



UNIVERSITA' DEGLI STUDI DI PADOVA

Facoltà di Scienze Statistiche

CORSO DI LAUREA IN STATISTICA, ECONOMIA E FINANZA

RELAZIONE FINALE DI LAUREA:

**IL SINDACO USCENTE CHE SI RICANDIDA
PARTE AVVANTAGGIATO SUL SUO
AVVERSARIO?
EVIDENZE DALLE ELEZIONI COMUNALI
ITALIANE A DOPPIO TURNO**

RELATORE: Ch.mo Prof. ENRICO RETTORE

LAUREANDO: DISION DAMIANO

Anno accademico 2008 - 2009

INDICE

INDICE.....	1
INTRODUZIONE.....	3
CAPITOLO 1	
COME VALUTARE L'IMPATTO DI UNA POLITICA?	
1.1	Introduzione.....6
1.2	Una formalizzazione del problema.....7
1.3	Il selection bias.....10
1.4	Metodi sperimentali
1.4.1	Proprietà della randomizzazione.....12
1.5	Metodi non sperimentali
1.5.1	<i>Il Regression Discontinuity Design</i>14
1.5.2	Due possibili situazioni.....15
1.5.3	Assunzioni <i>RDD</i>16
1.5.4	Specificazione del modello.....17
CAPITOLO 2	
TUTTO COMINCIA DA QUI: IL CASO AMERICANO	
2.1	Spiegazione generale.....19
2.2	Evidenza empirica.....20
2.3	L'analisi RDD.....21
2.4	I risultati dell'analisi di Lee (2008).....22
2.5	Test di specificazione.....24
CAPITOLO 3	
IL CASO ITALIANO	
3.1	introduzione.....26

3.2	Perché i ballottaggi?.....	27
3.3	La scelta dei comuni.....	28
3.4	Premesse fondamentali per l'analisi.....	31
3.5	L'analisi operativa attraverso l'RDD.....	33
3.6	Verifiche di compatibilità tra i gruppi.....	35
3.6.1	Probabilità di vittoria al tempo t-1.....	36
3.6.2	Differenza percentuali di voti al tempo t, rispetto alla popolazione totale residente nel comune.....	38
3.6.3	Differenze percentuali di voti al tempo t, rispetto al numero di famiglie residenti per ogni comune.....	41
3.6.4	Differenze percentuali di voti al tempo t, rispetto alla quota di residenti in possesso almeno del diploma.....	41
3.6.5	Differenze percentuali di voti al tempo t, rispetto al tasso di occupazione in ogni comune.....	44
3.6.6	Differenze percentuali di voti al tempo t, rispetto al tasso di disoccupazione in ogni comune.....	45
3.6.7	Differenze percentuali di voti al tempo t, rispetto al numero medio di componenti familiari.....	47
3.6.8	Differenze percentuali di voti al tempo t, rispetto alla quota di maschi residenti in ogni comune.....	48

CAPITOLO 4

CONSIDERAZIONI FINALI.....	50
-----------------------------------	-----------

BIBLIOGRAFIA.....	52
--------------------------	-----------

INTRODUZIONE

Con questo lavoro cerco di saperne di più, dal punto di vista statistico, sull'effetto causale derivante dall'aver governato una città sulla probabilità di rielezione successiva.

La cosa interessante della tesi è che poggia per intero su dati reali, infatti essi sono stati interamente forniti con gentilezza dal Ministero dell'Interno, perciò tale lavoro può interessare la platea dei singoli cittadini interessati, come anche i partiti politici e gli amministratori altrettanto bramosi di 'capire o prevedere' il proprio futuro.

La domanda di fondo che ci accompagnerà quindi è la seguente: "è davvero così automatico che governando per 5 anni sia più facile essere rieletti?".

"È il ruolo di amministratore, e i poteri che ne derivano, oppure sono altre le caratteristiche che rendono la probabilità di rielezione più alta?"

"Ammesso che vi sia qualche differenza nella probabilità di rielezione, dobbiamo tener conto delle differenze economiche e politiche proprie del territorio italiano (di zona in zona, da regione a regione, da nord a sud), oppure ciò non influisce sui risultati che si raggiungono?"

Queste sono alcune delle domande che mi sono posto prima ancora di partire con lo sviluppo dell'elaborato e che hanno dettato la strada.

Il lavoro di per sé non ha la pretesa di essere esaustivo o un punto fermo di questo tipo di ricerca, anche perché, come vedrà chi avrà la pazienza di arrivare in fondo a queste pagine, i risultati sono poco soddisfacenti dal punto di vista quantitativo prima ancora che qualitativo.

Lo schema logico con cui presento la tesi è dapprima una presentazione dei metodi per la valutazione di politiche, in particolare *il Regression Discontinuity Design*. Il secondo capitolo è dedicato al paper di Lee (2008), nel quale si studia il caso americano, che ha ispirato questa ricerca. Il terzo capitolo è invece il 'clou' dell'opera, poiché in esso viene spiegato tutto il lavoro di questi ultimi mesi con relative verifiche.

Comunque sia, posso già dire che l'ipotesi in questione non verrà confermata, essenzialmente perché i dati a disposizione non consentono stime sufficientemente precise. Ma il bello del partire con un lavoro di ricerca è che non si sa mai quale sarà l'arrivo...

CAPITOLO 1

COME VALUTARE L'IMPATTO DI UNA POLITICA?

1.1 Introduzione

Spesso è interessante sapere se una determinata decisione da parte di una autorità (sia essa una legge adottata dal Parlamento o un corso di qualificazione organizzato dal management aziendale) abbia avuto un qualche effetto sulla popolazione a cui si rivolgeva, detto in altri termini abbia raggiunto l'obiettivo.

Il 'ramo' della statistica che si occupa di ciò, è la cosiddetta 'valutazione d'impatto' e serve a misurare l'effetto di un intervento, definito come "la differenza tra ciò che osserviamo *dopo* che l'intervento è stato attuato e ciò che avremmo osservato, nello stesso periodo e per gli stessi soggetti, *in assenza* di intervento" (Martini, 2006). Definito in questo modo, si evita di considerare come risultati della politica quelli che si sarebbero ottenuti ugualmente anche senza alcun intervento, si identifica cioè il 'contributo netto' dell'intervento, separandolo dai molteplici fattori, estranei ad esso, che influenzano comunque i destinatari e le loro condizioni o comportamenti.

Definita la *variabile risultato* come la caratteristica dei soggetti appartenenti alla popolazione di riferimento rispetto alla quale si conduce la valutazione d'impatto (il più delle volte l'obiettivo della politica), l'effetto di una politica consiste nel confronto tra un evento fattuale e un evento controfattuale.

L'evento fattuale è il valore effettivo che la variabile risultato assume in un periodo di tempo e che è osservato; al contrario, il contro fattuale consiste nell'evento ipotetico, in particolare, per gli esposti è il valore che la variabile risultato avrebbe mostrato nello stesso periodo di tempo nell'ipotesi in cui non fossero stati esposti all'intervento, per i non esposti è il valore che la stessa variabile avrebbe mostrato qualora fossero stati esposti (vedi Figura 1.1).

Può sembrare ovvio, ma in effetti osservare ciò che sarebbe successo, in assenza di intervento, agli stessi soggetti esposti non è possibile.

Ecco perché si ricorre ad alcune tecniche statistiche, in evoluzione negli ultimi anni, per avvicinarci ad una stima sempre più verosimile di tale impatto.

Perché una valutazione d'impatto sia praticabile è necessario:

- 1) che sia ben definita la politica come insieme di azioni intraprese con un certo fine;
- 2) che sia ben definita la popolazione obiettivo;
- 3) che una o più caratteristiche dei componenti la popolazione obiettivo possano essere modificate dalla politica, tali caratteristiche devono essere ben definite anche in assenza della politica e devono essere misurabili;
- 4) che si disponga di dati (di natura amministrativa o raccolti ad hoc) che consentano di misurare accuratamente caratteristiche, comprese quelle modificabili dalla politica stessa, di individui coinvolti e non coinvolti nella politica.

1.2 Una formalizzazione del problema

Sia P la popolazione obiettivo costituita da N soggetti/unità, Y la variabile risultato per ogni soggetto.

In generale l'effetto della politica sull'i-esimo soggetto si rappresenta come :

$$\alpha_i = Y_i^T - Y_i^{NT} \quad (1.1)$$

Valore che la variabile risultato assumerebbe se questo fosse esposto alla politica

Valore che la variabile risultato assumerebbe se questo NON fosse esposto alla politica

Il problema è dato dal fatto che α_i **non è osservabile** poiché, come evidenziato sopra, un soggetto o è esposto o non è esposto ad un trattamento. Per superare questa

situazione indichiamo con I la variabile binaria che vale 1 se l' i -esima unità è stata esposta all'intervento e 0 altrimenti. Il valore assunto da I_i , vale a dire lo status esposto/non esposto alla politica, determina quale dei due esiti potenziali diventa l'evento fattuale e quale l'evento controfattuale (vedi figura 1.1).

	I	
	0	1
Y_i^T	evento controfattuale	Evento fattuale
Y_i^{NT}	evento fattuale	evento controfattuale

Figura 1.1: Distribuzione di eventi fattuali e contro fattuali in base allo status I

Una volta definita la variabile indicatrice, possiamo spostare l'attenzione dai singoli α_i alla distribuzione di α nella popolazione di riferimento.

- possiamo valutare l'impatto tra esposti e non esposti, dove il nostro parametro di interesse è : $E[\alpha | I = 1]$ (1.2)

Come esposto in precedenza possiamo osservare solo i risultati fattuali e perciò si possono calcolare i risultati medi da trattati per coloro che sono stati esposti e da non trattati per coloro che non sono stati esposti.

Ma il vero problema è che l'impatto medio per gli esposti (ATT¹):

$$E[\alpha | I = 1] = E[Y^T | I = 1] - E[Y^{NT} | I = 1] \quad (1.3)$$

coinvolge un termine non osservabile, quindi il parametro d'interesse risulta esso stesso non identificabile.

Per approssimare la quantità non identificabile, si utilizza una semplice tecnica matematica, sommare e sottrarre la stessa quantità $E[Y^{NT} | I = 0]$, per

¹Average treatment on the treated

cui si rende calcolabile l'equazione (1.4), che indica la **differenza tra esposti e non esposti** come **impatto della politica** più **selection bias** (ovvero la distorsione dovuta alla selezione che indica le differenze tra i due gruppi non dovute alla politica bensì al processo di selezione):

$$\begin{aligned} E[Y^T | I = 1] - E[Y^{NT} | I = 0] &= \\ &= E[\alpha | I = 1] + \left(E[Y^{NT} | I = 1] - E[Y^{NT} | I = 0] \right) \end{aligned} \quad (1.4)$$

Il tipico problema della valutazione d'impatto consiste proprio nel depurare la differenza tra le medie misurabili dal *selection bias*, in modo tale da mettere in luce solo l'impatto medio sugli esposti dovuto all'effetto della politica.

- allo stesso modo si determina il parametro di interesse se il confronto avviene *prima/dopo* : ovvero se abbiamo informazioni su un gruppo di esposti prima e dopo un programma e vogliamo valutarne l'efficacia al netto dell'evoluzione che gli stessi soggetti avrebbero avuto anche in assenza del programma (dinamica spontanea) , la base teorica è la stessa:

$$\begin{aligned} \text{Il parametro di interesse è: } E[\alpha_{t_2} | I = 1] &= E[Y_{t_2}^T | I = 1] - E[Y_{t_2}^{NT} | I = 1] \\ &(1.5) \end{aligned}$$

L'ultimo termine a destra non è osservabile, quindi il parametro di interesse non è identificabile.

Come prima, superiamo 'l'empasse' con un piccolo trucco matematico, sommando e sottraendo la quantità $E[Y_{t_1}^{NT} | I = 1]$, così risulta valida e calcolabile la seguente equazione:

$$\begin{aligned} (1.6) \quad E[Y_{t_2}^T | I = 1] - E[Y_{t_1}^{NT} | I = 1] &= \\ &= E[\alpha_{t_2} | I = 1] + \left(E[Y_{t_2}^{NT} | I = 1] - E[Y_{t_1}^{NT} | I = 1] \right) \end{aligned}$$

Ciò che è evidenziato rappresenta l'evoluzione media nel tempo della variabile risultato che gli esposti avrebbero ottenuto se non fossero stati esposti (dinamica o evoluzione spontanea). Quindi la differenza prima/dopo per gli esposti è data dall'impatto della politica più la distorsione dovuta alla dinamica spontanea.

Concludendo la panoramica generale sulla valutazione di politiche possiamo affermare che per stimare l'impatto si può procedere in due modi:

- istituire il confronto in modo tale che il termine di distorsione sia nullo (in assenza della politica le differenze osservate sono, mediamente, nulle).
- rilevare informazioni che consentono di controllare/eliminare al meglio il termine di distorsione.

1.3 Selection bias

Con il termine *selection bias* indichiamo una quantità che rappresenta una distorsione, infatti impedisce alla quantità calcolabile di coincidere con la quantità d'interesse. Esso in pratica ci indica che i soggetti esposti ($I=1$) otterrebbero dei risultati mediamente diversi dai soggetti non esposti ($I=0$), anche se non fossero coinvolti nel trattamento.

Viene chiamato '*bias* di selezione' poiché la differenza tra esposti e non esposti dipende in modo cruciale da come avviene il processo di selezione, infatti è uguale a zero se e solo se la selezione avviene casualmente e i due gruppi rappresentano la stessa popolazione.

Al contrario è diverso da zero ogniqualvolta il processo di selezione dà luogo ad un gruppo di soggetti esposti all'intervento sistematicamente diversi dai non esposti rispetto a certe caratteristiche X rilevanti per la variabile risultato.

Il valutatore dispone di due tipi fondamentali di strategie per eliminare il *selection bias*, rispetto alla modalità di valutazione esse sono:

- valutazioni sperimentali

- valutazioni non sperimentali

1.4 Metodi sperimentali

“L’approccio sperimentale: è una sorta di must della ricerca valutativa, esso rappresenta un pilastro di una fetta rilevante della teoria valutativa contemporanea, e alcuni dei maggiori autori la ritengono tuttora la regina degli approcci” (Bezzi, 2005²)

Quando la selezione avviene tramite assegnazione casuale o randomizzata ai due gruppi, vale la seguente condizione di ortogonalità:

$(Y^T, Y^{NT}) \perp I \implies$ posso affermare con certezza che

$$E[Y^{NT} | I = 1] = E[Y^{NT} | I = 0] \quad (1.7)$$

In pratica esposti e non esposti alla politica sono rappresentativi della stessa popolazione, per cui le differenze tra le medie dei risultati dei due gruppi ($I=1, I=0$) coincide con l’impatto medio dell’intervento nella popolazione.

Operativamente dapprima si estrae un campione casuale di soggetti, detto gruppo sperimentale, che avrà diritto a partecipare al programma, e un gruppo di controllo, costituito dai restanti soggetti, che non parteciperà alla politica. Il fatto che il gruppo di controllo ed il gruppo di trattamento siano tratti casualmente dalla stessa popolazione obiettivo garantisce, in condizioni ideali, la correttezza della stima ottenuta dall’impatto reale sui partecipanti al programma.

Questo metodo è utilizzato prevalentemente nell’ambito di progetti-pilota, cioè la messa in atto di una politica solo su scala ridotta, allo scopo di valutarne gli effetti ed eventualmente estendere tali risultati al resto della popolazione.

²Claudio Bezzi in: “Alberto Martini, 2006, Metodo sperimentale approccio controfattuale e valutazione degli effetti delle politiche pubbliche”

1.4.1 Proprietà della randomizzazione

Il metodo sperimentale è argomento molto dibattuto tra gli esperti del settore, infatti esso gode di molti pro e di alcuni contro che fanno sì che tale approccio vanti molti estimatori, ma non sia così facile da esportare in tutte le realtà della valutazione di politiche.

Di sicuro ‘la necessità di manipolare il processo di selezione’ mediante assegnazione casuale è al tempo stesso la grande forza e il grande limite del metodo sperimentale” (Martini, 2006). **Forza** perché l’assegnazione casuale garantisce che, salvo il caso, l’unica differenza tra gruppo sperimentale e gruppo di controllo sta nel fatto di essere o meno esposti all’intervento e che quindi tutto ciò che succede al gruppo di controllo riproduce *plausibilmente* ciò che sarebbe successo al gruppo sperimentale *se questo non fosse stato esposto* all’intervento.

Ne rappresenta anche il **Limite** fondamentale per due ragioni:

Non è sempre eticamente sostenibile, molto spesso far parte o meno di un gruppo di trattamento (pensiamo al caso di un vaccino di cui è conosciuto l’effetto benefico, o di un corso di formazione professionale) può davvero esser decisivo per la vita di un uomo...è sempre saggio e giusto lasciarla nelle mani del caso??

Spesso gli assegnatari ai due gruppi si ‘ribellano’, vanificando di fatto gli sforzi del gruppo di valutatori (ciò accade spesso in casi in cui non è certo l’esito dell’esperimento, si pensi ad un nuovo farmaco) che perdono di fatto il controllo dell’esperimento.

Le proprietà prettamente statistiche sono invece 2:

- i. Funziona solo con la media a causa della linearità della media, a differenza infatti degli altri indici come la mediana o la varianza per cui tale proprietà non viene rispettata, cosa necessaria, ad esempio, per calcolare l’impatto medio per gli esposti come:

$$E(\alpha | I = 1) = E(Y^T - Y^{NT} | I = 1) = E(Y^T | I = 1) - E(Y^{NT} | I = 1)$$

(1.8)

- ii. Consente lo studio delle distribuzioni marginali di (Y^T, Y^{NT}) , non della congiunta.

Quando tale manipolazione sia impossibile o comunque non sia proponibile, ma si desideri comunque quantificare l'effetto di un intervento, occorre ricorrere ad altri metodi per ricostruire la situazione controfattuale, metodi che i valutatori definiscono come *non-sperimentali*.

Qui ci troviamo di fronte a una grande varietà di metodi, la cui applicabilità dipende dalle caratteristiche dell'intervento, dai dati utilizzabili, dalle risorse a disposizione.

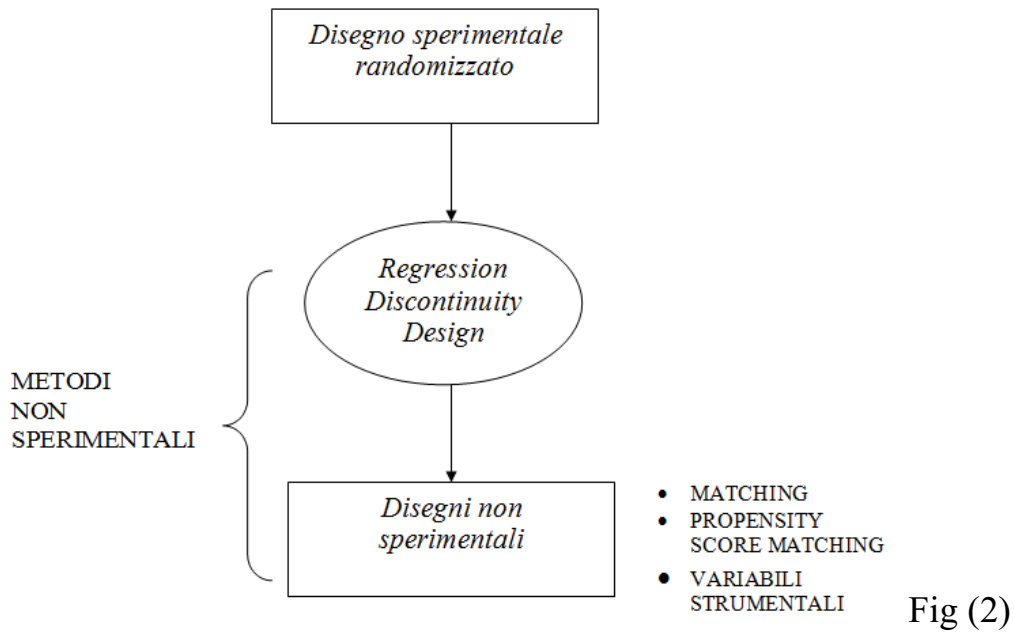
1.5 Metodi non sperimentali

In questo paragrafo si cerca di spiegare come procede il valutatore di un intervento se non ha la possibilità di selezionare casualmente i due gruppi, per cui si limita ad osservare quel che accade senza poter intervenire sulla selezione.

Il suo lavoro comincia subito dopo l'intervento (ex-post) e consiste nel lavorare su *observational data* per eliminare o controllare al meglio il *selection bias*. Egli raggiunge questo scopo se riesce ad ottenere un gruppo di confronto con caratteristiche simili al gruppo dei trattati, creando così uno scenario controfattuale.

In questi casi i gruppi di controllo e dei trattati sono scelti o tramite *autoselezione* o da una *autorità esterna*.

In generale l'RDD si colloca a metà tra i due metodi di valutazione, sperimentale e non; questo perché così come per il secondo metodo, non si lavora con una situazione 'da laboratorio', bensì il gruppo dei trattati è sistematicamente diverso da quello dei non trattati. Il punto forte che lo caratterizza, è la conoscenza della variabile che determina la selezione (in questo modo i due gruppi tornano ad essere identici salvo quella caratteristica), cosa che non avviene nei disegni classici non sperimentali, ove non si conosce il processo di selezione, ed il valutatore non sa rispetto a quali caratteristiche i due gruppi sono diversi.

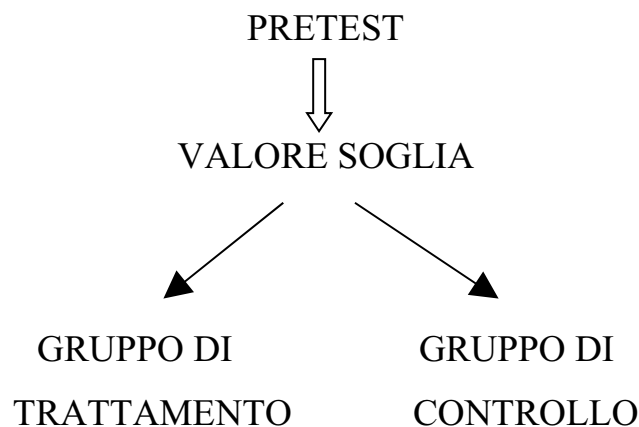


1.5.1 IL REGRESSION DISCONTINUITY DESIGN

Questo disegno fu introdotto per la prima volta nella letteratura nel campo della valutazione da Thistlethwaite e Campbell nel 1960, fatta eccezione di alcune pubblicazioni econometriche, solo alla fine degli anni '90 trovò un largo numero di applicazioni in studi economici e cominciò a svilupparsi.

Esso rientra nella categoria dei metodi di valutazione non sperimentali e si sviluppa su due gruppi di soggetti : *trattamento e controllo*, su entrambi i gruppi si dispongono di misure pre e post trattamento.

- I partecipanti sono sottoposti ad un PRE-TEST che misura una qualche variabile prima dell'inizio del programma



La divisione nei due gruppi avviene quindi a seconda che il valore del pre test sia inferiore o minore al valore soglia; ciò determina come prima cosa che i due gruppi non siano perfettamente comparabili; anzi, per costruzione, essi sono selezionati rispetto al valore della misura pre-test (*selezione nelle osservabili*) che evidentemente non è uguale per tutti gli individui.

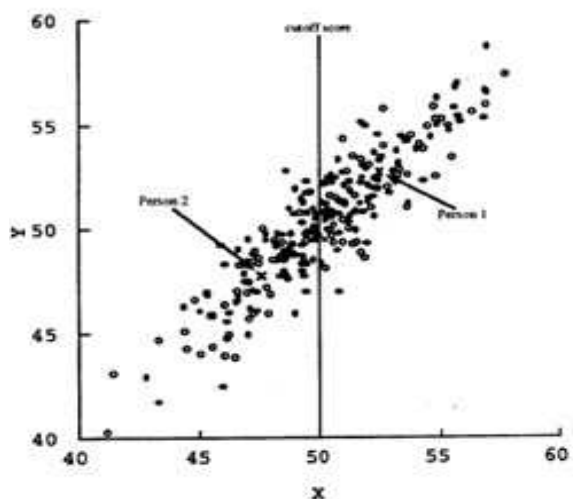
La cosa principale per la buona riuscita di un RDD è che il valutatore sappia con certezza quali siano le regole o le caratteristiche che determinano la selezione, infatti *il teorema fondamentale* su cui si basa questo particolare disegno non sperimentale dice che se i due gruppi vengono resi equivalenti rispetto ad x (la variabile che determina la selezione) non ci si deve più preoccupare delle altre variabili (vi è sicurezza che marginalmente al valore soglia, chiamiamolo X_c , i due gruppi sono identici rispetto a tutte le altre caratteristiche, salvo il caso), tale che :

$$(Y^T, Y^{NT}) \perp I \mid X = X_c \quad (1.9)$$

Questo è il motivo per cui si restringe l'attenzione a quello che capita attorno al valore soglia (*cutoff score*), infatti grazie alla (9) l'RDD permette di sfruttare i benefici della randomizzazione all'interno di una determinata sotto popolazione.

Proprio sul cutoff si dice che il valutatore 'va a caccia di discontinuità' (*discontinuity gap*) che coincide con la ricerca dell' EFFETTO CAUSALE.

1.5.2 Due possibili situazioni



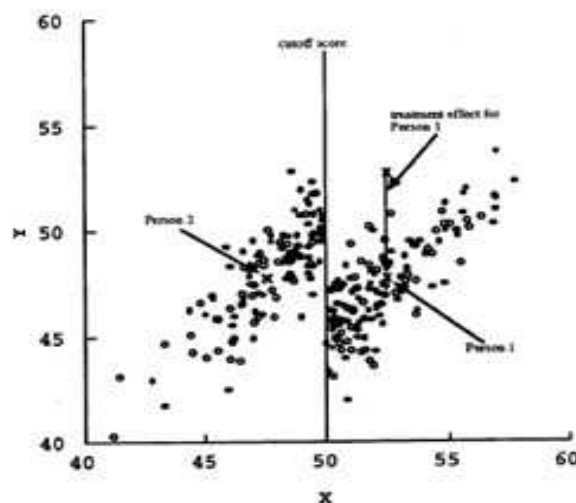
(fig 1.3)

a. DOPO LA SELEZIONE NESSUNO VIENE TRATTATO

Resta un'elevata correlazione tra prima e seconda misurazione, nulla di più.

b. DOPO LA SELEZIONE AVVIENE IL TRATTAMENTO

L'attenzione si sposta marginalmente al valore soglia...si nota che l'effetto causale del trattamento è stato quello di abbassare mediamente il valore della seconda misurazione per i trattati. (fig 1.4)



1.5.3 Assunzioni dell'RDD

Una volta che il valutatore ha determinato la variabile continua e osservabile, il Regression Discontinuity design è caratterizzato da una serie di proprietà:

- i. Perfetta assegnazione ai due gruppi (esposto/non esposto) attraverso il valore soglia, cioè oltrepassando il valore soglia la probabilità di far parte dell'intervento passa da 0 a 1 si parla di *sharp rdd*.
Si denota invece con *fuzzy rdd* quando la probabilità di ricevere il trattamento non cambia da 0 ad 1 obbligatoriamente sulla soglia e quindi la differenza tra la $\Pr [I(Xc^-)=1]$ e la $\Pr[I(Xc^+)=1]$ è minore di 1: persone che secondo la regola sarebbero escluse partecipano ugualmente al programma e persone che secondo la regola dovrebbero partecipare decidono di non essere trattate.
- ii. Corretta forma funzionale della variabile pre e post trattamento *solo in un'analisi parametrica*; i metodi non parametrici invece non poggiano su forti assunzioni sulla forma funzionale del modello matematico per approssimare il fenomeno in questione, sono in grado quindi di produrre inferenze 'robuste' (Rettore Trivellato e Martini, 2002) poiché il modello ipotizzato è solo un'approssimazione del vero modello;
- iii. Non ci sono altri fattori oltre l'effetto del trattamento che possano creare discontinuità, e ciò, oltre ad essere essenziale, sta nelle mani del buon valutatore verificarlo;

- iv. Trattati ricevono lo stesso ammontare di trattamento, in modo tale da poter considerare il trattamento come quantitativamente unico per tutti i trattati. .

1.5.4 Specificazione del modello

Y risultato post test

X risultato pre-test

I variabile trattato-non trattato (I=1 se $X > x_c$)

x_c valore soglia

Se la relazione tra X e Y è lineare:

$$(1.10) \quad Y_i = \beta_o + \beta_1 X_i + \beta_2 I_i + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2)$$

per $i=1 \dots n$

Si ha:

$$\begin{aligned} E[Y^T | I = 1, X] - E[Y^{NT} | I = 0, X] &= \\ &= (\beta_o + \beta_1 X + \beta_2) - (\beta_o + \beta_1 X) = \beta_2. \end{aligned} \quad (1.11)$$

Il coefficiente β_2 è l'**effetto del trattamento**; il modello impone che l'effetto sia uguale per ogni individuo e non vari con X.

La relazione può assumere forme più complesse, con interazioni tra variabile pre-test e variabile di assegnazione al trattamento e anche forme quadratiche.

La sua espressione più generale è:

$$Y_i = \beta_o + \beta_1 X_i + \beta_2 I_i + \beta_3 X_i^2 + \beta_4 X_i I_i + \dots + \beta_{p-1} X_i^s I_i + \beta_p X_i^s I_i + \varepsilon_i \quad (1.11)$$

In questo caso β_4 rappresenta l'effetto d'interazione, ovvero il grado con cui l'effetto del programma è collegato al valore delle misure pre-test.

Per arrivare ad una corretta specificazione bisogna tener conto che nel RDD le esplicative sono fissate a priori (X e trattamento), perciò la scelta della corretta forma

funzionale della relazione tra X e Y è l'assunto fondamentale del RDD (e il problema più complicato).

Infine in un RDD è meglio scegliere un numero di termini superiore al modello vero piuttosto che inferiore (sovrapparametrizzare).

CAPITOLO 2

IL CASO AMERICANO

2.1 Spiegazione generale

Da Lee (2001): “The electoral advantage to incumbency and voters’ valuation of politicians’ experience: a regression discontinuity analysis of elections to the U.S. House”

In questo studio l’autore verifica se vincere un seggio al Congresso possa avere un effetto causale sulla probabilità di essere rieletto alla successiva elezione.

Questo perché i deputati eletti al Congresso, secondo l’autore, agiscono in modo da aumentare la probabilità di essere rieletti, attraverso vari meccanismi:

- influenzare le tasse, la politica della spesa o la politica monetaria in modo da favorire i propri elettori;
- vendere favori politici in cambio di contributi per la campagna elettorale;
- votare leggi che riflettono l’ideologia o gli interessi economici dei propri elettori.

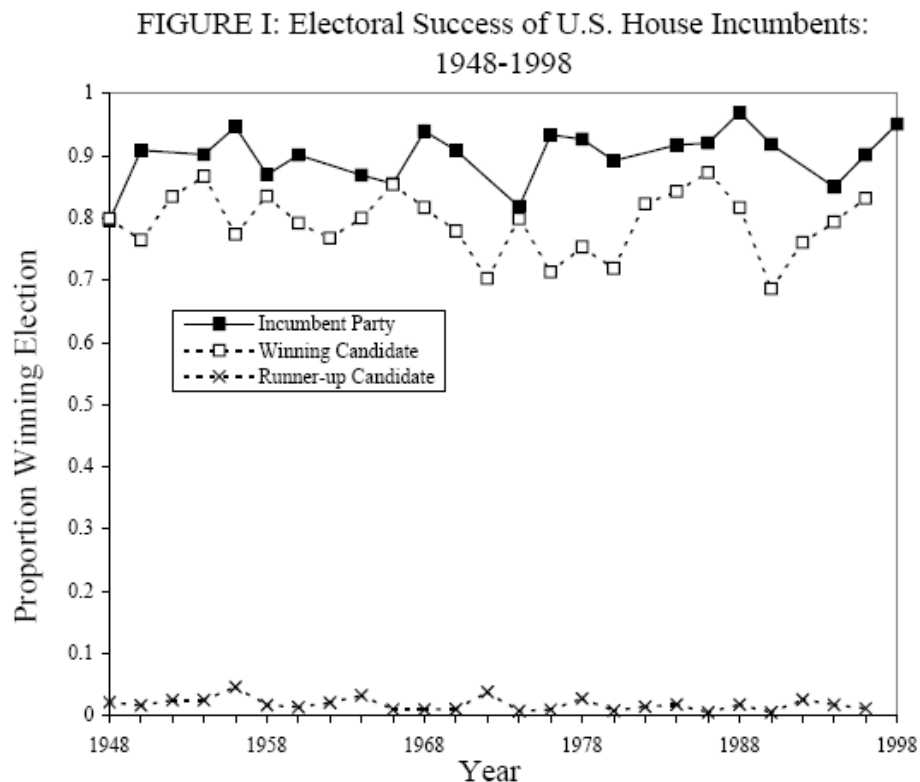
Il Congresso degli Stati Uniti d’America è un parlamento bicamerale, che ha sede nel Campidoglio a Washington, ed è composto da un Senato (Senate) e da una Camera dei Rappresentanti (House of Representatives). La Camera dei Rappresentanti è attualmente formata da 435 membri eletti in collegi uninominali ripartiti su base statale in proporzione alla popolazione dello Stato stesso, mentre il Senato è composto da 100 membri. Comunque, quale che sia la consistenza della popolazione dei singoli Stati, ad ognuno di essi è garantito almeno un rappresentante. Attualmente, in base alla ripartizione del 2000, hanno un solo rappresentante l’Alaska, il Delaware, il Montana, il Dakota del Nord, il Dakota del Sud, il Vermont e il Wyoming. Lo Stato che elegge più rappresentanti è invece la California che ne conta 53.

La principale funzione del congresso consiste nel produrre le leggi federali, destinate cioè a valere su tutto il territorio degli USA.

Il caso Americano rappresenta una situazione di studio abbastanza semplice, poiché la lotta politica avviene praticamente solo tra due schieramenti: il Partito Repubblicano e il Partito Democratico.

2.2 Evidenza empirica

La motivazione dello studio di Lee (2008) viene da una chiara evidenza, ovvero che vi è un tasso molto elevato di successi elettorali per i cosiddetti ‘*incumbents*’ (coloro che sono in carica), ciò viene evidenziato attraverso un grafico molto chiaro:



Note: Calculated from ICPSR study 7757. Details in Data Appendix. Incumbent party is the party that won the election in the preceding election in that congressional district. Due to re-districting on years that end with "2", there are no points on those years. Other series are the fraction of individual candidates in that year, who win an election in the following period, for both winners and runner-up candidates of that year.

Tre sono le evidenze che balzano subito all'occhio in figura 1:

- il partito che vince le elezioni all'occasione t le rivince all'occasione $t+1$ nel 90% dei casi;

- i candidati eletti hanno l'88% di probabilità di essere ricandidati e condizionatamente al fatto di essersi ricandidati hanno una probabilità del 90% di essere rieletti;
- i candidati avversari (runner-up) hanno una probabilità pari al 20% di ricandidarsi e una probabilità pari al 3% di essere eletti.

2.3 L'analisi e l'RDD

Quindi osservando l'evidenza empirica sarebbe facile affermare che l'effetto causale c'è ed è pure positivo. Ma come precedentemente esposto non abbiamo alcuna dimostrazione del fatto che i risultati elettorali di eletti e non eletti non siano dovuti, in parte o del tutto, a caratteristiche pre-elettorali che li rendevano diversi a priori:

“It is well known, however, that the simple comparison of incumbent and non-incumbent electoral outcomes does not necessarily represent anything about a true electoral advantage of being an incumbent.”
(David Lee³)

Ad esempio, in collegi a maggioranza democratica il candidato democratico avrà una probabilità più alta di farsi eleggere e quindi di farsi rieleggere, così come il candidato più ricco o più carismatico potrebbe vantaggi maggiori.

Per ovviare a questo, il metodo usato è il **Regression Discontinuity Design**, poiché si confrontano i risultati elettorali dei candidati che vincono le elezioni per un soffio con quelli che le perdono per un soffio, creando così quel valore soglia (50% dei voti), unica causa di selezione-separazione dei gruppi.

³2007, *Randomized experiments from non-random selection in U.S. House elections*

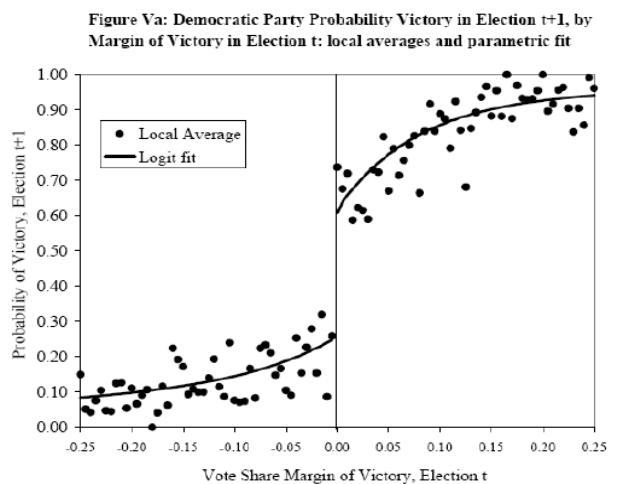
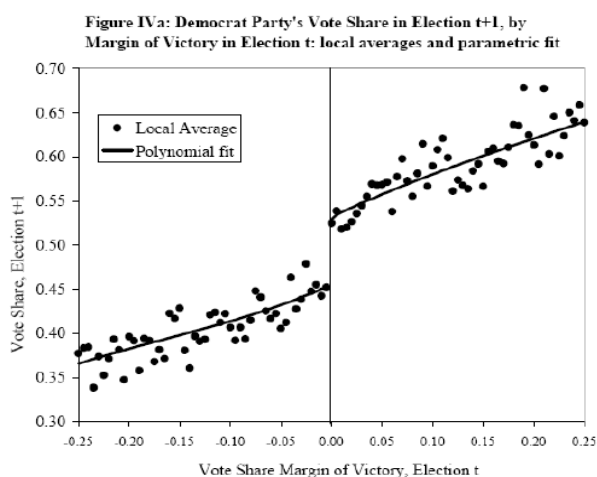
L’RDD è così formulato:

- VARIABILE PRE-TEST = RISULTATO ELEZIONI al tempo t
- VALORE SOGLIA: 50% DEI VOTI
- TRATTATI = ELETTI
- NON TRATTATI = NON ELETTI
- VALORE POST TEST = PROBABILITA’ ESSERE RIELETTO AL tempo t+1

Così facendo siamo sicuri che i due gruppi siano ex-ante in tutto e per tutto comparabili (avendo avuto risultati elettorali molto simili), tranne per il fatto che uno ha vinto le elezioni e l’altro no (quindi possiamo assimilare questo evento ad una estrazione casuale, in cui vincer o perdere dipende dal lancio di una moneta).

A questo punto è palese come l’eventuale variazione della probabilità dei vincitori di essere rieletti al tempo t+1 nei collegi prossimi al valore soglia, dipenda solo dal fatto di aver vinto le elezioni ed aver agito, durante la legislatura, per essere rieletti: quindi, la discontinuità che emerge dal grafico, della probabilità di vittoria all’elezione successiva tra coloro che hanno vinto rispetto ai perdenti al tempo t nel punto di soglia, può essere propriamente interpretata come l’effetto causale dell’aver vinto le elezioni al tempo t sulla probabilità di vincerle al tempo t+1.

2.4 I risultati dell’analisi di Lee (2008)

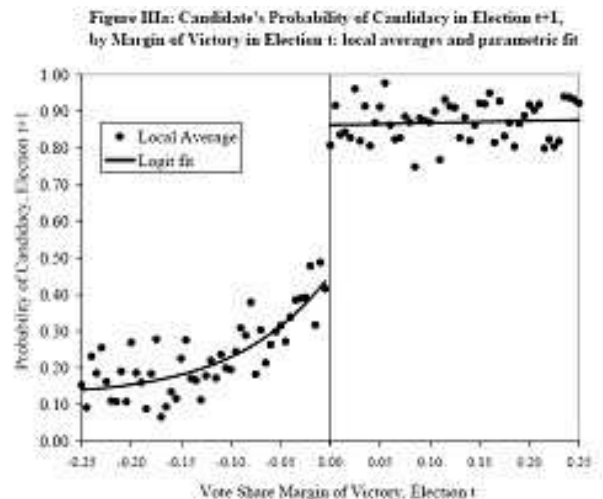
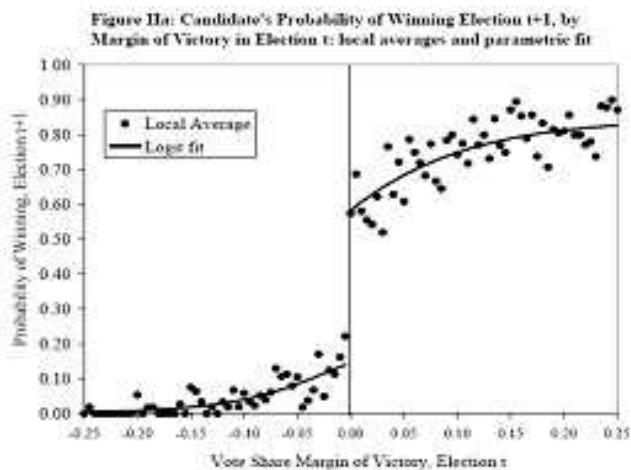


Le figure IVa e Va si concentrano più sul partito che sul candidato (così come la mia analisi in seguito). La prima a sinistra mostra come varia la percentuale di voti presi dal partito democratico alle elezioni al tempo $t+1$ al variare del margine di vittoria al tempo t . La seconda invece mostra l'analoga relazione per la probabilità di vittoria alle elezioni al tempo $t+1$.

Mi concentro più sulla seconda delle due poiché più simile alla ricerca da me condotta. Qui è palese come l'effetto causale dell'aver vinto al tempo t sia di circa 45 punti percentuali.

Ogni punto riportato nel grafico non è altro che una media locale del variabile risposta Y (binomiale 0-1), ed il modello stimato si adatta bene ai dati.

Le evidenze sono molto simili anche nelle analisi successive.



La figura IIa mostra come varia la probabilità di vincere le elezioni al tempo $t+1$ per il candidato in carica al variare della differenza di voti ottenuta alle elezioni precedenti.

La figura IIIa mostra invece la probabilità per il candidato di esser ricandidato all'elezione successiva, rispetto alla differenza percentuale di voti all'elezione precedente.

Anche in quest'ultime analisi, la discontinuità è presente significativamente, il che avvalora l'ipotesi di partenza.

2.5 Test di specificazione

Per confermare l'interpretazione della discontinuità come effetto causale dell'aver vinto al tempo t , è necessario provare che i due gruppi di candidati siano comparabili, cioè che non vi siano differenze ex-ante che li rendano sistematicamente diversi.

Lee (2008) considera altre caratteristiche dei due gruppi e verifica se sono diversi rispetto a tali caratteristiche (quest'ultime devono essere pre-determinate, quindi osservate prima, e non modificabili dal trattamento).

Se i due gruppi sono comparabili, queste caratteristiche non devono creare alcuna discontinuità sulla soglia del grafico pre-test / post-test.

Qualora così non fosse, avremmo trovato una caratteristica che rende i due gruppi sistematicamente diversi a priori (oltre alla variabile che crea la selezione) e ciò farebbe sì che l'effetto causale possa essere dovuto a tale diversità e non al fatto di aver vinto o meno al tempo t .

Insomma verrebbe meno l'assunto fondamentale*(rimando alla prima parte teorica: 1.9) del Regression Discontinuity Design.

Le variabili considerate sono:

1. numero di vittorie ad elezioni precedenti (fig. IIb)
2. numero di tentativi in elezioni precedenti (fig. IIIb)
3. percentuali di voti per il partito democratico al tempo $t-1$ (fig. IVb)
4. probabilità di vittoria del partito democratico all'elezione al tempo $t-1$ (fig. Vb)

Figure IIb: Candidate's Accumulated Number of Past Election Victories, by Margin of Victory in Election t: local averages and parametric fit

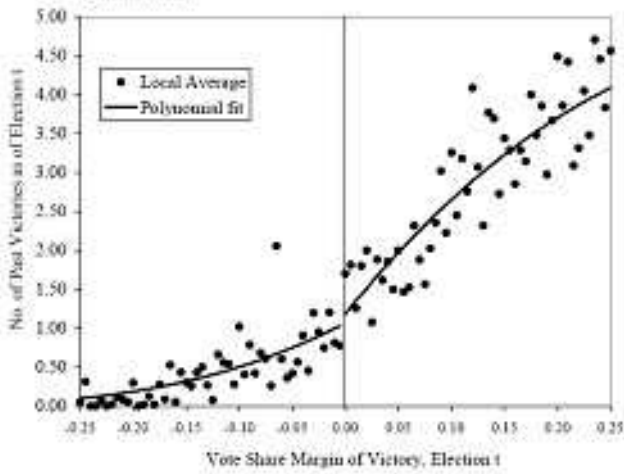


Figure IIIb: Candidate's Accumulated Number of Past Election Attempts, by Margin of Victory in Election t: local averages and parametric fit

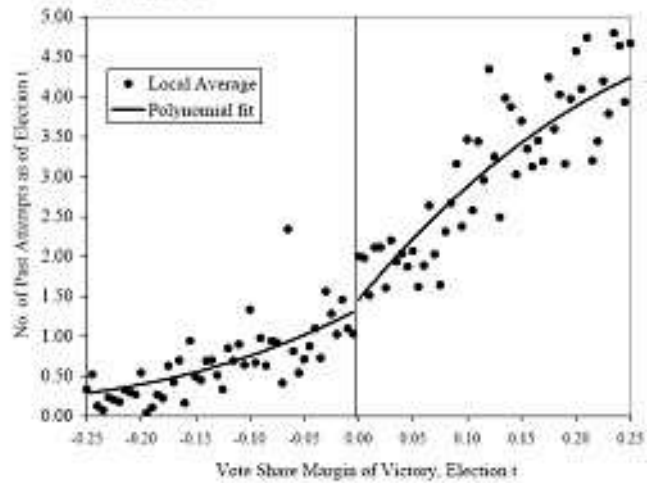


Figure IVb: Democratic Party Vote Share in Election t-1, by Margin of Victory in Election t: local averages and parametric fit

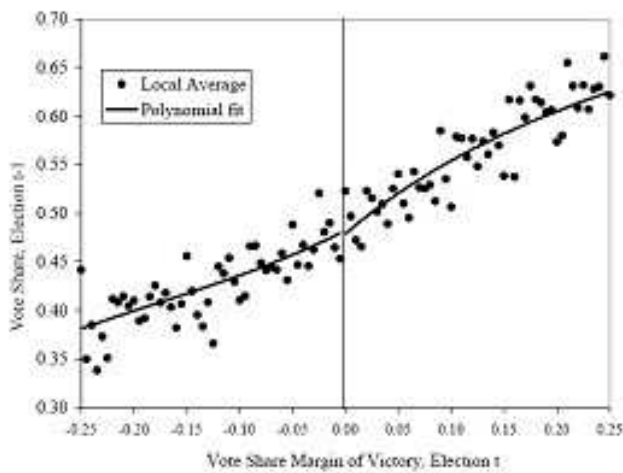
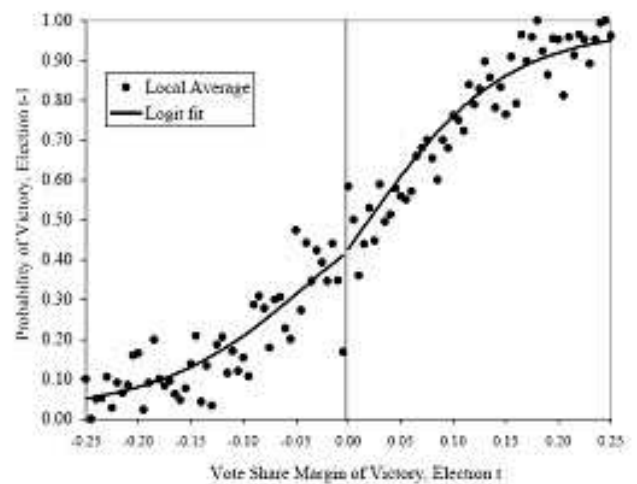


Figure Vb: Democratic Probability of Victory in Election t-1, by Margin of Victory in Election t: local averages and parametric fit



È chiaro che i due gruppi sono abbastanza comparabili in queste quattro dimensioni, infatti non vi sono discontinuità nel punto di soglia.

Tutto questo avvalora l'ipotesi che l'effetto causale sulle variabili obiettivo è dovuto proprio all'esperienza di una legislatura al congresso.

CAPITOLO 3

IL CASO ITALIANO

3.1 Introduzione

Lo scopo di questo lavoro è quello di verificare attraverso il *Regression Discontinuity Design*, l'effetto dell'aver vinto al ballottaggio (e quindi aver governato) sulla probabilità di essere rieletti all'elezione successiva.

Dal 1993 in Italia l'elezione dei sindaci dei comuni con almeno quindicimila abitanti e dei presidenti delle province ha luogo mediante ballottaggio fra i due candidati più votati al primo turno, qualora nessuno dei candidati abbia ottenuto la maggioranza assoluta dei voti validamente espressi (senza contare, dunque, le schede bianche e nulle), ovvero il 50% più uno.

Il mio lavoro prende spunto dall'analisi precedentemente esposta sulle elezioni al Congresso degli Stati Uniti. Infatti l'impianto teorico è lo stesso e poggia sull'ipotesi che essere eletti, nel caso americano al Congresso, nel nostro a sindaco, possa avere un effetto positivo sulla probabilità di essere rieletti all'elezione successiva e quindi mantenere la 'poltrona' (carica).

Questo perché colui che governa ed amministra il potere, può coltivare i propri consensi attraverso leggi, concessioni e restrizioni volte ad accrescere il favore dei propri elettori.

Ovviamente la realtà dei comuni italiani non è molto omogenea nelle caratteristiche, infatti vi sono zone più politicamente schierate di altre (pensiamo all'Emilia, alla Toscana, al Veneto....), zone molto diverse quanto a ricchezza, a densità abitativa, a possibilità di proseguire gli studi... tutto questo complica il nostro intento valutativo, in quanto se confrontiamo gruppi di comuni tra loro diversi a priori l'eventuale effetto del trattamento che troveremo potrebbe non coincidere con l'effetto causale da

noi cercato, per via del *bias* da selezione. Dobbiamo quindi utilizzare i cosiddetti metodi non sperimentali, che a parità di determinate condizioni rendono confrontabili i due gruppi (trattato/non trattato o meglio vincitore/vinto) eliminando il *selection bias* dato dal fatto che i due gruppi non sono confrontabili ex-ante.

Per cui si è scelto di utilizzare l’RDD, poiché attraverso esso andiamo a confrontare i successi elettorali di candidati che vincono le elezioni per un soffio, con quelli che le perdono per un soffio, considerando il governare per una legislatura come trattamento a cui viene sottoposto solo il gruppo che prende almeno il 50% più uno dei voti.

Al termine della legislatura andiamo a vedere se il candidato o il suo partito vengono riconfermati.

3.2 Perché i ballottaggi?

Per avere una situazione simile a quella degli Stati Uniti d’America, in cui 52 stati eleggono ognuno il proprio rappresentante e la lotta politica è, da anni, solo tra due schieramenti in Italia dobbiamo arrivare fin all’elezione del sindaco (e della provincia). Sono queste infatti le uniche elezioni nelle quali si riesce a ricondursi a una logica bipolare: centro-destra contro centro-sinistra.

In particolare a noi interessano le elezioni decise ‘al margine’, quelle cioè la cui vittoria può essere equiparata al lancio della moneta, e cioè casuale.

Questo lo possiamo riscontrare nei casi di ballottaggio alle elezioni comunali, infatti qui arrivano solo i comuni ‘in bilico’, quelli cioè in cui ex-ante entrambi i contendenti sono alla pari poiché hanno avuto un risultato elettorale, al primo turno, che non testimonia una prevalenza dell’uno sull’altro, possiamo quindi assumere che l’arrivare al ballottaggio e ancor più i risultati decisi al margine indicano che tra i due gruppi (esposti e non esposti per una manciata di voti) non vi sono grosse differenze se non il fatto che l’uno ha vinto le elezioni e l’altro no.

Per cui tramite il nostro disegno possiamo identificare al margine del valore soglia, l’effetto causale dell’aver vinto l’elezione precedente sulla probabilità di vittoria all’elezione successiva.

Nel caso di elezione al primo turno invece, non avremmo la possibilità di interpretare tale probabilità come effetto causale dell'aver governato, poiché il vincere al primo turno indica proprio una certa differenza tra il gruppo dei trattati (vincenti) e non trattati (sconfitti), molto spesso la caratterizzazione politica territoriale è complice di tali differenze.

Avremmo potuto valutare anche i ballottaggi provinciali, però sarebbe venuta meno una condizione essenziale su cui si basa l'analisi, ovvero che il voto sia espressione e conseguenza dell'operato di colui è stato eletto; l'esperienza italiana ci racconta che i cittadini vedono la provincia come un organo sostanzialmente inutile di cui più o meno nessuno conosce ruoli e competenze, perciò sarebbe stata un'imposizione abbastanza forte quella di vedere il voto come un giudizio 'sul mandato'.

3.3 La scelta dei comuni

La nostra scelta è stata quella di cercare i dati degli ultimi 15 anni solo per le elezioni dei comuni (infatti si presuppone che per l'elezione del sindaco, più che per altre elezioni, valga molto la conoscenza effettiva della persona e l'operato di essa piuttosto che l'appartenenza politica ed ideologica); e tra essi solo i comuni con più di 90.000 abitanti, cioè escludendo dall'insieme quella miriade di piccoli comuni in cui è più facile che i candidati non siano riconducibili ai due schieramenti destra – sinistra. In realtà inizialmente il lavoro si sviluppava solo sulle città capoluogo di regione, in esse infatti prevale molto tale divisione politica, essenziale per le analisi; successivamente ho allargato il campione mantenendo intatte le ipotesi di partenza, infatti ci siamo limitati ai 55 comuni con più di 90.000 abitanti⁴ (più L'Aquila e Potenza inclusi perché capoluoghi di regione).

In effetti con questi grandi comuni è stato quasi sempre possibile arrivare alla valutazione del ballottaggio come scontro tra centro-destra e centro-sinistra (a parte alcuni sporadici casi in cui vi erano appartenenti allo stesso schieramento come

Fonte ISTAT: 2001, *14° Censimento della popolazione italiana*

avversari, oppure liste civiche non rientranti in nessuna categoria politica che arrivavano al ballottaggio, questi dati non sono stati presi in considerazione).

3.4 Premesse fondamentali per l'analisi

Paradossalmente fin dall'inizio il vero problema di questa analisi non è stata tanto la modellazione dei dati e l'interpretazione dei risultati. L'impresa davvero ardua è stata quella di trovare tutti i dati di cui necessitavamo.

Fortunatamente dopo qualche tentativo di ricerca 'di comune in comune' (cosa che poteva destare qualche perplessità a livello di univocità ed autorevolezza delle fonti), il Ministero dell'Interno ci ha aiutati dandoci tutto il materiale che si chiedeva.

La tabella 3.1 riassume in modo dettagliato le informazioni importanti per l'analisi, estrapolate dai data set ministeriali.

Tabella 3.1: Per ogni ballottaggio vengono riportati il nome del comune, l'anno del ballottaggio, 'Vcd' la percentuale di voti andata al centro destra, 'Vcs' la percentuale di voti per il centro sinistra (ho raggruppati i candidati per schieramento), 'x' la differenza di voti Vcd-Vcs, ed infine 'Y' la variabile d'interesse che indica se all'elezione successiva, indipendentemente dal fatto che essa sia al primo turno o al ballottaggio, abbia vinto il centro destra (Y=1) o il centro sinistra (Y=0).

COMUNE	anno	Vcd	Vcs	x	Y
Andria	1993	24.8	75.2	-50.4	0
Ancona	1993	28.5	71.5	-43	0
...					
...					
...					
Verona	1994	61.5	38.5	23	1
Siracusa	1999	73.6	26.4	47.2	1

Così facendo i ballottaggi che ho preso in considerazione sono solo 80, una numerosità campionaria non molto ampia per il tipo di lavoro che vado a fare, però ho preferito essere stringente su alcuni casi limite non prendendoli in considerazione piuttosto che 'penalizzare' l'analisi con approssimazioni errate.

Inoltre mi sono fermato ad 80 casi poiché gli ultimi ballottaggi, quelli di cui conosco l'esito al tempo t, ma non so ancora l'esito dell'elezione al tempo t+1 (l'esempio più

famoso è Roma nell'ultima elezione del 2008 vinta da G. Alemanno), chiaramente non posso includerli nell'analisi.

Detta x la variabile su cui si basa l'assegnazione al gruppo di trattamento, ovvero la differenza percentuale di voti tra i due schieramenti, *voti al centro-destra meno voti al centro-sinistra*, se essa risulta positiva ($x > 0$) al tempo t la vittoria va al candidato del centro destra, viceversa se $x < 0$ la vittoria al tempo t va al candidato di centro sinistra; *zero* è il 'valore soglia', quello per "cui poco più in là si è trattati, poco più in qua no".

Considero il trattamento come l'aver governato da parte del centro destra, per cui utilizzo una nuova variabile z che mi indica lo status di trattato o meno:

$$z = \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x < 0 \end{cases} \quad (3.1)$$

A questo punto sono andato a verificare se esiste una discontinuità nel punto $x=0$ per la probabilità che il centro-destra vinca le elezioni al tempo $t+1$ dopo aver governato per una legislatura, nei comuni simili in tutto (implicazione che deriva proprio dall'essere arrivati al ballottaggio) ma che al tempo t hanno avuto un risultato appena al di sotto del 50% utile per governare, e quelli che il 50% l'hanno ottenuto per poco. (Il risultato non è molto soddisfacente, principalmente a causa della bassa numerosità campionaria che determina la non precisione della stima.)

3.5 L'analisi operativa attraverso l'RDD

Ho selezionato i casi di ballottaggio nei comuni italiani dal 1993 al 2004, di fatto l'ultimo anno in cui dopo un ballottaggio so già il risultato dell'elezione successiva, indipendentemente che essa sia finita al primo o al secondo turno.

Poiché i due gruppi sono ex-ante uguali -salvo il caso- l'eventuale discontinuità nel punto di soglia, può essere propriamente interpretata come l'effetto casuale dell'aver vinto le elezioni al tempo t sulla probabilità di vincerle al tempo $t+1$.

Y_{t+1} = risultato dell'elezione successiva
 $\begin{cases} 1 \text{ se vince centro-destra} \\ 0 \text{ se vince centro-sinistra} \end{cases}$

x_t = differenza di voti al tempo t in percentuale espressa come $V_{cd} - V_{cs}$;

z = variabile indicatrice che indica lo status di trattato o meno (vedi (3.1));

il modello usato è un glm (*generalized linear model*) di famiglia binomiale (la variabile dipendente ha solo 2 possibili valori), il link usato è il *probit* (per cui il predittore lineare è funzione della distribuzione normale).

Nell'uso del *regression discontinuity design* vi è da tenere conto di due importanti aspetti. Al fine di eliminare il selection bias è sufficiente includere tra le esplicative la variabile x , cioè la variabile che determina l'assegnazione al gruppo di trattamento. In secondo luogo è fondamentale specificare correttamente il modello, ovvero la forma della relazione tra la variabile dipendente e x .

Tutto ciò l'ho potuto riscontrare in questo lavoro, infatti si sono susseguite numerose specificazioni (*step*) per arrivare dal modello saturo a quello che si adatta meglio ai dati.

Infatti il modello saturo comprende interazioni e termini al quadrato: il *quadratic pretest coefficient* (il coefficiente quadratico della variabile pre-test) e il *quadratic interaction* (interazione quadratica): per cui il modello si presenta così:

$$y_i = \beta_0 + \beta_1 \tilde{X}_i + \beta_2 Z_i + \beta_3 \tilde{X}_i Z_i + \beta_4 \tilde{X}_i^2 + \beta_5 \tilde{X}_i^2 Z_i + e_i \quad (3.2)$$

Dove i β_i sono i parametri, e_i sono i termini d'errore.

Tabella 3.2: *Stima dei parametri del modello saturo*

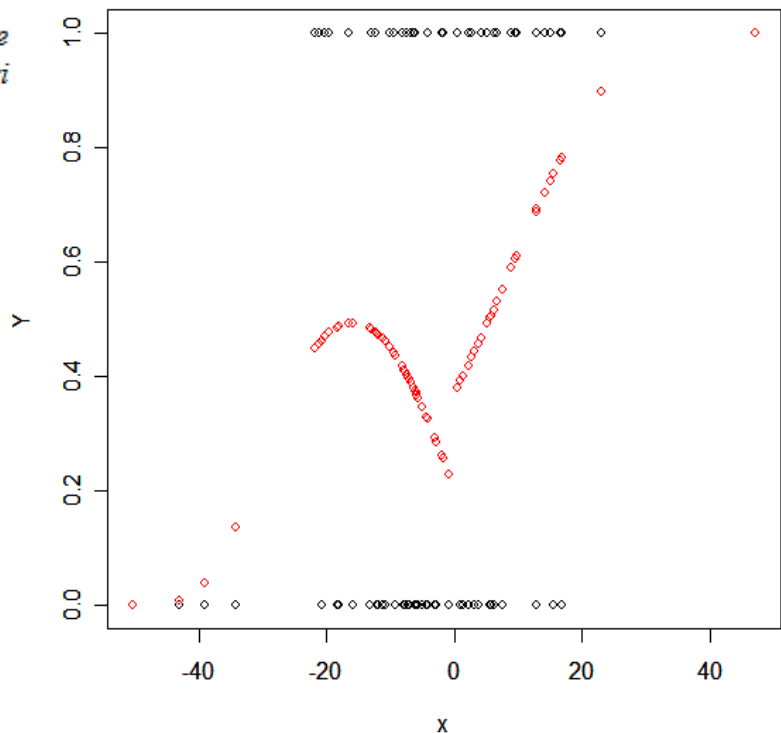
Coefficients:				
	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.827105	0.511206	-1.618	0.106
x	-0.102136	0.075763	-1.348	0.178
zTRUE	0.495642	0.785170	0.631	0.528
I(x^2)	-0.003216	0.002281	-1.410	0.159
x:zTRUE	0.158814	0.157333	1.009	0.313
zTRUE:I(x^2)	0.003771	0.007082	0.532	0.594

Null deviance: 110.453 on 79 degrees of freedom
Residual deviance: 99.208 on 74 degrees of freedom
AIC: 111.21

Si nota che il valore che ci interessa davvero, ovvero il coefficiente di z (poiché indica l'effetto del trattamento puro), ha un livello di significatività abbastanza elevato, tale da indurci a credere che nel caso dei comuni italiani l'effetto del governare sulla probabilità di essere rieletti non è poi così presente.

Grafico 3.1: *diagramma di dispersione dei dati osservati (nero), e dei valori stimati attraverso il modello saturo*

Il grafico 3.1 è costruito in modo da evidenziare sull'asse delle ascisse la differenza percentuale di voti tra i due schieramenti, sull'asse delle ordinate la probabilità di essere rieletti all'elezione successiva, per ogni livello di differenza di voti.



Si nota che come avevamo dedotto dagli output del modello, il grafico ci conferma l'ipotesi di non significatività della discontinuità al margine (il che implica che l'aver vinto di poco non incide sulla

elezione successiva): anche se vediamo che marginalmente alla differenza di voti pari a 0 una qualche discontinuità la troviamo, essa non è statisticamente significativa.

Procedendo con l'eliminazione di parametri dal modello, tenendo conto del principio della 'parsimonia' nelle esplicative, il modello migliore confrontando tra gli altri anche l'indice AIC risulta essere il seguente:

Figura 3.3: *Stime del modello ridotto. Selezionando il miglior modello, restano i coefficienti di z, della differenza di voti al quadrato e dell'interazione al quadrato.*

```
Call:
glm(formula = Y ~ z + I(x^2) + z:I(x^2), family = binomial(probit))

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.1829135   0.2124337  -0.861    0.389
zTRUE         0.0481530   0.3790838   0.127    0.899
I(x^2)       -0.0006073   0.0005120  -1.186    0.236
zTRUE:I(x^2)  0.0039310   0.0022826   1.722    0.085 .

Null deviance: 110.45  on 79  degrees of freedom
Residual deviance: 101.47  on 76  degrees of freedom
AIC: 109.47
```

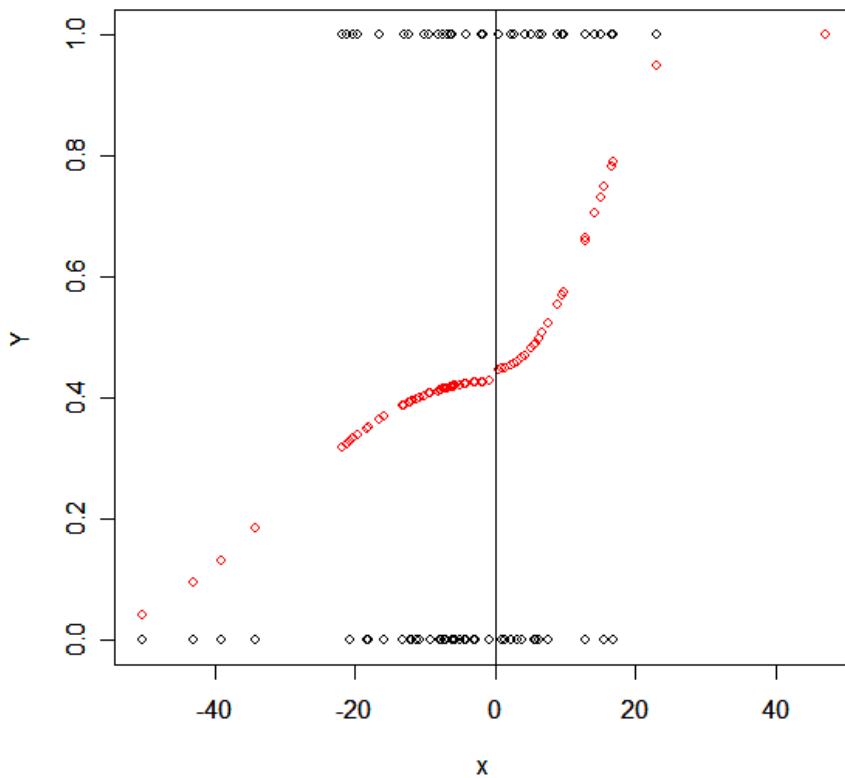
Si nota che il parametro di interesse per noi, ovvero il coefficiente di z, non è significativo, anzi l'ipotesi di nullità verificata mediante il test t è confermata.

In questo caso risulta marginalmente significativa solo l'interazione tra z e la differenza percentuale di voti al quadrato, il che potrebbe significare che non conta tanto il segno della differenza di voti al tempo t (ovvero se vince il centro-destra o il centro-sinistra), che al quadrato viene eliminato, ma è il vincere di poco o di tanto che conta, poiché più la differenza di voti è marcata più la probabilità di essere rieletti sale rapidamente, più si avvicina a zero, da una parte e dall'altra più si appiattisce attorno al 40% (vedi grafico 3.2).

Si può azzardare l'ipotesi che in Italia, chi governa una città non sia agevolato di per sé per la riconferma, cioè non basta vincere di un pelo un'elezione per governare in

modo tale da aumentare significativamente la propria probabilità di essere rieletti, bensì servono ‘i numeri per governare’ cioè una maggioranza solida in consiglio e un consistente consenso iniziale; ma poiché, attraverso l’RDD, il nostro confronto si limita al margine, ove la differenza è circa 0, poiché solo qui i due gruppi sono confrontabili, ci limitiamo a non rilevare alcuna differenza di probabilità per la rielezione.

Grafico 3.2: *diagramma di dispersione dei dati osservati (nero) e dei valori osservati attraverso il modello ridotto (rosso)*



Da questo grafico è ancora più evidente come l’effetto causale che cercavamo non sia presente, nel punto di soglia la discontinuità è assolutamente minima, questo indica che la vittoria al tempo t , seguita da circa 5 anni di governo, non ha alcun effetto sulla probabilità di vittoria al tempo $t+1$ per il centro destra. (Comunque l’andamento dei valori stimati indica una certa

continuità di voto laddove le differenze al tempo t sono più evidenti.)

3.6 Verifiche di compatibilità tra gruppi

Ho comunque voluto scongiurare eventuali differenze pre-trattamento tra i due gruppi tali per cui la stima dell'effetto dell'aver governato o meno venisse distorto.

Infatti per poter considerare l'eventuale discontinuità sulla soglia rispetto alla probabilità di vittoria al tempo $t+1$, come effetto causale dell'aver vinto al tempo t , i due gruppi di candidati devono avere le caratteristiche osservabili e non osservabili non sistematicamente diverse tra vincitori e perdenti marginali.

Il metodo usato è quello di esaminare altre caratteristiche dei due gruppi (queste sono pre-determinate quindi non modificabili dal trattamento) e vedere se i due gruppi sono diversi rispetto ad esse.

Per svolgere tale verifica ho raccolto dal censimento 2001 alcuni dati; con essi non c'è la pretesa di collocare su un ottica temporale l'analisi (non è aggiornata anno per anno o ballottaggio per ballottaggio), ma assumiamo che un singolo sindaco nell'arco della propria legislatura, durante gli ultimi 15 anni, non abbia potuto modificarli significativamente, così da poterli assumere come caratteristiche date.

Le variabili pre-determinate considerate sono:

- 1) Popolazione totale residente nel comune
- 2) Numero di famiglie
- 3) Quota residenti con almeno il diploma (laureati + diplomati)
- 4) Tasso di occupazione
- 5) Tasso disoccupazione
- 6) Numero medio di componenti per famiglia
- 7) Quota di maschi residenti rispetto al totale

Tabella 3.3: La seguente tabella riporta per ogni comune i dati ufficiali del censimento del 2001

Comune	pop_tot	num_fam	res_dipl	tasso_occ	tasso_disocc	media_comp	q_maschi
Ancona	100507	41175	44.1	46.3	5.9	2.4	0.47
Andria	95653	30736	19.9	35.4	21.9	3.1	0.49
Arezzo	91589	36478	37.4	48.2	5.2	2.5	0.47
...							
...							
...							
Verona	253208	109786	41	47.6	4.9	2.3	0.48
Vicenza	107223	44716	42.2	49.8	3.1	2.4	0.47
Udine	95030	43520	47.1	46.7	5.3	2.2	0.46

3.6.1 Probabilità di vittoria al tempo t-1

Seguendo Lee (2008), considero la differenza di voti tra i due schieramenti per ogni ballottaggio al tempo t ($V_{cd}-V_{cs}$) come variabile pre-test, però questa volta valuto se tale differenza è correlata con l'aver vinto al tempo t-1.

In pratica si rovescia l'analisi precedente, andando questa volta ad inserire nell'analisi anziché i risultati dell'elezione successiva, quella dell'elezione precedente (poiché nell'anno 1993 il sistema elettorale è cambiato, ho escluso dal campione quei ballottaggi il cui risultato precedente apparteneva al periodo precedente al 1993).

La scelta delle variabili esplicative non muta, mantengo le stesse specificazioni dell'esercizio precedente; anche nella modalità di valutazione nulla cambia, infatti andremo a valutare nel punto di soglia (laddove la differenza di voti al tempo t è circa zero e si vince o si perde per effetto del caso) se vi è o no una qualche discontinuità che determinerebbe la prova del fatto che i due gruppi non sono uguali al tempo t.

Perciò:

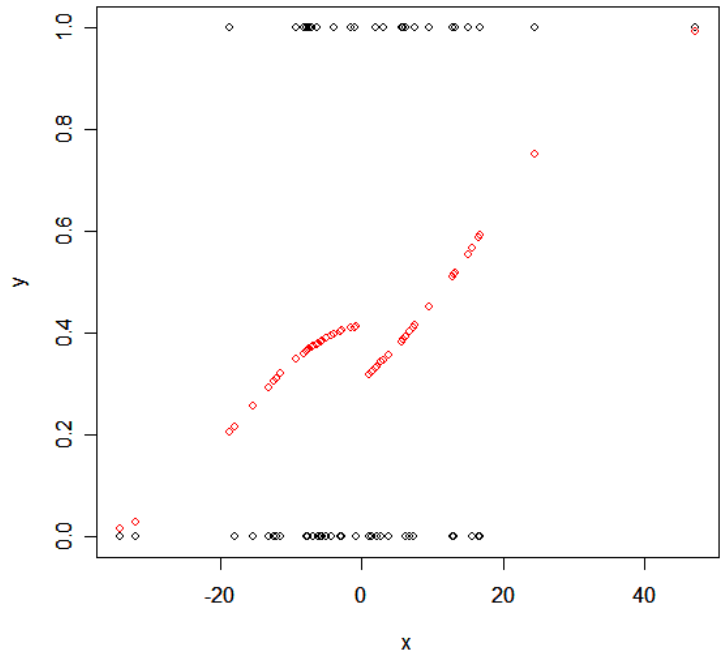
$$Y_{(t-1)} = \text{risultato dell'elezione precedente} \begin{cases} =1 & \text{se vince centro-destra} \\ =0 & \text{se vince centro-sinistra} \end{cases}$$

$X_{(t)}$ = differenza di voti al tempo t in percentuale espressa come $V_{cd}-V_{cs}$;

z = variabile indicatrice che indica lo status di trattato o meno.

Tabella 3.4: stime della regressione del modello completo, con variabile dipendente Y ovvero la vittoria del centro-destra al tempo t-1				
<code>glm(formula = Y ~ x + z + x * z + I(x^2) + I(x^2) * z, family = binomial(probit))</code>				
Coefficients:				
	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.216681	0.612004	-0.354	0.723
x	0.005346	0.112410	0.048	0.962
zTRUE	-0.288124	0.812986	-0.354	0.723
I(x^2)	-0.001474	0.004052	-0.364	0.716
x:zTRUE	0.028437	0.141720	0.201	0.841
zTRUE:I(x^2)	0.002077	0.005058	0.411	0.681

Grafico 3.3: *diagramma di dispersione della regressione in tabella 3.4; i valori osservati in colore nero, quelli stimati col modello saturo in rosso.*



Evidentemente il test per la significatività del coeff. di z accetta l'ipotesi di nullità, per cui non rinveniamo alcuna correlazione tra l'aver governato successivamente rispetto al risultato dell'elezione precedente.

Il Grafico 3.3 evidenzerebbe una certa discontinuità marginalmente al valore zero della differenza di voti percentuali al tempo t , ma essa non è statisticamente significativa.

Da notare inoltre che per effetto delle limitazioni 'temporali' la numerosità campionaria è solo di 58 unità (non ho potuto prendere in considerazione la fascia di ballottaggi la cui elezione precedente è avvenuta prima del 1993), abbastanza scarso per poter aver una stima precisa.

Con il modello ridotto le conclusioni non cambiano, infatti il coefficiente della variabile indicatrice del trattamento, z , rimane non significativo, e in generale non c'è un grande adattamento del modello ai dati, come dimostrano i seguenti grafici.

Tabella 3.5: *stime del modello di regressione ridotto con y risultato elezione al tempo $t-1$*

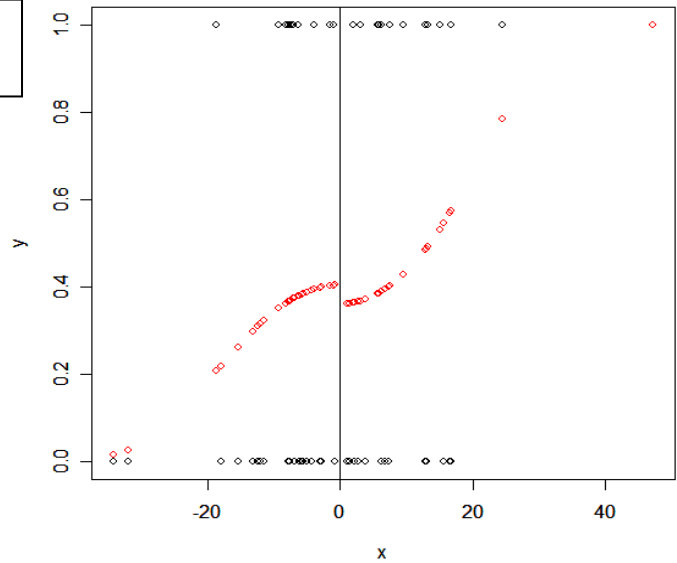
```
Call:
glm(formula = y ~ z + I(x^2) + z:I(x^2), family = binomial(probit))
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.242061	0.283952	-0.852	0.394
zTRUE	-0.114111	0.429531	-0.266	0.790
I(x^2)	-0.001657	0.001653	-1.002	0.316
zTRUE:I(x^2)	0.003584	0.002508	1.429	0.153

grafico 3.4: *diagramma di dispersione dei valori stimati dal modello ridotto in Tabella 3.5 (rosso) e valori osservati (nero).*

Il grafico avvalora la nostra ipotesi, infatti la discontinuità che c'è non è statisticamente significativa, ciò indica che non vi è correlazione tra l'aver governato al tempo t e l'esito delle elezioni precedenti a t .



3.6.2 Differenza percentuali di voti al tempo t , rispetto alla popolazione totale residente nel comune

Questa è la prima delle verifiche basata sul censimento, il modello applicato è quello di una relazione lineare tra la variabile dipendente 'numero di residenti' rispetto alla differenza di voti al tempo t (x) e la solita variabile z che indica l'essere esposti o meno al trattamento.

Parto sempre dal modello 'saturato' fino ad arrivare ad uno 'ridotto' cercando il miglior adattamento possibile ai dati.

Tabella 3.6: *stime della regressione del modello saturo con il numero di residenti totali per ogni comune come variabile dipendente.*

```
lm(formula = pop_tot ~ x + z + x * z + I(x^2) + I(x^2)*z)
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  334878.1   94451.9   3.545 0.000693 ***
x             11329.9   11737.9   0.965 0.337656
z1TRUE       -164118.0  136049.1  -1.206 0.231645
I(x^2)         135.2     249.2    0.543 0.589011
x:z1TRUE      -3147.1   17732.4  -0.177 0.859633
z1TRUE:I(x^2)  -334.3     391.0   -0.855 0.395488
Residual standard error: 263900 on 72 degrees of freedom
Multiple R-Squared:  0.0383,    Adjusted R-squared:  -0.02848
F-statistic: 0.5736 on 5 and 72 DF,  p-value: 0.72
AIC 2176.522
```

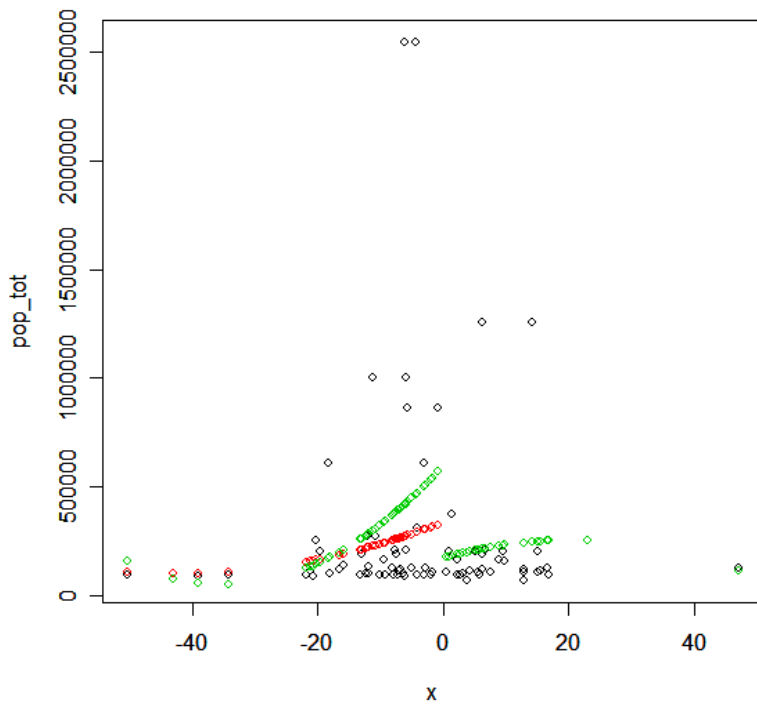


grafico 3.5: grafico di dispersione della popolazione totale contro la differenza di voti tra centro-dx e centro-sx per ogni ballottaggio, in nero i valori osservati, in verde le stime del modello completo con i dati di Roma, in rosso le stime escludendo Roma dal campione.

Nel grafico 3.5 si nota una discontinuità abbastanza elevata nel punto di soglia, cosa che dalle stime del modello non si evince, non essendo significativo il coefficiente del parametro z . Tale discontinuità è

creata dal numero di abitanti troppo elevati della città di Roma che vanta 2.546.804 abitanti rispetto agli altri comuni che spaziano dai 69.060 di Potenza al milione circa di Milano. È per questo che le stime indicate in colore verde evidenziano una discontinuità nel punto di soglia molto elevata.

Per ovviare a ciò ho stimato un ulteriore modello che non tiene conto dei valori della capitale, infatti la discontinuità si riassorbe.

Sempre nel grafico 3.5 il secondo modello (escludendo tali valori estremi) è in colore rosso, e varia rispetto al precedente solo nella parte relativa ai non trattati, infatti per i trattati le stime si sovrappongono poiché Roma non è presente. Nelle stime in figura 3.6 è riportato solo l'output del modello che non tiene conto dei valori della capitale.

Come prima stimo anche il modello ridotto e ne riporto il grafico, viene selezionato tenendo conto del principio della parsimonia, in modo da evitare l'inclusione nel modello di variabili con ridotto potere esplicativo; ma soprattutto valutando i minori valori dell'AIC (*Akaike Information Criterion*):

Tabella 3.7: stima della regressione 'ridotta', numero dei residenti totali come variabile dipendente
Call: lm(formula = pop_tot ~ z1 + I(x^2) + I(x^2) * z1)
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	254357.43	44089.86	5.769	1.74e-07 ***
z1TRUE	-32225.48	67415.75	-0.478	0.634
I(x^2)	-93.42	76.91	-1.215	0.228
z1TRUE:I(x^2)	64.48	142.36	0.453	0.652

Residual standard error: 262700 on 74 degrees of freedom
Multiple R-Squared: 0.0208, Adjusted R-squared: -0.0189
F-statistic: 0.5239 on 3 and 74 DF, p-value: 0.6672
AIC
[1] 2173.929

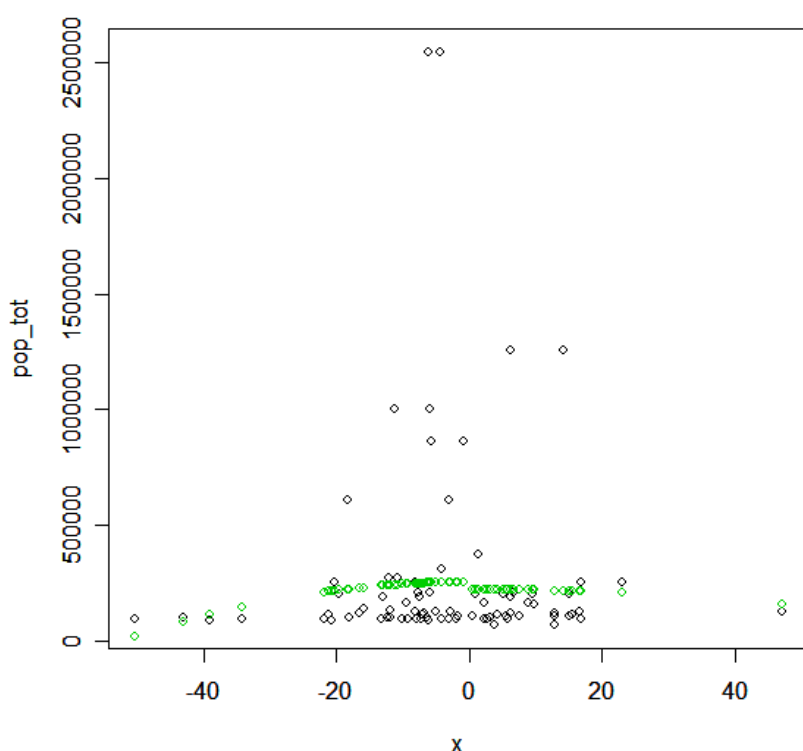


Grafico 3.6: grafico di dispersione della popolazione totale contro la differenza di voti tra centro-dx e centro-sx per ogni ballottaggio con il modello ridotto, in nero i valori osservati, in verde le stime del modello ridotto con i dati di Roma,

Come ci si aspettava il valore dell' R^2 peggiora, poiché ci si allontana dal modello saturo, però si nota una piccola diminuzione della devianza residua (*residual standar error*). Ma il valore importante è la piccola decrescita dell'AIC che seppur di poco ci dice che è preferibile mantenere il modello ridotto. Comunque sia il valore del coefficiente di z non risulta essere significativamneto diverso da zero, per cui le conclusioni non cambiano.

Il grafico 3.6 che come al solito esprime per ogni valore di x (sulle ascisse) il corrispettivo numero di abitanti che ha quel comune, le stime in verde oltre a non tener conto dei valori di Roma non evidenziano alcuna discontinuità.

Si può quindi concludere che tale caratteristica non rende certo diversi i due gruppi di vincenti e perdenti marginali.

3.6.3 Differenze percentuali di voti al tempo t, rispetto al numero di famiglie residenti per ogni comune

Se tengo come variabile dipendente il numero di famiglie residenti, il risultato della specificazione è lo stesso del modello precedente, essendo il numero di famiglie legato al numero di residenti in un comune.

Si nota che non essendoci alcuna discontinuità significativa nel punto di soglia, neanche questa caratteristica rende diversi i due gruppi al tempo t sul livello soglia, oltre appunto all'aver vinto o meno.

3.6.4 Differenze percentuali di voti al tempo t, rispetto alla quota di residenti in possesso almeno del diploma

Questa quota non è espressa esplicitamente nel censimento, ma l'ho calcolata sommando il numero di laureati e diplomati residenti nel comune, diviso il numero di residenti totali.

Mantengo sempre le stesse esplicative, dapprima con il modello saturo, poi con quello ridotto.

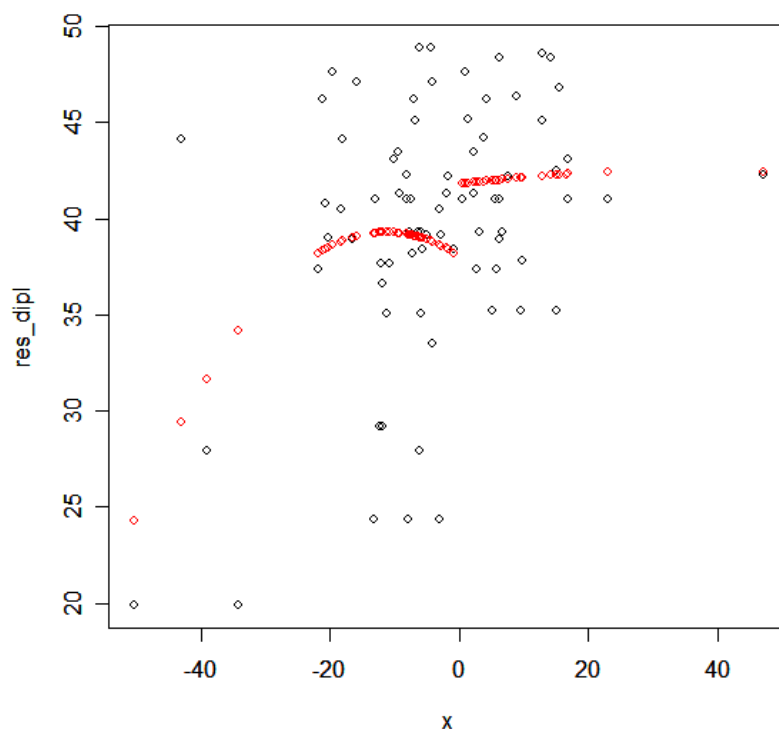
Tabella 3.8: *Stime della regressione con la quota di residenti almeno diplomati residente in ogni comune come variabile dipendente, modello saturo.*

```
Call:
lm(formula = res_dipl ~ x + z + x * z + I(x^2) + I(x^2) * z)

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  38.052669   2.079452  18.299  <2e-16 ***
x            -0.220561   0.262887  -0.839   0.4042
zTRUE        3.729964   3.041219   1.226   0.2239
I(x^2)       -0.009786   0.005608  -1.745   0.0852 .
x:zTRUE      0.261979   0.399811   0.655   0.5143
zTRUE:I(x^2) 0.009211   0.008838   1.042   0.3007

Residual standard error: 5.981 on 74 degrees of freedom
Multiple R-Squared: 0.1974, Adjusted R-squared: 0.1432
F-statistic: 3.64 on 5 and 74 DF, p-value: 0.005338
```

grafico 3.7: grafico di dispersione della quota di residenti in possesso almeno di diploma contro la differenza di voti tra centro-dx e centro-sx per ogni ballottaggio con il modello completo, in nero i valori osservati, in rosso le stime del modello.



Si può notare che si evidenzia una certa discontinuità nel momento in cui si passa da una differenza di voti positiva a una negativa presso lo zero.

Questa, almeno da questo modello non è dovuta al

parametro z anche se il rispettivo p -value del coefficiente stimato è 0.22, ancora comunque al di là di un livello utile per poter rifiutare l'ipotesi di nullità (penso a un 10%). In effetti con una specificazione migliore le cose cambiano, poiché con un modello che si adatta meglio ai dati è facile notare come la significatività dei parametri aumenti, compreso il coefficiente di z (al 10%), vedi figura 3.7 .

Tabella 3.9: Stime della regressione con la quota di residenti almeno diplomati residente in ogni comune come variabile dipendente., modello ridotto.

Call:

```
lm(formula = res_dipl ~ z + I(x^2) + z:I(x^2))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	39.592317	0.969764	40.827	< 2e-16	***
zTRUE	2.450335	1.505425	1.628	0.10773	
I(x^2)	-0.005313	0.001728	-3.075	0.00292	**
zTRUE:I(x^2)	0.005600	0.003210	1.745	0.08508	.

Residual standard error: 5.931 on 76 degrees of freedom

Multiple R-Squared: 0.1896, Adjusted R-squared: 0.1576

F-statistic: 5.925 on 3 and 76 DF, p-value: 0.001094

Grafico 3.8: grafico di dispersione della quota di residenti in possesso almeno di diploma contro la differenza di voti tra centro-dx e centro-sx per ogni ballottaggio con il modello ridotto, in nero i valori osservati, in verde le stime del modello.

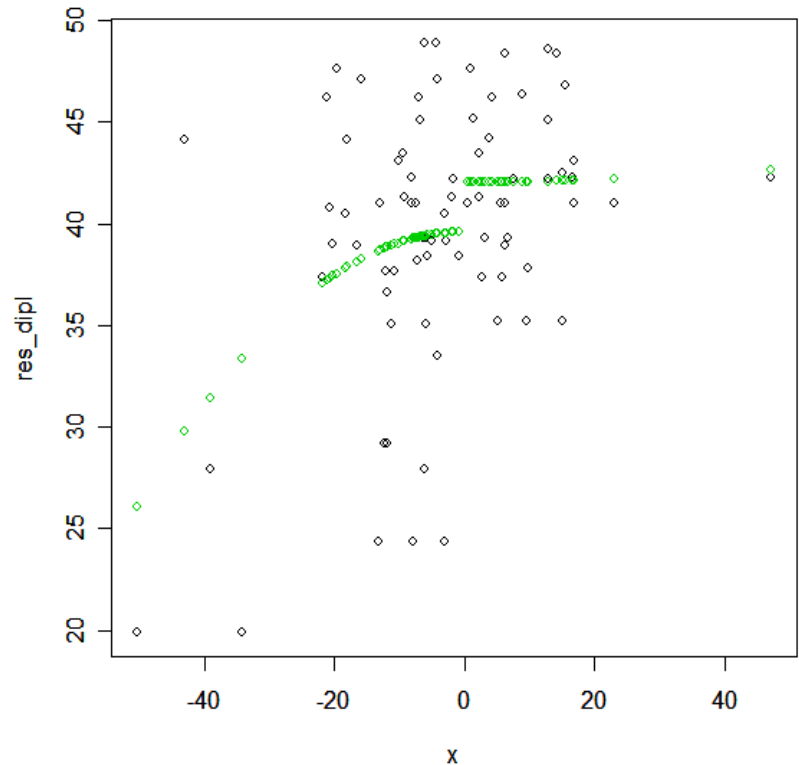
Non si nota un vero e proprio miglioramento

dell'adattamento del modello, infatti l' R^2 non subisce variazioni significative.

Però 3 parametri su quattro sono significativi: l'intercetta all'1%, x^2 al 5% e il termine

d'interazione al quadrato appena sotto il livello del 10%; ma quel che più conta è che il coefficiente di zeta arriva ad un livello che si può definire di 'border line' in cui non è scontato accettare l'ipotesi di nullità di tale parametro. Sembra quasi che dove abbia vinto al tempo t il ballottaggio di un soffio il centro destra la quota di residenti diplomati sia in media più alta (si passa da un circa 38% quando il centro-destra perde, a un 42% circa nel momento in cui al contrario vince), il che renderebbe i due gruppi non sistematicamente uguali.

Fatto importante, poiché questa caratteristica potrebbe inficiare i risultati ottenuti facendo uso del *Regression Discontinuity Design*, su cui si poggia questa analisi.



3.6.5 Differenze percentuali di voti al tempo t, rispetto al tasso di occupazione in ogni comune

I tassi di occupazione di ogni comune sono derivati dai dati del censimento come numero di occupati sulla popolazione attiva tra i 15 e 64 anni.

L'analisi che svolgo è sempre la solita, mantengo la specificazione con le stesse esplicative delle precedenti, ovvero la differenza di voti tra i due schieramenti al tempo t (variabile x che crea la selezione) e z (l'aver governato o meno), la variabile dipendente invece diventa proprio il tasso di occupazione di ogni comune.

La conclusione che traggio è che la discontinuità che si evidenzia dal grafico non è statisticamente significativa, in entrambe le specificazioni, del modello saturo e di quello ridotto; per cui possiamo affermare che anche rispetto a questa caratteristica i due gruppi, chi ha vinto e chi no, non sono diversi.

Tabella 3.10: Stime della regressione con il tasso di occupazione di ogni comune come variabile dipendente; modello saturo.

```
lm(formula = tasso_occupazione ~ x + z + x * z + I(x^2) +
I(x^2) * z)
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	42.8676859	2.2583264	18.982	<2e-16	***
x	0.1397996	0.2855003	0.490	0.626	
zTRUE	1.8472760	3.3028254	0.559	0.578	
I(x^2)	0.0004248	0.0060909	0.070	0.945	
x:zTRUE	-0.2624333	0.4342027	-0.604	0.547	
zTRUE:I(x^2)	-0.0028189	0.0095983	-0.294	0.770	

Tabella 3.10: Stime della regressione con il tasso di occupazione come variabile dipendente; modello ridotto..

```
Call:
lm(formula = tasso_occupazione ~ z + I(x^2) + I(x^2) * z)
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	41.891803	1.050766	39.868	<2e-16	***
zTRUE	2.053261	1.631171	1.259	0.212	
I(x^2)	-0.002410	0.001872	-1.287	0.202	
zTRUE:I(x^2)	-0.002533	0.003478	-0.729	0.469	

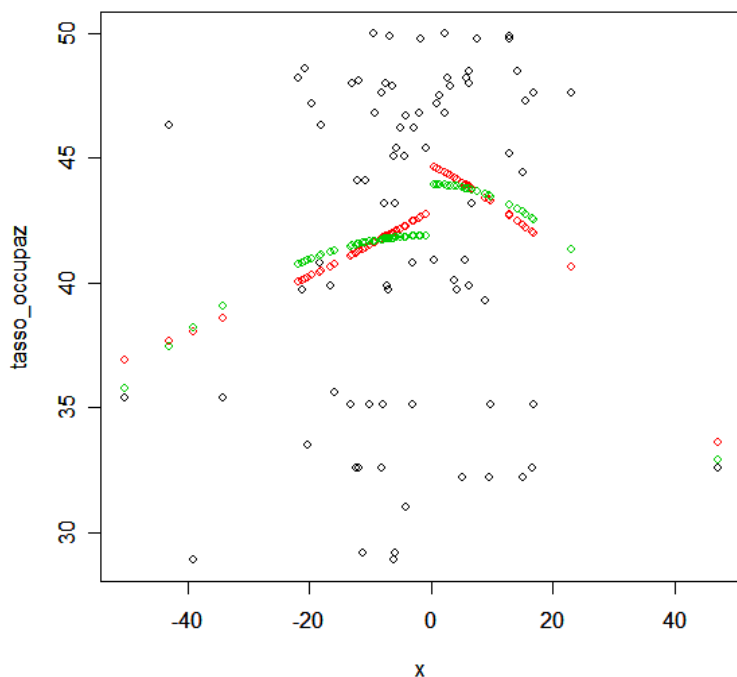


grafico 3.9: grafico di dispersione del tasso di occupazione contro la differenza di voti tra centro-dx e centro-sx di ogni comune per ogni ballottaggio, in nero i valori osservati, in rosso le stime del modello saturo, in verde quelle del modello ridotto..

3.6.6 Differenze percentuali di voti al tempo t, rispetto al tasso di disoccupazione in ogni comune

Il tasso di disoccupazione non è altro che il numero di disoccupati sul totale della forza lavoro.

Il discorso è analogo a quello sul tasso di occupazione.

Qui potrebbe sembrare che dove ha vinto il centro-sinistra ci sia un tasso di disoccupazione più elevato. Ciò non è certo legato al trattamento, governare o meno, che ha un valore del coefficiente di z tale per cui si accetta l'ipotesi di nullità.

Tabella 3.12: Stime della regressione con il tasso di disoccupazione come variabile dipendente; modello saturo..

```
lm(formula = tasso_disoccupazione ~ x + z + x * z + I(x^2) +
    I(x^2) * z)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	10.7337111	3.0135210	3.562	0.000648	***
x	-0.2017246	0.3809729	-0.529	0.598045	
zTRUE	-2.6604728	4.4073052	-0.604	0.547922	
I(x^2)	-0.0003799	0.0081277	-0.047	0.962846	
x:zTRUE	0.4796958	0.5794020	0.828	0.410381	
zTRUE:I(x^2)	0.0006581	0.0128080	0.051	0.959157	

Residual standard error: 8.668 on 74 degrees of freedom

Multiple R-Squared: 0.08041, Adjusted R-squared: 0.01827

F-statistic: 1.294 on 5 and 74 DF, p-value: 0.2757

Tabella 3.13: Stime della regressione con il tasso di disoccupazione di ogni comune come variabile dipendente; modello ridotto

Call:

```
lm(formula = tasso_disoccupazione ~ z + I(x^2) + I(x^2) * z)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	12.141867	1.405024	8.642	6.48e-13 ***
zTRUE	-2.323517	2.181108	-1.065	0.290
I(x^2)	0.003711	0.002503	1.482	0.142
zTRUE:I(x^2)	0.002346	0.004650	0.505	0.615

Residual standard error: 8.593 on 76 degrees of freedom

Multiple R-Squared: 0.07188, Adjusted R-squared: 0.03525

F-statistic: 1.962 on 3 and 76 DF, p-value: 0.1268

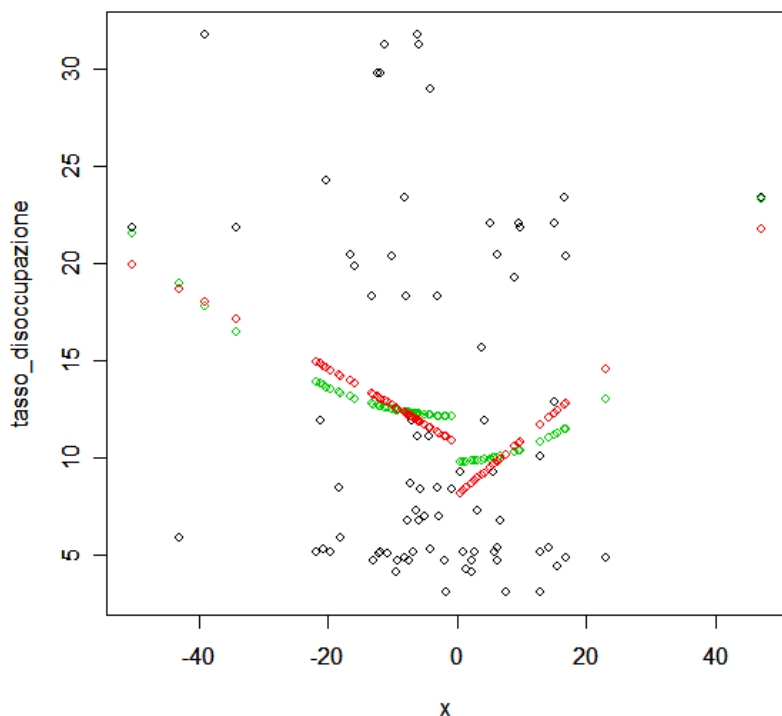


grafico 3.10: grafico di dispersione del tasso di occupazione contro la differenza di voti tra centro-dx e centro-sx di ogni comune per ogni ballottaggio, in nero i valori osservati, in rosso le stime del modello saturo, in verde quelle del modello ridotto.

Il grafico mostra sull'asse delle ascisse la variabile x causa di selezione, sulle ordinate invece vi è il tasso di disoccupazione per ogni livello differenza di voti al

ballottaggio dei vari comuni.

Si nota come una piccola discontinuità marginalmente al valore soglia è presente, ma tale differenza nel tasso non è riconducibile al trattamento, governare o meno, poiché il valore del coefficiente della variabile indicatrice nella regressione non è significativo.

3.6.7 Differenze percentuali di voti al tempo t, rispetto al numero medio di componenti familiari

La penultima caratteristica che esamino è il numero medio di componenti per ogni famiglia, per capire se i due gruppi di comuni al margine siano o o meno diversi quanto a numerosità media delle. Senza ombra di dubbio in figura 3.11, non si crea alcuna discontinuità sul punto di soglia.

Si può affermare con certezza che i gruppi presso il valore soglia non sono ex-ante differenti rispetto al numero medio di componenti per famiglia, ciò lo si vede con chiarezza già dalla regressione col modello saturo, per cui non riporto il modello ridotto che non migliorano l'analisi.

Tabella 3.14: Stime della regressione con il tasso di disoccupazione di ogni comune come variabile dipendente; output del modello completo.

```
Call:
lm(formula = media_comp ~ x + z + x * z + I(x^2) + I(x^2) * z)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.409e+00  1.171e-01  20.579  <2e-16 ***
x            -1.629e-02  1.480e-02  -1.101   0.275
zTRUE       -6.703e-04  1.712e-01  -0.004   0.997
I(x^2)      -1.094e-04  3.158e-04  -0.346   0.730
x:zTRUE      2.999e-02  2.251e-02   1.332   0.187
zTRUE:I(x^2) -7.364e-05  4.976e-04  -0.148   0.883
```

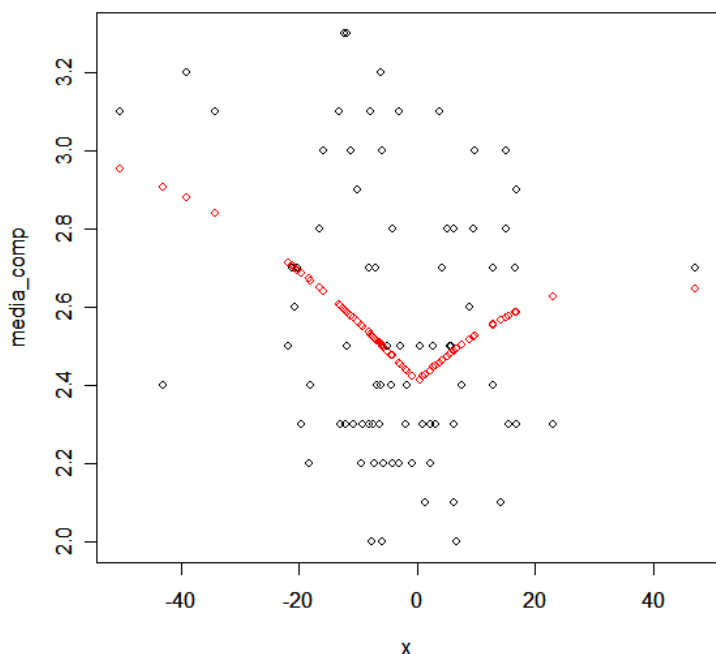


grafico 3.11: grafico di dispersione del numero medio di componenti familiari contro la differenza di voti tra centro-dx e centro-sx di ogni comune per ogni ballottaggio con il modello completo, in nero i valori osservati, in rosso le stime del modello.

3.6.8 Differenze percentuali di voti al tempo t, rispetto alla quota di maschi residenti in ogni comune

Per ovviare al problema dei numeri ‘assoluti’, si è preferito tenere in considerazione anziché il numero dei maschi residenti nelle varie città, il rapporto tra questi e la popolazione residente totale nei comuni.

Anche in questo caso il R^2 è un po’ basso, ma riporto sia il modello ‘saturo’, sia quello ridotto .

Tabella 3.15: Stime della regressione con la quota di maschi residenti rispetto al totale dei residenti di ogni comune come variabile dipendente; output del modello completo.

```
lm(formula = quota_maschi ~ x + z + x * z + I(x^2) + I(x^2) * z)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.751e-01  3.010e-03 157.818  <2e-16 ***
x            -1.601e-04  3.806e-04  -0.421   0.675
zTRUE       -2.234e-03  4.403e-03  -0.507   0.613
I(x^2)       2.418e-06  8.119e-06   0.298   0.767
x:zTRUE      4.048e-04  5.788e-04   0.699   0.486
zTRUE:I(x^2) -1.228e-06  1.280e-05  -0.096   0.924

Residual standard error: 0.008659 on 74 degrees of freedom
Multiple R-Squared: 0.1117, Adjusted R-squared: 0.05167
F-statistic: 1.861 on 5 and 74 DF, p-value: 0.1115
```

Tabella 3.16: Stime della regressione con la quota di maschi residenti rispetto al totale dei residenti di ogni comune come variabile dipendente; output del modello ridotto.

```
Call:
lm(formula = quota_maschi ~ z + I(x^2) + I(x^2) * z)

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.762e-01  1.402e-03 339.728  <2e-16 ***
zTRUE       -1.815e-03  2.176e-03  -0.834   0.4069
I(x^2)       5.664e-06  2.497e-06   2.268   0.0262 *
zTRUE:I(x^2)  6.150e-07  4.639e-06   0.133   0.8949

Residual standard error: 0.008573 on 76 degrees of freedom
Multiple R-Squared: 0.1058, Adjusted R-squared: 0.07048
F-statistic: 2.997 on 3 and 76 DF, p-value: 0.03588
```


La non significatività del coefficiente stimato di z ci indica che anche qui l'eventuale discontinuità sul grafico dei valori stimati non è da interpretare come effetto del trattamento, o meglio indica che la differenza nel rapporto tra maschi e residenti totali rispetto al vincere o meno le elezioni non è presente.

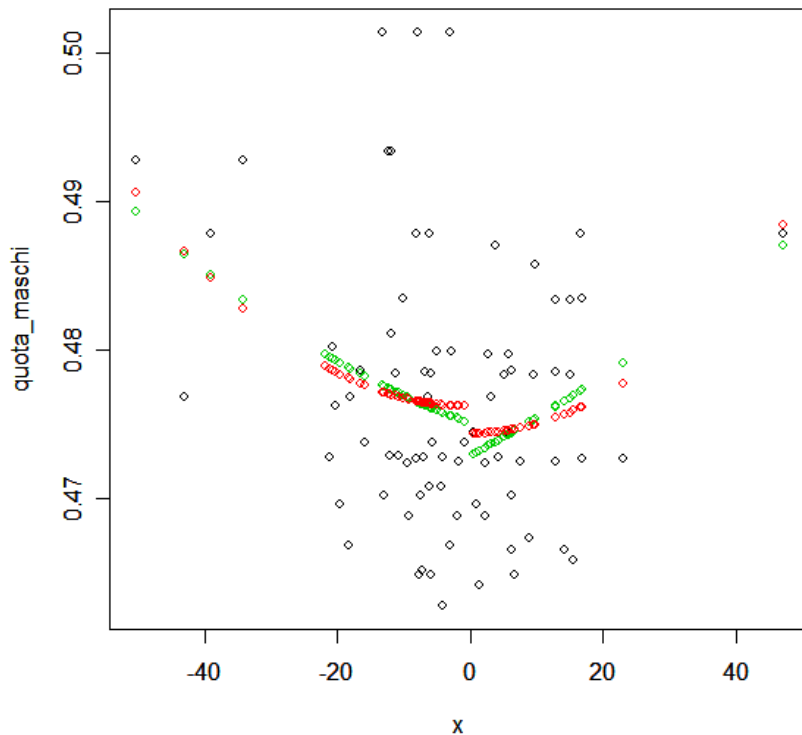


grafico 3.12: grafico di dispersione della quota di maschi rispetto al totale dei residenti contro la differenza di voti tra centro-dx e centro-sx di ogni comune per ogni ballottaggio, in nero i valori osservati, in verdi le stime del modello completo e in rosso le stime del modello ridotto.

A parte i 3 valori molto elevati sopra lo 0.50 per la quota di maschi (del comune di Barletta), per il resto del nostro campione, non si nota una elevata discontinuità al limite del valore soglia.

Si può concludere che nemmeno quest'ultima caratteristica esaminata renda diversi i due gruppi.

CAPITOLO 4

COMMENTI FINALI

Questo lavoro è, a mio avviso, interessante, di sicuro molto attuale visto il periodo storico in questa tesi è stata elaborata, prossimi al rinnovo di migliaia di amministrazioni comunali in tutta Italia. Purtroppo sarebbe stato ancor più interessante se l'analisi potesse comprendere anche questi dati, infatti il vero limite di quest'elaborato è proprio la limitata numerosità campionaria.

Lasciamo ai posteri interessati l'opera di aggiornamento.

Comunque qualche considerazione da mettere in risalto c'è, e la prima è forse il risultato più evidente, ovvero che dai dati fin qui raccolti, in Italia essere sindaco uscente non implica una maggior probabilità di vittoria all'elezione successiva per la coalizione che lo ha sostenuto. Questo un po' l'esperienza ce lo racconta poiché c'è una generalizzata delusione nei confronti di chi governa, e spesso l'alternanza politica è stata un'arma a doppio taglio per lo sviluppo ed il progresso delle nostre città, nel senso che da una parte portano una ventata di rinnovamento e di nuovi buoni propositi, dall'altra bloccano tutti quei progetti di crescita che hanno nella continuità di governo un perno fondamentale.

Come ho già accennato, non c'è abbastanza precisione nelle stime, questo soprattutto è causato da una numerosità campionaria molto limitata, infatti dal '93 ad oggi i casi di ballottaggio nei comuni con più di 90.000 abitanti sono circa un centinaio, ma per l'analisi se ne escludono più d'uno perché casi 'atipici' (ballottaggio destra vs destra, o viceversa, liste civiche slegate dai partiti tradizionali, ecc..). Inoltre per la prima analisi (test al tempo t e risultato al tempo $t+1$) andavano esclusi quei ballottaggi di cui non conosco il risultato all'elezione successiva, per la seconda (test al tempo t risultato al tempo $t-1$) andavano eliminati quelli la cui elezione precedente avviene prima del '93... insomma resta poco.

Ciò che è stato interessante capire era se i due gruppi a priori non avessero caratteristiche che li rendevano diversi al di là dall'aver superato o meno il 50% dei voti: generalmente la risposta è stata positiva, l'unica eccezione forse potrebbe essere la quota di diplomati rispetto al totale di residenti; rispetto a questa infatti una certa differenza tra i due gruppi precedente all'assegnazione del trattamento si nota...se così fosse ciò andrebbe tenuto in considerazione nell'analisi del RDD iniziale.

Infine un altro risultato di quest'analisi riguarda l'uso del *Regression Discontinuity Design* come strumento di valutazione di politiche, in quanto esso “risulta ancora poco diffuso e incompreso nell'ambito delle valutazioni” (Trochim, 2002). Invece nel nostro caso, che ha avuto come fondamento il lavoro esemplare di Lee (2008), è stato molto utile come metodo di valutazione, in quanto una volta capita quale poteva essere la caratteristica dalla quale i due gruppi sono separati, non ci siamo più dovuti interessare delle altre caratteristiche per arrivare a trarre delle conclusioni (se non per scrupolosa verifica).

Personalmente avrei preferito arrivare a delle conclusioni diverse, ovvero che il lavoro di chi governa la città, anche se con pochi voti di maggioranza, fosse riconosciuto e ricompensato con una rinnovata fiducia (come succede negli Stati Uniti d'America anche se Lee tende più a interpretare la maggiore probabilità in un ottica 'clientelare' o di favoritismi da parte di chi governa), ma forse in Italia non siamo ancora giunti ad una democrazia matura, forse la classe dirigente non è messa in grado di poter incidere davvero nella vita dei cittadini, oppure essa stessa non è in grado di farlo pur avendone i mezzi.

BIBLIOGRAFIA

- [1] Bloom H.S., 2002, *Using A Regression Discontinuity Design (RDD) to Measure Educational Effectiveness*, MDRC.
- [2] Faraway J.J., July 2002, *Practical Regression and Anova using R*, University of Bath.
- [3] Hahn H., Todd P. and Van der Klaauw W., 2001, *Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design*, *Econometrica*.
- [4] Istat, 2001, *14° Censimento della popolazione italiana - Popolazione e abitazioni*, Primi risultati.
- [5] Istat, 2001, *14° Censimento della popolazione italiana - Popolazione e abitazioni*, F. Provinciali.
- [6] Istat, 10-2001, *Conoscere il censimento - Il piano di rilevazione e il sistema di produzione*.
- [7] Imbens G.W., Lemieux T., 2008, *Regression Discontinuity Design: a guide to practice*, *Journal of Econometrics*.
- [8] Lee D.S., 2008, *Randomized experiments from non-random selection in U.S. House elections*, *Journal of Econometrics*.
- [9] Martini A., 2006, *Metodo sperimentale approccio controfattuale e valutazione degli effetti delle politiche pubbliche*, *Rassegna italiana di valutazione*.
- [10] Pace L., Salvan A., 2001, *Introduzione alla statistica II*, Padova, Cedam.
- [11] Pelagatti M. (2000), *L'approccio alla statistica robusta basato sulla funzione d'influenza: appunti per un seminario*, Working Paper QD 2001/1 – January.
- [12] Rettore E., Trivellato U., Martini A., *La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati?*, Working Paper n. 46 novembre 2002, Progetto di ricerca cofinanziato dal MIUR ‘Dinamiche e

persistenze nel mercato del lavoro italiano ed effetti di politiche (basi di dati, misura, analisi)

- [13] Rettore E., a.a. 2008-2009, *Dispense corso metodi statistici valutazione di politiche*, Scienze Statistiche, Università di Padova
- [14] Segreteria Organizzativa Federale, Maggio 2007, *Il sistema elettorale nei comuni*, Ufficio elettorale Milano.
- [15] Trivellato U., 2007, *Il monitoraggio delle politiche di sostegno monetario e la valutazione dei loro effetti*, Università di Padova.
- [16] Trochim W.M.K, 2002, *The research methods knowledge base*, Beverly Hills, Sage Publications.

SITI WEB dai quali ho tratto informazioni utili per lo svolgimento della tesi:

- http://it.wikipedia.org/wiki/Categoria:Elezioni_amministrative_in_Italia
- www.interno.it
- www.comune.roma.it
- http://it.wikipedia.org/wiki/Sistema_elettorale
- <http://www.repubblica.it/speciale/elezioni/index.html>

RINGRAZIAMENTI

A me stesso, ma soprattutto alla mia passione per la politica che mi ha spinto a continuare quest'opera con entusiasmo e voglia di capire, conoscere ed arricchirmi.

Alla mia famiglia, per avermi sempre fatto sentire il loro sostegno durante questi tre anni, pur non entrando mai nelle mie scelte, al contrario rispettandole senza imposizioni.

Ovviamente ringrazio gli amici, da quelli più stretti a quelli meno vicini, dai ragazzi del 'patro' ai compagni di facoltà, di mensa e di biblioteca (un particolare saluto va al mio 'Guru di R': il mitico E-Dinasty..), dagli amici del Partito ai colleghi della Fiera e a tutto lo staff di 'Pizza New' per tutte le serate in giro in motorino aspettando la pizza e la coca finale.

Un grazie va anche all'amico 'biancoscudato' Juan e alla sua tipografia, che con molta pazienza ha sopportato tutte le modifiche dell'ultima ora.

Un ringraziamento di cuore va all'Area II dell'Ufficio Consultazioni Elettorali Referendarie del Ministero dell'Interno, in particolare alla signora Maria Grazia (senza la quale probabilmente questo lavoro non sarebbe nemmeno iniziato) che si è spesa in maniera gratuita per farmi avere tutti i dati ordinati di città in città e completi di anno in anno.

Infine un sentito ringraziamento va al mio professore Relatore, che ha segnato con chiarezza e professionalità il percorso di questa tesi, accompagnandomi dall'inizio (l'idea parte da lui), fino all'ultima stesura.

Damiano Dision

