC

通常可以以兩個不同的方法估計模型參數，分別是均方差(Mean Square Deviation, MSD)(Erev and Roth, 1998)和最大對數概似量(Maximum Log Likelihood, MLL)(Camerer and Ho, 1999)。

將吸引值轉換成機率，這就是對於下一次選擇的預測，均方差的概念就類似迴歸的最小平方法，要找到使均方差(MSD)最小的參數，也就是 $\min\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{k=1}^{2}\frac{[P_i^k(t)-I(s_i^k,s_i(t))]^2}{T\cdot N\cdot 2}$，舉例來說，如果玩家的選擇是$(C,D,C)$，預測選擇合作機率為 $(0.7,0.2,0.6)$，則其對平方合的貢獻為 $2\cdot\left[(0.7-1)^2+(0.2-0)^2+(0.6-1)^2\right]$。

對數概似量 $LL=\sum_{t=1}^{T}\sum_{i=1}^{N}\ln\left(\sum_{k=1}^{2}I\left(s_i^k,s_j\left(t\right)\right)\cdot P_i^k\left(t\right)\right)$，以上例來說，$LL=\ln\left(0.7\right)+\ln\left(0.8\right)+\ln\left(0.6\right)=\text{-}1.09$。

最小平方法的重點在於使變異數及偏差平方之和為最小，而最大概似估計法則是需要正確的模式，然而重複賽局之下，各賽局的結果往往相依而使模式更加複雜。因此本文選擇使用最小平方法，除了比較不同模型的最小均方差以外，還可以比較 AIC、BIC，只是 AIC 及 BIC，都是建立在概似函數上，因此對 AIC 及 BIC 僅當作參考。

2. 交叉驗證

在統計學上經常使用這個技巧，將資料分割成兩部分，分別是估計資料(Training Data)與驗證資料(Testing Data)，將交叉驗證的技巧應用在學習模型上，主要是在看模型的預測能力，將資料的前部分的回合則設定為估計資料，後部分的回合設定為驗證資料，利用估計資料的配適結果預測驗證資料的結果，計算其均方差，若驗證資料的均方差越小，則代表模型的預測能力越好，則越傾向選擇該模型。

除了上述兩種判斷標準，本文尚提出 F^* 檢定以及驗證模型假設，增加選取模型的判斷標準。如下：

1. F^{*} 檢定

將迴歸的 F 檢定套用在學習模型裡，檢定方式如下：

H_0 : Reduce model
H_1 : Full model

$\alpha = 0.05$

$$if \quad F^{*} = \frac{\dfrac{SSR_F - SSE_R}{Df_R - Df_F}}{\dfrac{SSR_F}{Df_F}} > F_{0.05}\left(Df_R - Df_F, Df_F\right) \qquad , reject \quad H_0$$

舉個例子說明，假設一筆資料有 20 個玩家，分別以增強學習模型以及延伸的增強學習模型配適模型，初始值以平均報酬為設定，玩家的人數當作自由度，得到的殘差平方和(Residual Sum Of Square,RSS)分別是 $SSR = 95.45$(增強學習模型)、$SSE = 73.82$(延伸的增強學習模型)，可以將增強學習模型視為縮減模型(Reduce Model)，使用了 2 個參數，$Df_R = 18$，延伸的增強學習模型視為完整模型(Full Model)，使用了 3 個參數，$Df_F = 17$，則 F^{*} 檢定如下：

H_0 : Reduce model
H_1 : Full model

$$F^{*} = \frac{\dfrac{SSR_F - SSE_R}{Df_R - Df_F}}{\dfrac{SSR_F}{Df_F}} = \frac{\dfrac{95.45\text{-}73.82}{18-17}}{\dfrac{73.82}{17}} = 4.98 > F_{0.05}\left(1,17\right) = 4.45$$

$\alpha = 0.0$ 的標準判斷，拒絕 H_0 的虛無假設，表示延伸的增強學習模型是有顯著的效果。這裡必須修正一件事情，若 SSR 與 SSE 為卡方分佈，則資料之間彼此為獨立且符合常態，但是這裡用學習模型配適的資料是多人的賽局，人與人之間的選擇，直觀上來看必然不符合獨立條件，因此這裡的 F^{*} 值有可能不符合 F 分佈，不過可以用蒙地卡羅模擬去找出近似的臨界值，這部分在第四章會詳談。

2. 驗證模型假設

在過去學習理論的實證分析，不管是用最小均方差或者用最大對數概似量估計參數，僅能比較模型彼此之間的好壞，也就是僅能比較配適模型後，最小均方差或最大對數概似量的大小，或者以交叉驗證中驗證資料的預測結果好壞來判斷。但是對於適合度分析卻沒有太多著墨，就好像不同的迴歸配適模型，相較之下，總是會有 R^2 較高的模型，或者可以觀察 $R^2_{adj.}$ 、AIC、BIC，除了比較這些，還是必須觀察殘差，檢定是否來自常態、是否彼此之間獨立、是否有變異數同質性等等的假設，檢定結果必須符合種種假設才可以說這個迴歸模型是合理的。

賽局學習模型裡，較複雜的模型如：延伸的增強學習模型、加權經驗吸引模型，其配適結果的最小均方差勢必會比增強學習模型較好，或者交叉驗證的結果之中，必然有一個是之中最好的，但是光是以這樣的標準選取模型，似乎不夠謹慎，除了以此標準選取模型，還期望能夠驗證模型假設是否正確。

賽局學習模型並不如迴歸模型有一些常態分配的假設，勉強只能稱其為二項分布，而且玩家與玩家之間的選擇並不獨立，玩家上一次的選擇與下一次的選擇也不獨立，以至於此，二項分佈無法套用在賽局學習模型，為了解決分佈未知、不獨立的問題，所以使用蒙地卡羅模擬法。

本文希望利用蒙地卡羅的模擬驗證模型假設，若賽局實證資料來自某學習模型，根據此學習模型的設定產生 1000 次亂數，若假設為真，可將實證資料視為其中一筆模擬亂數，因此亂數特性的信賴區間(如：參數)也應包含實證資料，利用蒙地卡羅模擬法，希望可以檢查下列的信賴區間是否包含實際值：1. 外在的表現，如：合作機率、2. 最小均方差(MSD)、3. 參數估計值，如：ϕ、λ.....等。

縱觀以上的判斷準則，過去選取標準為：1.最小均方差、AIC、BIC 2.交叉驗證。以及本文提出的輔助判斷：1. F^* 檢定 2.驗證模型假設。因此挑選模型時，在符合模型假設的條件下，選取其餘三種判斷標準(MSD、交叉驗證和 F^* 檢

定)皆有較佳表現的模型，但若三種判斷結果皆異時，擇優先選取交叉驗證的判斷標準，因為最小均方差與模型複雜度有關，AIC 與 BIC 則是建立在概似量的函數上，前面提到概似量需要模型假設，因此賽局資料不太適用。

二、修正估計

本文主要修正估計參數的想法有兩個，第一個為設定初始值為參數，第二個則是對玩家分組，並給予不同初始值。第一種設定初始值為參數，在一開始配適學習模型時，嘗試了許多初始值的設定，結果有好有壞，較難挑選出較好的初始值設定方式，而且在下一節提到的給予玩家不同初始值的部分，如果僅以人為設定會有困難，因此本文將初始值設定為參數之一。

在模型介紹裡提到的學習模型，$A^k(0)$ 初始值選取往往都是以根據報酬多寡來給定，通常以該選項的平均報酬當作初始值，但是以報酬來決定初始值似乎不是那麼恰當，每個人對於報酬的吸引不會一致，根據式子推導：

$$P_i^C(t) = \frac{e^{\lambda \cdot A_i^C(t-1)}}{e^{\lambda \cdot A_i^C(t-1)} + e^{\lambda \cdot A_i^D(t-1)}} = \frac{1}{1 + e^{\lambda \cdot (A_i^D(t-1) - A_i^C(t-1))}}$$

可以看到會影響下一回合的機率值只決定於 $A^D(t) - A^C(t)$，所以在選定初始值時令 $A^D(t) - A^C(t)$ 為一參數，藉此找出最佳初始值以更能配適模型。

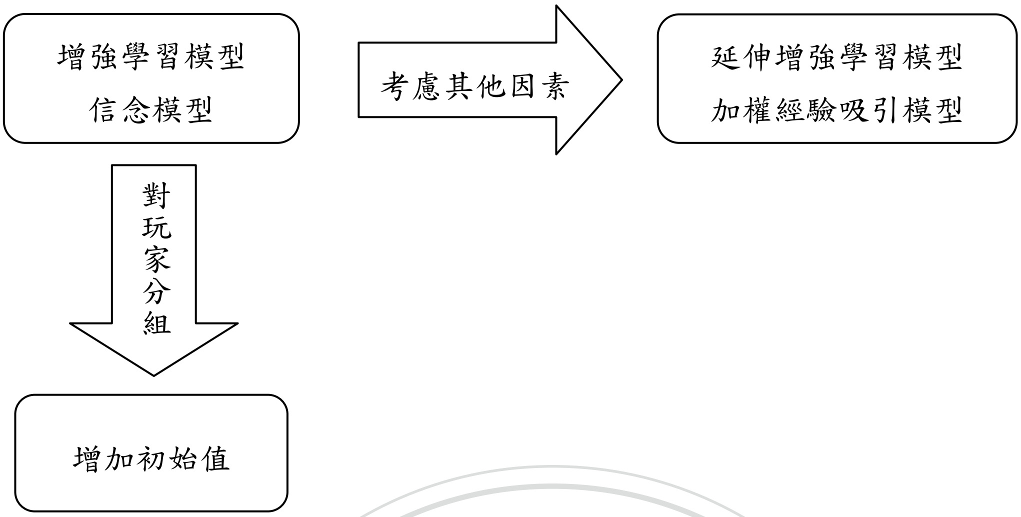


圖 3.4、模型流程圖

第二種增加初始值，過去學習理論希望有更好的配適結果，通常會考慮其他因素，如圖 3.4 模型流程圖所示，但是除了將右鍵的方向以外，本文認為可以將玩家分組，因為玩家的特性可能不一，有些玩家傾向選擇合作，有些玩家傾向選擇不合作，而學習模型的初始值，可將其視為玩家對於選擇策略的傾向，因此將選擇傾向不同的玩家做分組，應該會對模型配適有所幫助。

表 3.5、卡方獨立性檢定

Number of C		RM - Round 6~23				Chi-squared test
		0~2	3~6	7~18		P-value
Round 1~5	0~1	28	10	6	44	0.1458
	2~5	11	7	8	26	
		39	17	14	70	

Number of C		WH - Round 6~23				Chi-squared test
		0~2	3~6	7~18		P-value
Round 1~5	0~1	19	10	5	34	0.0545
	2~5	11	12	13	36	
		30	22	18	70	

Number of C		WH(c) - Round 6~23			Chi-squared test
		0~6	7~18		P-value
Round 1~5	0~2	22	12	34	<0.001
	3~5	5	31	36	
		27	43	70	

Number of C		WH(p) - Round 6~23			Chi-squared test
		0~6	7~18		P-value
Round 1~5	0~1	31	7	38	<0.001
	2~5	9	23	32	
		40	30	70	

前面提到將玩家分群可能有助估計參數，如果分群是有幫助的，則玩家一開始的選擇，和之後的選擇應該會相關，所以將玩家前 5 回合的選擇的合作次數以及後 18 回合的合作次數分組，表 3.5 是將玩家分組，檢定 WH(c)、WH(p)時，由於有區塊的數字小於 5，因此將 round 6~23 的 0~2 與 3~6 合併。分別檢定不同情境下，玩家在前面的選擇與之後的選擇是否有關，WH(c)和 WH(p)是顯著的，認為在這兩個遊戲設定下玩家的表現特性不一，一開始選擇合作的玩家到後期選擇合作的意願也較高，給予不同的初始值應有助估計參數。

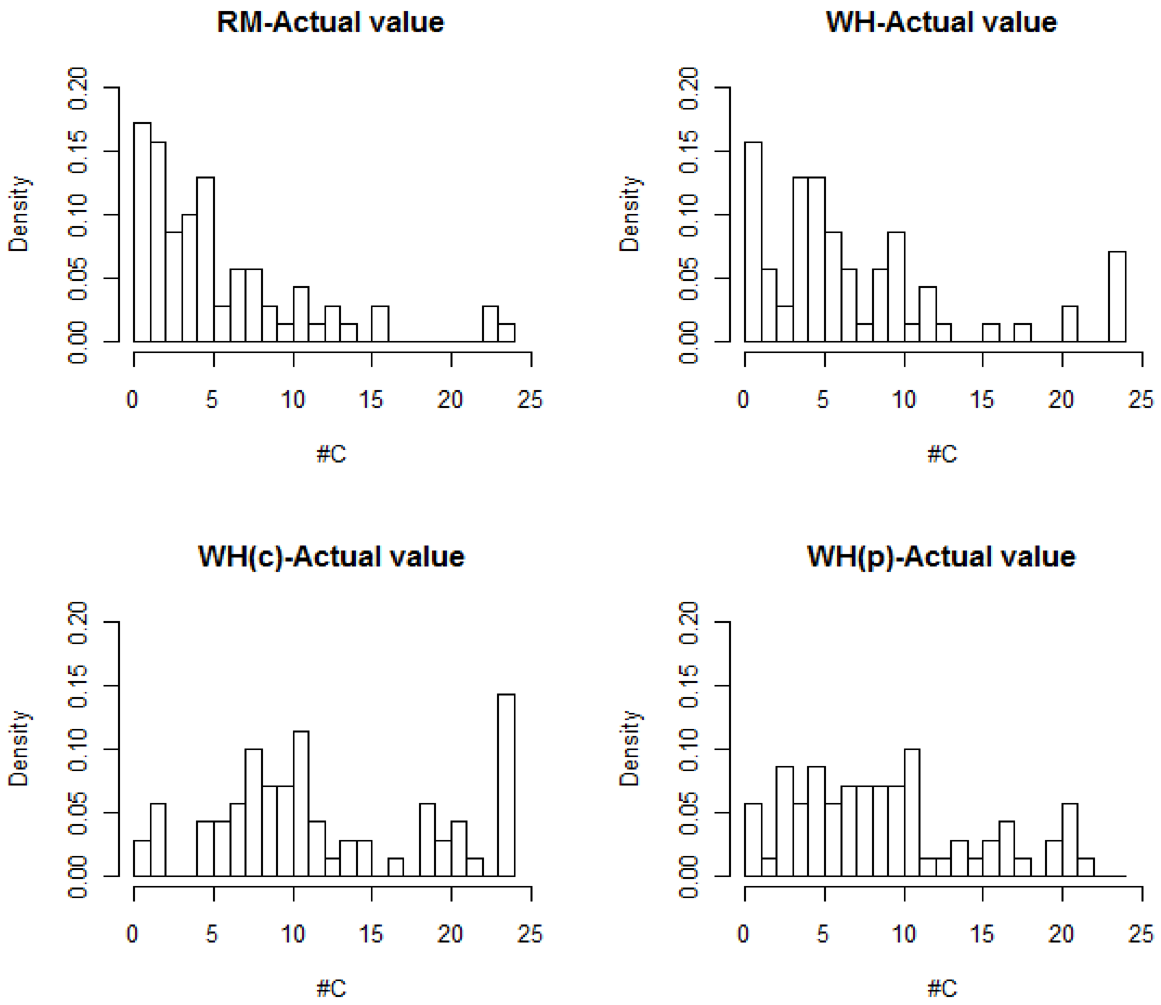


圖 3.5、合作次數直方圖

圖 3.5 為玩家的合作次數直方圖，可以看到在不同規則下玩家的表現是有差異的，為了方便看出差異，將其修勻以後，畫成機率密度圖，如下；

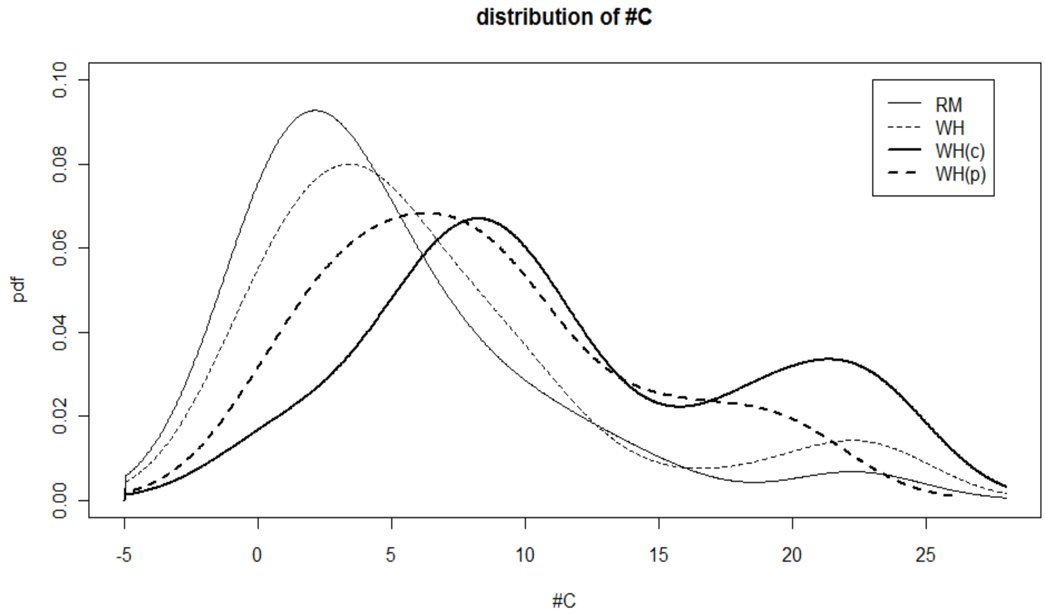


圖 3.6、合作次數機率密度

圖 3.6 為修勻後的合作次數密度圖，越往右邊的玩家越傾向選擇合作，越往左邊的玩家越傾向選擇不合作，可以看到在 RM 裡，幾乎所有玩家都傾向選擇不合作，只有少部分的人仍然持續選擇合作，這少部分的人可能天生善良，有利他的行為，但是這畢竟是少數，大多數的玩家還是會選擇不合作，從這也可以知道有些玩家並不是完全理性的，甚至如果根據賽局理論來推論，不應該會出現選擇任何一次合作的玩家，若是嚴格來說，這次參與 RM 遊戲規則的玩家，大部分都是不理性的，因此也可以看出純粹的賽局理論在這裡似乎是不適用的。

再看到 WH 的曲線，可以看到已經有一小部分的人願意選擇合作，代表這設定下的玩家，的確會因為遊戲規則願意選擇合作；再看 WH(c) 的曲線，很明顯可以看到雙峰的樣子，右峰的人傾向選擇合作，左峰的人傾向選擇不合作，但是也可以注意，這遊戲規則下大部分的人的合作意願都有增加，整個曲線比起 RM、WH，有較往右移動的趨勢；在 WH(p)下，就比較沒有完全選擇合作的玩家存在，雖然很多人選擇不合作，但是也可以看到有一部分的人聚集在選擇次數 10~20 附近，這些人可能稍微傾向選擇合作，但是偶爾還是會欺騙對手。而整個從這個圖看起來，RM 跟 WH 大部分的玩家都傾向不合作，所以可以預期對這兩

種資料分組效果應該不大，也可以發現 WH(c)似乎很明顯可以分成兩組，預期對 WH(c)的分組效果會較明顯，而 WH(p)就處在分組與不分組的界線，或許分組效果會不錯，或許效果不彰。而本文對於玩家分組的依據為玩家在遊戲中的合作次數，合作次數較高的人分在一組，合作次數低的人分至另一組，以此畫分不同傾向的玩家。

第四章　實證資料分析

上一章介紹了基本資料特性，以及如何選取及驗證模型，本章節將介紹實證資料的分析，包括模型配適結果和模擬結果，根據上一章的判斷準則的標準選取模型，在假設為正確的條件下，選取最佳的模型。在此之前，配適賽局資料時經常直接使用學習模型配適，但是學習模型是否真的會配適的較好，或許像是以平均概念的最佳反應均衡配適模型就已足夠，因此一開始先比較最佳反應均衡以及最簡單的學習模型。

第一節　學習模型的必要性

在第二章介紹的模型包括 QRE 與其他學習模型，或許像是使用 QRE 這種以平均的方式配適資料就足以描述的資料，因此接下來試著比較 QRE 以及最簡單的學習模型 RL。

表 4.1、RM 配適結果

RM	QRE	QRE 2	RL	RL 2
ϕ	—	—	0.8149	0.9472
λ	0.5404	0.8143&-0.1996	0.1137	0.0369
$A^D(0)-A^C(0)$	—	—	7.7674	40.90 & -23.75
#parameter	1	2	3	4
MSD	0.1673	0.1303	0.1331	**0.1265**

註：畫底線者為該項最佳模型。

表 4.1 中，QRE 2 代表隨機最佳反應均衡的模型將人分成兩組，RL 2 則是增強學習模型將玩家分別給予兩組不同初始值，因 QRE 沒有初始值的問題，直接將玩家分成兩組配適模型，所以會有兩個 λ 值。

可以看到結果 QRE 2 的 MSD 較 RL 小，代表將人分成兩組，即使是無學習

效果的模型，也能有效降低 MSD，雖然 RL 的 MSD 較大，但是此模型將玩家一開始的特性都設定為一樣，而最後得到與分組差不多的結果，代表使用有學習效果的模型仍然是有效果的，而在 RM 資料裡，分組的學習模型就較無有效降低 MSD。

表 4.2、RM 交叉驗證

RM	QRE	QRE 2	RL	RL 2
ϕ	—	—	0.8162	1.0000
λ	0.5442	0.8195 & -0.1902	0.1162	0.0215
$A^D(0)-A^C(0)$	—	—	7.5487	63.30& -24.77
MSD-training 1~18 round	0.1665	0.1301	0.1354	0.1263
MSD-testing 19~23 round	0.1700	0.1311	**0.1207**	0.1323

註：畫底線者為該項最佳模型。

交叉驗證以資料的 1~18 回合當做估計資料(Training Data)，19~23 回合當作驗證資料(Testing Data)，觀察模型在驗證資料的預測能力如何。從表 4.2 交叉驗證分析來看，QRE 和 QRE2 在驗證資料的預測上沒有太大的幅度改善，而 RL 在驗證資料上的 MSD 則是縮減了許多，反而 RL 2 在交叉驗證上的表現較 RL 差，但不管怎樣 RL 模型的表現會比 QRE 來的好。

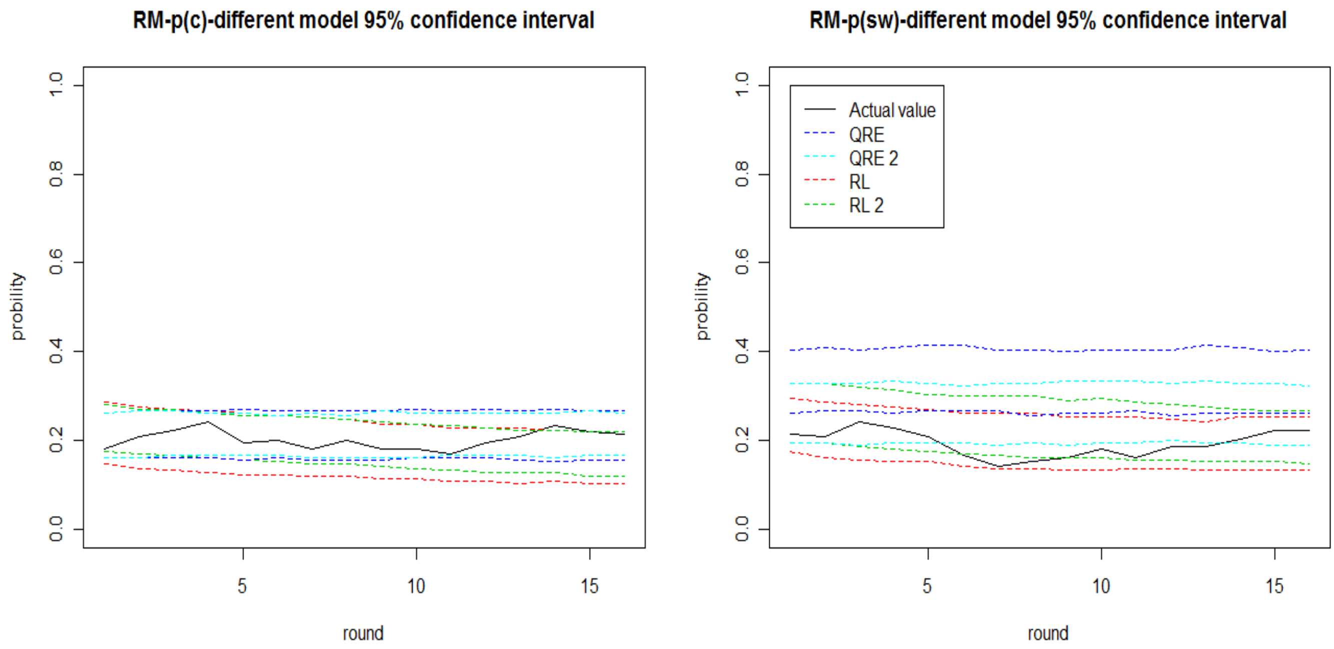


圖 4.1、RM-95%信賴區間

從模擬的角度來看，雖然 QRE 和 QRE2 從 p(C)的角度來看皆表現不錯，但是看 p(sw)的圖，QRE 的選擇機率有始至終都一樣，所以 p(sw)也都一樣，可以看到 QRE 2 的 p(sw)會比 QRE 低，因為 QRE 2 將玩家分成兩組，合作機率不同組別分別的 p(sw)都會較低，舉例來看，如果分成一組的合作機率是 0.4，p(sw)就會是 0.48，如果分成兩組的合作機率分別是 0.8、0.2，p(sw)就分別會是 0.32、0.32，所以平均下來分組的 p(sw)會較低，但即使是 QRE 2 在 p(sw)的模擬信賴區間仍不包含實際值。

在符合模型假設的條件下(信賴區間應包含實際值)選取模型，這裡僅有 RL 模型大致上符合，因此認為四種模型比較下來，RL 模型會較合適，也說明了學習模型在 RM 資料裡的確有其必要性。不過也可以注意到一點，RL 2 雖然配適結果 MSD 較小，但是模擬結果反而沒有比較好，這也可以說明，只以 MSD 選取模型似乎不是那麼恰當的，也可以看到 RL 的模擬特性 p(C)與 p(sw)都會隨著時間遞減，p(sw)遞減代表資料有收斂的趨勢，而 p(C)遞減代表玩家選擇漸漸收斂至不合作，接著探討學習模型在 WH 資料型態是否是有必要的。

表 4.3、WH 配適結果

WH	QRE	QRE 2	RL	RL 2
ϕ	—	—	0.8454	0.9720
λ	0.3563	0.5858 & -0.6270	0.0873	0.0283
$A^D(0)-A^C(0)$	—	—	6.4812	50.38 &-14.72
#parameter	1	2	3	4
MSD	0.2041	0.1478	0.1470	**<u>0.1398</u>**

註：畫底線者為該項最佳模型。

在條件配對的資料裡，QRE 2 的 MSD 與 QRE 比起來的確有減少許多，但 QRE 2 的配適結果和 RL 已相差不多，分組的效果和學習模型的效果幾乎一樣，RL 2 依舊是 MSD 最小的模型，只是學習模型分成兩組似乎沒有使 MSD 降低太多。

表 4.4、WH 交叉驗證

WH	QRE	QRE 2	RL	RL 2
ϕ	—	—	0.8332	1.0000
λ	0.3409	0.5708 & -0.5131	0.0871	0.0221
$A^D(0)-A^C(0)$	—	—	6.7125	50.26-26.67
MSD-training 1~18 round	0.2071	0.1543	0.1582	0.1474
MSD-testing 19~23 round	0.1935	0.1300	**<u>0.1069</u>**	0.1234

註：畫底線者為該項最佳模型。

從表 4.4 交叉驗證結果來看，QRE 與 QRE 2 的預測能力都不怎麼理想，RL 預測能力則是之中最好的，RL 2 的預測能力在 WH 的表現和在 RM 一樣都不如 RL。接著觀察模擬的結果。

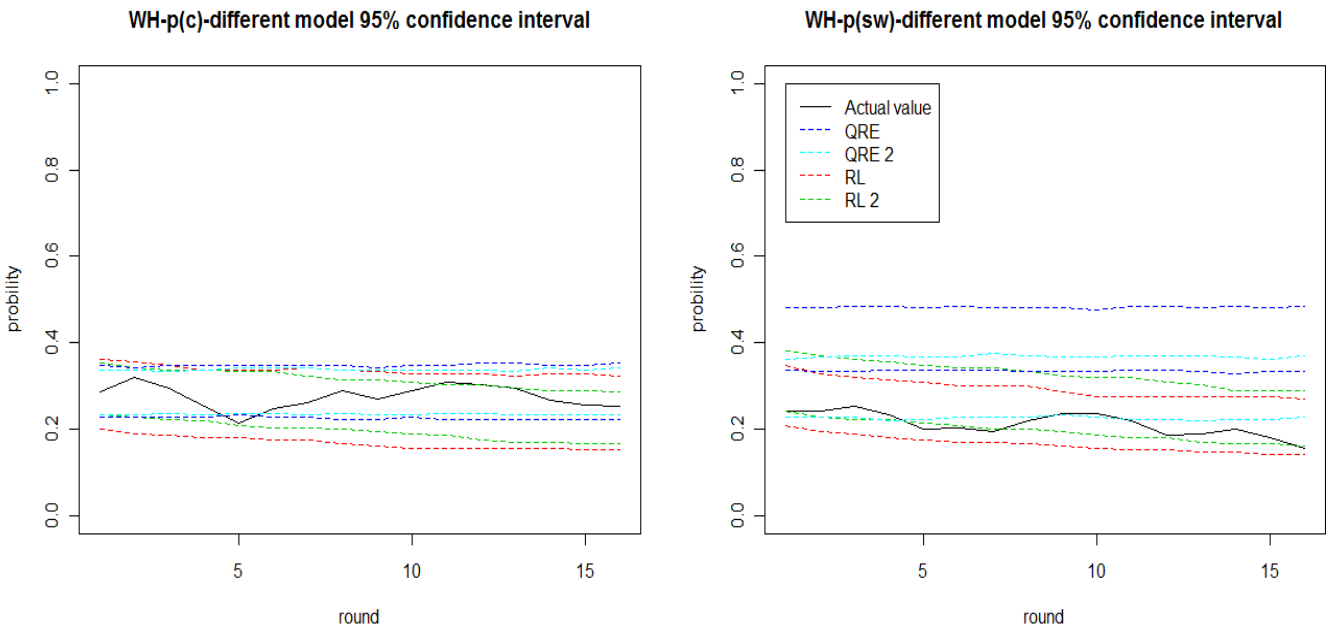


圖 4.2、WH-95%信賴區間

從圖 4.2 觀察 QRE 與 QRE 2 的模擬結果，在 p(C)依舊是有很不錯的結果，但是在 p(sw)也和在 RM 一樣不甚理想，RL 2 在這裡的模擬結果也是不如 RL，RL 的模擬則是很漂亮的包含了 p(C)、p(sw)的實際值，相同地，在符合模型假設的條件下選取最佳的模型，因此在 WH 資料裡，RL 模型應會比其他三種適合，所以不管是 RM 亦或是 WH，使用學習模型的確是有助配適資料，而接下來的WH(c)以及 WH(p)的資料型態用 QRE 配適的結果非常不理想，所以就不討論，而從以上的分析發現不管是隨機配對或者是條件配對的情況，學習模型比起沒有學習效果的模型都會是較佳的選擇。

第二節　　情境資料-RM

表 4.5、RM 配適結果

RM	BB	RL	RL 2	ERL	EWA
ϕ	—	0.8149	0.9472	0.7776	0.7881
λ	—	0.1137	0.0369	0.0754	0.0621
γ	—	—	—	0.3875	—
ρ	—	—	—	—	0.1442
δ	—	—	—	—	-0.6038
a	41.82	—	—	—	—
b	155.06	—	—	—	—
$A^D(0)-A^C(0)$	—	7.7674	40.90 & -23.75	14.1331	18.6041
Initial p(C)	0.2124	0.2925	0.181 & 0.706	0.2562	0.2395
#parameter	2	3	4	4	5
MSD	0.1675	0.1331	**<u>0.1265</u>**	0.1307	0.1319
AIC	1671.335	1400.724	**<u>1328.768</u>**	1370.96	1381.060
BIC	1675.832	1407.469	**<u>1337.762</u>**	1379.954	1392.303

註：畫底線者為該項最佳模型。

RM 資料的配適結果，從初始合作機率來看，可以發現在 RM 裡一開始大都還是傾向選擇不合作的，代表在 RM 裡選擇合作相較之下是比較沒有優勢的，而從分組配適結果來看，總共 70 個玩家，大約有 12 人是屬於傾向合作的，這些人的初始合作機率高達 0.7，可以將這些人看成利他主義者，較不會欺騙對手，所以初始合作機率較高，但是這畢竟不多，大部分的人還是會傾向選擇不合作。

從配適結果的 MSD 看來，RL 2 是所有模型中表現最好的，即使是較複雜的模型，如：ERL、EWA，其 MSD 還是會比 RL 2 大，從 AIC、BIC 的角度判斷，亦是 RL 2 較佳，總觀配適結果，RL 2 似乎會是最後選擇的模型。

表 4.6、RM 交叉驗證

RM		BB	RL	RL 2	ERL	EWA
ϕ		—	0.8162	1.0000	0.7737	0.7769
λ		—	0.1162	0.0215	0.0709	0.0629
γ		—	—	—	0.6124	—
ρ		—	—	—	—	0.1797
δ		—	—	—	—	-0.5476
a		41.70	—	—	—	—
b		155.83	—	—	—	—
$A^D(0)-A^C(0)$		—	7.5487	63.30& -24.77	15.0206	17.6430
MSD-training	1~18 round	0.1676	0.1365	0.1263	0.1362	0.1354
	6~18 round	0.1577	0.1207	0.1104	0.1171	0.1194
MSD-testing	19~23 round	0.1715	0.1211	0.1323	0.1233	**<u>0.1207</u>**

註：畫底線者為最佳。

從表 4.6 交叉驗證的角度來看，情境 RM 下，EWA 在驗證資料(Testing Data)上的表現最佳，不過也可以看到 RL 與 ERL 的結果也還不錯。

過去經濟學家希望決定出最佳的模型，通常會以配適結果或者是交叉驗證的結果來判斷，以 RM 來看，配適結果的 MSD、AIC 與 BIC 都判斷 RL 2 會較佳，而如果以交叉驗證判斷，則是會選擇 EWA 的模型，但就如前面提到，這樣的判斷方式非常依賴資料特性，有可能同樣的設定下，實驗兩次配適的結果或者交叉驗證的結果可能會不同，所以除了以這種方式判斷，希望可以用蒙地卡羅模擬亂數，以模型假設產生亂數，如果資料的確來自此模型，可以將實際值視為模擬亂數的其中一筆，而模擬亂數的信賴區間也應該包含實際值，其中可以去看基本性質、局部基本性質、參數等特性。

表 4.7、RM-基本性質信賴區間

RM	p(C)		p(CC)		p(sw)	
Actual value	0.202		0.046		0.193	
BB	0.1881	0.2365	0.0222	0.0508	0.2962	0.3708
RL	0.1420	0.2254	0.0142	0.0587	0.1731	0.2294
RL 2	0.1690	0.2214	0.0190	0.0540	0.1983	0.2614
ERL	0.1524	0.2556	0.0175	0.0699	0.1605	0.2210
EWA	0.2952	0.4111	0.0762	0.1794	0.2622	0.3303

註：網底者表信賴區間包含實際值。

表 4.7 為整體的基本性質信賴區間，表中各個模型在三種機率中會有兩個數字，分別代表模擬 95%信賴區間的上下界，BB 在 p(C)和 p(CC)還不錯，但是卻在 p(sw)差比較多，用過去對手的來調整自己的選擇似乎能解適合作機率的部分，但是對於轉換機率卻無法描述的很好；RL 與 ERL 全部的信賴區間包含實際值，RL 2 則是僅差了一點而以；而 EWA 雖然是當中最複雜的模型，但卻與實際值差異最大，由此也可以看到，EWA 的模型假設似乎不符合 RM 的資料。而除了整體的基本資料，也試著去看局部的基本資料特性。

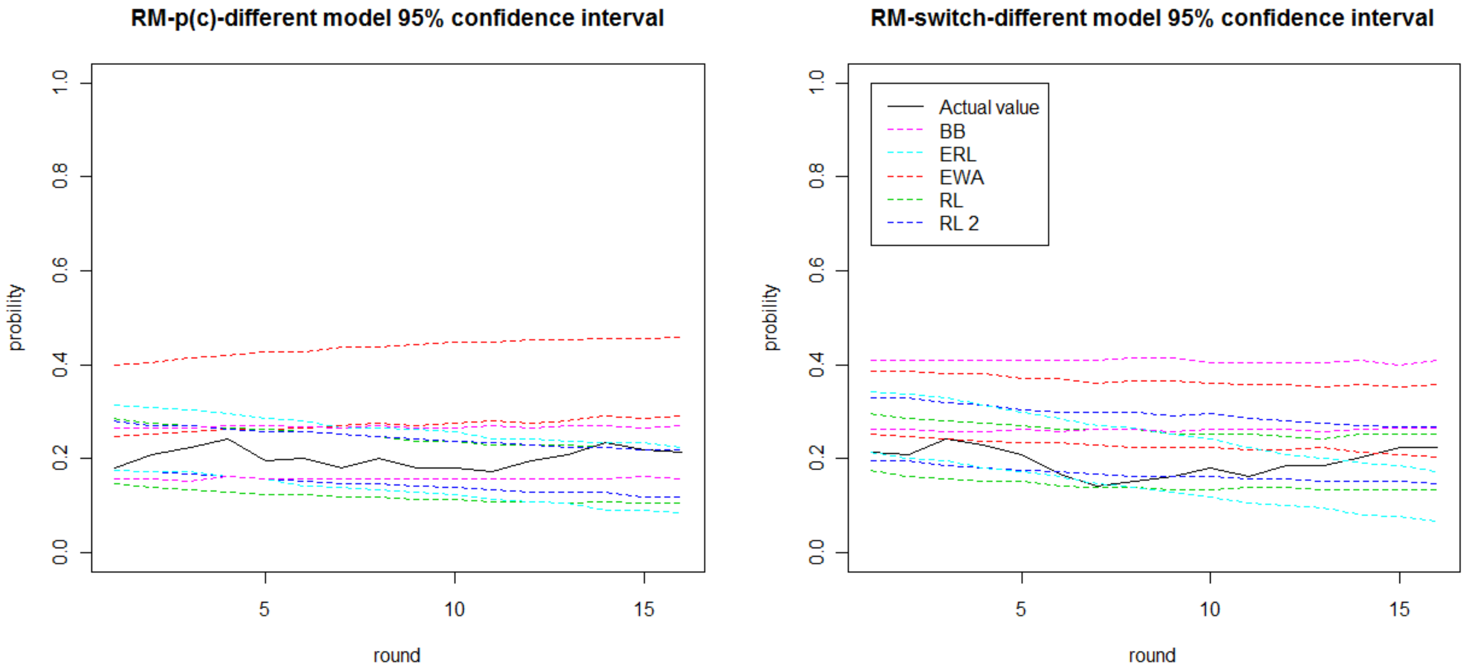


圖 4.3、RM-局部基本性質信賴區間

觀察圖 4.3 局部基本性質的信賴區間，可以發現 BB 在合作信賴區間裡還不錯，EWA 模型則是差異最大，與表 4.7 觀察到的一致；從細部觀察可以看到 ERL 與 RL 2 的模擬值，在 p(C)的表現還算可以，但在 p(sw)就已經有點偏離實際值，所以認為這兩個模型並不是很合適，最後看到 RL 的信賴區間，不管在 p(C)或是 p(sw)，RL 的模擬值幾乎都涵蓋實際值，用 RL 模擬結果去驗證模型的基本資料以及局部基本資料特性，都算是有很合理的結果，因此 RL 模型似乎是比較適合當作配適 RM 資料的模型。接著再觀察參數信賴區間。

表 4.8、 RM-參數信賴區間

Model：RL	RM		
percent	2.5%	Actual	97.5%
ϕ	0.7638	0.8149	0.8693
λ	0.0865	0.1137	0.1440
$A^D(0)-A^C(0)$	4.6998	7.7674	12.4157
MSD	0.1142	0.1331	0.1379

表 4.8 為以 RL 模型為假設產生亂數，再將亂數重新估計參數，所得到的參數信賴區間。如同之前所述，如果此筆資料來自 RL 模型，可以將其視為產生的其中一筆亂數，其參數值也可視為其中一筆亂數所產生的，如此產生的參數信賴區間應包含實際參數值，而從上表我可以看到，以 RL 模型為假設的亂數信賴區間也都有包含實際值，表示 RL 的模型假設是合理的。

對於增加初始值是否較好，之前的章節提到可以用 F^* 檢定為另一個判斷標準，下表則是 RM 資料的 F^* 檢定。

表 4.9、RM - F^* 檢定

模型	RL	RL 2	模擬 F 值($\alpha=0.05$) =3.544
殘差平方和	214.33	203.68	
自由度	67	66	$F_{0.05}(1,66)$=3.9862
F^*	--	3.4510	

表 4.9 的 F^* 即為之前公式所計算出的值，模擬 F 值是以 RL 為合理模型的假設下，產生 1000 筆亂數，再分別以 RL 以及 RL 2 模型配適亂數資料，如此便會得到 1000 個 F 值，如果模型假設為正確，實際值所計算出的 F^* 應落在亂數產生 F 值的 95%信賴區間內，如果落在信賴區間內，變有更充足的證據說 RL 是較合適的模型，反之，則比較偏向 RL 2 的模型，這裡模擬的 F 值為 3.544，可以看到和 $F_{0.05}(1,66)$ 也相差不多，代表應該算是一個合理的值，而檢定結果 $F^* = 3.451 < 3.$[illegible]，所以接受 RL 是較合適的模型，與前面的結果一致。

第三節　情境資料-WH

表 4.10、WH 配適結果

WH	BB	RL	RL 2	ERL	EWA
ϕ	—	0.8454	0.9720	0.7664	0.7551
λ	—	0.0873	0.0283	0.0645	0.0580
γ	—	—	—	0.4908	—
ρ	—	—	—	—	0.1991
δ	—	—	—	—	-0.5049
a	1.1773	—	—	—	—
b	2.8870	—	—	—	—
$A^D(0)-A^C(0)$	—	6.4812	50.38 &-14.72	10.6908	12.8937
Initial p(C)	0.2897	0.3621	0.194 & 0.603	0.3341	0.3213
#parameter	2	3	4	4	5
MSD	0.1697	0.1470	**<u>0.1398</u>**	0.1438	0.1454
AIC	1660.122	1500.958	**<u>1424.718</u>**	1459.735	1490.195
BIC	1664.619	1507.703	**<u>1433.712</u>**	1468.729	1501.438

註：畫底線者為該項最佳模型。

配適 WH 的初始機率較 RM 高了一點，分組後較傾向選擇合作的人也增加至 25 人，而這裡傾向選擇合作的機率約 0.6，卻反而比 RM 還低，這是因為在配

適模型時，是以合作次數做為分組依據，理所當然合作次數高的就分配至傾向選擇合作的，合作次數低的就分配至傾向選擇不合作，而 RM 分組的界線為 9，WH 為 6，也就是說在 RM 裡總共選擇合作次數低於 9 的玩家，就將他們分到傾向不合作區，反之，高於 9 的玩家，將其分至傾向合作區，而 WH 是以 6 為界線，所以傾向合作區的合作機率會稍低於 RM，當然還必須考慮玩家選擇合作的比例，這裡就不詳談，只是 RM 傾向合作的機率大於 WH 是有可能的。

WH 的配適結果，不論是從 MSD、AIC 或 BIC 判斷，都是 RL 2 的模型較佳，到這裡也不難發現，RL 2 的 ϕ 都會比其他模型大，代表對過去經驗累積較難遺忘，因為在分組的時候給予不同初始值，而可以看到 RL 2 的初始值比起其他模型都來的大，對於不同特性的完家，一開始就給予相當高的差異，玩家對於自己的特性相對來說也較難遺忘。

表 4.11、WH 交叉驗證

WH		BB	RL	RL 2	ERL	EWA
ϕ		—	0.8332	1.0000	0.7513	0.7804
λ		—	0.0871	0.0221	0.0655	0.0485
γ		—	—	—	0.6014	—
ρ		—	—	—	—	0.1812
δ		—	—	—	—	-0.6981
a		1.3932	—	—	—	—
b		3.3212	—	—	—	—
$A^D(0)-A^C(0)$		—	6.7125	50.26&-26.67	10.4645	16.8309
MSD-training	1~18 round	0.1795	0.1582	0.1474	0.1571	0.1582
	6~18 round	0.1595	0.1407	0.1287	0.1387	0.1407
MSD-testing	19~23 round	0.1230	0.1069	0.1234	**0.1041**	0.1056

註：畫底線者為最佳。

WH 的交叉驗證，認為 ERL 會是較好的模型，而這裡可以看到不管是哪種模型，其驗證資料的 MSD 都會減小許多，這是因為學習模型對於收斂的資料配適結果會較好,從之前 WH 的 p(sw)會漸漸遞減看來,玩家的選擇會越來越穩定，所以驗證資料的 MSD 會較小。

表 4.12、WH-基本性質信賴區間

WH	p(C)		p(CC)		p(sw)	
Actual value	0.273		0.152		0.207	
RL	0.1825	0.3159	0.0429	0.1825	0.1975	0.2605
RL 2	0.2198	0.2984	0.0603	0.1524	0.2286	0.2924
BB	0.2389	0.3452	0.0397	0.1032	0.3298	0.4422
ERL	0.2087	0.3468	0.0683	0.2096	0.1782	0.2445
EWA	0.3770	0.5040	0.1603	0.3095	0.3000	0.3765

註：網底者表信賴區間包含實際值。

WH 的模擬資料與 RM 的結果幾乎一樣，只有 RL 與 ERL 是全部包含實際值的，而 RL 2 則是 p(sw)沒有落在信賴區間內，其他的 BB 在合作機率上依舊表現不錯，但是其他還是不符合，EWA 則是相差太多，從整體信賴區間看來似乎是僅有 RL 與 ERL 符合假設，而 RL 2 則是介於邊界。

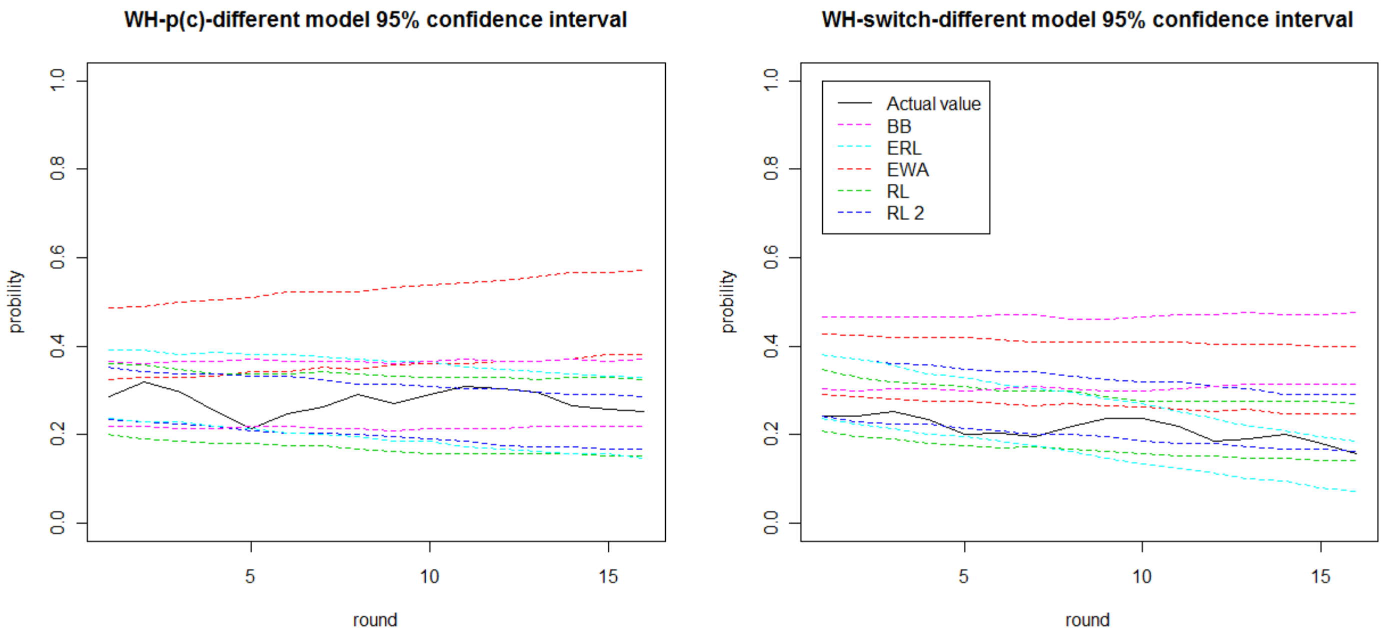


圖 4.4、WH-局部基本性質信賴區間

觀察圖 4.4WH 局部信賴區間，BB 描述合作機率仍然有不錯的效果，EWA 偏離實際值太多，暫時不討論，RL 2 在 p(sw)雖然大部分都符合，但是仍有少部分的回合偏離信賴區間，而也能看到 RL 與 ERL 的 95%信賴區間皆符合假設，但是 ERL 在 p(sw)下降速度有點略快，而且 RL 的參數較少，屬於較簡單的模型，所以在假設條件都符合的情況下，會傾向選擇較簡單、參數較少的模型，所以 RL 是較適合描述 WH 資料的模型。

表 4.13、WH-參數信賴區間

Model：RL	WH		
percent	2.5%	Actual	97.5%
ϕ	0.7979	0.8454	0.8871
λ	0.0689	0.0873	0.1084
$A^D(0)-A^C(0)$	2.9668	6.4812	10.8983
MSD	0.1294	0.1470	0.1566

以 RL 模型為假設所產生的參數信賴區間，其實際值也全部都落在 95%的信賴區間內，表示模型假設是可以被接受的。

表 4.14、WH - F^* 檢定

模型	RL	RL 2	模擬 F 值($\alpha = 0.05$) =3.740
殘差平方和	236.670	225.078	
自由度	67	66	$F_{0.05}(1,66)$=3.9862
F^*	--	3.399	

F 檢定結果 $F^* = 3.399 < 3.740$，即不拒絕 RL 為較佳模型的假設，與前面得到的結果也一致，在 WH 資料裡，分組的效用也不明顯，使用一個初始值的 RL 即足夠描述資料。

第四節　　情境資料-WH(c)

表 4.15、WH(c)配適結果：

WH(c)	BB	RL	RL 2	ERL	EWA
ϕ	—	0.8882	0.9005	0.8932	0.9058
λ	—	0.0453	0.0332	0.0456	0.0608
γ	—	—	—	-0.0239	—
ρ	—	—	—	—	-0.0953
δ	—	—	—	—	0.1632
a	1.5343	—	—	—	—
b	1.6896	—	—	—	—
$A^D(0)-A^C(0)$	—	-6.3403	37.81 & -34.92	-6.3984	-4.3610
Initial p(C)	0.4759	0.5714	0.222 & 0.761	0.5724	0.5659
#parameter	2	3	4	4	5
MSD	0.2132	0.1858	**0.1735**	0.1858	0.1846
AIC	1976.544	1767.405	**1686.052**	1769.235	1756.553
BIC	1981.041	1774.150	**1695.046**	1778.229	1767.796

註：畫底線者為該項最佳模型。

從 WH(c)的配適結果可以看到，初始值都是負的，表示其合作機率大於 0.5，代表在 WH(c)裡玩家一開始選擇合作的意願是較高的，而分組的效果，分組界線為 8，傾向合作組的人數高達 42 人，合作機率為 0.761，比 RM 與 WH 都來的高，傾向不合作組的機率與 RM 與 WH 比起來則是偏高一點，但相差不多。從 MSD、AIC 與 BIC 的準則可以判斷 RL 2 會較佳，雖然這裡 ERL 與 RL 的 MSD 皆為 0.1858，但 ERL 的 MSD 的確略低於 RL 的，只是四捨五入後數字都為 0.1858，ERL 中的 γ 幾乎趨近於 0，代表玩家過去選擇特性在 WH(c)裡幾乎是沒有效用的，而 EWA 的估計結果與 RL 幾乎一樣，所以模型複雜化此筆資料說是較無效果的。

表 4.16、WH(c)交叉驗證

WH(c)		BB	RL	RL 2	ERL	EWA
ϕ		—	0.8983	0.9428	0.9202	0.9064
λ		—	0.0474	0.0236	0.0491	0.0764
γ		—	—	—	-0.1275	—
ρ		—	—	—	—	-0.1334
δ		—	—	—	—	0.2529
a		1.7093	—	—	—	—
b		1.8056	—	—	—	—
$A^D(0)-A^C(0)$		—	-6.1431	27.29&-65.22	-6.3353	-3.8021
MSD-training	1~18 round	0.2189	0.1886	0.1734	0.1870	0.1886
	6~18 round	0.2073	0.1802	0.1717	0.1795	0.1802
MSD-testing	19~23 round	0.1929	0.1768	**0.1760**	0.1777	0.1768

註：畫底線者為最佳。

從交叉驗證的結果看來是 RL 2 較佳，RL、ERL 和 EWA 其驗證資料的 MSD 都非常接近，所以如果只以交叉驗證來判斷模型的話，也很難非常肯定哪種模型是最好的。

表 4.17、WH(c)-基本性質信賴區間

WH(c)	p(C)		p(CC)		p(sw)	
Actual value	0.494		0.319		0.324	
RL	0.3627	0.4968	0.1532	0.3071	0.3084	0.3857
RL 2	0.4190	0.5190	0.2039	0.3341	0.3058	0.3874
BB	0.3984	0.5532	0.1190	0.2714	0.4138	0.4979
ERL	0.3603	0.4968	0.1492	0.3063	0.3109	0.3840
EWA	0.2817	0.4397	0.1143	0.2857	0.2336	0.3118

註：網底者表信賴區間包含實際值。

整體基本性質的信賴區間，僅 RL 2 全部包含實際值，RL 與 ERL 雖然 p(CC)的信賴區間不包含實際值，不過也只差一點而已，所以 RL 與 ERL 可以再觀察看看，而 BB 雖然 p(C)包含實際值，但是 p(sw)與實際值相差太多，EWA 則是三個都不在信賴區間內，暫時不考慮此模型。

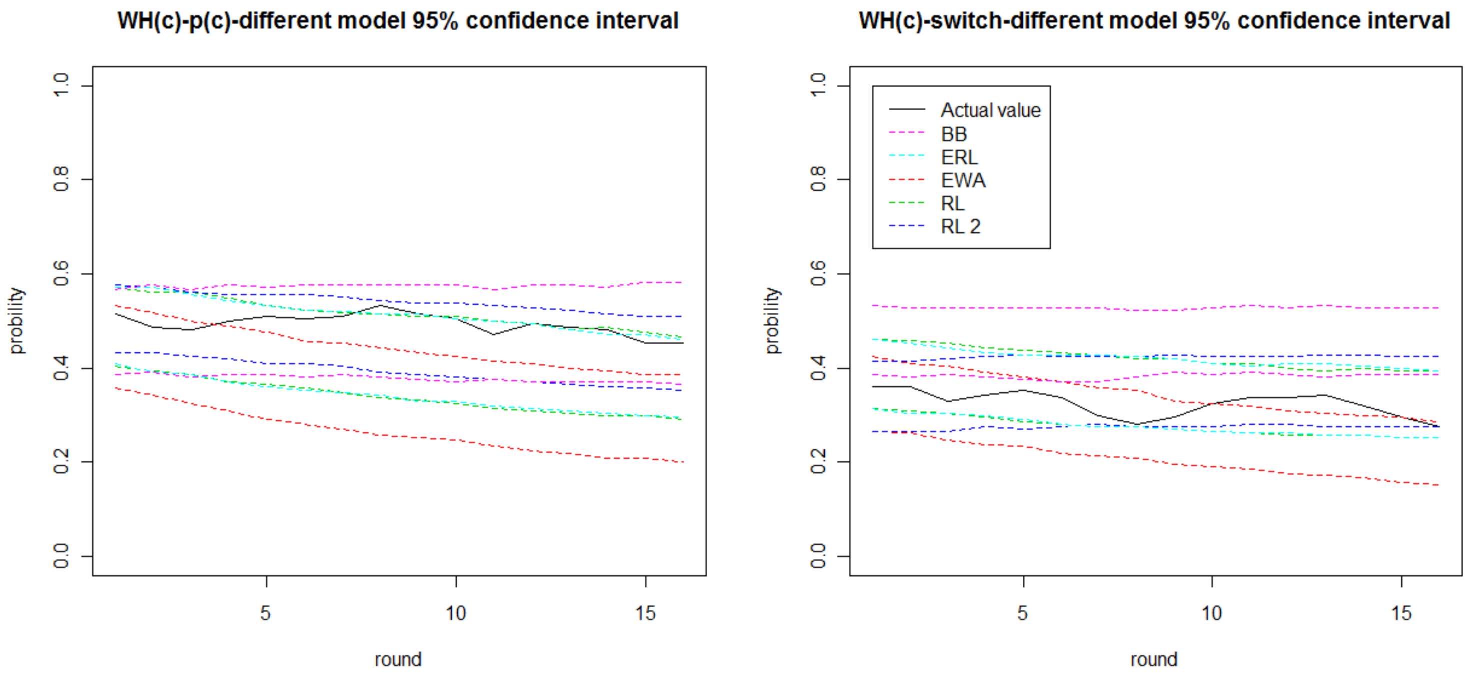


圖 4.5、WH(c) -局部基本性質信賴區間

觀察 WH(c)的局部基本性質信賴區間，BB 僅在 p(C)表現得不錯，EWA 則是都偏離太多，仔細觀察可以發現 RL 與 ERL 信賴區間的線幾乎是重疊的，先前的配適結果兩個模型的 MSD 幾乎一樣，在這裡模擬結果也幾乎一樣，更能說明 ERL 的玩家過去選擇特性在 WH(c)幾乎是沒有效果的，不過也可以發現，雖然 RL 與 ERL 大部分都包含了實際值，但沒有 RL2 來的好，從這兩張圖看起來以及之前的分析，RL 2 應當是 WH(c)表現最好的模型。

表 4.18、WH(c)-參數信賴區間

Model：RL 2	WH(c)		
percent	2.5%	Actual	97.5%
ϕ	0.8887	0.9005	0.9537
λ	0.0176	0.0332	0.0351
$A^D(0)-A^C(0)$	27.42	37.81	63.54
	-86.21	-34.92	-33.51
MSD	0.1685	0.1735	0.1977

以 RL 2 為模型假設產生亂數，所有參數的信賴區間都包含實際值，連傾向合作組與傾向不合作組的初始值也都包含在裡面，表示 RL 2 是個可以被接受的模型，其模型假設也都合理。

表 4.19、WH(c) - F^* 檢定

模型	RL	RL 2	模擬 F 值($\alpha=0.05$) =3.476
殘差平方和	299.138	279.335	
自由度	67	66	$F_{0.05}(1,66)$=3.9862
F^*	--	4.679	

模擬的 F 值為 3.476，在 RM 與 WH 模擬的 F 值分別為 3.544 與 3.740，可以看到模擬的 F 值是非常接近的，三種不同實驗設定下的模擬 F 值都非常接近，或許這意味著真的有一個臨界值在，只是目前理論無法推導出此值為何。

從 F 檢定結果來看，$F^*=4.679>3.476$，即拒絕 RL 為較佳模型的假設，在 WH(c)的資料裡，選擇 RL 2 的模型是較好的，與前面的分析結果一致。這裡因為 EWA 與 BB 的模擬結果差異太大，所以不考慮分組的 EWA 與 BB，而 ERL 則是因為與 RL 幾乎一樣，故分組效果應會與 RL 一致。

第五節　情境資料-WH(p)

表 4.20、WH(p)配適結果：

WH(p)	BB	RL	RL 2	ERL	EWA
ϕ	—	0.7844	0.8880	0.7476	0.5738
λ	—	0.0719	0.0332	0.0637	0.4372
γ	—	—	—	0.1866	—
ρ	—	—	—	—	-0.5195
δ	—	—	—	—	0.5492
a	1.0212	—	—	—	—
b	1.7830	—	—	—	—
$A^{D}(0)-A^{C}(0)$	—	8.3763	60.88 & -17.31	10.0110	0.5754
Initial p(C)	0.3642	0.3538	0.1170 & 0.6398	0.3458	0.4374
#parameter	2	3	4	4	5
MSD	0.2003	0.1922	**0.1832**	0.1914	0.1895
AIC	1902.226	1849.162	**1777.326**	1844.449	1841.517
BIC	1906.723	1855.908	**1786.320**	1853.443	1852.759

註：畫底線者為該項最佳模型。

WH(p)的起始合作機率則是傾向不合作，也可以看到 WH(p)的配適結果是四種資料中最不好的，分組之後傾向合作組有 30 人，合作機率只有 0.64，從配適結果看來，RL、ERL 和 EWA 的 MSD 都非常接近，只有 RL 2 有明顯較好，其 AIC 與 BIC 也是最低的。

表 4.21、WH(p)交叉驗證

WH(p)		BB	RL	RL 2	ERL	EWA
ϕ		—	0.7526	0.8895	0.7152	0.7051
λ		—	0.0850	0.0311	0.0768	0.0457
γ		—	—	—	0.1960	—
ρ		—	—	—	—	0.2215
δ		—	—	—	—	-0.8613
a		1.0283	—	—	—	—
b		1.8416	—	—	—	—
$A^D(0)-A^C(0)$		—	7.0793	59.58 & -27.96	8.1268	18.6030
MSD-training	1~18 round	0.2000	0.1896	0.1781	0.2107	0.1891
	6~18 round	0.1973	0.1906	0.1835	0.2057	0.1887
MSD-testing	19~23 round	0.2014	0.2026	0.1987	0.2102	**0.1979**

註：畫底線者為最佳。

這裡可以發現一件事情，前面 RM、WH、WH(c)的資料通常驗證資料的 MSD 都會較低，表示對未來的預測會較好，另一個角度來看就是玩家在後面的選擇會較穩定，但是 WH(p)似乎較後面的策略則是較有變化，選擇比較不同。若從交叉驗證的角度選取模型，則是 EWA 模型的表現較佳。

表 4.22、WH(p)-基本性質信賴區間

WH(p)	p(C)		p(CC)		p(sw)	
Actual value	0.367		0.187		0.326	
RL	0.2730	0.3817	0.0777	0.1985	0.2857	0.3521
RL 2	0.2944	0.3730	0.0921	0.1794	0.3235	0.3899
BB	0.2984	0.4270	0.0683	0.1604	0.3771	0.4758
ERL	0.2833	0.3921	0.0921	0.2063	0.2765	0.3420
EWA	0.0159	0.0572	0.0000	0.0302	0.0286	0.0597

註：網底者表信賴區間包含實際值。

WH(p)整體基本性質的信賴區間，EWA 模型產生的信賴區間都不包含實際值，BB 則是有捕捉到合作的信賴區間，但其他就不甚理想，RL 2 則是在 p(C)的信賴區間差了一點，但差異不大，所以暫時將其列入考慮，RL 與 ERL 則是非常漂亮都包含了實際值。

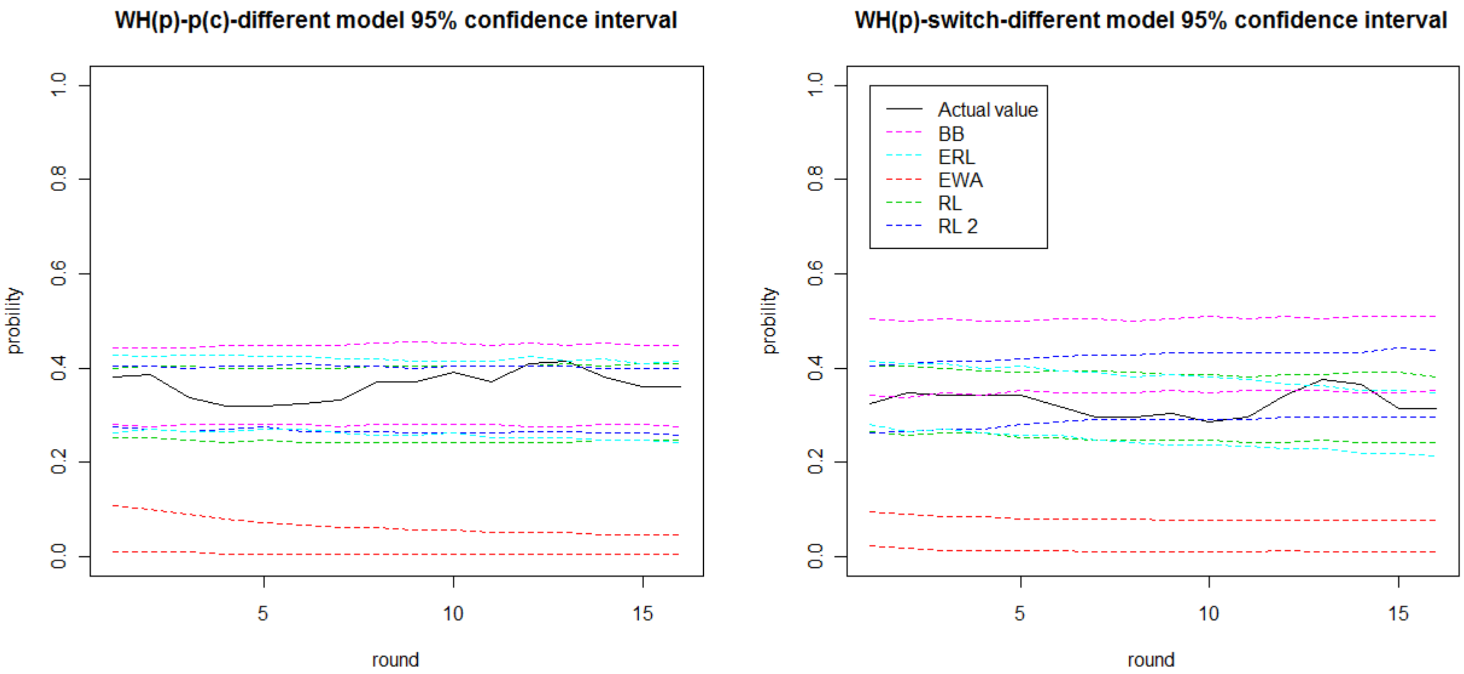


圖 4.6、WH(p)-局部基本性質信賴區間

觀察局部的信賴區間，BB 在 p(sw)表現依舊不理想， EWA 則是與實際值相差甚遠，其餘 RL、RL 2 與 ERL 都非常接近，沒有哪個模型是完全包含實際值的，但是這三個模型的表現都非常不錯，不管是在 p(C)或是 p(sw)的信賴區間上，大部分的信賴區間都包含了實際值，只有少部分幾回合有稍微偏離以外，考慮有可能產生的誤差，因此認為此三種模型都有可能是符合模型假設的，接著從三種模型中挑選較佳的模型。

表 4.23、WH(p) - F^* 檢定

模型	RL	RL 2	模擬 F 值($\alpha = 0.05$) =2.250
殘差平方和	309.4420	294.9520	
自由度	67	66	$F_{0.05}(1,66)$=3.9862
F^*	--	3.2423	

如果以模擬的 F 值做檢定，則檢定結果會是拒絕，即代表 RL 是較不佳的模型，因此三種符合假設的模型中，RL、RL2 和 ERL，在 MSD、AIC 和 BIC 的表現上，或是交叉驗證，或是 F^* 檢定上，皆顯示 RL 2 模型有較佳的表現，因此認為 WH(p)情境下，RL2 是較佳的。

表 4.24、WH(p)-參數信賴區間

Model：RL 2	WH(c)		
percent	2.5%	Actual	97.5%
ϕ	0.7232	0.8880	0.9235
λ	0.0273	0.0332	0.07868
$A^D(0)-A^C(0)$	13.79	60.88	73.96
	-28.98	-17.31	14.13
MSD	0.1715	0.1832	0.1960

觀察 RL2 的信賴區間，模擬的信賴區間皆符合實際值，因此認為 RL2 是合適的模型。

第六節　敏感度分析

前面五節的模擬皆根據模型假設產生亂數，觀察其模型特色信賴區間，是否包含實證資料，實證資料符合模型信賴區間，才較能說明資料是來自此模型，但難免會出現失誤，因此本節將探討不同學習模型的敏感度分析。

其想法如下：若資料來自 A 學習模型，但因為判斷失誤，選擇了 B 學習模型，若是發生這種情況，其誤差會有多大，是否在能接受的範圍內，因此根據模型假設產生亂數，再由其他模型配適結果，並觀察其差異。而該以什麼做為判斷標準？由於較複雜的模型估計結果必然較好，且AIC、BIC需要資料分佈的假設，因此使用交叉驗證做為判斷標準。判斷方法將每一次亂數分成估計資料以及驗證資料，以估計資料估計模型參數後，觀察驗證資料的最小均方差的變化，以模型的預測能力做為判斷標準。

觀察不同模型在其他模型假設下，是否會穩健(Robust)，如果模型在其他模型假設下仍然表現不錯，那麼即使實際資料來自別的學習模型，依舊可以使用該模型估計，也能得到不錯的結果。以下分兩個主要項目觀察，一個是驗證資料的平均 MSD 以及 MSD 的變異數。

一、RM 情境

以 RM 的配對方式產生亂數之後，對亂數做估計其結果如下：

表 4.25、RM 亂數驗證資料結果

驗證資料平均 MSD		配適模型(B)				
		QRE	BB	RL	ERL	EWA
實際模型(A)	QRE	0.1658	0.1668	0.1673	0.1695	0.1685
	BB	0.1678	0.1680	0.1688	0.1812	0.1714
	RL	0.1178	0.1374	0.1071	0.1115	0.1086
	ERL	0.1023	0.1348	0.0787	0.0739	0.0781
	EWA	0.2661	0.2367	0.1675	0.1688	0.1648

註：網底表示該組最佳　　(單位:MSD)

表 4.25 以實際的五種模型(A)假設產生亂數之後，再利用配適模型(B)估計亂數，並觀察其驗證資料的平均 MSD。可以看到除了 BB 以外，其餘以該模型產生的亂數，其結果就會越佳，頗符合常理，接著觀察模型的相對效率性。

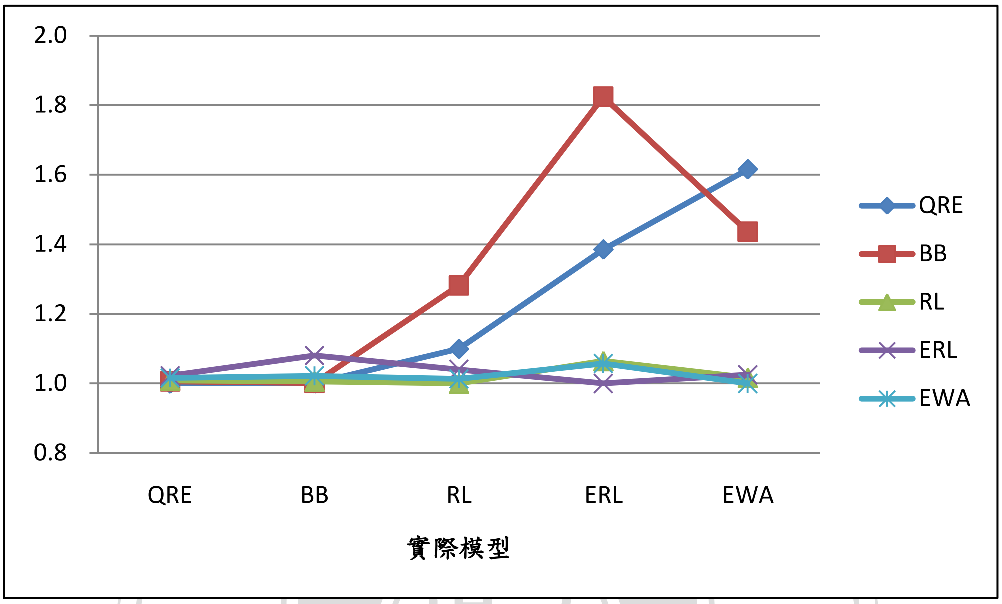


圖 4.7、RM-相對效率性(驗證資料)

圖 4.7 是將驗證資料的平均 MSD 除以該組最佳的平均 MSD 所畫的圖，不同的亂數假設下，最佳模型比值為 1，其他模型若是與最佳模型相差越大，比值則會越大，可以看到 QRE 以及 BB 在 RL、ERL、EWA 的模型假設下，其相對效率性表現不佳，而 EL、ERL 與 EWA，其相對效率性皆有不錯的表現，接著觀察驗證資料 MSD 的變異。

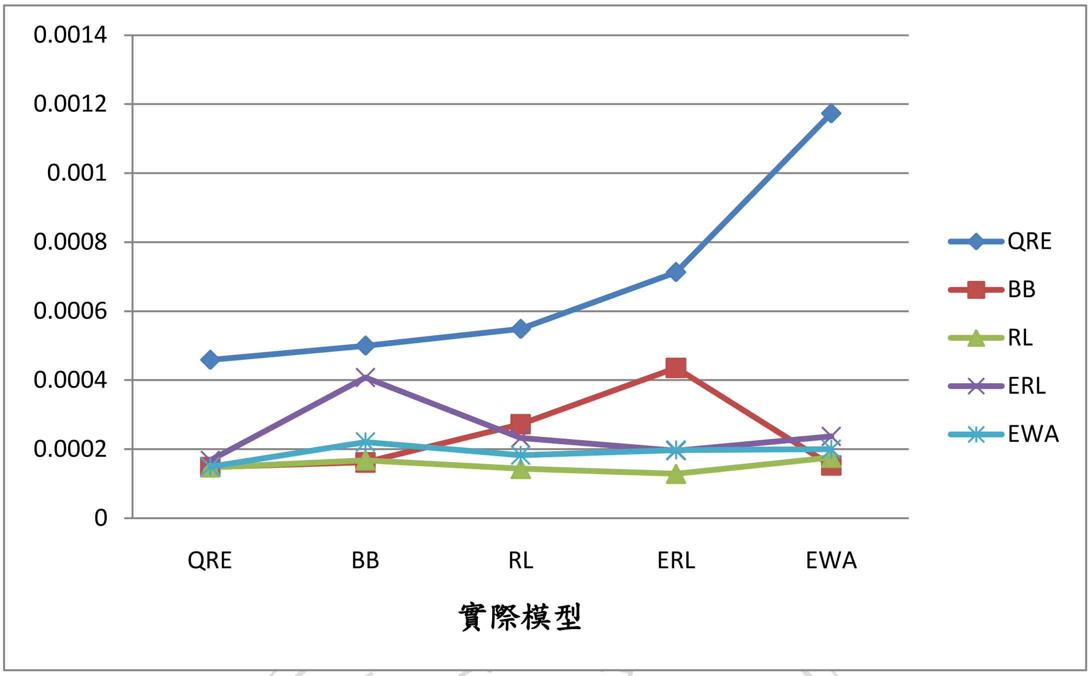


圖 4.8、RM-MSD 變異數(驗證資料)

在 MSD 變異數的表現上，QRE、BB、ERL 的變異數在某些模型都稍嫌太大，而 RL 與 EWA 皆表現得不錯，但是 RL 在這裡的表現會比 EWA 較好，因此在 RM 情境下，即使資料不是來自 RL 模型，RL 都可以表現得很穩健，和實際的模型差異皆不會很大。

二、WH 情境

表 4.26、WH 亂數驗證資料結果

驗證資料 平均 MSD		配 適 模 型(B)				
		QRE	BB	RL	ERL	EWA
實際模型(A)	QRE	0.2062	0.2054	0.2060	0.2078	0.2071
	BB	0.2047	0.1967	0.2052	0.2133	0.2071
	RL	0.1613	0.1553	0.1154	0.1168	0.1176
	ERL	0.1478	0.1492	0.0854	0.0785	0.0836
	EWA	0.2987	0.2437	0.1935	0.1943	0.1917

註：網底表示該組最佳 (單位:MSD)

在 WH 情境下，和 RM 情境的結果差不多，除了 QRE 以外，其餘以該模型產生的亂數，其結果就會越佳。

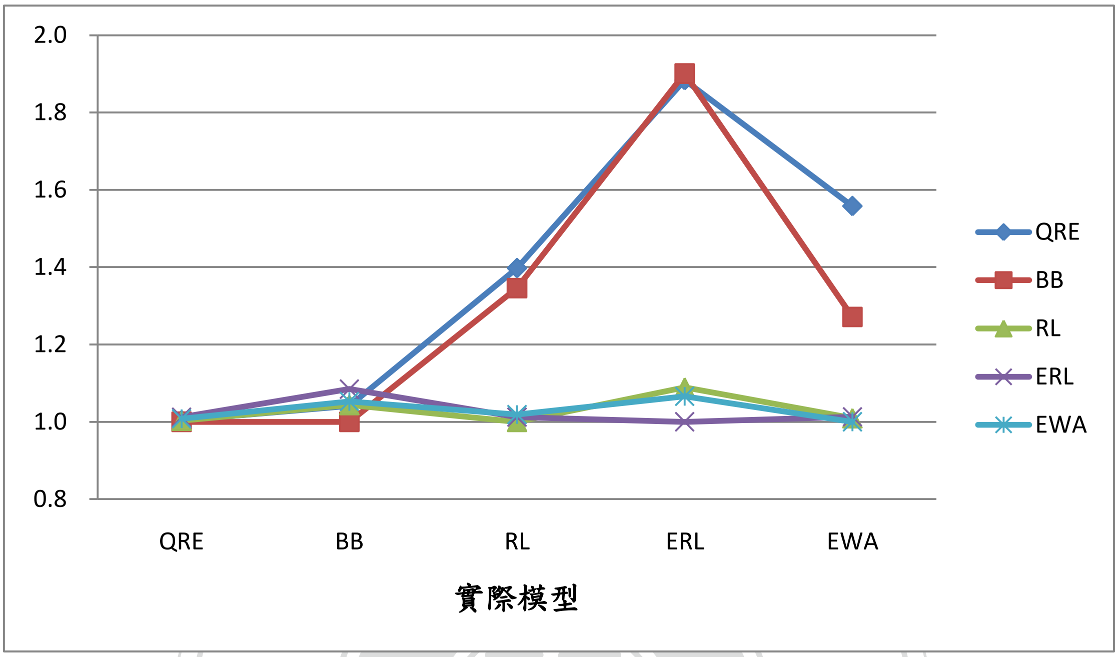


圖 4.9、WH-相對效率性(驗證資料)

WH 情境下各種模型的相對效率性，也是 QRE 與 BB 會與最佳的差異很大，其與 RL、ERL、EWA，則是非常接近。

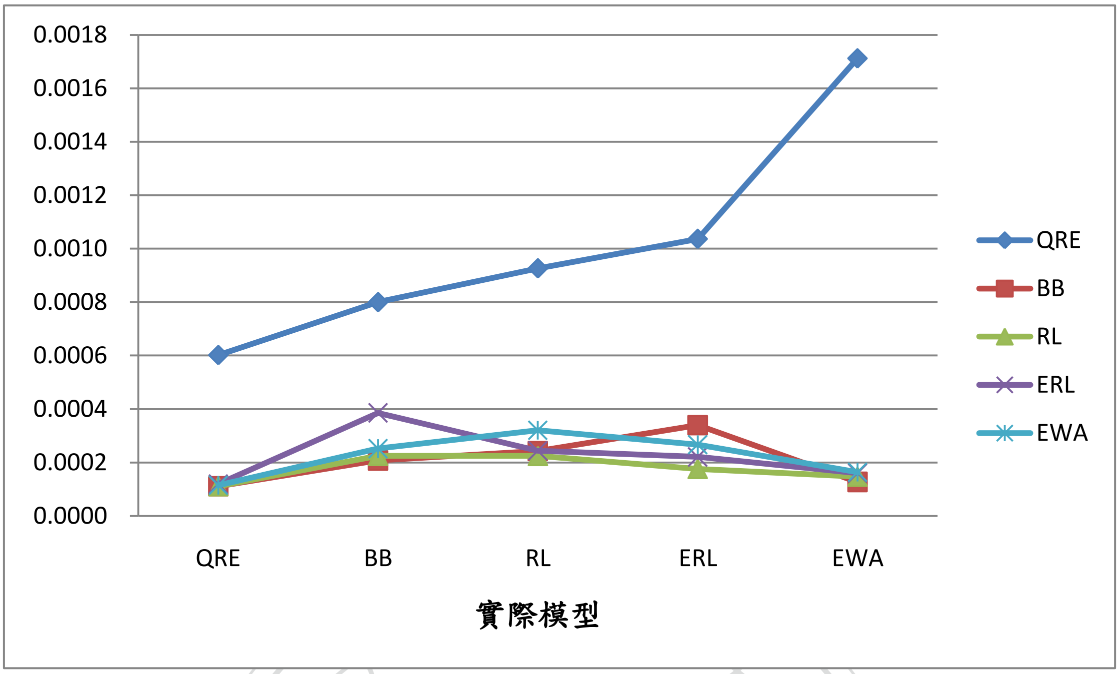


圖 4.10、WH-MSD 變異數(驗證資料)

WH 情境下的 MSD 變異數，QRE 在每個模型下產生的亂數變異皆非常大，除了 QRE，其餘的模型都有不錯的表現，若是仔細觀察，ERL 在 BB 的假設下以及 BB 在 ERL 的假設下，變異皆稍嫌大了一點，而觀察 EWA 與 RL，整體來說 RL 的變異較小，因此在 WH 情境下，即使資料不是來自 RL 模型，RL 也還是可以表現得很穩健，與 RM 情境相同。

不管是 RM 情境亦或是 WH 情境，RL 在各種模型的假設下，其預測能力，皆有不錯的表現，RL 相較於其他模型的穩健性會較好。從模型的配適結果、模擬驗證結果、模型穩健性的表現上，RL 皆會優於其他模型，因此 RL 似乎非常適合描述囚犯困境的賽局資料。

第五章　結論及建議

第一節　　結論

一開始賽局的發展著墨於理論較多，從單一回合的賽局，通常可以從理論證明均衡狀態，接著經濟學家以賽局實驗驗證理論，發現多回合的重複賽局不如理論所述，因此使用學習模型解釋人的學習行為，爾後也漸漸使用電腦模擬，探討各種學習模型的特性等等。

在重複的賽局問題裡，經濟學家以學習模型解釋人的學習行為，而過去決定模型通常以模型配適結果(最小均方差或最大概似估計量)，或者是以模型預測能力(交叉驗證)來判斷模型好壞，但這樣僅能比較模型之間的好壞，無法說明模型是合適的，有可能配適結果是最好的，但是資料卻不符合模型的假設。就如同迴歸模型，除了比較 $R^2_{adj.}$ 或者是 AIC、BIC 以外，還會對其殘差做多種檢定，如：常態檢定、變異數同質性檢定、獨立性檢定等。

因此本文希望引進統計概念驗證模型假設，提出輔助選取模型的標準，本文提出輔助的標準有：F^* 檢定以及驗證模型假設。而賽局資料裡，驗證資料是否符合模型假設會有困難，因為沒有好的分佈可以解釋賽局資料，而且資料彼此不獨立，對於這樣的資料，較難提出分佈假設驗證。因此本文使用蒙地卡羅模擬驗證這件事情，如果資料可以用某個學習模型解釋，則實證資料應可視為其中一筆亂數，模擬亂數特性的信賴區間也應包含實際值特性，用此種方法驗證模型假設，雖然不能說符合信賴區間就是正確的模型，但藉由這樣的方法至少可以挑出不正確的模型，進而從剩下的模型中挑出較佳的。而分析之後的結果得到幾個結論：

一、不同遊戲規則的資料，給予分組有助於配適結果

不同的遊戲規則，玩家選擇的策略也不同，在 RM 的規則下，大部分玩家都會選擇不合作；WH 的規則下，玩家會跟去過去五回合的記錄配對，雖然大部分的玩家仍然不願意先伸出合作的手，但是已經有少數的人願意為了達到長期合作的關係而選擇合作；在 WH(c)裡多加入了兩個固定選擇的玩家，使得玩家的分群更明顯，分成合作與不合作兩派，更多人願意選擇合作；WH(p)告知玩家的報酬排名，使得玩家開始想要轉換策略，而原本願意選擇合作的玩家，也不完全選擇合作，或許是當他們知道排名不是最好的時候，就萌生起選擇不合作的念頭。

在學習模型中，可以將初始值視為玩家的特性，如果一開始選擇合作的意願較高，則後續就會較傾向選擇合作，反過來說，選擇不合作的意願也是如此，概括來看，如果玩家的合作意願是明顯不同的，分組並給予不同的初始值則是會有助配適結果的。

二、選取模型以及檢驗模型假設

RM 跟 WH 的規則下，大部分的玩家都是傾向選擇不合作的，分組對其較無效果，模擬驗證及 F^* 檢定的結果，都顯示一個初始值的增強學習模型是較好的；WH(c)的玩家則是有比較明顯分成合作與不合作兩群，將 WH(c)情境的玩家分組，其模擬驗證及 F^* 檢定，都顯示分成兩組是有效果的；而 WH(p)則是比較沒有明顯的分群，但也沒辦法完全將玩家歸在同一類，是比較難判斷的情境，從模擬結果看來，增強學習模型、增強學習模型 2 (分成 2 組)、延伸增強學習模型，資料都有符合模型假設，而最後的判斷是增強學習模 2 的模型較佳。四種不同的情境，根據其不同的特性可以將玩家分組，對於某些規則的資料，分組的確是有助配適結果的。

模擬結果以及檢定結果，RM 與 WH 的情境下，是增強學習模型會較佳，WH(c)與 WH(p)的情境下，則是以將玩家分兩組的增強學習模型較佳。另外也觀

察到較複雜的模型其模擬結果未必就會較好，以 RM 與 WH 情境來看，加權經驗吸引模型的模擬結果就不如增強學習模型。

三、學習模型的穩健性

在第四章的敏感度分析中，發現增強學習模型比起其他模型都來的穩健，因此在囚犯困境的賽局資料中，即使實證資料不是來自增強學習模型(RL)，亦可以用此模型配適，從模擬結果看來，其配適結果會與實際模型非常接近，此結果與之前的分析有一致的結果，顯示增強學習模型非常適合描述囚犯兩難的賽局資料。

第二節　後續發展與建議

本文主要針對修正估計參數以及選取學習模型做探討，本文使用的資料都僅有一組數據，因此無法確定資料是否具有樣本代表性，但本文主要希望提出輔助方法幫助選取模型，因此在樣本代表性較少討論，而之後建議可實驗多組數據，以增加驗證模型假設的證據。另外，實證資料僅使用囚犯兩難賽局做實驗，或許可以嘗試不同類型的賽局，探討學習模型是否也能捕捉其他不同賽局的資料特性。

本文主要考慮的資料特性為合作機率以及轉換機率，但是僅以這兩種機率描述賽局資料的特性似乎仍有點不足，或許有不同的特性更能描述賽局資料。而現有的學習模型通常偏向漸漸收斂的效果，但是有些人在某些規則下似乎顯得震盪(較游移不定)，或許可以嘗試找出新的模型解釋某些人的震盪行為。

參考文獻

中文部分

張宮熊(2009):"賽局：又稱博奕論"高雄市:玲果國際文化出版。

劉常勇(2008):"賽局理論中的雙贏策略"。

取自 http://cm.nsysu.edu.tw/~cyliu/paper/paper10.html

英文部分

Anderson C. and Camerer C. F. (2000) , “Experience-weighted Attraction Learning in Sender-receiver Signaling Games,” Economic Theory, 16(3), 689-718

Arifovic J. and Ledyard J. (2004), “Scaling up learning models in public good games,” Journal of Public Economic Theory 6, 205–238

Bendor J., Diermeier D., and Ting M.(2003), “The Predictive Power of Learning Models in Social Dilemma Research,” Stanford University working paper.

Cabrales A. and Garcia-Fontes W. (2000), “Estimating Learning Models with Experimental Data,” University of Pompeu Febra working paper 501.

Cabrales A., Garcia -Fontes W. and Motta M. (2000) , “Risk Dominance Selects the Leader. An Experimental Analysis,” International Journal of Industrial Organization, 18:137-162.

Camerer, C. F., and Ho, T. (1999) , “Experience-weighted attraction learning in normal form games,” Econometrica 67, pp. 827–874.

Camerer C.F.(2003), “Behavioral Game Theory: Experiments in Strategic Interaction,” Princeton Univ. Press, Princeton, NJ.

Cournot, A. (1960), “Recherches sur les Principles Mathematiques de la Theories des Richesses,” Translate into English by N. Bacon as Researches in the

Mathematical Principle of the Theory of Wealth. London: Haffner

Erev I. and Roth A. E.(1998), "Predicting how people play games with unique, mixed-strategy equilibria," Amer. Econom. Rev., vol. 88, pp. 848 - 881.

Friedman D.(1998), "Evolutionary economics goes mainstream: A review of the theory of learning in games," Journal of Evolutionary Economics, 8: 423-432.

Jacob K. Goeree, Charles A. Holt, and Thomas R. Palfrey(2002), "Quantal response equilibrium and overbidding in private-value auctions," Journal of Economic Theory, 104(1):247-272

Joy Woller(1996) , " The Basics of Monte Carlo Simulations," Retrieved from http://www.chem.unl.edu/zeng/joy/mclab/mcintro.html

Lai Y.H.(2005), "A Study of Learning models for analyzing prisoners' dilemma data," Working Paper, Chengchi university.

McKelvey, R.D. and Palfrey, T.R.(1995), "Quantal Response Equilibria for Normal Form Games," Games and Economic Behavior, Vol. 10, pp. 6–38.

Sutton R.S., Barto A.G.(1998), "Reinforcement Learning: An Introduction," MIT Press, Cambridge, MA.

Yang C.-L., Yue J.C. and Yu I.-T. (2007) , "The Rise of Cooperation in Correlated Matching Prisoners Dilemma: An Experiment," Experimental Economics 10, 3-20.

Yang C.-L. and Yue, C.J.(2010), "Assortative Matching, Information and Cooperation: An Experiment," Economics Bulletin (SSCI), vol. 30(1), 414-420.