

執政優勢對選舉連任優勢的迴歸分析



財金學系三年級 B08303006 林姝延
經濟學系三年級 B08303024 張翔竣
經濟學系三年級 B08106043 翁紹洋
經濟學系三年級 B08303052 朱恆昀

緒論

在今年的 1 月 14 日，中央選舉委員會宣布了今年的九合一大選將訂在 11 月 26 日舉行投票。選舉是人們自由意志的展現，各個政黨憑著自己的能力、在不同議題的立場，說服選民把手上珍貴的選票、以及這個國家的未來，一同寄託給所認同的政黨。

然而，這個看起來公開公平的競爭，真的不會受到其他因素影響嗎？我們注意到在連選得連任的職位，例如總統，基本上都是連任兩屆後，依規定無法再選才將執政權交給其他人，於是我們有個猜想，執政黨相較在野黨更有優勢。

由於執政黨相較於在野黨，擁有更多資源，於在任期間也會宣導和自己立場比較相符的議題，如此一來是否會因為長期的推廣和教育，影響了中立選民對於某些議題的看法，進而讓執政黨在下次選舉時取得更多優勢，進而連任呢？

相關研究

現任中山大學企管系專任副教授的佘健源教授，曾經有針對台灣 2008 年以後，立委改為「單一選區兩票制」對於現任者競選優勢的影響進行討論，發表《台灣立委選制改革對現任者競選優勢之影響》，但是這篇論文相較而言比較像是單一事件分析（event study），且在文中的分析模型過於複雜，遠超我們目前的程度。

而這篇論文的結論是現任者的確有競選優勢，若要改制必須趁早開始。

方法論

• 資料蒐集及統整

為了解決以上的問題，我們決定針對兩黨制國家的資料來做討論，因為如此一來就會是 binary 的結果（A 黨選上或是 B 黨選上），在做迴歸分析時，可以有更直觀的解釋。

然而，由於台灣是自 1990 年民主自由化之後才開始公平選舉，而以**總統及副總統由公民直接選舉**為例，也是直到 1996 年才第一次由人民親自選出，至今為止，也只舉行過 7 屆真正意義上的民主選舉，所以資料的樣本數遠遠不夠，不足以讓我們跑出一個有足夠解釋力的迴歸分析結果。

於是我們把目光放在另一個兩黨制且選舉歷史相對悠久的國家——美國，我們蒐集了美國總統大選在各州的選舉結果（美國是選舉人制）各年間的資料，分析某個候選人在當選之後，是否會因為執政的緣故，而更容易在下次的選舉中連任。

- 模型一（OLS）

首先，我們決定以最簡單直觀的OLS迴歸分析討論：

$$win_dem_t1_i = \alpha + \beta win_dem_i + u_i$$

#如果在第 i 期當選， $win_dem = 1$ ；反之則 $win_dem = 0$ 。如果在第 $i+1$ 期當選， $win_dem_t1 = 1$ ；反之則 $win_dem_t1 = 0$ 。

α 是截距項。 u_i 是殘差項。而 β 即為我們想要觀察的效果，代表了執政黨在當選時，有 β 的機率，下一次也會當選。

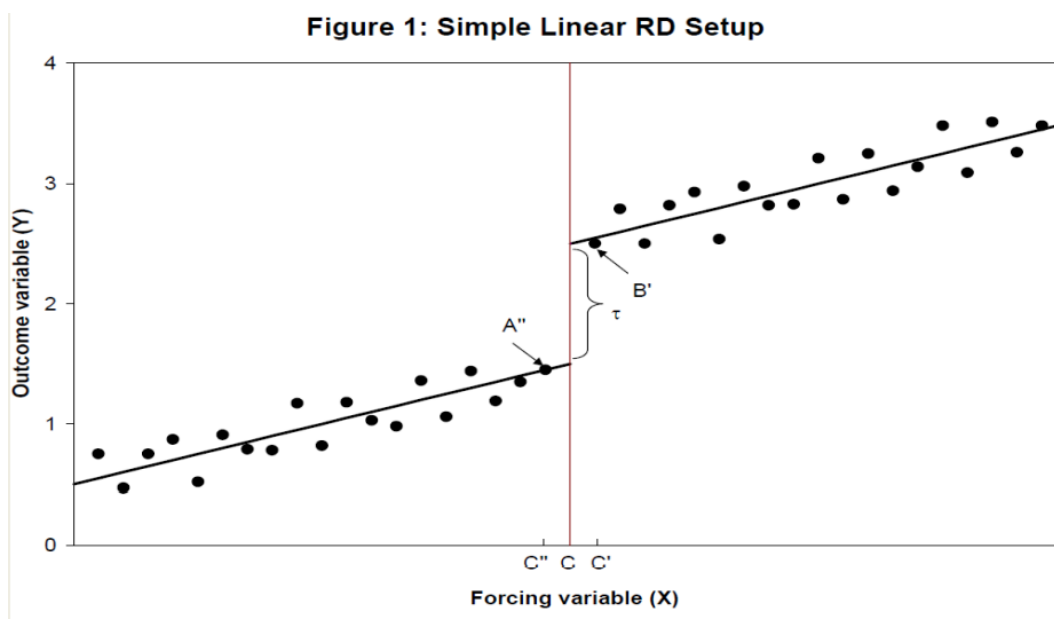
然而，該模型的結果雖然清楚好懂，但卻隱含許多沒估計到的遺漏變數。

其中最主要的遺漏變數，就是假設某地區（州）的選民意識形態特別偏向某個政黨，換言之，就是某個政黨的「鐵票區」，則當選的政黨是否執政其實不會影響他們的選舉結果（never taker / always taker），則如果我們將這些地區的資料也納入考慮，會大大高估執政帶來的優勢。所以即便簡單的 OLS 看起來相當直觀，我們也不能直接把結果當成執政優勢的效果。

- 模型二（RD - 線性）

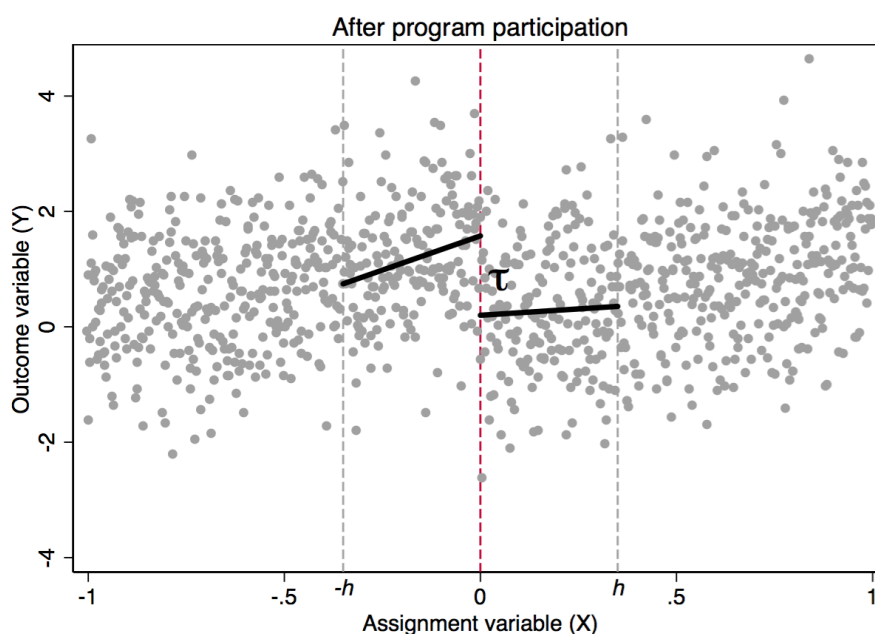
為了解決在模型一遇到的遺漏變數，我們決定套用在做迴歸時一個比較特殊的技巧，叫做 RD（Regression Discontinuity design）。

RD 的概念是把某個點設為斷點（cut-off point），而在斷點前和斷點後的樣本是受到不同的對待的（如圖一），因為在斷點前後的樣本原本並沒有太大差別，只差在有沒有受到 treated，所以如果在 treated 之後發現明顯斷點，我們就可以認為 treated 是有效果的。以我們的模型為例，我們將第 i 期的得票率，從 0 到 100% 列出，而由於是兩黨制，所以得票率超過 50% 即可執政，所以我們可以相信得票率大於 50% 的樣本，和得票率小於 50% 的樣本是受到不同對待的（treated），因此我們把 cutoff 設在 50%，並在模型中加入一個 dummy variable D ，當第 i 期的得票率小於 50% 時， $D = 0$ ；當第 i 期的得票率大於 50% 時， $D = 1$ 。



圖一。C 為斷點，小於 C 的是 Control 組，大於 C 的是 Treatment 組。

接著，我們會在斷點前後取一範圍（bandwidth）（如圖二），這是因為我們比較倚重斷點前後的資料，以我們的題目為例，原本第 i 期得票率在 50% 左右的話，代表兩黨之間原本是勢均力敵的，假設第 i 期得票率從小於 50% 到大於 50% 時，在第 $i+1$ 期當選機率有明顯的跳升，應該就是執政優勢的體現，且取 bandwidth 剔除掉得票率太高或太低的樣本，就可以規避掉在模型一遇到的「某地區（州）的選民意識形態特別偏向某個政黨，則其實當選的政黨是否執政其實不會影響他們的選舉結果」的問題了。



圖二。0 為斷點，h 為 bandwidth。τ 就是 treatment 的效果。

實際模型：

$$D_i = \begin{cases} 1 & \text{if } demvoteshare_i \geq 50\% \\ 0 & \text{if } demvoteshare_i < 50\% \end{cases}$$

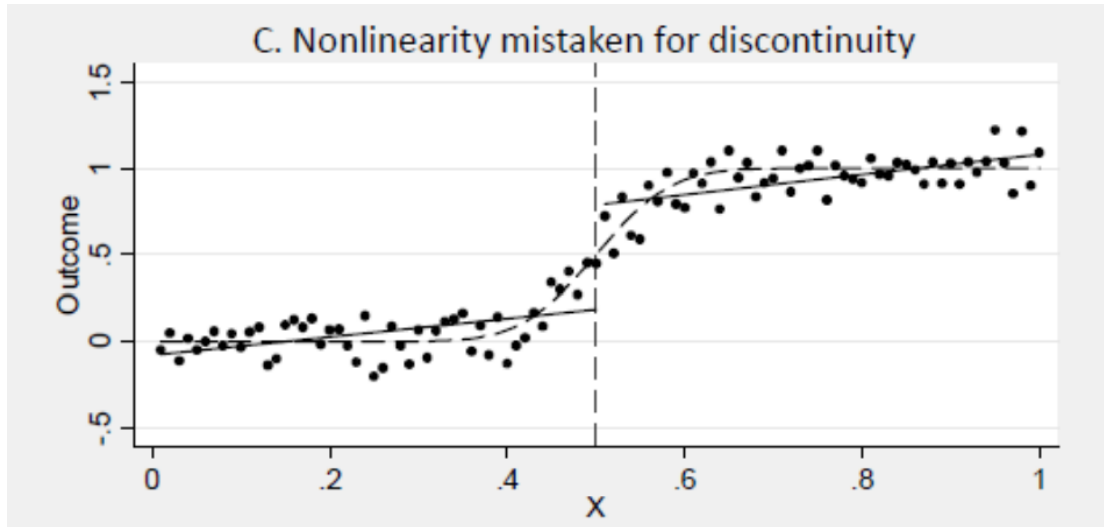
$$win_dem_t1_i = \alpha + \beta_{01}demvoteshare_i + \rho D_i + \beta_1^*(D_i \times demvoteshare_i) + u_i$$

如果在第 $i+1$ 期當選， $win_dem_t1 = 1$ ；反之則 $win_dem_t1 = 0$ ； $demvoteshare_i$ 代表在第 i 期的得票率。

α 是截距項。 u_i 是殘差項。而 ρ 即為我們想要觀察的效果，反映了第 i 期得票率從小於 50% 到大於 50% 時，在第 $i+1$ 期當選機率跳升的幅度。

- 模型三（RD - 非線性）

然而，RD 分析其實不太能單純倚靠 ρ 的大小判斷是否有斷點、以及 treatment 效果的大小，可能會因為用錯誤的模型估計（nonlinearity mistaken），導致看起來似乎有斷點但實際上沒有斷點（如圖三）。



圖三。X 是斷點。以線性估計非線性，但其實沒有斷點。

所以我們接下來會以更高次方的模型做迴歸，以免造成模型錯誤造成誤判了 treatment 的效果，並且以圖片輔助判斷。

實際模型：

$$D_i = \begin{cases} 1 & \text{if } demvoteshare_i \geq 50\% \\ 0 & \text{if } demvoteshare_i < 50\% \end{cases}$$

$$win_dem_t1_i = \alpha + \beta_{01}demvoteshare_i + \beta_{02}demvoteshare_i^2 + \rho D_i + \beta_1^*(D_i \times demvoteshare_i) + \beta_2^*(D_i \times demvoteshare_i^2) + u_i$$

如果在第 $i+1$ 期當選， $win_dem_t1 = 1$ ；反之則 $win_dem_t1 = 0$ ； $demvoteshare_i$ 代表在第 i 期的得票率。

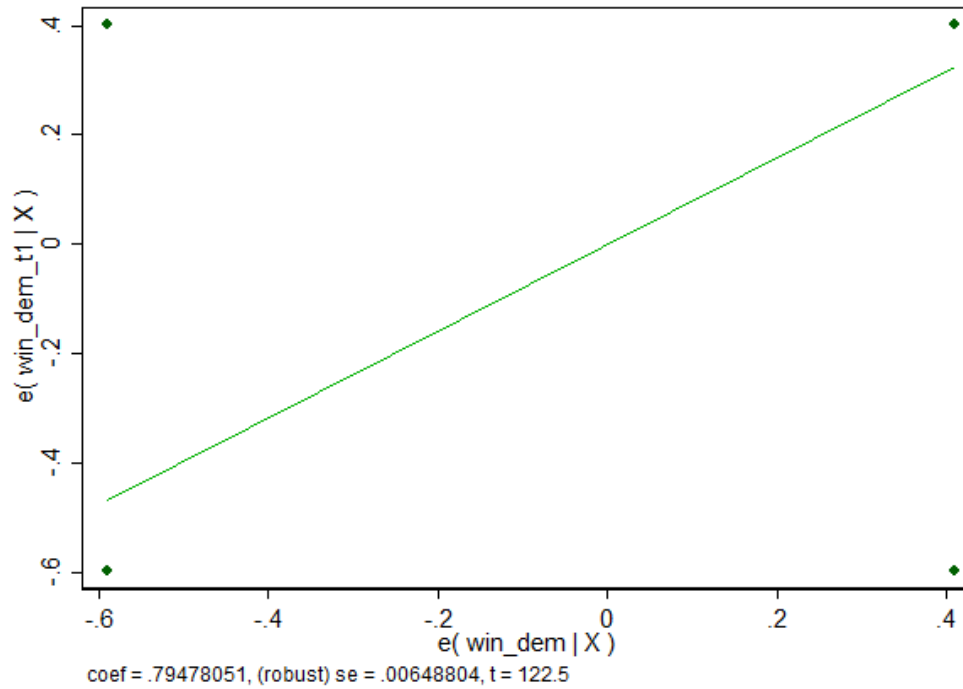
α 是截距項。 u_i 是殘差項。而 ρ 即為我們想要觀察的效果，反映了第 i 期得票率從小於 50% 到大於 50% 時，在第 $i+1$ 期當選機率跳升的幅度。

研究結果

- 模型一 (OLS)

OLS Regression Results						
=====						
Dep. Variable:	win_dem_t1	R-squared:	0.635			
Model:	OLS	Adj. R-squared:	0.635			
Method:	Least Squares	F-statistic:	1.628e+04			
Date:	Thu, 20 Jan 2022	Prob (F-statistic):	0.00			
Time:	14:19:48	Log-Likelihood:	-1886.2			
No. Observations:	9345	AIC:	3776.			
Df Residuals:	9343	BIC:	3791.			
Df Model:	1					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

Intercept	0.1287	0.005	26.898	0.000	0.119	0.138
win_dem	0.7948	0.006	127.608	0.000	0.783	0.807
=====						
Omnibus:	1072.062	Durbin-Watson:	2.151			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	10712.481			
Skew:	-0.046	Prob(JB):	0.00			
Kurtosis:	8.244	Cond. No.	2.89			
=====						
Warnings:						
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.						



$$win_dem_t1_i = 0.1287 + 0.7948 win_dem_i + u_i$$

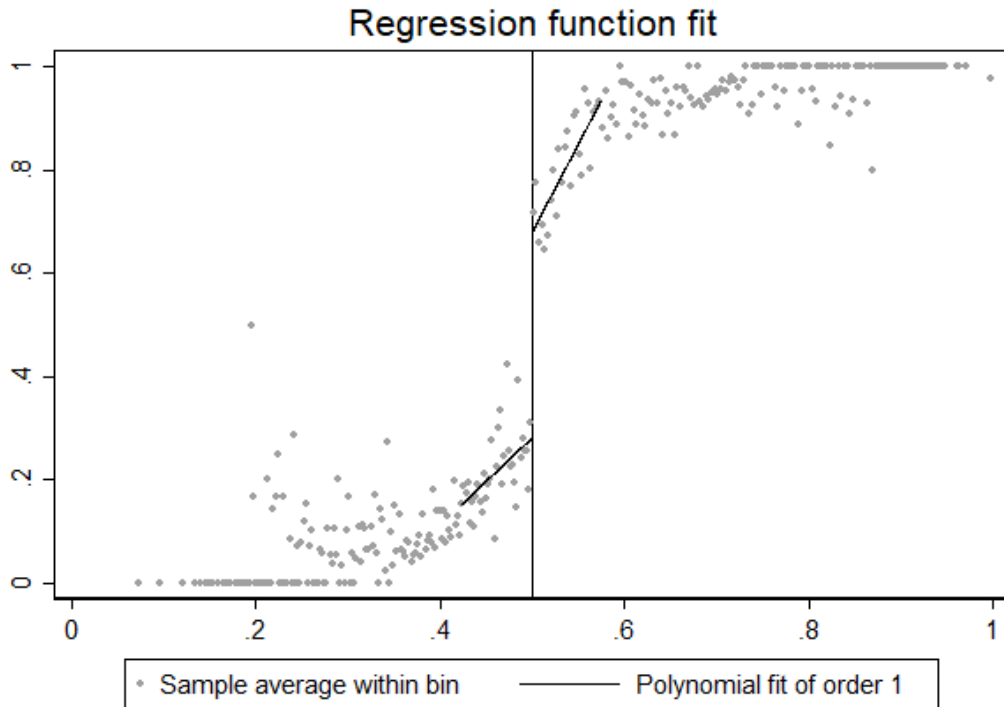
(0.005) (0.006)

結果發現，這次選舉如果當選，則會有 79.48% 的機率下次也會當選，標準差為 0.006，在 5% 的顯著水準下具有顯著性。

- 模型二（RD - 線性）

OLS Regression Results						
=====						
Dep. Variable:	y	R-squared:	0.621			
Model:	OLS	Adj. R-squared:	0.620			
Method:	Least Squares	F-statistic:	1257.			
Date:	Thu, 20 Jan 2022	Prob (F-statistic):	0.00			
Time:	16:36:46	Log-Likelihood:	-213.69			
No. Observations:	1801	AIC:	435.4			
Df Residuals:	1797	BIC:	457.4			
Df Model:	3					
Covariance Type:	HC1					
=====						
	coef	std err	z	P> z	[0.025	0.975]

Intercept	-0.6439	0.108	-5.973	0.000	-0.855	-0.433
C(Group)[T.Treatment]	0.3591	0.190	1.886	0.059	-0.014	0.732
x	1.8494	0.253	7.300	0.000	1.353	2.346
C(Group)[T.Treatment]:x	0.1711	0.373	0.459	0.647	-0.560	0.903
=====						
Omnibus:	101.922	Durbin-Watson:	1.988			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	394.387			
Skew:	-0.071	Prob(JB):	2.29e-86			
Kurtosis:	5.288	Cond. No.	91.1			
=====						
Warnings:						
[1] Standard Errors are heteroscedasticity robust (HC1)						



模型二：有加截距項，做出 sharp RD 的效果。線性。

$$\begin{aligned} win_dem_t1_i = & -0.6439 + 1.8494 demvoteshare_i + 0.3591 D_i \\ & (0.108) \quad (0.253) \quad (0.190) \\ & + 0.1711(D_i \times demvoteshare_i) + u_i \\ & (0.373) \end{aligned}$$

結果發現，由圖中可以看到跳升的現象非常明顯，第 i 期得票率從小於 50% 到大於 50% 時，在第 $i+1$ 期會突然增加 35.91% 的機率能夠當選，標準差為 0.190，在 5% 的顯著水準下具有顯著性。

- 模型三（RD - 非線性）

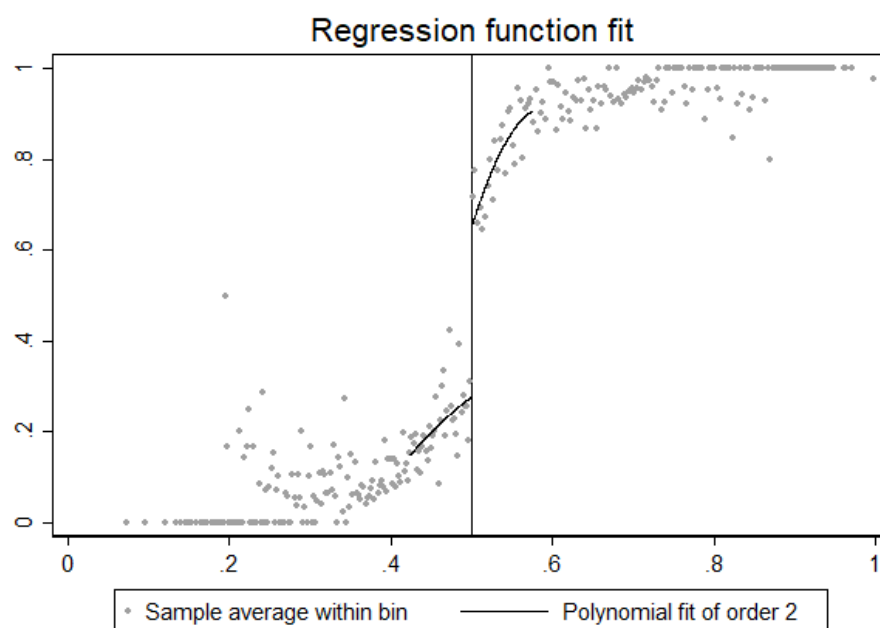
OLS Regression Results						
=====						
Dep. Variable:	y	R-squared:	0.474			
Model:	OLS	Adj. R-squared:	0.472			
Method:	Least Squares	F-statistic:	299.2			
Date:	Thu, 20 Jan 2022	Prob (F-statistic):	7.17e-220			
Time:	16:40:24	Log-Likelihood:	-437.31			
No. Observations:	1449	AIC:	886.6			
Df Residuals:	1443	BIC:	918.3			
Df Model:	5					
Covariance Type:	HC1					
=====						
	coef	std err	z	P> z	[0.025	0.975]

Intercept	-10.4791	6.095	-1.719	0.086	-22.426	1.467
C(Group)[T.Treatment]	0.3060	10.174	0.030	0.976	-19.635	20.247
x	44.5190	26.494	1.680	0.093	-7.408	96.446
C(Group)[T.Treatment]:x	-6.8309	40.210	-0.170	0.865	-85.640	71.978

x2	-46.2070	28.733	-1.608	0.108	-102.524	10.110
C(Group)[T.Treatment]:x2	14.1764	40.151	0.353	0.724	-64.518	92.871

Omnibus:	27.825	Durbin-Watson:	2.018
Prob(Omnibus):	0.000	Jarque-Bera (JB):	52.698
Skew:	0.072	Prob(JB):	3.60e-12
Kurtosis:	3.923	Cond. No.	1.01e+04

Warnings:
[1] Standard Errors are heteroscedasticity robust (HC1)



模型三：非線性

$$\begin{aligned}
 win_dem_t1_i = & -10.4791 + 44.5190 demvoteshare_i - 46.2070 demvoteshare_i^2 \\
 & \quad (6.095) \quad (26.494) \quad (28.733) \\
 & + 0.3060 D_i - 6.8309(D_i \times demvoteshare_i) \\
 & \quad (10.174) \quad (40.210) \\
 & + 14.1764(D_i \times demvoteshare_i^2) + u_i \\
 & \quad (40.151)
 \end{aligned}$$

結果發現，由圖中可以看到跳升的現象非常明顯，第 i 期得票率從小於 50% 到大於 50% 時，在第 $i+1$ 期會突然增加 30.60% 的機率能夠當選。

研究討論／結論

1. 無法單純的以簡單的 OLS 迴歸的結果下定論，因為有無法去除的遺漏變數，而為了要去除這個問題，我們決定使用 RD 模型。且為了避免使用錯誤的模型估計，我們使用多組模型跑迴歸，結果發現非線性的模型 fit 數據的效果較好，所以我們討論以下結論時，會以模型三的結果討論。

2. 由RD模型，在 50% 的 cutoff 可以看到明顯的斷點，且無論是線性或是高次方迴歸都存在，大約跳升了 30.60%，這顯示執政優勢的確存在。在原本勢均力敵的狀況下，有執政的政黨在下次選舉時多了 30.60% 的機率選上連任。
3. 執政優勢的原因，我們認為是因為執政黨在任期間擁有影響力，會對於和自己政黨觀點或立場比較相同的議題多加著墨，例如若是執政黨傾向右派，則在任期間可能比較多和反墮胎、反安樂死、反外遇相關的講座、教育推廣、以及遊行；若執政黨傾向左派，則在任期間可能比較多和同性婚姻、女權、種族平等有關的講座、教育推廣、以及遊行，或者甚至進行相關議題的改革和法規重構。

在長期的教育環境下，中間選民的立場會慢慢的淺移默化，逐漸偏向執政黨，而這個改變可能連人民本身都沒有發現，此消彼長之下，在下一次的選舉時，執政黨會容易獲得更多中間選民的選票，因此更容易連任。

4. 在一直以來政治學家對於兩黨制的國家，政黨對於議題的立場究竟是因為人民影響政黨，或是政黨影響人民，而眾說紛紜。前者認為政黨為了獲取更多選票，應該會把自己的立場更趨向大眾，導致兩黨的訴求漸趨相同；後者認為由於執政黨擁有影響力，會對於自己關注的議題更多加推廣、教育，或是對於不滿意的制度進行改革，當人民在有傾向的政治環境下，中間選民的思想有可能會容易傾向和執政黨相同，如此一來政黨會更堅守自己的立場。

而我們以RD為模型的實驗，恰巧能說明後者的解釋力更大一些，由於執政優勢的確存在，而產生執政優勢的原因恰巧為後者的解釋路徑，是因為中間選民的立場因為潛移默化產生改變，所以我們有理由相信，一個國家政黨的立場，是能夠影響那個國家中人民的思想的。