



**UNIVERSITÀ POLITECNICA DELLE MARCHE  
FACOLTÀ DI ECONOMIA “GIORGIO FUÀ”**

---

Corso di Laurea Magistrale o Specialistica in  
DATA SCIENCE PER L'ECONOMIA E LE IMPRESE

**ANALISI ECONOMETRICA DELLE ELEZIONI  
POLITICHE ITALIANE DEL 2018 E DEL 2022**

**ECONOMETRIC ANALYSIS OF ITALIAN  
POLITICAL ELECTIONS OF 2018 AND 2022**

Relatore: Chiar.mo  
Prof. Riccardo Lucchetti

Tesi di Laurea di:  
Daniele Torregiani

Anno Accademico 2022 – 2023

# Sommario

L'elaborato presenta un'analisi statistica multivariata relativa alle elezioni politiche italiane del 2018 e del 2022.

Lo scopo del lavoro è rilevare i principali fattori sociali, economici e demografici che hanno portato agli esiti elettorali osservati a livello comunale, inserendo un confronto tra le due elezioni.

Si inizia con una contestualizzazione descrittiva dei risultati elettorali tramite alcune visualizzazioni, un'analisi delle corrispondenze tra gruppi politici a livello regionale e un approfondimento delle spese sostenute dai partiti per inserzioni sulle piattaforme Meta.

Segue l'analisi inferenziale, parte centrale dell'elaborato, nella quale viene applicato il modello teorico ideato da Katz & King per dati composizionali aggregati.

In aggiunta, è stata realizzata un'analisi controfattuale per mezzo delle simulazioni statistiche.

# Indice

<b>Elenco delle figure</b>	v
<b>Elenco delle tabelle</b>	vi
<b>1 Introduzione</b>	1
<b>2 Le elezioni politiche in Italia nel 2018 e nel 2022</b>	3
2.1 Storia delle elezioni politiche del 2018 e del 2022 . . . . .	5
2.1.1 Le elezioni politiche del 2018 . . . . .	5
2.1.2 Le elezioni politiche del 2022 . . . . .	12
2.2 I dati elettorali . . . . .	20
2.3 Analisi descrittiva dei risultati elettorali . . . . .	27
2.3.1 Proporzioni di voto . . . . .	27
2.3.2 Analisi delle corrispondenze . . . . .	31
2.3.3 Inserzioni su Meta per la campagna elettorale del 2022 . . . .	35
<b>3 Modello statistico per le elezioni multipartito</b>	44
3.1 Definizione generale del modello KK . . . . .	46
3.1.1 Stimatore di massima verosimiglianza . . . . .	51
3.2 Le variabili indipendenti nel modello . . . . .	54
3.3 Risultati del modello applicato alle elezioni politiche italiane . . . .	60
3.3.1 Interpretazione dei coefficienti del modello . . . . .	62

3.3.2	Analisi delle correlazioni dei residui . . . . .	70
3.3.3	Analisi predittiva e controfattuale degli esiti elettorali . . . .	72
<b>4</b>	<b>Conclusioni</b>	<b>87</b>
<b>A</b>	<b>Statistiche descrittive</b>	<b>89</b>
A.1	Mappe dei risultati elettorali . . . . .	89
A.2	Analisi delle corrispondenze . . . . .	94
A.3	Inserzioni su Meta . . . . .	96
<b>B</b>	<b>Modello inferenziale</b>	<b>99</b>
B.1	Covariate . . . . .	99
B.2	Analisi controfattuale . . . . .	103
	<b>Riferimenti bibliografici</b>	<b>110</b>

# Elenco delle figure

2.1	Partiti con più voti per comune . . . . .	30
2.2	Analisi delle corrispondenze - 2018 . . . . .	33
2.3	Analisi delle corrispondenze - 2022 . . . . .	34
2.4	Grafico a dispersione per spesa sui social e voto ottenuto a livello regionale - Elezioni 2022 . . . . .	42
3.1	Grafico delle densità dei voti attesi simulati a Roma - Elezioni 2022.	76
3.2	Proporzioni di voto in media nazionale al variare degli over60 - Elezioni 2018 . . . . .	79
3.3	Andamento delle proporzioni di voto in media nazionale al variare degli over60 - Elezioni 2022 . . . . .	80
A.1	Proporzione dei voti per comune - Lega . . . . .	89
A.2	Proporzione dei voti per comune - Forza Italia (FI) . . . . .	90
A.3	Proporzione dei voti per comune - Fratelli d'Italia (FDI) . . . . .	90
A.4	Proporzione dei voti per comune - Partito Democratico (PD) . . . . .	91
A.5	Proporzione dei voti per comune - Movimento 5 Stelle (M5S) . . . . .	91
A.6	Proporzione dei voti per comune - Partiti locali . . . . .	92
A.7	Proporzione dei voti per comune - Altri partiti . . . . .	92
A.8	Proporzione dei voti per comune - Azione - Italia Viva (AZ-IV) . . . . .	93

# Elenco delle tabelle

2.1	Il sistema elettorale in vigore in Italia nel 2018 . . . . .	7
2.2	Risultati elezioni politiche 2018 . . . . .	12
2.3	Variazioni al sistema elettorale nel 2022 rispetto al 2018. . . . .	13
2.4	Risultati elezioni politiche 2022 . . . . .	19
2.5	Esempio struttura dataset elettorale estratto da Eligendo con dati 2022. . . . .	24
2.6	Proporzioni di voto nel 2018 . . . . .	28
2.7	Proporzioni di voto nel 2022 . . . . .	29
2.8	Quota di non votanti sul totale degli elettori (in % per zona) . . . .	29
3.1	Coefficienti stimati per le elezioni 2018. . . . .	63
3.2	Coefficienti stimati per le elezioni 2022. . . . .	64
3.3	Correlazioni dei residui per elezioni 2018 . . . . .	71
3.4	Correlazioni dei residui per elezioni 2022 . . . . .	71
3.5	Statistiche descrittive e voti reali a Roma - Elezioni 2022 . . . . .	77
3.6	Tasso di inclusione RdC/PdC provinciale 2021 e differenziali M5S 2022. . . . .	86
A.1	Analisi delle corrispondenze (righe) - Elezioni 2018 . . . . .	94
A.2	Analisi delle corrispondenze (colonne) - Elezioni 2018 . . . . .	94
A.3	Inerzia - Elezioni 2018 . . . . .	94

A.4	Analisi delle corrispondenze (righe) - Elezioni 2022 . . . . .	95
A.5	Analisi delle corrispondenze (colonne) - Elezioni 2022 . . . . .	95
A.6	Inerzia - Elezioni 2022 . . . . .	95
A.7	Distribuzione regionale delle spese per inserzioni su Meta - Elezioni 2022 . . . . .	96
A.8	Spese per inserzioni su Meta ogni mille elettori - Elezioni 2022 . . .	97
A.9	Distribuzione delle spese per inserzioni su Meta in percentuale rispetto al totale di ogni classe - Elezioni 2022 . . . . .	98
B.1	Riepilogo delle covariate usate nel modello . . . . .	99
B.2	Statistiche descrittive delle covariate - Elezioni 2018 e 2022 . . . . .	100
B.3	Matrice delle correlazioni delle covariate - Elezioni 2018 . . . . .	101
B.4	Matrice delle correlazioni delle covariate - Elezioni 2022 . . . . .	102
B.5	Voti reali ( $V$ ), attesi ( $\hat{E}(V)$ ) e scostamenti ( $\hat{E}(V) - V$ ) - Elezioni 2018	103
B.6	Voti reali ( $V$ ), attesi ( $\hat{E}(V)$ ) e scostamenti ( $\hat{E}(V) - V$ ) - Elezioni 2022	104
B.7	Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con -5% in "Over 60"	105
B.8	Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con +5% in "Stranieri"	106
B.9	Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con +5% in "Laureati"	107
B.10	Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con circa +5% in "Reddito medio da lavoro dipendente" . . . . .	108
B.11	Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con +5% in "Tasso di disoccupazione" . . . . .	109

# Capitolo 1

## Introduzione

Nel sistema politico italiano, le elezioni politiche rappresentano il vertice dell'espressione democratica.

Stiamo parlando di un fenomeno universale, che coinvolge in modo diretto la quasi totalità dei cittadini in quanto chiamati a votare per esprimere la propria preferenza. Ma indirettamente, tutti ne vengono influenzati, anche chi per legge ne rimane escluso. Il voto espresso alle elezioni politiche, infatti, determina la composizione del Parlamento, l'organo che detiene il potere legislativo e sul quale si regge l'intera struttura democratica del nostro Paese.

Il fenomeno elettorale si sostanzia nell'esistenza di più partiti politici in competizione tra loro. Alcuni accomunati da orientamenti ed ideali condivisi, altri in completo distacco. Si creano coalizioni, nascono e si estinguono le varie entità in continuazione. Da anni in Italia il panorama politico si presenta in modo frammentato ed in perpetuo mutamento.

Inoltre, un qualsiasi risultato politico deriva dalla commistione di innumerevoli fattori tra loro eterogenei: basti pensare alla crescente complessità insita nei sistemi sociali in cui viviamo. Possiamo infatti immaginare come gli esiti elettorali vengano influenzati, ad esempio, da fattori demografici come l'età degli elettori, oppure economici come il reddito da essi percepito.



Un fenomeno universale come quello elettorale, per essere analizzato, richiede quindi l'uso di approcci multidimensionali, capaci di tenere in considerazione sia la frammentarietà dei partiti politici che la complessità dei possibili fattori determinanti.

In questo elaborato, cercheremo di studiare gli esiti delle elezioni politiche italiane del 2018 e del 2022 sotto una lente multidimensionale.

Procedendo in modo graduale, il capitolo 2 è dedicato ad una prima contestualizzazione del fenomeno. Tratteremo inizialmente la storia che ha portato alle elezioni politiche del 2018 e del 2022, l'effettiva realtà politica e gli esiti elettorali che si sono verificati. Dal punto di vista statistico, verrà presentato un primo studio descrittivo integrato con mappe ed analisi delle corrispondenze. Si parlerà anche di social network e di spese sostenute dai vari partiti per le inserzioni sulle piattaforme Meta (Facebook ed Instagram).

Il capitolo 3 è il cuore di tutto l'elaborato. In esso si sviluppa l'analisi inferenziale, attraverso la quale cercheremo di individuare quali sono state le principali determinanti degli esiti elettorali del 2018 e del 2022. Per questo scopo, è stato adattato il modello ideato da Katz e King per le elezioni multipartito (1999) [10] ai due casi italiani.

## Capitolo 2

# Le elezioni politiche in Italia nel 2018 e nel 2022

Il presente capitolo contiene tre paragrafi.

Il paragrafo 2.1 introduce gli eventi che portarono alle elezioni politiche del 2018 e del 2022, assieme alla composizione degli schieramenti politici ed ai punti principali dei programmi elettorali.

Per chiarire maggiormente il contesto, è stata aggiunta una sintesi sul sistema elettorale vigente in Italia in quegli anni.

Il paragrafo 2.2 è dedicato ai dati elettorali. In essa sono indicate le fonti e le successive fasi di pulizia e di trasformazione dei dataset per adattarli alle esigenze di analisi.

I dati elettorali a livello comunale, opportunamente elaborati ed integrati con un insieme di variabili indipendenti (fattori demografici, socio-economici e geografici), saranno l'input del modello statistico definito nel capitolo successivo.

Infine, il paragrafo 2.3 raggruppa le principali statistiche descrittive dei dati elettorali di partenza.

In questo caso, si cerca di fornire una sintesi grafica ed intuitiva dei dati di partenza, prima di procedere con l'analisi inferenziale.

## 2.1 Storia delle elezioni politiche del 2018 e del 2022

In Italia le elezioni politiche definiscono la composizione delle due camere del Parlamento: la Camera dei Deputati ed il Senato della Repubblica.

La Costituzione prevede che le elezioni politiche debbano avvenire a suffragio universale diretto ogni cinque anni, oppure a seguito di uno scioglimento anticipato delle Camere stabilito dal Presidente della Repubblica prima della fine della legislatura.

A livello normativo, oltre agli articoli contenuti nel Titolo I della Parte Seconda della Costituzione, l'ordinamento italiano definisce anche il proprio sistema elettorale, ovvero la legge ordinaria specifica che definisce come devono essere tradotti i voti espressi dai cittadini italiani in seggi all'interno del Parlamento.

Alla data odierna, le votazioni più recenti per il rinnovo della Camera e del Senato risalgono al 2018 ed al 2022. Anche se entrambe sono state caratterizzate dallo stesso sistema elettorale, il *Rosatellum*, presentano chiaramente degli aspetti distintivi l'una dall'altra.

### 2.1.1 Le elezioni politiche del 2018

A seguito dello scioglimento delle Camere deliberato il 28 dicembre 2017 dal Presidente della Repubblica, il 4 marzo 2018 i cittadini italiani hanno votato per eleggere i 630 deputati e i 315 senatori della XVIII legislatura.

La situazione politica prima del voto si presentava in modo estremamente frammentato.

Dal 2013 al 2017 sono caduti tre governi (Letta, Renzi e Gentiloni) e si sono registrati forti cambiamenti nella composizione e nelle figure di *leadership* dei partiti

italiani.

In particolare, il centro-destra decise di presentarsi alle urne sotto un'unica coalizione composta da Forza Italia di Silvio Berlusconi, Lega di Matteo Salvini, Fratelli d'Italia di Giorgia Meloni e Noi con l'Italia - UdC.

Il centro-sinistra invece, tenne separata la coalizione maggiore formata dal Partito Democratico di Matteo Renzi, +Europa, Civica Popolare e Italia Europa Insieme, dalla lista Liberi & Uguali.

Il Movimento 5 Stelle, rappresentato al tempo da Luigi di Maio, decise di non coalizzarsi.

Oltre a quanto già menzionato, si presentarono altre diciotto liste rappresentanti partiti minori e locali.<sup>1</sup>

Come già accennato, il sistema elettorale in vigore alle politiche del 2018 era il cosiddetto *Rosatellum* (*legge n. 165/2017*): si tratta di un sistema elettorale misto in cui, dei 630 seggi previsti per la Camera,<sup>2</sup> 232 venivano assegnati con metodo maggioritario nei collegi uninominali, 386 con metodo proporzionale nei collegi plurinominali e i restanti 12 dalla circoscrizione Estero sempre con metodo proporzionale. Per i 315 seggi al Senato, 116 venivano assegnati con metodo maggioritario nei collegi uninominali, 193 con metodo proporzionale nei collegi plurinominali e i restanti 6 dalla circoscrizione Estero.

La tabella 2.1 sintetizza gli aspetti principali della legge elettorale per il 2018.

---

<sup>1</sup>Per partiti locali intendiamo quelli iscritti soltanto in alcune circoscrizioni elettorali.

<sup>2</sup> Nel 2018 la Camera era composta da 630 seggi mentre il Senato da 315 seggi elettivi. Il referendum costituzionale del 2020 decretò in seguito la riduzione del numero di parlamentari, passando a 400 deputati e 200 senatori eletti.

Tabella 2.1: Il sistema elettorale in vigore in Italia nel 2018

	Camera	Senato
Circoscrizioni	28	20
Collegi uninominale (seggi)	232 (232)	116 (116)
Collegi plurinominali (seggi)	63 (386)	33 (193)
Seggi circoscrizione estero	12	6
Totale seggi elettivi	630	315
Eletti con maggioritario (%)	36% (collegi uninominali)	
Eletti con proporzionale (%)	64% (coll. plur. + circ. estero)	
Listini bloccati	1 candidato nei collegi uninominali; da 2 a 4 candidati nei collegi plurinominali	
Pluricandidature	Max in 5 collegi uninominali	
Parità di genere	60% uomini e 40% donne, o viceversa	
Voto disgiunto	Non previsto	
Soglie di sbarramento	Liste singole: 3% - Coalizioni: 10%	
Premio di maggioranza	Non previsto	
Ballottaggio	Non previsto	

## Fatti precedenti al 2018 e programmi politici dei principali partiti

**Lega** La Lega Nord venne costituita a inizi anni Novanta come movimento di centro-destra a favore dell'autonomia del Nord.

Alle elezioni politiche del 2013 ottenne all'incirca il 4%. Da quel momento, ci furono importanti cambiamenti all'interno del partito.

La segreteria infatti passa da Roberto Maroni a Matteo Salvini (dicembre 2013), cercando di estendere il proprio consenso anche nel centro e sud Italia già durante

le regionali del 2015.

Con lo stesso obiettivo, la Lega si presenta alle politiche del 2018 con lista e simbolo rinnovati, "Lega - Salvini Premier", rimuovendo in essi i riferimenti espliciti al nord, presenti invece fin dalle sue origini.

Dal programma uniformato tra le varie liste della coalizione di centro-destra, le proposte della Lega nel 2018 si focalizzarono in particolare sui seguenti punti:

- Riforma del sistema tributario con un'unica aliquota fiscale ipotizzata al 15% per famiglie e imprese (*Flat tax*);
- Riforma previdenziale con sostituzione integrale della legge Fornero;
- Pace fiscale per tutti i piccoli contribuenti che si trovano in condizioni di difficoltà economica;
- Abolizione del limite all'uso del contante;
- Piano per il Sud, per uno sviluppo infrastrutturale e industriale;
- Meno vincoli dall'Europa e tutela del *Made in Italy*;
- Blocco degli sbarchi e rimpatrio in tema di immigrazione;
- Rafforzamento delle autonomie locali.

**Forza Italia** Forza Italia (FI) viene fondata nel 1994 da Silvio Berlusconi con orientamento di centro-destra ed ideologia liberale.

Nel 2009 viene sciolto a favore del Popolo della Libertà (Pdl), sotto la guida dello stesso Berlusconi. Il partito partecipa alle politiche del 2013 ottenendo circa il 22% dei voti nazionali.

Dopo pochi mesi il Pdl viene sciolto per confluire nella nuova Forza Italia, creata sempre da Berlusconi nel novembre 2013 utilizzando la stessa denominazione del partito originario.

FI, da sempre identificato nel suo leader, inizia a perdere consenso a partire dalla crisi del 2008-2009 e con la decadenza da senatore di Silvio Berlusconi dopo la condanna in via definitiva per frode fiscale nel 2013.

All'interno del programma della coalizione di centro-destra, FI propose:

- Riforma del sistema tributario con un'unica aliquota fiscale ipotizzata al 23% per famiglie e imprese (*Flat tax*) e futura riduzione dell'aliquota (a differenza del 15% proposto dalla Lega);
- Cancellazione dell'IRAP;
- Riforma previdenziale con revisione della legge Fornero (senza abolizione integrale);
- Aumento delle pensioni minime a 1.000 euro per 13 mensilità;
- Reddito di dignità di circa 1.000 euro per azzerare la povertà assoluta.

**Fratelli d'Italia** Partito fondato nel 2012 con orientamento di destra, Fratelli d'Italia (FdI) è sotto la guida di Giorgia Meloni dal 2014.

All'interno della coalizione di centro-destra, alle politiche del 2013 ottenne quasi il 2% dei voti nazionali.

Nuovamente schierato insieme al centro-destra, FdI aggiunse i seguenti punti al programma comune della coalizione:

- Famiglia come primo e fondamentale nucleo della società;
- Piano straordinario per la natalità con asili nido gratuiti e consistenti assegni familiari più che proporzionali al numero dei figli;
- Meno vincoli dall'Europa e tutela del *Made in Italy* (di comune accordo con la Lega);



- Investimenti in personale, mezzi e tecnologie per le forze dell'ordine e le forze armate;
- Contrasto all'immigrazione irregolare e no allo *Ius Soli* (cittadinanza italiana a tutti i nati sul territorio nazionale, indipendentemente dalla cittadinanza dei genitori).

**Partito Democratico** Il Partito Democratico (PD), orientato nel centro-sinistra, ha subito vari cambiamenti nel corso delle ultime due legislature.

Costituito nel 2007, il PD guidato da Pier Luigi Bersani ottenne la maggioranza alle elezioni politiche del 2013 con la coalizione di centro-sinistra (quasi 30% dei voti nazionali).

Successivamente, cambia la segreteria del partito così come il premier in carica, passando da Enrico Letta a Matteo Renzi nel febbraio 2014.

Lo stesso Renzi, dopo il "no" del referendum costituzionale del 2016 da lui promosso, si dimette da capo del governo e da segretario del PD. Nonostante ciò, viene rieletto segretario del partito nell'aprile 2017 e conduce il PD alle elezioni. Di seguito i punti principali del loro programma:

- 1 milione di nuovi posti di lavoro con la riduzione di 4 punti percentuali del cuneo contributivo per i contratti a tempo indeterminato;
- Salario minimo legale;
- Estensione bonus 80 euro anche a partite Iva e autonomi con redditi fino a 26 mila euro annui, e alle famiglie con figli fino a 18 anni;
- Riduzione IRES al 22%;
- Legge sullo *Ius Soli*.

**Movimento 5 Stelle** Fondato nel 2009 dall'ex comico Beppe Grillo e dall'imprenditore Gianroberto Casaleggio, il M5S non si collocava a destra e neanche a sinistra. Nato con l'idea di opporsi alle logiche dei partiti tradizionali, partecipò alle sue prime elezioni politiche nel 2013, registrando circa il 25% dei voti nazionali. Dietro l'indirizzo dato dal fondatore Beppe Grillo e con la nuova carica di capo politico assegnata a Luigi Di Maio, il Movimento 5 Stelle focalizza il suo programma per le politiche del 2018 su:

- Reddito di cittadinanza (RdC) come misura di politica attiva contro la povertà e la disoccupazione, e pensione di cittadinanza (PdC);
- Abbassamento delle tasse comprendendo la riduzione delle aliquote IRPEF, del cuneo fiscale e dell'IRAP per le PMI, più una *no tax area* per redditi fino a 10.000 euro;
- Taglio ai costi legati alla politica ed alle pensioni superiori ai 5.000 euro netti;
- Riduzione delle leggi e della burocrazia;
- Superamento della Legge Fornero con Quota 41 e Opzione donna;
- Rimpatri per gli immigrati irregolari e 10.000 nuove assunzioni nelle commissioni territoriali per l'immigrazione;
- Incremento della spesa pubblica per istruzione e sanità.

## Risultati elettorali del 2018

I risultati delle elezioni sono riassunti nella tabella 2.2, dove emerge un dato di affluenza in calo rispetto al 2013 (passando dal 75% al 73% circa per Camera e Senato), il primato dei voti al M5S (circa 32%) non sufficiente ad ottenere una

maggioranza assoluta in nessuna delle due Camere, seguito dal PD (19%) e dalla Lega (17%).<sup>3</sup>

Tabella 2.2: Risultati elezioni politiche 2018

	Camera	Senato
Affluenza	72,9%	73,0%
Lega	17,4%	17,6%
Forza Italia	14%	14,4%
Fratelli d'Italia	4,4%	4,3%
Movimento 5 Stelle	32,7%	32,2%
Partito Democratico	18,7%	19,1%
Locali	0,6%	0,5%
Altri	12,2%	11,9%

### 2.1.2 Le elezioni politiche del 2022

Anche la XVIII legislatura fu segnata da molteplici instabilità politiche.

Il debole accordo trovato tra M5S e Lega si concluse nel settembre 2019 con la caduta del governo Conte I.

Il seguente esecutivo Conte II, appoggiato da M5S, PD, Liberi e Uguali ed Italia Viva (partito fondato da Matteo Renzi nel 2019 e seguito da 26 deputati e 15 senatori), rassegnò le proprie dimissioni ad inizio 2021 per una nuova crisi di consenso. Mario Draghi, inizialmente sostenuto da un'ampia maggioranza parlamentare, compose un governo di unità nazionale formato da esponenti delle varie forze politiche. Tuttavia, anche questo esecutivo si concluse con una crisi nel luglio 2022.

---

<sup>3</sup>La categoria "Locali" raggruppa al suo interno tutte le liste presenti in meno di 40 collegi uninominali (rispetto ai 232 in tutta Italia nel 2018). "Altri", invece, raccoglie in sé i restanti partiti.

Il dettaglio delle due categorie verrà poi articolato nel paragrafo 2.2.

Dopo le dimissioni del governo Draghi, il presidente della Repubblica Sergio Mattarella scioglie anticipatamente le Camere per la mancanza di una maggioranza politica in grado di governare, e indice le elezioni politiche per settembre 2022.

Alle elezioni del 25 settembre, le liste Lega, Forza Italia, Fratelli d'Italia e Noi Moderati si unirono sotto la stessa coalizione di centro-destra; il Partito Democratico, Impegno Civico, Alleanza Verdi e +Europa definirono la coalizione di centro-sinistra. Il M5S, Azione-Italia Viva e le altre liste minori decisero invece di non unirsi ad alcuna coalizione.

Anche se il referendum costituzionale del 2020 portò alla riduzione dei seggi in Parlamento, il sistema elettorale rimase lo stesso (*Rosatellum*): per questo non ci furono modifiche alle modalità di assegnazione dei seggi (sistema misto con quote proporzionale e maggioritaria). Le uniche variazioni rispetto al 2018, schematizzate nella tabella 2.3, furono necessarie per adattare il numero di collegi elettorali al numero (ridotto) dei seggi in Parlamento.

Tabella 2.3: Variazioni al sistema elettorale nel 2022 rispetto al 2018.

	Camera	Senato
Collegi uninominali (seggi)	da 232 a 147	da 116 a 74
Collegi plurinominali	da 63 a 49	da 33 a 26
N. seggi coll. plurinominali	da 386 a 245	da 193 a 122
Seggi circoscrizione estero	da 12 a 8	da 6 a 4

## Fatti precedenti al 2022 e programmi politici dei principali partiti

**Lega** La Lega ricevette il proprio miglior risultato alle politiche del 2018.

Tuttavia, non fu abbastanza per avere una maggioranza al Parlamento e soltanto dopo tre mesi venne formato il nuovo governo guidato da Giuseppe Conte in alleanza tra Lega e M5S. Salvini ricevette la carica di vicepremier (assieme a Luigi di Maio) e di ministro dell'Interno. Fu lo stesso leader della Lega ad aprire una crisi di governo nell'agosto 2019. Dalla crisi nascerà nel settembre dello stesso anno un nuovo esecutivo, il Conte II, formato dall'alleanza M5S e PD.

La Lega rimase quindi all'opposizione fino alla formazione del governo Draghi. Alle successive elezioni politiche, la coalizione di centro-destra si presenta con un programma comune di quindici punti condivisi dalle quattro liste. In aggiunta ad esso, la Lega evidenziò nel proprio programma le seguenti proposte:

- *Flat Tax* per le partite Iva con fatturato fino a 100.000 euro, con estensione futura a tutte le persone fisiche e giuridiche, senza limiti di reddito;
- Nuova Pace Fiscale;
- Contrasto all'immigrazione irregolare con il ripristino dei "nuovi decreti Sicurezza";
- Autonomia differenziata, incrementando l'autonomia amministrativa dei territori;
- Modifica al Superbonus in ottica di un contenimento degli oneri a carico dello Stato;
- Riforma del RdC;
- Riforma delle pensioni con Quota 41 (pensione anticipata di anzianità con 41 anni di contributi) e Opzione Donna strutturale.

**Forza Italia** Dopo le elezioni politiche del 2018, FI resta sotto la guida di Berlusconi, ma viene affiancato dal nuovo vicepresidente del partito Antonio Tajani nel luglio 2018. FI, rimasto all'opposizione nei governi Conte I e II, decise di appoggiare l'esecutivo di Mario Draghi.

Oltre al programma unificato del centro-destra, FI propose:

- *Flat tax* in tre fasi, per arrivare ad una tassa unica al 23% per famiglie ed imprese (al pari della Lega);
- Nessuna imposta patrimoniale su casa, risparmi e sui capitali privati;
- Aumento delle pensioni minime a 1.000 euro al mese per 13 mensilità;
- Nuovo patto europeo per l'immigrazione e l'asilo.

**Fratelli d'Italia** FdI rimase all'opposizione nei tre esecutivi della legislatura (Conte I, Conte II e Draghi).

Tuttavia, si inizia a vedere un aumento dei consensi per il partito di Giorgia Meloni già dalle elezioni regionali in Abruzzo nel 2019. In quella occasione viene eletto presidente di regione un esponente di FdI per la prima volta. Successivamente, anche nelle Marche viene nominato un altro loro rappresentante come presidente di regione (settembre 2020).

Di seguito, i punti principali del programma integrativo di FdI:

- Sostegno alla natalità e alla famiglia con un sistema di tassazione rapportato al numero di componenti del nucleo familiare, aumenti degli importi dell'assegno unico ed asili nido gratuiti;
- Efficiente utilizzo di risorse del PNRR e dei fondi europei;
- Abolizione del RdC;
- *Flat tax* in tre fasi come il resto della coalizione e riduzione dell'IVA per i beni di prima necessità;

- Sostegno al sistema imprenditoriale italiano;
- Promozione e protezione del *Made in Italy*.

**Partito Democratico** Le elezioni politiche del 2018 segnarono il peggior risultato mai ottenuto dal PD a partire dalla sua fondazione (19% dei voti nazionali). Dopo le dimissioni di Renzi, la segreteria passò prima a Maurizio Martina e poi a Nicola Zingaretti.

Il PD, finito all'opposizione durante il Conte I, rientrò all'interno della maggioranza dei governi Conte II e Draghi.

Anche Zingaretti si dimise e venne eletto Enrico Letta nuovo segretario nel marzo 2021. Sarà Letta a portare il partito alle politiche del 2022.

Riepilogando il loro programma elettorale del 2022:

- Introduzione del salario minimo contrattuale, ricalibrazione del RdC, promozione dello *smart working* ed abolizione degli stage gratuiti;
- Effettiva applicazione della legge sulla parità salariale;
- Approvazione della legge contro l'omolesbobitansfobia (DDL Zan) e introduzione del matrimonio egualitario;
- Trasporti pubblici gratuiti in base all'ISEE;
- Approvazione dello *Ius scholae*, una legge che riconosca la cittadinanza italiana per chi è nato o entrato in Italia prima dei 18 anni, ha risieduto legalmente ed ha studiato nel nostro Paese;
- Piano strategico nazionale contro la violenza nei confronti delle donne e la violenza domestica;

- Dotazione di 10.000 euro, erogata al compimento dei 18 anni sulla base dell'ISEE familiare, per coprire le spese relative alla casa, all'istruzione e all'avvio di un'attività lavorativa. Misura coperta prevalentemente dalla modifica dell'imposta sulle successioni e donazioni superiori ai 5 milioni di euro.

**Movimento 5 Stelle** Il M5S uscì dalle politiche del 2018 come il partito con maggiori consensi in Italia (circa 32% dei voti nazionali).

Tuttavia, nel corso della XVIII legislatura il M5S fu segnato da numerose espulsioni e dimissioni dei propri membri.

A seguito delle due crisi di Governo, nell'agosto 2022 il partito decise di eleggere l'ex premier Giuseppe Conte come presidente del M5S.

Senza allearsi in alcuna coalizione, il partito presentò un programma incentrato su:

- *Cashback* fiscale per l'immediato accredito su conto corrente delle spese detraibili sostenute con strumenti elettronici;
- Salario minimo legale a 9 euro lordi l'ora;
- Rafforzamento del RdC e monitoraggio delle misure antifrode;
- Riscatto gratuito della laurea e pensione garanzia giovani per aiutare i giovani con carriere intermittenti;
- Riforma delle pensioni evitando il ritorno alla legge Fornero e proroga di Opzione Donna per l'uscita anticipata dal lavoro;
- Consolidamento della cessione dei crediti fiscali, confermando il Superbonus e gli altri bonus edilizi, estendendo la misura ad altre forme di investimento agevolato.

**Azione - Italia Viva** La lista formata dai due partiti Azione di Carlo Calenda e Italia Viva di Matteo Renzi è stata un'ulteriore novità rispetto alle precedenti



elezioni.

L'alleanza venne definita anche "Terzo Polo", con l'idea di presentarsi come un'alternativa di centro rispetto alle due coalizioni di centro-destra e di centro-sinistra. L'obiettivo dichiarato dal Terzo Polo in vista delle elezioni era quello di proseguire l'agenda e il "metodo Draghi", cercando di riportare lo stesso ex capo del governo al vertice dell'esecutivo.

Tra i vari aspetti trattati dal loro programma, i punti principali furono:

- Azzeramento delle tasse per i giovani che avviano un'attività imprenditoriale;
- Introduzione del salario minimo e modifica al RdC (annullamento del sostegno dopo il primo rifiuto di un'offerta di lavoro congrua e riduzione dopo due anni di benefici);
- Raggiungimento dell'indipendenza dal gas russo nel breve periodo, riduzione del 55% le emissioni di CO<sub>2</sub> entro il 2030 con fonti rinnovabili ed inclusione del nucleare nel mix energetico per l'obiettivo "emissioni zero" nel 2050;
- Ripristino e rafforzamento dell'Industria 4.0 attraverso l'aggiornamento della lista dei beni agevolati e l'aumento del tetto massimo per gli investimenti;
- Riforma dell'IRPEF, abolizione dell'IRAP, e passaggio ad un sistema a due aliquote IVA.

## Risultati elettorali del 2022

I risultati del 2022 segnarono la vittoria del centro-destra, che ottenne a livello di coalizione il 59,25% dei seggi alla Camera e il 57,5% di quelli elettivi al Senato. La coalizione di centro-sinistra ebbe rispettivamente il 21% e il 22% dei seggi alla Camera e ad Senato; il M5S ricevette il 13% ed il 14% tra le due Camere. I restanti vennero assegnati ad Azione-Italia Viva, SVP e ad altri partiti minori.

In termini di voti, gli esiti del 2022 diedero la vittoria della coalizione di centro-destra con il 44% dei voti nazionali e il primato di FdI con il 26%. Il PD ottenne il 19%, seguito dal M5S con il 15%. L'astensione salì notevolmente a quasi 36% (+9% rispetto al 2018). Di seguito i risultati schematizzati nella tabella 2.4.

Tabella 2.4: Risultati elezioni politiche 2022

	Camera	Senato
Affluenza	63,9%	63,9%
Lega	8,8%	8,9%
Forza Italia	8,1%	8,3%
Fratelli d'Italia	26%	26%
Movimento 5 Stelle	15,4%	15,6%
Partito Democratico	19,1%	19%
Azione - Italia Viva	7,8%	7,7%
Locali	1,4%	1,4%
Altri	13,4%	13,1%

## 2.2 I dati elettorali

I dati relativi alle elezioni politiche del 2018 e del 2022 sono disponibili in formato *open* su [Eligendo](#), una piattaforma online *ad hoc* per la diffusione dei risultati elettorali gestita dal Dipartimento per gli Affari Interni e Territoriali del Ministero dell'Interno.

In assenza di dati a livello individuale sulla popolazione, le analisi statistiche articolate in seguito (descrittiva nel paragrafo 2.3 ed inferenziale nel capitolo 3) si baseranno su dataset a livello comunali.

In questo modo cercheremo di indagare sulle caratteristiche della popolazione e sulle determinanti principali del voto sfruttando la maggiore granularità disponibile in formato *open*.

Nell'analisi svolta sono stati utilizzati soltanto i dati sugli esiti registrati per la Camera dei Deputati. Non sono stati considerati i voti per il Senato. I risultati delle liste tra Camera e Senato sono pressoché identici, con l'unica discriminante in tema di età dell'elettorato attivo (18 anni per votare la Camera, 25 per il Senato). Inoltre, utilizzando dati aggregati a livello di comune, il valore delle covariate è di fatto lo stesso. Da qui la scelta di concentrarsi esclusivamente sulla Camera, l'organo parlamentare eletto dal più ampio bacino di elettori.

Va detto che la Valle d'Aosta non è stata considerata nell'analisi. Le liste presentate in Valle d'Aosta alle politiche del 2018 e del 2022 sono in parte differenti da quelle a livello nazionale e per lo più a trazione locale. Inoltre, tale regione per legge<sup>4</sup> costituisce una circoscrizione a sé con un unico collegio elettorale: i circa centomila elettori valdostani, che rappresentano appena lo 0,2% dell'elettorato passivo in Italia, possono eleggere soltanto un deputato ed un senatore a maggioranza dei voti validi nell'unico collegio uninominale.

Quindi, i dati utilizzati considerano 7.884 comuni per il 2018 e 7.830 per il 2022

---

<sup>4</sup>Riferimento alla *legge costituzionale 26 febbraio 1948, n. 4*.

(totale dei comuni italiani al netto dei 74 paesi valdostani).

I due sottoparagrafi successivi chiariscono la logica dietro alla creazione delle classi politiche, oggetto centrale in tutto lo studio, ed il processo di preparazione dei dati alla struttura più adatta per svolgere l'analisi.

### **Definizione delle classi politiche**

In ognuna delle due elezioni in esame si sono presentate oltre venti liste politiche. Com'è intuibile, molte di esse hanno ricevuto in percentuale pochi voti a livello nazionale e talvolta zero seggi in Parlamento. Per evitare di avere dataset con un numero molto elevato di liste, è sembrato opportuno un loro raggruppamento in specifiche classi di interesse.

Nel concreto, ognuna delle liste che ha ricevuto più voti è stata assegnata ad una classe *ad hoc* (Forza Italia, Lega, Fratelli d'Italia, Partito Democratico e Movimento 5 Stelle nel 2018, alle quali è stata aggiunta Azione - Italia Viva nel 2022). Le liste politiche presenti solo in alcune aree o regioni d'Italia sono state raggruppate sotto la stessa classe "Partiti locali". Per essere precisi, quest'ultima include tutte le liste presentate in meno di 40 collegi uninominali nel 2018 (rispetto ai 232 in tutta Italia), o in meno di 30 nel 2022 (rispetto ai 147 totali). Le liste restanti, che hanno ricevuto la minoranza dei voti e un numero molto ridotto di seggi in Parlamento, sono state etichettate come "Altri partiti".

Schematizzando le classi politiche definite dal il raggruppamento delle liste elettorali:

### **Elezioni politiche 2018**

1. Forza Italia (FI)
2. Fratelli d'Italia (FdI)
3. Lega
4. Partito Democratico (PD)
5. Movimento 5 Stelle (M5S)
6. Partiti locali
  - SVP/PATT
  - Autodeterminazione
  - Blocco Nazionale per le Libertà
  - Grande Nord
  - Italia nel Cuore
  - Patto per l'Autonomia
  - Siamo
  - Rinascimento MIR
7. Altri partiti
  - Liberi e Uguali
  - +Europa
  - Civica Popolare
  - Italia Europa Insieme
  - Noi con l'Italia
  - CasaPound Italia
  - Potere al Popolo
  - Il Popolo della Famiglia
  - Italia agli Italiani
  - Partito Valore Umano
  - Partito Comunista
  - Partito Repubblicano Italiano
  - 10 Volte Meglio
  - Lista per la Costituzione
  - Per una Sinistra Rivoluzionaria

### **Elezioni politiche 2022**

1. Forza Italia (FI)
2. Fratelli d'Italia (FdI)
3. Lega
4. Partito Democratico (PD)
5. Movimento 5 Stelle (M5S)
6. Azione - Italia Viva (AZIV)
7. Partiti locali
  - SVP/PATT
  - Free
  - Forza del Popolo
  - Partito della Follia Creativa
  - Partito Comunista Italiano
  - Partito Animalista
  - Sud Chiama Nord
8. Altri partiti
  - Noi Moderati
  - Impegno Civico -  
Centro Democratico
  - Alleanza Verdi e Sinistra
  - +Europa
  - Alternativa per l'Italia -  
No Green Pass
  - Mastella Noi di Centro
  - Vita
  - Italexit per l'Italia
  - Unione Popolare
  - Italia Sovrana e Popolare

## **Fasi nel processo di preparazione dei dati**

I dataset elettorali per la Camera, così come estratti da Eligendo, si presentano con le seguenti colonne al loro interno:

- Circoscrizione
- Collegio plurinomiale
- Collegio uninominale
- Comune
- Elettori
- Elettori maschi
- Votanti
- Votanti maschi
- Schede bianche
- Voti lista
- Descrizione lista
- Cognome del candidato
- Nome del candidato
- Luogo di nascita del candidato
- Data di nascita del candidato
- Sesso del candidato
- Voti candidato

Tra gli attributi territoriali elencati, il più aggregato è riferito alla circoscrizione, poiché ogni circoscrizione è composta da 1 o più collegi plurinominali (relazione 1 a N). Si scende poi a livello di collegio plurinomiale, ognuno di essi contenente a sua volta 1 o più collegi uninominali (relazione 1 a N). Si arriva infine al livello di comune: nella maggior parte dei casi, ogni collegio uninominale comprende 1 o più comuni. Ad esempio, nel 2018 il collegio uninominale "04 - Ancona" raggruppava Jesi, Fabriano, Camerano, la stessa Ancona insieme ad altri venticinque comuni di dimensioni ridotte. Tuttavia, per le città italiane più grandi vale il contrario: Roma, Milano, Torino, Napoli, Palermo e Genova sono divise in più collegi uninominali con un numero di elettori mai superiore a duecentocinquantemila.

Quindi, ogni riga del dataset iniziale contiene i dati relativi ad un comune, in riferimento ad uno specifico collegio uninominale, definiti in un collegio plurinomiale

ed in una circoscrizione elettorale. Un esempio con dati del 2022 nella tabella 2.5.

Tabella 2.5: Esempio struttura dataset elettorale estratto da Eligendo con dati 2022.

Circoscrizione	Collegio Plurinominale	Collegio Uninomiale	Comune	...
...	...	...	...	...
MARCHE	MARCHE - P01	MARCHE - U03 (ANCONA)	ANCONA	...
MARCHE	MARCHE - P01	MARCHE - U03 (ANCONA)	JESI	...
...	...	...	...	...
LAZIO 1	LAZIO 1 - P01	LAZIO 1 - U01 (ROMA: MUNICIPIO I)	ROMA	...
LAZIO 1	LAZIO 1 - P01	LAZIO 1 - U02 (ROMA: MUNICIPIO III)	ROMA	...
...	...	...	...	...

Il rapporto che c'è tra lista e candidato è invece il seguente: ogni lista o coalizione può presentare un unico candidato per collegio uninominale (vincolo necessario dal momento che l'assegnazione dei seggi per gli uninominali è di tipo maggioritaria), e da 2 a 4 candidati per collegio plurinominale (nei quali l'assegnazione dei seggi è proporzionale su base nazionale per la Camera). Inoltre, l'elettore può decidere nella scheda elettorale se votare esclusivamente il singolo candidato, la singola lista/coalizione o entrambi: in caso di voto per la singola lista/coalizione, esso confluisce anche sui voti per il candidato di riferimento, ma non vale il contrario (poiché le liste che appoggiano il candidato sono più di una in caso di coalizione).

Di conseguenza, a livello comunale, i voti riferiti ai candidati uninominali sono sempre pari o di poco superiori ai voti riferiti alle liste.

Da ciò deriva la scelta di analizzare i voti ottenuti dalle singole liste anziché quelli dei singoli candidati, nei quali non si potrebbe distinguere il voto allo specifico partito in caso di coalizione.

La struttura dei dataset di partenza, così come sono estraibili in formato *open*,

non è perfettamente adatta allo svolgimento delle successive analisi di questo elaborato. Pertanto, sono state eseguite alcune trasformazioni alla struttura delle tabelle di partenza:

1. Aggregazione pivot per disporre di un singolo comune in ogni riga, e di ogni lista in una specifica colonna (i valori del campo "Descrizione lista" diventano le denominazioni delle nuove colonne, valorizzate dalla somma delle quantità in "Voti lista");
2. *Join* delle denominazioni dei comuni con i rispettivi Codici Istat, assenti nei dataset originali, in modo da avere un identificativo riconoscibile;
3. Raggruppamento delle liste "minori" e locali nelle apposite classi politiche, così come definite in precedenza (sottoparagrafo in 2.2), sommando i voti per comune di ogni lista associata alla classe.



Le colonne dei dataset per le elezioni 2018 e 2022, dopo le varie trasformazioni, diventano:

- Codice Istat
- Comune
- Collegio plurinomiale
- Circoscrizione
- Regione
- Elettori
- Elettori maschi
- Votanti
- Votanti maschi
- Schede bianche
- Lega
- Forza Italia
- Fratelli d'Italia
- Partito Democratico
- Movimento 5 Stelle
- Azione - Italia Viva (solo elezioni 2022)
- Partiti locali
- Altri partiti
- Totale voti (somma dei voti di tutte le classi, da Lega ad altri partiti)
- Non voto (differenza tra elettori e totale voti)

In sintesi: le tabelle finali dei dati elettorali della Camera dei Deputati sono due, una per il 2018 e l'altra per il 2022.

I dati al loro interno sono *cross-section*, in cui ogni osservazione si riferisce ad uno specifico comune italiano.

Escludendo la Valle d'Aosta, le tabelle hanno 7.884 righe per il 2018 e 7.830 per il 2022.

Le colonne numeriche valorizzano il numero di elettori e di votanti, le schede bianche ed i voti ottenuti delle singole classi (così come definite in precedenza nello specifico sottoparagrafo) in ogni comune italiano.

Il totale dei voti e dei "non voti" sono ottenibili con semplici operazioni algebriche dai dati già disponibili.

## **2.3 Analisi descrittiva dei risultati elettorali**

Una sintesi numerica e grafica dei dati originali è indispensabile per comprendere al meglio il fenomeno oggetto di studio.

Essa dà un apporto sostanziale anche alla comprensione ed all'interpretazione dei successivi risultati inferenziali.

La sintesi di cui stiamo parlando, che precede lo studio inferenziale, è l'analisi descrittiva.

In questo caso, un riassunto preliminare permette di avere una visione semplice ed immediata degli esiti elettorali del 2018 e del 2022, fornendo osservazioni interessanti a livello territoriale.

L'analisi descrittiva, approfondita nei tre sottoparagrafi successivi, comprende:

1. Statistiche sulle proporzioni di voto e visualizzazione grafica mediante mappe in 2.3.1;
2. Analisi delle corrispondenze (CA) tra coalizioni e circoscrizioni elettorali in 2.3.2;
3. Analisi regionale delle inserzioni sulle piattaforme Meta durante la campagna elettorale del 2022 in 2.3.3.

### **2.3.1 Proporzioni di voto**

Il paragrafo 2.1 contiene i numeri principali a livello nazionale dei risultati e dei seggi assegnati dopo le elezioni del 2018 e del 2022.

Qui di seguito, invece, sono riportati i valori con le percentuali di voto ottenute nel 2018 e nel 2022 dalle classi prese in esame rispetto alle cinque zone geografiche italiane definite dall'Istat. Le tabelle 2.6 e 2.7 derivano dalle elaborazioni dei dati

*open* per la Camera, sulla base delle considerazioni fatte nel precedente capitolo (voti per lista al netto della Valle d'Aosta).

Le due tabelle sono utili per confrontare, a distanza di 4 anni, come sono cambiate le proporzioni di voto all'interno di ogni macro-area.

Oltre alle tendenze più evidenti (aumento in tutte le zone d'Italia dei voti a FdI, calo di Lega e M5S) si può notare come, sul totale dei voti al nord, il M5S abbia perso una quota rilevante di elettori (dal 24,71% al 7,42% in nord-est, dal 23,31% all'8,69% al nord-ovest). Tale quota è stata assorbita in parte dal centro-destra (in particolare da FdI) e in parte dal Terzo Polo (Azione - Italia Viva, non presente nel 2018).

È interessante osservare anche com'è cambiata nel tempo la quota di voti "locali" nelle isole: nel 2022, infatti, hanno contribuito al 7,67% del totale. Questa cifra raccoglie in sé gli oltre 200.000 voti di "Sud chiama nord" in Sicilia, lista non presente nel 2018 che è riuscita ad eleggere due suoi membri all'interno del Parlamento.

Tabella 2.6: Proporzioni di voto nel 2018

(in % rispetto al totale dei voti validi per la zona)

Zona	Lega	FI	FdI	PD	M5S	Locali	Altri
Centro	15,93	11,73	6,27	23,00	30,02	0,02	13,04
Isole	6,72	19,46	3,81	12,10	47,04	0,57	10,30
Nord-est	26,06	10,25	3,94	20,43	24,71	2,29	12,33
Nord-ovest	26,11	13,81	4,07	20,77	23,31	0,11	11,83
Sud	6,31	18,28	3,92	13,47	45,99	0,01	12,01
Italia	17,66	14,24	4,43	18,64	32,42	0,56	12,06

Tabella 2.7: Proporzioni di voto nel 2022

(in % rispetto al totale dei voti validi per la zona)

Zona	Lega	FI	FDI	PD	M5S	AZIV	Locali	Altri
Centro	6,77	6,49	29,49	21,83	13,28	8,34	0,27	13,53
Isole	5,44	10,66	20,43	13,40	26,60	4,86	7,67	10,95
Nord-est	11,03	6,25	28,65	21,13	7,42	8,16	2,13	15,23
Nord-ovest	12,30	7,79	27,91	19,55	8,69	9,46	0,01	14,30
Sud	5,58	11,27	20,93	15,66	29,66	5,21	0,37	11,32
Italia	8,84	8,20	26,22	18,95	15,44	7,63	1,34	13,39

In ogni caso, indipendentemente dalle motivazioni, nelle riflessioni sui risultati vanno sempre considerati i non votanti sul totale degli elettori. Il termine "non votanti" comprende infatti sia gli astenuti che i voti inespresi. Un "astenuto" è un potenziale elettore che decide di non recarsi alle urne per votare. Il totale degli astenuti viene calcolato come differenza tra gli elettori ed i votanti.

Nel computo totale dei votanti, tuttavia, rientrano anche i cosiddetti "voti inespresi", ovvero tutte le schede bianche e le schede nulle di una votazione, legate ad un errore dell'elettore, o ad un più frequente voto di protesta.

Nonostante le quote già molto alte di non votanti nel 2018, dalla tabella 2.8 è evidente un ulteriore aumento generalizzato per tutte le aree d'Italia, con picchi di

Tabella 2.8: Quota di non votanti sul totale degli elettori (in % per zona)

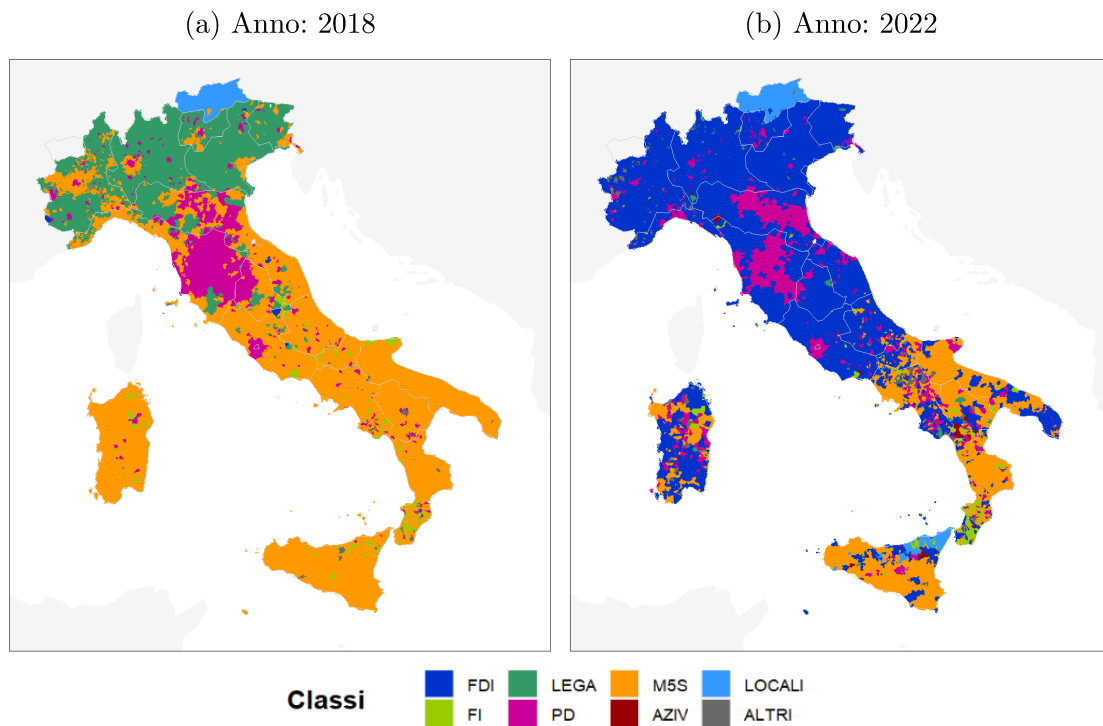
Zona	2018	2022
Centro	29,81	38,03
Isole	40,93	49,51
Nord-Est	27,77	35,42
Nord-Ovest	29,58	36,83
Sud	36,44	49,18
Italia	32,23	41,18

quasi 50% in Sicilia, Sardegna e nel resto del sud.

Ad ogni modo, per avere una rappresentazione più immediata degli esiti elettorali può essere utile una visualizzazione grafica.

Appena sotto vengono riportate le mappe colorate in base alla classe che ha ricevuto più voti a livello di comune.

Figura 2.1: Partiti con più voti per comune



Le due visualizzazioni nella figura 2.1 mostrano quali sono state le classi che hanno ricevuto più voti in ogni comune: oltre ad apparire una forte relazione di prossimità (com'è attendibile, località vicine fra loro che votano per lo stesso partito), è molto marcato il passaggio degli elettori del M5S e della Lega a FdI non soltanto al nord ma anche al sud della penisola. Salta all'occhio anche il calo del PD in Umbria e Toscana, così come il maggiore interesse verso i partiti locali in provincia di Messina (lista "Sud chiama nord", dov'era appunto in carica il leader

del partito come sindaco del capoluogo di provincia tra il 2018 ed il 2022).

L'Appendice A.1 raccoglie altre mappe con il dettaglio per ogni classe politica. La saturazione dei colori nelle mappe è legata alla percentuale di voti ottenuti dalla classe nel singolo comune. Intuitivamente, l'intervallo di colori applicati (dal più al meno intenso) varia in base allo stesso range di valori (da 0% a 100%).

Da esse, le dinamiche nel tempo sottostanti alla figura 2.1 appaiono in modo ancora più evidente.

### **2.3.2 Analisi delle corrispondenze**

L'analisi delle corrispondenze (CA) è una tecnica molto elegante per visualizzare una sintesi dei dati all'interno di una tabella di contingenza ed estrarre conoscenza utile.

Tra gli statistici che hanno contribuito a definirla e svilupparla rientra sicuramente M.J. Greenacre<sup>5</sup>: tecnicamente, egli presenta la CA come un metodo statistico multivariato di riduzione della dimensionalità basato sulla decomposizione ai valori singolari (SVD).

Per poter eseguire la CA è necessario che i dati siano della stessa scala (ad esempio come quelli contenuti all'interno delle tabelle di contingenza, che riportano i valori delle frequenze assolute o relative tra i caratteri di due variabili categoriche).

Questa tecnica conduce ad un risultato grafico, definito come mappa simmetrica o "nuvola" di punti nello spazio multidimensionale, che mostra le associazioni tra le variabili qualitative in esame.

Ogni punto della mappa rappresenta un profilo riga o colonna (pari alla probabilità condizionata della riga o colonna), mentre l'origine corrisponde al baricentro, ovvero al profilo medio.

---

<sup>5</sup>Molti i suoi contributi sul tema. Ad esempio, si veda Greenacre (2010) [4].

All'interno dello spazio multidimensionale è possibile calcolare le distanze tra due profili riga/colonna o tra un profilo riga/colonna e il baricentro<sup>6</sup>.

Moltiplicando le rispettive frequenze marginali alle distanze tra i punti e il baricentro si ottiene l'inerzia totale, invariante tra righe e colonne: l'inerzia è l'indicatore principale di tutta l'analisi, poiché misura la deviazione dei singoli profili rispetto al profilo medio, ovvero la variabilità rispetto alla condizione di uguaglianza dei profili (indipendenza delle variabili).<sup>7</sup>

Per quanto riguarda l'interpretazione del grafico derivante dall'analisi delle corrispondenze, se due punti relativi allo stesso profilo (riga o colonna) sono vicini tra loro, allora tali righe/colonne avranno quasi le stesse proporzioni. Invece, se ad essere vicini sono due punti di un profilo diverso (uno riga e l'altro colonna), allora la frequenza congiunta della specifica riga con la colonna contigua sarà più elevata rispetto alle altre colonne, e viceversa. Inoltre, se i punti si collocano vicino all'origine degli assi, allora quei profili non avranno delle proporzioni molto distanti da quelle del baricentro. E di conseguenza, minore sarà la distanza dei punti rispetto all'origine, minore sarà la dipendenza tra le due variabili in esame.

Dopo questa breve introduzione della tecnica, i due successivi paragrafi sono dedicati all'applicazione della CA ai casi delle elezioni del 2018 e del 2022. Con lo scopo di favorire ulteriormente la capacità di sintesi della CA, ridurre le etichette nella mappa simmetrica e mostrare in modo più chiaro le associazioni tra voto ai partiti e territorio, solo in questo caso è stato scelto di unire le liste FI, FDI e Lega all'interno di un'unica classe, il centro-destra (CDX).

---

<sup>6</sup>Nell'analisi delle corrispondenze si utilizza il quadrato della distanza  $\chi^2$ . La distanza  $\chi^2$  è pari a quella euclidea pesata per i rispettivi profili medi.

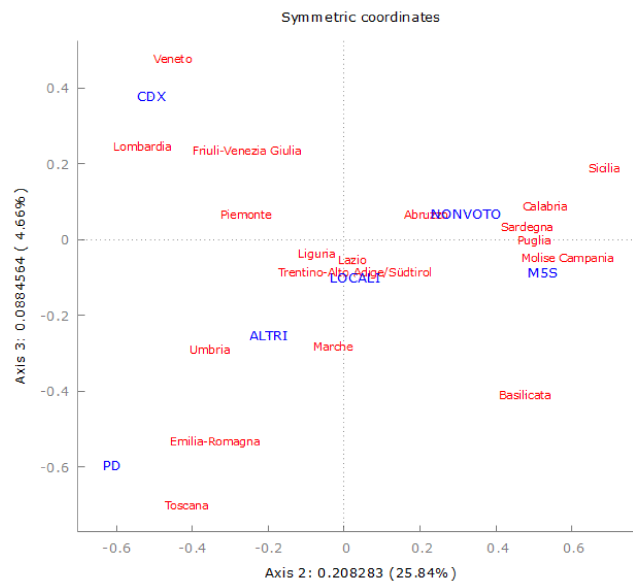
<sup>7</sup>Non a caso l'inerzia totale è anche pari al rapporto tra la statistiche  $\chi^2$  e il numero totale di osservazioni.

## Elezioni 2018

L'analisi delle corrispondenze applicata alla tabella a doppia entrata con i dati delle elezioni politiche del 2018 aggregati a livello regionale (prima variabile categorica: Regione) rispetto alle classi politiche (seconda variabile categorica: Classi politiche) riesce a sintetizzare il 98,66% del contenuto informativo in appena tre nuove dimensioni (dettaglio in Appendice A.3).

La variabilità contenuta nel primo asse viene catturata quasi esclusivamente dai caratteri "Trentino-Alto Adige/Südtirol" (65,91% di inerzia delle colonne) e "Partiti locali" (67,50% di inerzia delle righe). Il tutto motivato dalla differenza molto evidente rispetto alle altre regioni in termini di proporzioni di voto: Südtiroler Volkspartei (SVP), partito locale rappresentativo dei gruppi linguistici tedesco e ladino dell'Alto Adige, ottiene rispettivamente il 24% ed il 23% nelle ultime due elezioni nella circoscrizione trentina.

Figura 2.2: Analisi delle corrispondenze - 2018



Ponendo l'attenzione sul secondo e terzo asse, depurati dall'effetto SVP (Figura 2.2), si possono osservare aspetti sottili del legame tra territori e classi politiche. Innanzitutto, viene confermata la maggiore inclinazione di alcune regioni verso orientamenti politici specifici (come già visto nelle mappe, Veneto, Lombardia e Friuli orientate a centro-destra, Toscana ed Emilia-Romagna a centro-sinistra).



Poi, c'è un altro particolare molto rilevante che appare dalla disposizione delle regioni del sud e delle isole con M5S e non voto. La vicinanza di questi caratteri nella mappa simmetrica risalta il fatto che l'alternativa al M5S (lista con più voti in quelle zone) non sembra essere il voto ad un altro partito, ma è piuttosto l'astensione o il voto di protesta (36.44% al sud e 40.93% nelle isole).

Nel complesso, gli assi sembrano ben riprodurre le alternative di voto:

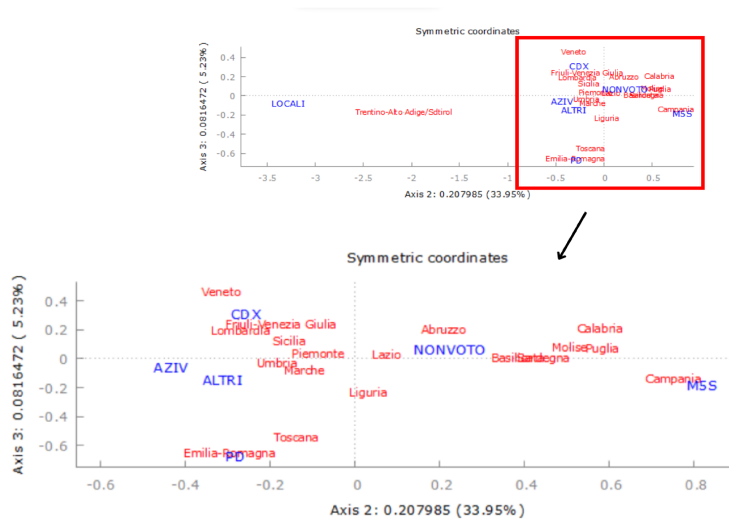
- l'asse "3" divide nettamente gli orientamenti di destra e di sinistra, mostrando poi che gli altri partiti minori si avvicinano come alternativa alla sinistra;
- l'asse "2" divide invece il M5S ed il non voto dal resto dei partiti presenti da più tempo nel panorama politico italiano. Un sintomo della scarsa tenuta dei partiti tradizionalmente a destra ed a sinistra nelle regioni del sud Italia.

## Elezioni 2022

Il contenuto del primo asse non cambia nella sostanza rispetto al 2018: i casi d'eccezione dei partiti locali (oltre a SVP in Alto Adige, nel 2022 subentra anche Sud chiama nord in Sicilia) vengono catturati quasi interamente dalla prima dimensione, per questo non riportata graficamente.

Anche per il 2022, la CA riesce a sintetizzare quasi il 98% dell'intero contenuto in soli tre assi (inerzia cumulata asse 3 = 97,89%).

Figura 2.3: Analisi delle corrispondenze - 2022



La figura 2.3 mostra le dimensioni 2 e 3: parte dello scostamento dei partiti locali e del Trentino rispetto al baricentro rimane anche nel secondo asse. Tuttavia, le lente di dettaglio sulle altre categorie nella mappa mostrano delle differenze sostanziali rispetto alla precedente elezione.

Innanzitutto, diverse regioni si sono avvicinate al centro-destra (Sicilia, Umbria e Marche). Poi, ad eccezione della Campania, le altre regioni del sud si sono allontanate dal M5S per avvicinarsi alla categoria "non voto". E tutto ciò è coerente con l'altissima incidenza di astenuti, schede nulle e non valide della zona (quasi il 50% sul totale degli elettori al sud e nelle isole).

È possibile notare nella mappa la vicinanza del Terzo Polo alla categoria Altri in termini di proporzioni di voto, senza essere associato ad una particolare regione.

Infine, il caso delle Marche mostra un cambiamento significativo rispetto ai risultati di soli quattro anni precedenti. Si può infatti osservare l'avvicinamento ad altre regioni geograficamente più lontane (come il Piemonte), distaccandosi al tempo stesso da Toscana ed Emilia-Romagna, storicamente più affini nell'orientamento al voto. Un risultato che in ogni caso conferma l'avvicinamento degli elettori marchigiani al centro-destra, già avviato con le elezioni regionali del 2020.

### **2.3.3 Inserzioni su Meta per la campagna elettorale del 2022**

Prima di passare all'analisi inferenziale del capitolo 3, chiudiamo lo studio descrittivo dei dati elettorali di partenza trattando la presenza dei partiti politici nei social network, e come essa possa collegarsi ai punti principali dei programmi politici (riassunti precedentemente nel paragrafo 2.1).

Integrare i dati sulle spese sostenute dai partiti politici nelle piattaforme social all'interno di un modello inferenziale per studiare il loro impatto sui risultati elettorali potrebbe fornire indicazioni davvero interessanti in uno studio multidimensionale

del fenomeno elettorale. Purtroppo, per motivi di indisponibilità dei dati e di comparabilità dei risultati, non è stato possibile inserire questa variabile all'interno del modello inferenziale. Tuttavia, per dare una maggiore completezza all'analisi sviluppata nel capitolo successivo, parliamo ora delle inserzioni prodotte dai partiti sulle piattaforme Meta (Facebook e Instagram) per la campagna elettorale del 2022.

Come sappiamo, in vista delle elezioni, ogni partito è tenuto a definire un programma elettorale in linea con il proprio orientamento. Per comunicare i propri obiettivi e rafforzare l'immagine verso l'elettorato, i partiti sostengono ingenti spese elettorali. In esse rientrano, ad esempio, i costi volti ad acquisire spazi su radio e televisioni private, organizzare eventi di propaganda, produrre affissioni e manifesti. Nell'ultimo decennio, tuttavia, è subentrata una nuova dimensione, sempre più impattante nelle nostre vite<sup>8</sup> e fortemente legata alla comunicazione di informazioni e di idee. Si tratta della dimensione dei social network.

Conoscere le spese reali sostenute dai partiti politici in tutte le piattaforme social per pubblicizzare le proprie campagne elettorali, così come il numero e la distribuzione geografica degli utenti che hanno visualizzato i loro contenuti definirebbe un patrimonio informativo preziosissimo per l'analisi dei risultati elettorali. Tuttavia, il limite principale resta la raccolta dei dati, sia per motivi legati alla disponibilità delle informazioni, sia per la complessità nello svolgimento di questo compito.

I dati che invece sono disponibili e che possono contribuire come valida integrazione all'elaborato ce li fornisce Meta attraverso la sua *Ad Library*. Attualmente, la Libreria Inserzioni di Meta permette di ricercare tutti i contenuti a pagamento pubblicati da qualsiasi utente o pagina su Facebook e Instagram a partire da aprile

---

<sup>8</sup>Il numero di utenti dei social è in forte crescita. I report internazionali più noti sul tema, come [Digital 2022 Global Overview Report](#) e [Demand Sage - Social Media Users](#), stimavano circa 4,6 miliardi di utenti in tutto il mondo nel 2022 (58% della popolazione globale), di cui 42 milioni in Italia (71,6% della popolazione italiana).

2019. Per ogni inserzione è disponibile la distribuzione regionale delle visualizzazioni, il costo sostenuto, la data e la durata della pubblicazione, indicando anche chi sono il beneficiario ed il pagante (che non coincidono necessariamente).

Come accennato poche righe fa, bisogna considerare un aspetto non banale nella raccolta di questi dati: la trasparenza nelle inserzioni pubblicate nelle piattaforme Meta da parte delle pagine ufficiali di ogni lista o leader politico non garantisce la completezza dei contenuti a sostegno di un certo partito. Esistono infatti numerose pagine non ufficiali gestite in modo anonimo che producono continuamente contenuti a pagamento a favore di alcuni partiti. Inoltre, anche se Facebook ed Instagram sono ampiamente usati dai partiti come strumento per fare campagna elettorale, non sono gli unici social impiegati con questo scopo. Basti pensare ai video caricati su YouTube e TikTok, o ai *tweet* in Twitter.

Raccogliere anche queste informazioni con i mezzi disponibili diventa particolarmente proibitivo; per questo è stato scelto di estrarre esclusivamente i dati sulle inserzioni promosse dalle pagine ufficiali dei partiti politici e dei rispettivi leader. L'intento è di misurare le spese sostenute direttamente dalle liste a livello regionale (non sono disponibili dati meno aggregati) nella campagna elettorale del 2022. Avendo a disposizione soltanto dati regionali e per la sola elezione del 2022, non è stata inserita alcuna covariata relativa ai social all'interno del modello inferenziale. Per questo, almeno in questa sede, non è stato possibile stimare i possibili impatti delle campagne social sui voti ottenuti dai partiti. Nonostante ciò, i dati estratti sono stati impiegati per integrare l'elaborato con l'analisi descrittiva riepilogata in questa sede.

### **Analisi descrittiva delle inserzioni rilevate**

Le inserzioni su Facebook ed Instagram sono i contenuti che un qualsiasi utente può pubblicare a pagamento nelle due piattaforme. Come è noto, finora il modello di business dei social network accessibili in modo gratuito è stato alimentato

quasi esclusivamente da questi contenuti a pagamento, che permettono al soggetto pagante di raggiungere uno specifico *target* di utenti. Infatti, in base al costo sostenuto dal pagante, Meta permette lui di scegliere un beneficiario (che potrebbe non coincidere con lo stesso soggetto), il quale potrà creare un contenuto di elementi testuali e/o visuali da pubblicare in una o entrambe le piattaforme. Comparirà ad un bacino di utenti ben definiti, collocati in una specifica zona geografica e per un periodo di tempo limitato.

Da anni ormai i social vengono sfruttati dai partiti politici per integrare gli strumenti "tradizionali" della campagna elettorale, come le affissioni, le comunicazioni televisive o i comizi. Le caratteristiche intrinseche delle inserzioni permettono ai partiti di raggiungere gli elettori in modo diretto e con contenuti personalizzati, ovunque essi si trovino. Lo scopo? Provare a consolidare il rapporto con i propri elettori, ma soprattutto conquistarne di nuovi. Ovviamente, le elezioni politiche del 2022 non fanno eccezione.

Ai fini di questo approfondimento, sono stati raccolti i dati sulle inserzioni pubblicate dalle pagine ufficiali e dai leader delle classi politiche definite in questa analisi. Volendo concentrare l'attenzione sui contenuti a pagamento creati specificamente per la campagna elettorale delle elezioni politiche, sono state considerate soltanto le inserzioni prodotte a partire dal 21 luglio 2022, ovvero dal momento in cui le dimissioni del precedente governo Draghi sono state accolte dal Presidente della Repubblica, fino al giorno dell'elezione, avvenuta il 25 settembre 2022.

Escludendo dallo studio le inserzioni di importo inferiore a 99€, sono stati analizzati<sup>9</sup> i contenuti generati a partire dal 21 luglio fino al 25 settembre 2022 dei seguenti

---

<sup>9</sup>La scelta di non considerare le inserzioni al di sotto dei 99€ dipende da due motivi. Attualmente la Libreria Inserzioni non fornisce un intervallo di spesa chiaro al di sotto dei 99€, che quindi potrebbe variare da pochi euro fino al limite di quasi un centinaio. Poi, queste inserzioni, già di per sé poco frequenti, generano delle *impression* (numero di visualizzazioni complessive di un certo post, conteggiando anche quelle ripetute da uno stesso utente) estremamente basse (poche migliaia di visualizzazioni). Il loro impatto è ancora più marginale se le mettessimo in relazione

account (in grassetto le classi politiche associate):

1. Lega

- Lega - Salvini Premier
- Matteo Salvini

2. Forza Italia

- Forza Italia
- Silvio Berlusconi

3. Fratelli d'Italia

- Fratelli d'Italia
- Giorgia Meloni

4. Movimento 5 Stelle

- Movimento 5 Stelle
- Giuseppe Conte

5. Partito Democratico

- Partito Democratico
- Enrico Letta

6. Azione - Italia Viva

- Azione
- Italia Viva
- Carlo Calenda

7. Altri partiti

- +Europa
- Emma Bonino (+Europa)
- Luigi Di Maio (Impegno Civico)
- Noi Moderati
- Maurizio Lupi (Noi Moderati)
- Gianluigi Paragone (Italexit)
- Catenò de Luca (Sud Chiama Nord)
- Partito Comunista
- Unione Popolare
- SVP - Südtiroler Volkspartei

Nella classe "altri partiti" non sono stati inseriti tutti gli altri account di pagine ufficiali delle liste o dei loro leader in cui non comparivano inserzioni a pagamento nel periodo considerato.

Come prima indicazione, analizziamo la distribuzione aggregata per regione delle spese per le inserzioni di ogni classe politica, raggruppando quindi tutti gli account ad essa collegati (tabella A.7). Ordinando le colonne da sinistra a destra in base alla classe che ha speso di più, e le regioni nelle righe in ordine decrescente, troviamo nella prima cella numerica in alto a sinistra la spesa di FdI in Sicilia, pari a 19.025€. Il dato dimostra l'interesse di Fratelli d'Italia verso l'isola, numericamente al quarto

---

con gli altri contenuti più costosi, che possono generare da soli diverse milioni di visualizzazioni.

posto nel 2022 per numero di elettori (3,9 mln). Ma come FdI, anche gli altri partiti mostrano una particolare attenzione nella propria campagna social verso la Sicilia. Ad esempio, per la Lega è la terza regione in ordine di spesa (11.730€) e la seconda per il M5S (7.373€). Ricapitolando le proporzioni di voto nella regione, sul totale degli elettori troviamo un 46% di non votanti, il 12,3% di elettori a favore del M5S e il 9,9% per il FdI. Una quota rilevante di elettori siciliani, pari al 12,5%, ha votato per gli altri partiti. Soprattutto a favore della lista locale "Sud Chiama Nord".

Sulla base dei dati raccolti e rispetto alle sette classi analizzate, il partito di Giorgia Meloni è in testa in quasi tutte le regioni. È infatti la prima classe politica per spesa su Meta in tutta Italia, ad eccezione del Lazio e della Campania, della Lombardia (dove pure ha delle quote rilevanti di spesa, superata solo dal blocco degli "altri partiti"), del Trentino-Alto Adige e del Molise. Dalla tabella A.7 notiamo poi come il Terzo Polo abbia deciso di stanziare una quota di spesa per la campagna social su Meta più alta del Partito Democratico. Osserviamo anche la spesa relativamente contenuta del M5S, incentrata soprattutto verso le regioni del sud. Forza Italia, invece, concentra oltre la metà del suo impegno economico complessivo in Lombardia (8.538€, oltre il 50% della spesa totale su Meta). La classe "Altri" spicca in Trentino per le spese di SVP, ma anche in Lazio per la campagna di +Europa e Noi Moderati, in Campania con le quote stanziare da Impegno Civico e +Europa.

Proseguendo con l'analisi, per ogni classe politica possiamo distinguere la quota spesa dall'account ufficiale della lista e dal proprio leader. Ad esempio, sia la Lega che il M5S decidono di finanziare la campagna a favore delle inserzioni pubblicati dai rispettivi leader: l'account di Matteo Salvini spende il 97% dei 100.039€ associati alla Lega, mentre l'account di Giuseppe Conte spende la totalità dei 37.123€ del proprio partito. FdI e FI, invece, stanziavano circa l'80% della propria spesa creando inserzioni nell'account ufficiale della lista e il restante 20% per i contenuti dei loro capi di partito.

Se invece volessimo valutare il dato di spesa su Meta rispetto al "peso" di ogni regione in termini di elettorato? La tabella A.7 definisce le spese in valore assoluto: può quindi sembrare scontato che le quote di spese maggiori della campagna social vengano destinate nelle regioni più grandi. Per questo, è opportuno integrare quanto visto finora con un'altra metrica: la spesa su Meta ogni mille elettori.

Con questo secondo indicatore, riusciamo a visualizzare i dati di spesa sotto una stessa scala di confronto, indicando quanti euro di spesa hanno sostenuto le classi politiche su Meta ogni mille elettori tra le varie regioni. La tabella A.8 ci conferma l'alto interesse dei partiti verso la Sicilia, dove prevale FdI con 4,85€ di spesa sulle piattaforme Meta ogni mille elettori. Considerato l'alto numero di elettori in Lombardia, possiamo notare che in realtà la spesa sostenuta dai partiti non è così elevata, in relazione alle altre regioni. Appare quindi in modo più evidente che l'interesse dei partiti nella campagna social si sposta principalmente verso le regioni del sud: approssimando all'unità i valori della colonna "Totale" della tabella A.8, i partiti hanno speso complessivamente quasi 18€ ogni mille elettori in Sicilia ed in Campania, circa 15€ in Calabria ed in Puglia, 13€ in Sardegna e 11€ in Sardegna. Al centro-nord prevale il Lazio con 18€ e la Toscana con 11€. In Lombardia il dato complessivo di spesa è poco superiore ai 10€.

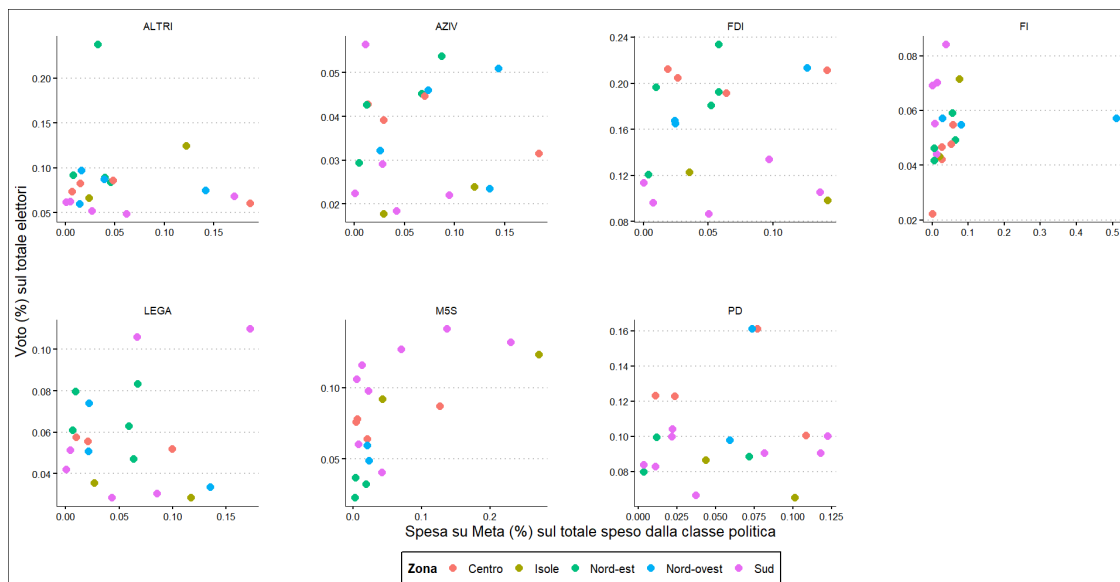
Dal momento che il dato regionale sulla spesa dei partiti su Facebook ed Instagram non è stata inserita nel modello inferenziale, non è stato possibile stimarne un suo effetto potenziale nel voto ottenuto da ogni partito. In ogni modo, ciò che possiamo comunque fare a livello descrittivo è visualizzare le variabili "Spesa su Meta (%)" sul totale speso dalla classe politica" e "Proporzione di voti ottenuti sul totale degli elettori" all'interno del piano cartesiano<sup>10</sup> (figura 2.4).

---

<sup>10</sup>Decidendo di considerare la "proporzione di voti ottenuti sul totale degli elettori", il dato terrà conto anche della quota di elettori non votanti. Il dato mostrato sarà quindi più basso della



Figura 2.4: Grafico a dispersione per spesa sui social e voto ottenuto a livello regionale - Elezioni 2022



La figura 2.4 contiene un grafico a dispersione per ognuna delle classi politiche analizzate; i punti al loro interno corrispondono alle regioni italiane, colorati in base alla rispettiva zona geografica e definiti sulle due dimensioni "Spesa su Meta (%) sul totale speso dalla classe politica" (asse  $x$ ) e "Voto (%) sul totale elettori" (asse  $y$ ). I grafici ci permettono di visualizzare quali siano le zone di maggiore interesse nella campagna social di ogni classe, e se esse corrispondono anche alle aree in cui i partiti riescono a ricevere una quota di voti più alta. Ad esempio, possiamo notare come la campagna del M5S sia ben allineata con i risultati ottenuti: il partito ha speso di più dove effettivamente ha ricevuto più voti, ovvero al sud e nelle isole. Anche la Lega mira alle regioni del sud Italia (soprattutto Sicilia, Campania, Puglia e Calabria), ma ottiene tra l'1 ed il 3% dei consensi in quella zona. PD ed AZIV presentano situazioni simili alla Lega: i partiti spendono relativamente di più nelle aree dove ottengono minori consensi (visivamente, molte regioni si collocano in

proporzione di voto ottenuta rispetto al totale dei votanti.

basso a destra nei rispettivi grafici). Ma possiamo dire che la loro campagna social su Meta è stata inefficace? O che hanno investito male le loro risorse? Ovviamente no. Attraverso questo grafico non possiamo azzardare una conclusione del genere. In altre parole, non ci permette di comprendere se la spesa sostenuta su Meta ha permesso la Lega o il PD di ricevere comunque qualche voto in più rispetto a quanto si è verificato. O viceversa, se la spesa è stata del tutto irrilevante.

L'auspicio è quello di riuscire a raccogliere in futuro dei dati sui social network più completi, in modo tale da poterli integrare al modello inferenziale, al fine di stimare una potenziale relazione tra spese per inserzioni e voti ricevuti dal partito. Perché no, magari anche per analizzare i possibili effetti della spesa di una certa lista nei confronti dei voti ottenuti da un altro partito.

L'analisi descrittiva si conclude qui.

Il prossimo capitolo, come già annunciato, sarà dedicato completamente allo studio inferenziale tra i risultati elettorali ed i possibili fattori determinanti.

## Capitolo 3

# Modello statistico per le elezioni multipartito

Le statistiche riportate nel capitolo precedente si limitano ad analizzare il voto sulla base degli stessi esiti elettorali, declinati per zona geografica.

Abbiamo potuto osservare il cambio di orientamento dell'elettorato tra le due elezioni a favore del centro-destra in tutte le macro aree della penisola, tenendo sempre in considerazione l'ulteriore crescita della quota di elettori non votanti, già molto elevata nel 2018.

Ora, si passa alla fase successiva di questo studio.

Nel presente capitolo si cercherà di comprendere quali sono stati i reali fattori ad impattare sulle scelte di voto degli elettori italiani. Analizzeremo se e come sono cambiate queste determinanti tra le due elezioni.

A tale scopo, va prima individuato un insieme di fattori demografici, socio-economici, culturali e geografici (covariate) che potenzialmente producono effetti nei voti espressi dagli elettori (variabili dipendenti).

Identificare la corretta relazione tra le variabili dipendenti e le covariate scelte è oggetto dell'analisi inferenziale, che porta alla definizione di uno specifico modello

statistico.

In letteratura, sono stati sviluppati vari approcci modellistici per l'analisi dei dati elettorali aggregati, tra i quali ricordiamo Gibson *et al.* (1995) [3] e Tomz *et al.* (2002) [14]). Uno fra i più noti, che ha ispirato centinaia di lavori successivi, è il modello proposto da Katz e King (1999) [10], in seguito "KK".

Il modello di KK, ideato alla fine del secolo scorso ed ampliato dagli studi successivi di Honacker *et al.* (2002) in [5] e [6], è un notevole strumento per spiegare quanto i fattori esterni impattino sui risultati elettorali in un sistema politico multipartito<sup>1</sup>. KK arricchiscono ulteriormente il valore del proprio modello proponendo un metodo di previsione degli esiti di voto basato su simulazioni di Monte Carlo.

Per anticipare il contenuto successivo:

- il paragrafo 3.1 è dedicato alla definizione generale del modello di KK;
- la scelta ed il dettaglio delle covariate, affiancate dalle rispettive fonti, saranno il tema del paragrafo 3.2;
- nel paragrafo 3.3 troveremo infine l'applicazione del modello alle elezioni politiche italiane, l'analisi predittiva e controfattuale con l'interpretazione complessiva dei risultati ottenuti.<sup>2</sup>

---

<sup>1</sup>Un sistema politico multipartito è caratterizzato dalla presenza di tre o più partiti in competizione tra loro; si differenziano così dal bipartitismo in cui sono soltanto due i gruppi principali a contendersi effettivamente la vittoria delle elezioni (come negli Stati Uniti per il Congresso e la presidenza). Il sistema italiano è chiaramente multipartito, data l'elevata frammentazione delle forze politiche.

<sup>2</sup>Il codice con il quale sono stati ottenuti i risultati di questa analisi è disponibile pubblicamente al presente [link](#).

### 3.1 Definizione generale del modello KK

I dati elettorali aggregati possono essere espressi in termini assoluti, dichiarando il numero effettivo di voti ottenuti da un certo partito nell'unità considerata (comune, collegio, circoscrizione, regione, ecc...), oppure in termini proporzionali rispetto al totale. Da questo momento in poi, i dati elettorali verranno trattati in termini proporzionali.

Definendo  $V_{i,j}$  la proporzione dei voti ottenuti dal partito "j" con  $j = 1, \dots, J$  nell'aggregato "i" con  $i = 1, \dots, N$ , i dati elettorali multipartito<sup>3</sup> hanno due caratteristiche fondamentali.

La prima proprietà indica che, trattandosi di dati proporzionali rispetto al totale, ogni proporzione dei voti deve ricadere nell'intervallo compreso tra 0 ed 1. Formalmente:

$$V_{i,j} \in [0, 1] \text{ per tutti "i" e "j"} \quad (3.1)$$

La seconda proprietà prevede che la somma delle proporzioni dei voti ottenuti da tutti i partiti nel j-esimo aggregato debba necessariamente sommare ad 1, ovvero che:

$$\sum_{j=1}^J V_{i,j} = 1 \text{ per tutti "i" e "j"} \quad (3.2)$$

Sono due facili requisiti da intuire e verificare: un partito non può procurarsi un valore negativo di voti nell'aggregato considerato (nel caso peggiore otterrà zero voti) e non può nemmeno ottenere più del 100% dei voti. Nonostante l'immediatezza delle due proprietà, saranno di fondamentale importanza nella costruzione del modello statistico e nella verifica dei risultati prodotti.

I dati che incorporano tali caratteristiche sono detti "dati composizionali" ed hanno

---

<sup>3</sup>Nel nostro caso specifico, "j" si riferirà alla classe politica come definita precedentemente nel paragrafo 2.2, mentre "i" sarà relativo all'entità "comune". Nel caso originariamente trattato da KK [10], l'entità considerata era il distretto elettorale (disponendo così tra le 455 e le 522 osservazioni per ognuna delle elezioni politiche avvenute dal 1959 al 1992 nel Regno Unito).

una vasta gamma di applicazioni. Ad esempio in geologia servono ad analizzare la composizione chimica, minerale e fossile dei campioni di rocce; nelle analisi della domanda dei consumatori, possono permettere l'investigazione sulle proporzioni di ogni potenziale categoria di spesa sul budget individuale.

Tuttavia, i vincoli intrinseci dei dati composizionali creano problemi nella modellazione attraverso variabili casuali con spazi campionari ristretti. Per questo, la soluzione più celebre per risolvere questa tipologia di problemi è quella di Aitchinson (1982) [1], che propone di applicare ai dati originali una trasformazione per associarli alla distribuzione normale.

Analiticamente, definendo  $Y_{ij}$  come il rapporto dei logaritmi (o *log-ratio*) della proporzione di voto del partito  $j$ -esimo e del partito " $J$ " usato come base:

$$Y_{ij} = \ln(V_{ij}/V_{iJ}), \text{ con } j = 1, \dots, J - 1 \quad (3.3)$$

Aitchison assume che il vettore di variabili casuali  $Y_i = (Y_{i1}, \dots, Y_{i(J-1)})$  di dimensione  $J - 1$  si distribuisca come una normale multivariata con media  $\mu$  e varianza  $\Sigma$ .

La trasformazione inversa che permette di tornare dal *log-ratio* ( $Y_{ij}$ ) alle proporzioni di voto osservate ( $V_{ij}$ ) è la logistica multivariata, definita come

$$V_{ij} = \frac{\exp(Y_{ij})}{1 + \sum_{j=1}^{J-1} \exp(Y_{ij})} \quad (3.4)$$

Ciò che notano gli autori in Katz *et al.* (1999) [10] è la base per un cambiamento sostanziale rispetto alla soluzione iniziale di Aitchinson: i loro dati elettorali, divisi per partito a livello di distretto, presentano valori molto frequenti attorno alla media e con poche eccezioni circoscritte soltanto ad alcuni distretti in tutto il Regno Unito, molto differenti dagli altri risultati. Questa osservazione porta KK ad associare i dati trasformati non più alla distribuzione normale multivariata teorica, ma alla  $t$  di Student multivariata.

Quindi, KK propongono un nuovo modello basato sulla distribuzione logistica additiva  $t$  di Student (LT), che nel loro caso ha mostrato una miglior aderenza alle singole proporzioni di voto. Di conseguenza, essi assumono che il *log-ratio*  $Y_i$  si distribuisca come una  $t$  di Student multivariata; applicando ad  $Y_i$  la trasformazione inversa (3.4) si ritorna alla proporzione di voto  $V_i$  con distribuzione LT così definita:

$$\begin{aligned}
 P(V_i|\mu, \Sigma, v) &= LT(V_i|\mu, \Sigma, v) = \\
 &= T[\ln(V_i/V_{iJ})|\mu_i, \Sigma] / \prod_{j=1}^{J-1} V_{ij} = \\
 &= \frac{\Gamma[(v+J-1)/2]|\Sigma|^{-1/2}}{\Gamma(v/2)v^{(J-1)/2}\pi^{(j-1)/2} \left( \prod_{j=1}^{J-1} V_{ij} \right)} \left( 1 + \frac{1}{v} (Y_i - \mu_i)' \Sigma^{-1} (Y_i - \mu_i) \right)^{(v+J-1)/2}
 \end{aligned} \tag{3.5}$$

Il valore atteso del rapporto di logaritmi  $Y_{ij}$  è  $\mu_{ij}$  e la sua varianza è  $\Sigma v/(v-2)$ , dove  $v$  è il parametro per i gradi di libertà ed è quindi strettamente positivo, mentre  $\Sigma$  è una matrice definita positiva di dimensione  $(J-1) \times (J-1)$ .

Il modello sarà diverso da quello ottenuto sotto ipotesi di distribuzione normale logistica additiva nel caso di  $v < \infty$ , e la loro divergenza cresce al ridursi del numero  $v$  riferito ai gradi di libertà.

Grazie all'assunzione di distribuzione LT dei dati composizionali, KK osservano che il cluster centrale di dati elettorali più addensati aderisce maggiormente alla distribuzione teorica, e che le code si adattano meglio ai pochi distretti molto differenti al caso medio. Ciò accade appunto perché nei dati elettorali UK usati dagli autori, la maggioranza dei collegi ha proprietà simili tra loro, e solo pochi di essi si distanziano di molto.

Nel caso in esame, KK considerano la media dei  $J-1$  *log-ratios* pari alla combinazione lineare tra il vettore delle covariate ed i parametri da stimare con il modello, ovvero:

$$\mu_{ij} = X_{ij}\beta_j$$

dove  $X_{ij}$  è il vettore delle covariate osservate di dimensione  $p_j \times 1$ , e  $\beta_j$  sono i

parametri da stimare.

KK non fanno assunzioni in merito all'Indipendenza delle Alternative Irrilevanti (IIA), necessarie per i modelli costruiti su dati a livello di singolo elettore (basati su sondaggi). In altre parole, il loro modello permette l'ingresso e l'uscita di un partito dalla competizione all'interno di un distretto (vedremo in seguito con l'inserimento di Azione - Italia Viva, altrimenti inglobato nella classe "Altri partiti").

Per quanto riguarda la stima dei parametri del modello, KK propongono una metodologia basata sulla massima verosimiglianza. Tuttavia, essi concentrano parte della loro ricerca sul problema dei dati mancanti legati a quei partiti che non ottengono voti o che decidono di non presentarsi in un certo distretto. La prima soluzione in Katz *et al.* (1999) [10] è la stima con approccio "*Direct Likelihood*". Gli sviluppi successivi in King *et al.* (2001) [12], Honaker *et al.* (2001) [7], Honaker *et al.* (2010) [8] e Honaker *et al.* (2011) [9] sono stati dedicati quasi esclusivamente all'imputazione dei dati mancanti per limitare i costi computazionali. Precedentemente alla scelta del metodo di stima nel caso delle elezioni politiche italiane 2018 e 2022, è stata fatta la seguente riflessione: ad eccezione della Valle d'Aosta, non considerata nell'analisi, tutti i partiti principali si sono presentati in ognuno dei collegi uninominali designati (e di conseguenza ci sono dati disponibili per tutti i comuni italiani); esclusi pochi casi di paesi con dimensioni molto ridotte, tutti i partiti di rilievo hanno ottenuto almeno un voto; le liste locali, per definizione, sono presenti soltanto in pochi collegi. Per questo è stato deciso di non imputare dati ai maggiori partiti dove effettivamente gli elettori hanno scelto di non votarli, oppure aggiungere valori alla classe delle liste locali anche nei collegi dove in realtà mancavano.<sup>4</sup>

---

<sup>4</sup>Tenendo sempre presente che le liste locali, come definite nel paragrafo 2.2, hanno ottenuto in media nazionale appena lo 0,6% nel 2018 e l'1,4% nel 2022.



In questo modo, al fine di limitare al massimo la perdita di informazioni (ed il successivo rischio di distorsione nelle stime), nel modello la classe "Locali" è stata accorpata all'interno della classe "Altri".

Il modello statistico costruito stima con il metodo della massima verosimiglianza (ML - *Maximum Likelihood*) i relativi coefficienti per ognuna delle variabili dipendenti ( $Y_{ij}$ ) e per un dato insieme di covariate ( $X_{ij}$ ). In base a quanto detto in precedenza, alla base di questo modello con ML assumiamo che:

$$Y_{ij} = \log(V_{ij}/V_{iJ}) \sim t(\mu, \Sigma v/(v - 2), v)$$

Già dal capitolo precedente è stata evidenziata la sempre più rilevante quota di non votanti nel caso italiano. Per tenerne conto anche all'interno del modello, è stato deciso di aggiungere come ultima classe  $V_{iJ}$  la proporzione di non votanti sul totale degli elettori.

Così facendo, tutte le  $J - 1$  variabili dipendenti esprimono il valore del rapporto dei logaritmi tra ognuna delle classi definite in 2.2 rispetto ai non votanti.

Questo approccio, non adottato da KK ma suggerito da altri studi come Bekkouche *et al.* (2022) [2], porta con sé vari vantaggi. Innanzitutto, si dispone di un'equazione specifica per ogni classe politica oggetto di studio, senza escluderne alcuna dovendola usare come denominatore di riferimento nel *log-ratio*. Poi, l'interpretazione dei coefficienti, seppure non immediata, sarà standardizzata per ogni classe in esame, e risponderà alla domanda: quanto incide la variabile sul voto a favore di una specifica forza politica, rispetto alla decisione di non votare? Data la quota molto elevata di non votanti sul totale degli elettori in Italia nel 2018 e nel 2022, si riesce così a tenere conto anche di questo fenomeno all'interno del modello.

Prima di passare all'interpretazione dei risultati veri e propri, di seguito viene proposto un breve sottoparagrafo sul metodo ML per la stima dei parametri applicato al modello KK per le elezioni politiche italiane.

### 3.1.1 Stimatore di massima verosimiglianza

La stima di massima verosimiglianza richiede, come input principali, una realizzazione di una variabile casuale con "n" osservazioni  $y = y_1, \dots, y_n$ , ed un'ipotesi distributiva sul campione osservato.

Considerando un campione proveniente da una variabile casuale continua<sup>5</sup>, si può definire la densità congiunta di osservare l'intera realizzazione dei dati. La funzione di densità congiunta varierà così in base all'assunzione distributiva ed ai parametri della distribuzione scelta ( $\theta$ ). Si considera che tali parametri, ovviamente ignoti, rientrino all'interno di un certo spazio parametrico  $\Theta$ .

Questa funzione di densità congiunta è detta funzione di verosimiglianza ( $\mathcal{L}(\theta)$ ): essa rappresenta il valore di probabilità di osservare congiuntamente tutti i dati noti ( $y$ ) per un certo valore dei parametri della distribuzione. Lo scopo dello stimatore ML è perciò quello di individuare il valore dei parametri ignoti della distribuzione selezionata che permette di massimizzare la "verosimiglianza" di osservare quella specifica realizzazione campionaria.

Sotto condizioni non eccessivamente restrittive, ad esempio che le osservazioni siano *iid*, la probabilità congiunta diventa pari al prodotto delle singole probabilità. Inoltre, sfruttando la trasformazione in logaritmo della funzione di verosimiglianza (di seguito definita come "log-verosimiglianza"), il prodotto delle singole probabilità diventa pari alla sommatoria di ogni log-probabilità e si riesce così a semplificare gli aspetti analitici senza alterare il punto di massimo della funzione originaria.

Formalizzando quanto detto finora avremo che:

$$l(\theta) = \log(\mathcal{L}(\theta)) = \log f(y; \theta) = \sum_{i=1}^n l_i(\theta) \quad (3.6)$$

$$\hat{\theta}(y) = \operatorname{argmax}_{\theta \in \Theta} l(\theta) \quad (3.7)$$

---

<sup>5</sup>Tutti i discorsi successivi valgono analogamente al caso discreto purché ci si riferisca a funzioni di massa di probabilità  $p(y)$  anziché di densità  $f(y)$ .

Al verificarsi di specifiche condizioni, tra le quali disporre di un'ampiezza campionaria che tende ad infinito e ovviamente aver ipotizzato la corretta distribuzione sottostante ai dati osservati, lo stimatore ML gode di proprietà molto attraenti. In particolare, i parametri stimati  $\hat{\theta}$  saranno stime consistenti, efficienti ed asintoticamente normali dei veri parametri ignoti  $\theta$ .

Generalmente, per trovare i valori di  $\theta$  che massimizzano la log-verosimiglianza (3.7) si adottano tecniche numeriche: iterativamente, i metodi numerici scelgono una certa valorizzazione dei parametri all'interno dello spazio parametrico ( $\theta_0 \in \Theta$ ), definiscono il vettore gradiente (o vettore *score*) nei parametri selezionati ( $g(\theta_0)$ ) e ne valutano il comportamento. Considerando  $g(\theta)$  una funzione continua e differenziabile, se  $g(\theta_0)$  è prossimo allo zero, allora potremmo essere vicini al punto di massimo di  $\hat{\theta}(y)$  e quindi l'algoritmo si ferma. Nel caso in cui il gradiente è ancora lontano dal criterio di stop<sup>6</sup>, si procede a calcolare un vettore direzione funzione dello stesso gradiente che serve ad aggiornare il valore dei parametri. Fino a quando il vettore gradiente risulta diverso da zero (o non si è giunti ad un altro criterio di stop prefissato), l'algoritmo numerico va avanti nella ricerca. Tra i più noti metodi di ottimizzazione numerica basati sul gradiente troviamo, ad esempio, il BFGS ed il Newton-Raphson.

Un aspetto da non trascurare è l'inizializzazione dei parametri ( $\theta_0$ ): chiaramente, riuscire a scegliere dei valori di  $\theta_0$  iniziali il più possibile "vicini" al punto di massimo aiuta l'algoritmo di ottimizzazione (qualsiasi di essi) a raggiungere effettivamente  $\hat{\theta}(y)$ .

Riprendendo ora l'applicazione dello stimatore ML al caso in oggetto, l'ipotesi tale per cui  $Y_{ij} \sim t(\mu, \Sigma v/(v-2), v)$ , proposta da KK nel loro studio, è stata

---

<sup>6</sup>In realtà, ci sono vari criteri di stop degli algoritmi numerici per la massimizzazione di una funzione oltre all'azzeramento del gradiente. Ad esempio, per evitare costi computazionali troppo elevati, si potrebbe decidere di fermare il calcolo una volta giunto ad un certo numero predefinito di iterazioni, oppure quando il gradiente smette di variare da una certa iterazione in poi.

mantenuta anche per le elezioni politiche italiane. Alla base di ciò vi è il rifiuto dell'ipotesi nulla in ognuno dei test di normalità Doornik e Hansen sulle  $Y_{ij}$ .

Il senso è il seguente: nonostante un'ampiezza del campione di circa 8.000 unità, l'eterogeneità dei comuni italiani si manifesta anche in alcuni risultati molto differenti dalla media nazionale. Pertanto, il valore atteso condizionale di  $Y_{ij}$ , definito come  $\mu_{ij} = X_{ij}\beta_j$ , è modellato per ipotesi da una combinazione lineare tra le covariate ed i rispettivi coefficienti da stimare. Lo scarto tra il valore osservato delle variabili dipendenti ed il valore atteso condizionale corrisponde al termine di errore  $\varepsilon_{ij}$ : di conseguenza seguirà anch'esso, per ipotesi, una distribuzione  $t$  di Student multivariata. In questo modo, un termine di errore distribuito come  $t$  (anziché come normale multivariata) aderirà più coerentemente alla situazione realmente verificata, nella quale registriamo degli esiti di voto in alcuni paesi molto distanti dal loro valore atteso condizionale.

Definita l'essenziale ipotesi distributiva alla base della stima di massima verosimiglianza, i parametri  $\theta$  da massimizzare in funzione dei dati osservati saranno:

- $\beta_j$ , considerando la relazione lineare per definire il valore atteso condizionale del *log-ratio* pari a  $\mu_{ij} = X_{ij}\beta_j$ ;
- $\Sigma$ , matrice definita positiva, dalla quale viene calcolata la matrice varianza-covarianza di  $Y_{ij}$  pari a  $\Sigma v/(v-2)$ ;
- $v$ , ovvero il numero di gradi di libertà caratteristici della  $t$  multivariata e distintivi rispetto alla distribuzione normale (le distribuzioni coincidono per  $v \rightarrow +\infty$ ).

Nel caso in esame, la funzione di log-verosimiglianza sarà:

$$l(\theta) = \log(\mathcal{L}(\theta)) = \sum \log(T(Y_{i1}, \dots, Y_{iJ-1}|\theta)) \quad (3.8)$$

con  $\theta = \{\beta, \Sigma, v\}$  e  $T(Y_{i1}, \dots, Y_{iJ-1})$  pari alla densità congiunta  $t$  multivariata dei rapporti dei logaritmi di ogni classe politica rispetto al non voto.

Per favorire la convergenza dell'algoritmo di ottimizzazione numerica (BFGS), il valore dei parametri usati per l'inizializzazione ( $\theta_0$ ) provengono da una stima OLS multivariata.

## 3.2 Le variabili indipendenti nel modello

La selezione delle covariate è un aspetto di fondamentale importanza, al pari della scelta modellistica più adeguata al caso di studio.

Dietro un'ipotesi iniziale, vanno individuati dei fattori che potrebbero influenzare in qualche modo i valori osservati delle variabili dipendenti. Le covariate, dette anche variabili indipendenti o regressori, rappresentano quei fattori in modo numerico o testuale.

Uno degli obiettivi dell'analisi è senz'altro quello di ricercare quali sono state le principali determinanti di voto in Italia a livello nazionale nel 2018 e nel 2022.

Tenendo conto che non è possibile inserire nel modello tutte i possibili fattori impattanti sulle scelte di voto degli elettori (sia per mancanza di dati, che per problemi di correlazione tra le stesse covariate), occorre procedere con un'accurata selezione. Dopo aver raccolto dati sulle principali variabili demografiche, socio-economiche, culturali e geografiche disponibili a livello di comune, la scelta delle covariate da inserire nel modello è avvenuta in modo tale da:

- Riuscire a rappresentare le caratteristiche principali della popolazione su base comunale;
- Mantenere soltanto i regressori non altamente correlati tra loro per limitare i problemi nell'interpretazione dei coefficienti  $\hat{\beta}_j$  stimati;
- Integrare quelle variabili strettamente legate ai punti principali dei programmi politici dei partiti o delle coalizioni;

- Garantire in ogni caso una buona capacità predittiva del modello dato un set limitato di covariate.

In ognuno dei capoversi successivi verrà fornita una descrizione sintetica delle variabili indipendenti inserite nel modello. In questo modo, cercheremo di dare una contestualizzazione alla situazione generale dell'Italia in prossimità delle elezioni, segnalando quali sono stati i cambiamenti avvenuti in media nazionale dal 2018 al 2022. All'interno di ogni capoverso saranno contenute alcune statistiche e le rispettive fonti. Il dettaglio delle principali statistiche descrittive e la matrice delle correlazioni viene riportato nelle tabelle B.2, B.3 e B.4 in Appendice B.1. Si ricorda che ogni dato è relativo al valore rilevato in uno specifico comune italiano.

**Quota di votanti maschi sul totale:** questa prima covariata corrisponde al rapporto tra il numero di votanti di genere maschile ed il numero totale di votanti in un certo comune nell'elezione politica di riferimento.

Ovvero:  $\text{maschi}_i = \frac{\text{votanti maschi}_i}{\text{votanti}_i}$  con  $i = 1, \dots, n$  riferito all' $i$ -esimo comune.

La variabile ha un'alta densità attorno alla media (eccesso di curtosi pari a 10,12 e C.V. pari a 0,043). Tuttavia, la forte asimmetria positiva (+2,07) suggerisce la presenza di alcuni comuni in cui i maschi hanno votato in proporzione parecchio superiore alla media di circa 51%. Infatti, si trovano picchi oltre il 65-70% in piccoli paesi, a differenza della quota di femmine votanti che non supera mai il 60% dei votanti totali.

Fonte: *Ministero dell'Interno - Eligendo*

**Quota di stranieri sul totale della popolazione:** pari al rapporto tra il numero di residenti di cittadinanza straniera ed il totale della popolazione residente per comune al 1° gennaio.

$\text{stranieri}_i = \frac{\text{residenti stranieri}_i}{\text{residenti}_i}$  con  $i = 1, \dots, n$

Fonte: *Istat - Demo.statistiche demografiche*

**Reddito medio da lavoro dipendente (in logaritmo):** pari al logaritmo del reddito medio dichiarato ai fini IRPEF dai lavoratori dipendenti in ogni comune. La scelta di trasformare il dato grezzo in logaritmo è per ridurre la scala originaria (altrimenti di circa quattro ordini di grandezza in più rispetto alle variabili dipendenti) e la leptocurtosi, limitare l'asimmetria positiva e favorire l'interpretazione dei coefficienti in termini di elasticità.<sup>7</sup> Questa variabile è stata inserita come *proxy* della ricchezza media della classe dipendente nel comune di riferimento. Il dato si riferisce alle dichiarazioni fiscali relative all'anno d'imposta (a.i.) precedente all'anno delle elezioni.<sup>8</sup>

Fonte: *MEF - Statistiche fiscali*

**Popolazione residente (in logaritmo ed in forma quadratica):** pari al logaritmo applicato al numero di cittadini (italiani e stranieri) residenti. Il dato è un indicatore della dimensione demografica del comune. La motivazione dietro alla scala logaritmica è analoga al caso del reddito medio da lavoro dipendente, e questa volta con problematiche nei dati grezzi ancora più accentuate. La forma quadratica, invece, serve ad introdurre un elemento di non linearità nella relazione tra la covariata e le variabili dipendenti,<sup>9</sup> capace di differenziare ulteriormente i comuni

---

<sup>7</sup>In questa situazione, poiché sia la covariata che le variabili dipendenti sono espresse in logaritmo, il valore del coefficiente  $\beta$  associato risponderà alla domanda: "di quanto varia il rapporto tra proporzione di voti sul non voto ( $\frac{V_{i,i}}{V_{i,j}}$ ) in termini percentuali, al variare di un 1% del reddito medio da lavoro dipendente?"

<sup>8</sup>Quindi anno d'imposta 2017 per elezioni 2018 ed anno d'imposta 2021 per elezioni 2022.

<sup>9</sup>Soltanto la variabile  $\log(\text{popolazione})$  è stata inserita anche in forma quadratica, poiché la non linearità di altre covariate, anche diverse dal quadrato, venivano rilevate spesso in modo statisticamente non significativo dal modello.

altamente popolati dai piccoli paesi italiani. Il dato si riferisce ai residenti registrati al 1° gennaio.

Fonte: *Istat - Demo.statistiche demografiche*

**Quota di persone over 60 anni sul totale:** variabile pari al rapporto tra popolazione con più di 60 anni sul totale.

Quindi  $over60_i = \frac{\text{popolazione over60}_i}{\text{popolazione}_i}$  con  $i = 1, \dots, n$

Il dato, stimato dall'Istat con i suoi Censimenti permanenti validi per l'intera popolazione, è integrato nel modello per tenere conto di vari aspetti. Uno dei principali, per quantificare la quota di potenziali interessati alle misure previdenziali.

Le statistiche di sintesi ci segnalano un aumento nella media di over 60 nel 2022, passando da 31,8% a 33,4% in quattro anni (mantenendo quasi inalterati gli indici di variabilità).

Fonte: *Istat - Censimenti permanenti*

**Quota di popolazione con titolo di studio terziario sul totale:** variabile che comprende la proporzione dei residenti in Italia con un titolo di studio terziario (laurea triennale, magistrale, dottorato di ricerca, ...) o un diploma di tecnico superiore ITS.

L'idea dietro a questa variabile è di poter misurare la presenza o meno di residenti che hanno proseguito gli studi anche successivamente alla scuola secondaria di secondo grado, con tutte gli effetti che ne conseguono (maggiore occupabilità, ingresso successivo nel mondo del lavoro, interesse verso la ricerca, accesso più frequente a posizioni manageriali, ...).

Fonte: *Istat - Censimenti permanenti*



**Tasso di disoccupazione:** espresso come rapporto tra persone in cerca di occupazione e le corrispondenti forze di lavoro.<sup>10</sup>

Formalmente, tasso disoccupazione<sub>*i*</sub> =  $\frac{\text{cerca occupazione}_i}{\text{forze lavoro}_i}$  con  $i = 1, \dots, n$

Quindi, il tasso di disoccupazione rappresenta la quota di persone che non lavora ma sarebbe disponibile a lavorare, sul totale dei lavoratori e di coloro che stanno appunto cercando lavoro. Per questo, è un valido indicatore sullo stato dell'economia e ovviamente del mercato del lavoro. Ma è anche strettamente correlato alla povertà (assoluta e relativa) che per questo non è stata inserita nell'analisi.

Generalmente, a contrasto della disoccupazione e della povertà vengono messe in atto delle misure di politica attive e sussidi specifici, che di rado hanno un carattere strutturale. La misura più nota a riguardo negli anni di riferimento è il Reddito di Cittadinanza, promosso dal M5S a partire dalle elezioni del 2018. Viene ricordato per la vasta platea di beneficiari (tra i 2,5 ed i 3,8 milioni all'anno) e di conseguenza per il suo impatto sulla spesa pubblica (circa 30 miliardi di euro in 4 anni).

Fonte: *Istat - Censimenti permanenti*

**Zona geografica:** per tenere conto delle variazioni intrinseche in ognuna delle diverse aree geografiche italiane, sono state aggiunte delle variabili *dummy* (dicotomiche) associate ai cinque gruppi di regioni della codifica europea *NUTS 1* (Nord-ovest, Nord-est, Centro, Sud, Isole). Per evitare la "*trappola delle dummy*" ed i problemi di collinearità che ne conseguono, la variabile Nord-est è stata esclusa dal modello e sono state inserite le quattro restanti.

Fonte: *Istat*

---

<sup>10</sup>In breve, riprendendo le definizioni Istat, le "forze di lavoro" comprendono le persone occupate e quelle disoccupate (ovvero quelle in cerca di occupazione, che altrimenti rientrerebbero tra gli "inattivi").

**Grado di urbanizzazione:** l'Istat definisce una classificazione dei comuni basata sul criterio della contiguità geografica e su soglie di densità di popolazione minima della griglia regolare con celle da 1km<sup>2</sup>.

Seguendo questo criterio, i comuni vengono ripartiti tra:

- "Città" o "zone densamente popolate";
- "Piccole città o sobborghi" o "Zone a densità intermedia di popolazione";
- "Zone rurali" o "Zone scarsamente popolate".

Ricadono dentro la categoria delle "Città" circa 260 comuni con 80.000 abitanti in media e in "Sobborghi" oltre 2.600 comuni di 11.000 abitanti in media. Sono invece circa 5.050 le "zone rurali", mediamente popolate da 2.000 abitanti.

Questo raggruppamento aiuta a distinguere ulteriormente la varietà di comuni italiani, tutt'altro che omogenei tra loro.

Di nuovo, per evitare la *"trappola delle dummy"*, non è stata inserita la variabile "Città").

Fonte: *Istat*

**Mare e/o montagna:** come ultime covariate troviamo una *dummy* per indicare se il comune è situato in montagna, e/o si trova in una zona costiera. L'Istat ripartisce il territorio nazionale e classifica i comuni in base a valori soglia altimetrici, distinguendo tra montagna, collina e pianura. L'indicazione se il comune si trova in montagna (sia essa interna o litoranea) è stata inserita nel modello (circa 2.500 paesi). Collina e pianura no, poiché considerate troppo vaghe e non discriminanti in questo contesto.

Inoltre, l'Istat individua circa 1.170 comuni in zone costiere, collocati sulla costa o aventi almeno il 50% delle superficie ad una distanza dal mare inferiore a 10 km.

Per riassumere, le covariate usate nei modelli del 2018 e del 2022 saranno:

$x$	Descrizione	Dummy
$x_0$	Costante o intercetta	-
$x_1$	Quota di votanti maschi sul totale	-
$x_2$	Quota di stranieri sulla popolazione totale	-
$x_3$	Logaritmo del reddito medio da lavoro dipendente	-
$x_4$	Logaritmo della popolazione residente	-
$x_5$	Logaritmo della popolazione residente al quadrato	-
$x_6$	Quota di persone con più di 60 anni sulla popolazione totale	-
$x_7$	Quota di persone con titolo di studio terziario sulla popolazione totale	-
$x_8$	Tasso di disoccupazione	-
$x_9$	Nord-ovest	$d$
$x_{10}$	Centro	$d$
$x_{11}$	Isole	$d$
$x_{12}$	Sud	$d$
$x_{13}$	Sobborghi o zone a densità intermedia di popolazione	$d$
$x_{14}$	Zone rurali o scarsamente popolate	$d$
$x_{15}$	Comune in montagna	$d$
$x_{16}$	Zona costiera	$d$

### 3.3 Risultati del modello applicato alle elezioni politiche italiane

Riadattiamo brevemente quanto detto finora in merito al modello KK al caso delle elezioni politiche italiane del 2018 e del 2022.

Le variabili dipendenti  $Y_{i1}, \dots, Y_{iJ-1}$  corrispondono al rapporto dei logaritmi della proporzione di voto ottenuta da ogni classe politica rispetto alla quota di elettori non votanti. Per ipotesi, motivata dalla presenza di risultati molto distanti rispetto alla media nazionale, le variabili dipendenti si considerano provenire da una distribuzione  $t$  di Student multivariata.

Ognuna delle "j" variabili dipendenti è associata ad un'equazione che mette in relazione il *log-ratio* con una combinazione lineare tra il vettore delle covariate  $\mathbf{x}_{ij}$  elencate nel paragrafo 3.2 ed i coefficienti  $\beta_j$ .

Avremo così sei equazioni distinte per le elezioni 2018, e sette per il 2022 (avendo aggiunto AZIV, che altrimenti sarebbe confluito all'interno della classe "Altri partiti"). Esplicitamente, le equazioni appariranno in questa forma:

$$\log\left(\frac{V_i \text{ Lega}}{V_i \text{ NonVoto}}\right) = \alpha_1 + \mathbf{x}'_i \beta_1 + \varepsilon_{i1}$$

$$\log\left(\frac{V_i \text{ FI}}{V_i \text{ NonVoto}}\right) = \alpha_2 + \mathbf{x}'_i \beta_2 + \varepsilon_{i2}$$

$$\log\left(\frac{V_i \text{ FdI}}{V_i \text{ NonVoto}}\right) = \alpha_3 + \mathbf{x}'_i \beta_3 + \varepsilon_{i3}$$

$$\log\left(\frac{V_i \text{ PD}}{V_i \text{ NonVoto}}\right) = \alpha_4 + \mathbf{x}'_i \beta_4 + \varepsilon_{i4}$$

$$\log\left(\frac{V_i \text{ M5S}}{V_i \text{ NonVoto}}\right) = \alpha_5 + \mathbf{x}'_i \beta_5 + \varepsilon_{i5}$$

$$\log\left(\frac{V_i \text{ Altri}}{V_i \text{ NonVoto}}\right) = \alpha_6 + \mathbf{x}'_i \beta_6 + \varepsilon_{i6}$$

$$\log\left(\frac{V_i \text{ AZIV}}{V_i \text{ NonVoto}}\right) = \alpha_7 + \mathbf{x}'_i \beta_7 + \varepsilon_{i7}$$

dove  $\alpha_j$  indica l'intercetta,<sup>11</sup> i vettori  $\beta_j$  sono i coefficienti di ogni equazione associati allo stesso insieme di covariate  $\mathbf{x}_i$ , ed  $\varepsilon_{ij}$  corrisponde al termine di disturbo. Il metodo ML stimerà congiuntamente i parametri di ogni equazione ( $\alpha$ ,  $\beta$ ,  $\Sigma$ ,  $v$ )

---

<sup>11</sup>Finora l'intercetta, o costante, non era stata esplicitata. In questo caso, è stato deciso di inserirla nel modello per agevolare la sua capacità predittiva, dal momento che la sua interpretazione non è di per sé molto intuitiva (in quanto corrisponderebbe al *log-ratio* della singola classe politica rispetto al non voto, lasciando pari a zero il valore delle altre covariate).

che permetteranno di individuare la significatività delle relazioni tra variabili dai valori e dai test su  $\beta_j$ . Dai parametri stimati  $\Sigma$  e  $v$  potremo anche analizzare la correlazione dei residui delle differenti equazioni.

Alcune precisazioni.

Dal momento che le variabili dipendenti  $Y_{ij}$  hanno subito una trasformazione rispetto alle proporzioni iniziali  $V_{ij}$ , l'interpretazione dei coefficienti stimati  $\hat{\beta}_j$  potrebbe non essere di immediata chiarezza. Infatti, un valore positivo (negativo) di uno specifico coefficiente  $\hat{\beta}$  indicherebbe una relazione positiva (negativa) tra la covariata ed il *log-ratio*. Tutto ciò potrebbe tradursi come: una variazione positiva della covariata aumenta la propensione degli elettori a votare per quella specifica classe politica rispetto alla decisione di non votare nel caso in cui  $\hat{\beta} > 0$  ed è statisticamente significativo. In questo modo, il valore del coefficiente stimato potrà dare un'indicazione sull'importanza della covariata rispetto alle altre nel determinare la propensione a votare una certa classe politica piuttosto che non votare. Ma sarà molto utile anche nel definire quale sarà la classe politica a beneficiare di più (o di meno) di una certa situazione socio-economica o demografica, rappresentata proprio dai regressori del modello.

Per quanto riguarda la correlazione tra i residui di equazioni differenti (anche detta *cross-correlazione*), la sua presenza può indicare una relazione complessa tra le variabili dipendenti che il modello non riesce a catturare completamente. Per questo, una correlazione positiva si associa ad una maggiore vicinanza tra le due classi politiche in tutti quei fattori non considerati all'interno modello (e che per questo finiscono nei residui). Viceversa, una correlazione negativa suggerisce una risposta opposta di due classi politiche rispetto ai fenomeni non osservati.

### 3.3.1 Interpretazione dei coefficienti del modello

Le tabelle 3.1 e 3.2 mostrano i coefficienti stimati ( $\alpha$  e  $\beta$ ) dal modello statistico rispettivamente per le elezioni del 2018 e del 2022.

Tabella 3.1: Coefficienti stimati per le elezioni 2018.

	LEGA	FI	FDI	PD	M5S	ALTRI
Intercetta	-0,71 (0,1720)	-1,09 *** (0,0079)	-2,27 *** (2,52e-5)	-5,11 *** (2,84e-31)	-3,34 *** (2,70e-22)	-0,33 (0,4099)
Maschi	-2,04 *** (5,33e-13)	-0,89 *** (2,19e-5)	-2,18 *** (2,91e-15)	-2,61 *** (1,68e-28)	-0,33 * (0,0683)	-2,69 *** (2,38e-39)
Stranieri	1,10 *** (4,79e-18)	0,15 (0,1223)	0,56 *** (0,0001)	1,20 *** (6,08e-27)	-0,97 *** (1,09e-29)	-0,16 (0,1207)
log (Reddito)	0,26 *** (1,12e-8)	0,07 ** (0,0491)	0,19 *** (0,0001)	0,46 *** (2,00e-29)	0,34 *** (4,00e-27)	0,05 (0,1877)
log (Pop.)	0,00 (0,9087)	0,07 ** (0,0132)	0,00 (0,9843)	0,32 *** (2,21e-29)	0,17 *** (6,10e-13)	0,02 (0,5331)
log (Pop.) <sup>2</sup>	0,00 ** (0,0371)	-0,01 *** (0,0008)	0,00 (0,1927)	-0,02 *** (2,89e-29)	-0,01 *** (5,61e-13)	0,00 * (0,0773)
Over60	-1,92 *** (2,84e-81)	-1,76 *** (4,03e-111)	-1,63 *** (3,20e-52)	0,63 *** (1,77e-13)	-1,87 *** (9,89e-166)	-0,62 *** (6,06e-16)
Laureati	-2,42 *** (5,67e-38)	-0,34 ** (0,0233)	1,89 *** (4,68e-20)	1,65 *** (1,04e-22)	0,34 *** (0,0080)	4,00 *** (1,54e-160)
Disoccupazione	-4,34 *** (2,18e-237)	0,44 *** (9,66e-5)	-1,35 *** (1,80e-19)	-2,83 *** (7,02e-119)	-0,30 *** (0,0023)	-2,18 *** (4,05e-088)
Nord ovest	0,20 *** (1,66e-27)	0,17 *** (5,41e-36)	-0,29 *** (1,07e-52)	-0,33 *** (8,48e-116)	-0,40 *** (9,48e-283)	-0,14 *** (9,18e-24)
Centro	-0,95 *** (0,0000)	0,03 ** (0,0244)	-0,50 *** (2,79e-117)	-0,32 *** (4,50e-75)	0,03 ** (0,0304)	-0,05 *** (0,0012)
Isole	-1,01 *** (0,0000)	-0,16 *** (5,15e-15)	-0,58 *** (3,28e-89)	-0,51 *** (7,36e-106)	-0,10 *** (1,31e-06)	-0,22 *** (4,93e-26)
Sud	0,27 *** (2,74e-45)	-0,05 *** (7,08e-5)	-0,22 *** (6,50e-27)	-0,34 *** (1,34e-117)	-0,27 *** (6,44e-108)	0,03 * (0,0578)
Borgo	0,10 *** (5,29e-5)	0,00 (0,8823)	0,10 *** (0,0018)	-0,04 (0,1121)	-0,04 ** (0,0342)	0,04 (0,1192)
Zona rurale	0,04 (0,1872)	0,00 (0,9947)	0,08 (0,0260)	-0,08 *** (0,0098)	-0,05 ** (0,0253)	0,02 (0,5091)
Montagna	0,01 (0,2236)	-0,10 *** (4,95e-32)	-0,06 *** (3,16e-8)	-0,08 *** (4,94e-16)	-0,16 *** (1,45e-112)	-0,01 (0,1217)
Costa	-0,06 *** (2,55e-6)	0,01 (0,2881)	-0,03 * (0,0538)	-0,11 *** (7,72e-19)	0,00 (0,7493)	-0,08 *** (1,02e-11)

Il *p-value* viene riportato tra parentesi.

\*\*\*, \*\* e \* segnalano rispettivamente un *p-value* inferiore a 1%, compreso tra 1% e 5%, compreso tra 5% e 10%.

Tabella 3.2: Coefficienti stimati per le elezioni 2022.

	LEGA	FI	FDI	PD	M5S	AZIV	ALTRI
Intercetta	-4,73 *** (2,00e-21)	-3,68 *** (2,81e-15)	-4,24 *** (1,71e-28)	-8,81 *** (6,24e-93)	-8,27 *** (6,86e-76)	-1,04 *** (1,43e-96)	-2,57 *** (2,87e-11)
Maschi	-0,69 *** (0,0072)	-0,03 (0,9132)	-0,97 *** (4,04e-7)	-2,32 *** (7,21e-23)	0,27 (0,2803)	-1,70 *** (1,71e-11)	-1,23 *** (4,48e-10)
Stranieri	0,91 *** (3,68e-16)	0,68 *** (4,23e-9)	0,75 *** (2,18e-18)	1,66 *** (1,75e-57)	-0,02 (0,8148)	1,13 *** (1,48e-22)	-0,21 ** (0,0275)
log (Reddito)	0,44 *** (9,56e-23)	0,18 *** (1,38e-5)	0,49 *** (1,07e-44)	0,73 *** (1,60e-77)	0,57 *** (1,31e-43)	0,83 *** (4,21e-74)	0,18 *** (2,85e-7)
log (Pop)	-0,01 (0,7368)	0,09 *** (0,0026)	0,00 (0,9843)	0,21 *** (5,58e-13)	0,17 *** (5,51e-10)	0,16 *** (6,89e-7)	-0,02 (0,4897)
log (Pop) <sup>2</sup>	0,00 (0,1181)	-0,01 *** (7,06e-5)	0,00 ** (0,0405)	-0,01 *** (5,06e-12)	-0,01 *** (1,25e-6)	-0,01 *** (2,14e-6)	0,00 (0,3863)
Over60	-1,02 *** (9,95e-27)	-1,11 *** (5,55e-34)	-0,84 *** (1,08e-29)	1,79 *** (8,96e-98)	-0,75 *** (3,52e-17)	-0,05 (0,6085)	-0,69 *** (1,30e-18)
Laureati	-2,91 *** (9,95e-78)	-0,70 *** (4,55e-6)	-0,74 *** (3,28e-9)	2,74 *** (5,24e-79)	1,41 *** (9,66e-21)	4,02 *** (8,02e-134)	4,24 *** (1,85e-232)
Disoccupazione	-4,47 *** (8,40e-103)	-1,53 *** (4,42e-14)	-2,88 *** (1,33e-69)	-4,07 *** (7,77e-97)	2,27 *** (2,18e-29)	-4,60 *** (3,18e-116)	-2,97 *** (2,47e-80)
Nord ovest	0,38 *** (9,07e-114)	0,15 *** (1,22e-22)	-0,16 *** (1,23e-35)	-0,28 *** (2,14e-85)	-0,51 *** (2,42e-236)	0,03 * (0,0732)	0,02 (0,1491)
Centro	-0,46 *** (4,99e-144)	0,14 *** (1,24e-14)	-0,59 *** (0,0000)	-0,24 *** (2,62e-44)	0,27 *** (6,79e-42)	-0,44 *** (3,89e-133)	-0,45 *** (1,95e-202)
Isole	-0,46 *** (6,36e-72)	-0,01 (0,5568)	-0,58 *** (7,03e-183)	-0,21 *** (4,90e-18)	0,06 ** (0,0276)	-0,37 *** (1,53e-48)	0,08 *** (6,64e-6)
Sud	0,36 *** (2,54e-98)	-0,04 *** (0,0045)	-0,06 *** (1,59e-6)	-0,23 *** (5,61e-64)	-0,65 *** (0,0000)	-0,01 (0,3921)	0,19 *** (6,84e-43)
Borgo	0,08 *** (0,0003)	0,04 * (0,0535)	0,12 *** (1,02e-9)	0,00 (0,9632)	-0,07 *** (0,0039)	0,03 (0,4153)	0,02 (0,4506)
Zona rurale	0,00 (0,9127)	0,05 * (0,0589)	0,10 *** (6,11e-6)	-0,01 (0,7648)	-0,06 ** (0,0319)	-0,04 (0,1875)	-0,03 (0,2143)
Montagna	0,02 * (0,0942)	-0,12 *** (8,76e-33)	-0,06 *** (8,26e-14)	-0,08 *** (1,84e-15)	-0,20 *** (1,12e-96)	-0,02 ** (0,0423)	0,02 * (0,0607)
Costa	-0,09 *** (2,06e-11)	-0,03 *** (0,0070)	-0,04 *** (0,0005)	-0,07 *** (2,75e-9)	-0,02 (0,2605)	-0,12 *** (1,81e-16)	0,02 ** (0,0476)

Il *p-value* viene riportato tra parentesi.

\*\*\*, \*\* e \* segnalano rispettivamente un *p-value* inferiore a 1%, compreso tra 1% e 5%, compreso tra 5% e 10%.

Iniziando con la variabile "quota di votanti maschi sul totale", troviamo un primo aspetto interessante nella variazione dal 2018 al 2022 per FdI. Nonostante il partito di Giorgia Meloni fosse l'unico tra i maggiori partiti ad essere guidato da una leader donna, il coefficiente cala nel 2022, sia confrontato alla stima precedente che alle altre classi politiche nello stesso anno. Considerato il primato del partito in termini di voti ottenuti a livello nazionale, il calo del coefficiente da -2,18 a -0,97 è un indicatore dell'interesse generalizzato attorno a FdI nel 2022. In altre parole, nonostante FdI fosse l'unico tra i principali partiti ad avere una leader donna, non si registra una distinzione marcata nel sesso dei suoi votanti.

Proseguendo con le altre covariate, si può notare come la maggiore presenza di stranieri a livello comunale determini nel 2018 una propensione favorevole a votare i partiti della coalizione di centro-destra (soprattutto Lega con +1,10) e PD (+1,20), rispetto al non voto. Situazione opposta per il M5S, poiché il modello stima un coefficiente negativo (-0,97) e statisticamente significativo. Risultati più o meno analoghi nel 2022, confermando un impatto positivo sul voto per la coalizione (Lega +0,91; FI +0,68; FdI +0,75) e per il PD (+1,66). Nel 2022, coefficiente positivo anche per Azione - Italia Viva (+1,13).

Come detto in precedenza nel dettaglio delle covariate (paragrafo 3.2), il coefficiente associato al logaritmo del reddito medio da lavoro dipendente può essere interpretato nei termini di variazione percentuale.<sup>12</sup>

Concretamente (e paradossalmente) il modello stima dei coefficienti più alti per le classi di centro e centro-sinistra sia per il 2018 (PD +0,46 e M5S +0,34 contro +0,26 di Lega, +0,07 di FI e +0,19 di FdI) che per il 2022 (+0,73 per il PD, 0,57

---

<sup>12</sup>Grazie alla derivata del logaritmo,  $\beta_j = \frac{d \log(y)}{d \log(x)} = \frac{dy}{dx} \frac{x}{y} \simeq \frac{\Delta y / y}{\Delta x / x}$ , dove  $y = \frac{V_{ij}}{V_{iJ}}$  e  $x$  corrisponde al reddito medio per lavoro dipendente. Ad esempio, se  $\beta_j = 0,5$ , allora una variazione di +1% del reddito medio porterebbe, secondo la stima del modello, ad un +0,5% a favore della proporzione di voto tra il partito  $j$ -esimo rispetto al non voto.



per il M5S e +0,83 per AZIV, maggiori del +0,44 di Lega, +0,18 di FI e +0,49 di FdI). Il significato dietro a questi numeri potrebbe essere: lavoratori dipendenti con redditi dichiarati più elevati (come dirigenti, quadri o impiegati specializzati) sarebbero più inclini a votare per il centro-sinistra; lo stesso potrebbe invece non essere valido per i subordinati con redditi più bassi, come gli operai o la manodopera meno specializzata.

Anche la variabile "popolazione residente" è espressa in forma logaritmica: vale quindi la stessa interpretazione a variazioni relative adottata per il reddito medio. Tuttavia, aggiungendo un termine non lineare ( $\log(p)^2$ ), l'effetto marginale della covariata sulle variabili dipendenti sarà funzione sia del coefficiente che del regressore stesso.<sup>13</sup> In altre parole, potenzialmente ogni osservazione può avere un effetto stimato differente dalle altre, proprio perché è funzione del dato osservato. La presenza di coefficienti statisticamente significativi sulla base delle osservazioni disponibili aiuterà il modello a migliorare la capacità predittiva dei voti attesi, a maggior ragione data l'ampia eterogeneità nella dimensione dei comuni italiani. Tuttavia, in questa prima fase di interpretazione vale la pena sottolineare che per le classi Lega, FdI ed altri partiti, entrambi i coefficienti non risultano statisticamente diversi da zero. Questo significa che la preferenza per tali classi rispetto al non voto, a differenza delle altre, non è influenzata dal numero di abitanti di certo comune. Escludendo per un attimo la classe degli altri partiti, molto variabile al suo interno, si potrebbe quindi osservare come Lega e FdI ottengano preferenze indipendentemente dal numero di abitanti del comune.

La variabile "quota di persone con più di 60 anni sul totale" ha coefficienti significativi per tutte le classi e per entrambe le elezioni (unica eccezione Azione - Italia

---

<sup>13</sup>Formalmente,  $\frac{dy_{ij}}{dx_{ij}} = \beta_{4j} + 2\beta_{5j}x_{ij}$ , con  $\beta_{4j}$  e  $\beta_{5j}$  riferiti rispettivamente ai coefficienti di  $\log(p)$  e di  $\log(p)^2$  nella  $j$ -esima equazione.

Viva nel 2022). La proporzione di over 60 nei comuni italiani negli anni considerati è elevata ed in crescita, mediamente superiore al 30%.

Dai valori stimati, l'unica classe che fa eccezione con  $\beta > 0$  è il PD (+0.63 nel 2018 e +1.79 nel 2022). Il dato conferma sostanzialmente le analisi condotte da vari istituti di ricerca a partire dai propri sondaggi,<sup>14</sup> sottolineando l'orientamento a sinistra della popolazione più anziana. Il tutto valido sia nel 2018 che nel 2022, nonostante i vari tentativi della coalizione di centro-destra di introdurre misure *ad hoc* nel proprio programma condiviso (ad esempio, la proposta di FI per aumentare la pensione minima a 1.000 euro su 13 mensilità inserita all'interno del programma del centro-destra nel 2018). Semplificando, questo risultato ci potrebbe indicare che, all'aumentare del numero di persone con più di sessant'anni, aumenta la tendenza a non votare. Nonostante questo, l'unico partito che sembra beneficiarne è il PD.

Il dato sui laureati (o con titolo di studio terziario) segna una netta demarcazione tra la coalizione di centro-destra ed il resto dei partiti. Tutti i coefficienti di ogni elezione sono statisticamente significativi e di segno negativo soltanto per Lega, FI e FdI. Ciò suggerisce che, indipendentemente dallo stato lavorativo, dall'età e dal sesso dei votanti, l'interesse dei laureati è spostato a favore del centro-sinistra, dei partiti minori e di quelli locali.

Il tasso di disoccupazione è un'altra variabile che merita un approfondimento, per vari motivi. Innanzitutto, ogni stima in entrambe le elezioni è significativamente diversa da zero. Poi, possiamo osservare come, in valore assoluto, la magnitudo dei coefficienti sia più elevata in confronto alle altre variabili (anche rispetto a quelle con pari scala di grandezza, ad esempio le quote di laureati o di stranieri). La valorizzazione particolarmente elevata dei coefficienti relativi al tasso di disoccupazione

---

<sup>14</sup> Esempi di studi basati su sondaggi che trattano il tema: [Ipsos-2018](#), [CISE-2018](#) e [SWG-2022](#).

potrebbe essere motivata dall'ampia percentuale di disoccupati che ha deciso di astenersi al voto (conferma nei sondaggi riportati in nota 14). In questo modo, un denominatore più elevato nei comuni ad alta densità di disoccupati può enfatizzare ulteriormente il suo effetto.

Dai singoli valori del 2018, sembra che un'alta disoccupazione sia sfavorevole per Lega (-4,34), FdI (-1,35), PD (-2,83), altri partiti (-2,18) e, in misura minore, anche per M5S (-0,30); l'unico partito con una partecipazione più elevata da parte dei disoccupati sembra essere Forza Italia (+0,44). Due sono le possibili motivazioni. La prima attinente al leader di FI, da sempre impegnato in slogan come *"un milione di nuovi posti di lavoro per gli italiani"*<sup>15</sup>; la seconda legata alla proposta avanzata dal partito di introdurre un reddito di dignità di circa mille euro per superare la così detta "soglia di dignità".

Il dato ancora più rilevante lo troviamo nelle variazioni intercorse tra il 2018 ed il 2022: i coefficienti stimati per la seconda elezione sono tutti di segno negativo, ad eccezione del M5S (+2,27). Questo valore conferma chiaramente che una misura come il RdC, usufruita da milioni di elettori, ha sostenuto in modo consistente il voto per il M5S. Ma ci dice anche un'informazione in più: l'efficacia di una misura nell'influenzare il voto a favore di un partito è tanto maggiore nei casi in cui effettivamente tale misura è già stata attivata, a differenza delle proposte elettorali di rado attuate. Il reddito di cittadinanza rientrava nel programma del M5S già nel 2018. I coefficienti stimati, tuttavia, mostrano che nel 2018 la quota di disoccupati ha preferito favorire il partito di Silvio Berlusconi, potenzialmente per i motivi appena descritti (slogan storici del leader e reddito di dignità). La situazione cambia completamente nel 2022, dal momento in cui il M5S è riuscito effettivamente ad attuare una misura di politica attiva a contrasto della disoccupazione come il RdC (introdotto nel 2019). Nel frattempo, milioni di elettori ne hanno beneficiato, il M5S

---

<sup>15</sup>Link alla [fonte](#).

ne ha proposto un rafforzamento (come accennato nel dettaglio dei programmi nel sottoparagrafo 2.1.2) e gli altri partiti minacciano la sua revisione a ribasso (Azione - Italia Viva, PD e Lega) o la sua abolizione (FdI).

Per concludere l'analisi dei coefficienti, rimangono le variabili di controllo per i fattori territoriali e geografici.

La maggioranza dei coefficienti legati alle *dummy* nord-ovest, centro, isole e sud hanno segno negativo.<sup>16</sup> Di per sé, sembra contrastare con l'intuizione delle mappe fornite in 2.3.1, dove appaiono in modo chiaro le preferenze territoriali per le varie classi politiche nelle due elezioni. Tuttavia, le mappe non tengono conto di un elemento fondamentale che determina buona parte dei valori negativi dei coefficienti: la quota di elettori non votanti, che è la classe di riferimento al denominatore di ogni equazione del modello. Se mettessimo a confronto la quota di voti ottenuti da ogni classe rispetto alla percentuale di non votanti per singola zona geografica, scopriamo effettivamente che soltanto raramente i partiti riescono a superare il non voto sul totale degli elettori. Ad esempio, è vero che FdI ha ottenuto il primato di voti nazionali nel 2022 (26,22%) con un buono scarto rispetto al secondo partito (PD: 18,95%). Ma è altrettanto vero che la quota di non votanti è stata del 41,18%, e che in ognuna delle cinque macro-aree i non votanti hanno superato i voti a favore di FdI. Va poi fatta un'altra osservazione: le quattro aree inserite come variabili dicotomiche nel modello vanno relazionate anche rispetto al dato del nord-est, utilizzato come ulteriore parametro per il problema delle *dummy*.

Le due successive variabili dicotomiche sono "borgo" e "zona rurale". Messe in relazione rispetto ad una terza categoria non inserita nel modello (città), suggeriscono se ci sono delle preferenze specifiche nei piccoli sobborghi o nelle zone scarsamente popolate rispetto alle grandi città ad alta densità di popolazione. Ad esempio,

---

<sup>16</sup>Come detto in 3.2, si ricorda l'assenza dell'altra classe di aggregazione territoriale *NUTS1*, il nord-est, appunto per evitare la "trappola delle *dummy*".

sembra che in entrambe le elezioni il M5S non sia riuscito ad ottenere un interesse particolare dalle zone a bassa densità rispetto agli altri partiti ed ai propri risultati nelle grandi città. Diverso il caso di FdI nel 2022, in cui conferma l'ipotesi di una crescita generalizzata del suo consenso, anche nei paesi italiani a minore densità di popolazione.

Le ultime due *dummy* sono "costa" e "montagna": l'idea è quella di vedere eventuali variazioni nel voto legate a fattori turistici e logistici (ad esempio, la minore accessibilità o disponibilità di infrastrutture e servizi nelle zone montane), aiutando il modello in fase predittiva. Anche in questo caso, valori negativi per i coefficienti in "montagna" e "costa" possono essere motivati da un maggiore astensionismo degli elettori registrato in quelle zone (rispettivamente del 34% e 36% in media nel 2018 contro il 32% nazionale, salito anche a 42% e 46% nel 2022, rispetto al 41% nazionale).

### 3.3.2 Analisi delle correlazioni dei residui

La correlazione tra i residui delle diverse equazioni stimate (*cross-correlazione*) segnala una possibile relazione lineare tra le variabili dipendenti che tuttavia resta parzialmente nascosta al modello. Quindi, i residui di ogni equazione, che dipendono necessariamente dalle covariate scelte, possono fornire un'utile indicazione in più. Semplificando: se due equazioni hanno residui tra loro correlati significa che, nonostante l'ipotesi di linearità della media condizionata delle  $Y_{ij}$  pari a  $\mu_{ij} = X_{ij}\beta_j$ , ci sono dei fattori non osservati che portano il modello a "sbagliare" nella stessa direzione, sia essa negativa o positiva.

Da questa informazione possiamo estrapolare le vicinanze di due classi politiche rispetto alle preferenze di voto dei votanti, indipendentemente dall'orientamento effettivo dei singoli partiti.

Le due tabelle 3.3 e 3.4 contengono i dati sulle *cross-correlazioni* dei residui.

Tabella 3.3: Correlazioni dei residui per elezioni 2018

	Lega	FI	FDI	PD	M5S	Altri
Lega	1	0,41	0,42	0,17	0,03	0,11
FI	0,41	1	0,38	0,06	0,12	-0,01
FDI	0,42	0,38	1	0,08	0,05	0,10
PD	0,17	0,06	0,08	1	0,30	0,32
M5S	0,03	0,12	0,05	0,30	1	0,21
Altri	0,11	-0,01	0,10	0,32	0,21	1

Tabella 3.4: Correlazioni dei residui per elezioni 2022

	Lega	FI	FDI	PD	M5S	AZIV	Altri
Lega	1	0,36	0,56	-0,02	-0,18	0,26	0,02
FI	0,36	1	0,44	-0,06	0,01	0,24	0,00
FDI	0,56	0,44	1	0,06	-0,10	0,33	0,04
PD	-0,02	-0,06	0,06	1	0,37	0,25	0,29
M5S	-0,18	0,01	-0,10	0,37	1	0,02	0,23
AZIV	0,26	0,24	0,33	0,25	0,02	1	0,22
Altri	0,02	0,00	0,04	0,29	0,23	0,22	1

Ci sono delle evidenti correlazioni positive e di maggiore magnitudo tra i partiti della coalizione di centro-destra in entrambe le elezioni. È una situazione attendibile: indica che il sostegno a Lega, FI e FdI si modifica quasi linearmente rispetto alle altre variabili non osservate dal modello.

Guardando al dato del M5S nel 2018 si può notare la bassa correlazione rispetto alle altre classi individuate, con una lieve vicinanza al PD (+0,30) e alla classe degli altri partiti (+0,21). Ricordiamo che nel 2018 la formazione del primo governo

Conte vide l'alleanza tra Lega e M5S. L'analisi delle correlazioni ci segnala però una chiara divergenza tra la risposta reale dei due partiti, culminata nell'alleanza di governo, e le preferenze degli elettori. Sembra infatti che non ci sia alcuna relazione tra l'andamento dei voti della Lega con quelli del M5S in risposta a fattori non noti nel 2018 (+0.03). Questi dati confermerebbero nella pratica la posizione del M5S come partito esterno agli orientamenti tradizionali di destra o di sinistra. La situazione sembra invece cambiare nel 2022: il M5S appare più spesso come alternativa agli elettori di sinistra (correlazione con residui PD/M5S pari a +0,37) e in opposizione ai voti del centro-destra (-0,18 e -0.10 con Lega e FdI).

Altro dato interessante quello di Azione - Italia Viva nel 2022: il Terzo Polo, definito dai suoi stessi leader Renzi e Calenda come un'alternativa alla destra ed alla sinistra, in realtà dai fattori non osservati del modello appare completamente distaccato dal M5S e più vicino proprio alle due coalizioni.

### 3.3.3 Analisi predittiva e controfattuale degli esiti elettorali

Dopo aver analizzato i coefficienti stimati dal modello e le correlazioni dei residui, resta ancora da rispondere ad alcune domande:

1. Qual è la capacità predittiva del modello?
2. Possiamo definire altri esiti elettorali in presenza di fattori sociali, economici o demografici diversi da quelli effettivamente osservati?

Il primo punto rientra all'interno dell'analisi predittiva. Essa permette di verificare le potenzialità del modello nel prevedere il voto ottenuto dai singoli partiti, a partire dai soli valori noti delle covariate.

Il secondo punto invece corrisponde all'analisi controfattuale. Con essa si cerca di valorizzare quali sarebbero stati i possibili esiti elettorali nel caso in cui una covariata avesse presentato un valore diverso da quello effettivamente realizzato.

Nel nostro caso, ad esempio, può essere interessante comprendere l'effettiva capacità predittiva del modello sulla base delle covariate scelte. A tale scopo, sono stati selezionati otto comuni: Roma, Milano, Venezia, Alba, Ancona, Loreto, Cagliari e Vasto.<sup>17</sup> Su di essi è stata svolta l'analisi predittiva a parità di covariate sia per il 2018 che per il 2022. Il metodo di calcolo ed i risultati verranno definiti nell'approfondimento "Analisi predittiva del voto atteso".

Invece, per quanto riguarda l'analisi controfattuale, osserveremo i risultati che il modello propone in presenza, ad esempio, di differenti valori del tasso di disoccupazione, o della percentuale di persone con più di sessant'anni in un certo comune. Inoltre, dal momento che i modelli del 2018 e del 2022 sono stati costruiti a partire dalle stesse covariate, potremo confrontare gli scenari controfattuali attesi per gli stessi comuni nelle due differenti elezioni. La descrizione dettagliata è rimandata in seguito in "Analisi controfattuale degli esiti elettorali".

### **Analisi predittiva del voto atteso**

Si può definire "voto atteso"  $E(V_{ij})$  del partito " $j$ " nel comune " $i$ " il valore atteso dei voti predetti condizionali ad un insieme di variabili indipendenti  $\mathbf{x}_i$ .

Il calcolo del voto atteso richiede quindi, in primo luogo, di stimare i coefficienti  $\beta$  che evidenziano la relazione tra le variabili dipendenti e le covariate (i cui valori sono stati riportati nelle tabelle 3.1 e 3.2). In seguito, moltiplicando i coefficienti ottenuti con i valori noti delle covariate, si riesce ad definire una previsione di voto per uno specifico partito.

Tuttavia, va tenuto a mente che nello studio di un processo stocastico (l'esito di voto,  $Y$  in questo caso), si è in presenza di due forme di incertezza:

---

<sup>17</sup>La scelta delle otto città suddette è motivata dalla volontà di verificare i risultati predetti dal modello per comuni grandi e ridotti sia per il nord, che per il centro e sud Italia, oltre alla Capitale ed al capoluogo lombardo.



- Incertezza nella stima, dovuta all'incertezza presente nei parametri  $\theta$  da stimare.<sup>18</sup> Poiché non si può disporre di infinite osservazioni del fenomeno oggetto di studio, non si conosceranno mai i reali valori di  $\theta$ . Tuttavia, questa forma di incertezza viene in genere valorizzata attraverso gli errori standard dei parametri stimati  $\hat{\theta}$ ;
- Incertezza fondamentale, legata appunto alla casualità del fenomeno in oggetto  $Y$ , espresso sotto forma di variabile aleatoria con una specifica funzione di probabilità.

Per considerare entrambi i tipi di incertezza nel calcolo del voto predetto  $\hat{V}_{ij}$  e del voto atteso  $\hat{E}(V_{ij})$  si può ricorrere alle simulazioni statistiche.

Le simulazioni statistiche, dette anche simulazioni stocastiche o di Monte Carlo, sono uno strumento altamente flessibili e applicabile in un'ampia gamma di discipline. Nel nostro caso, le simulazioni si concretizzano inizialmente nella generazione di una serie di numeri pseudo-casuali estratti da una specifica distribuzione di probabilità ( $f(y)$  nel caso continuo). La serie di numeri generata permette di approssimare quantità teoriche di interesse con un grado di precisione crescente al crescere del numero di simulazioni effettuate. Ad esempio, se volessi calcolare il valore atteso di una qualche variabile aleatoria  $y \sim f(y)$ , anziché calcolare l'integrale  $E(y) = \int_{-\infty}^{+\infty} yf(y) dy$ , possiamo approssimare la quantità  $E(y)$  calcolando la media di una serie di numeri estratta in modo casuale da  $f(y)$ , ovvero  $\hat{E}(y) = \sum_{k=1}^M y_k/M$  (con  $M$  pari al numero di simulazioni).

Concretamente, per ottenere una valorizzazione del voto atteso  $\hat{E}(V_{ij})$  di un certo partito "j" in un certo comune "i", è stato seguito l'approccio impiegato in

---

<sup>18</sup>Ricordiamo che nel nostro studio  $\theta = \{\beta, \Sigma, v\}$ . La stima avviene in modo tale da massimizzare la funzione di verosimiglianza 3.8.

Katz (1999) [10]:

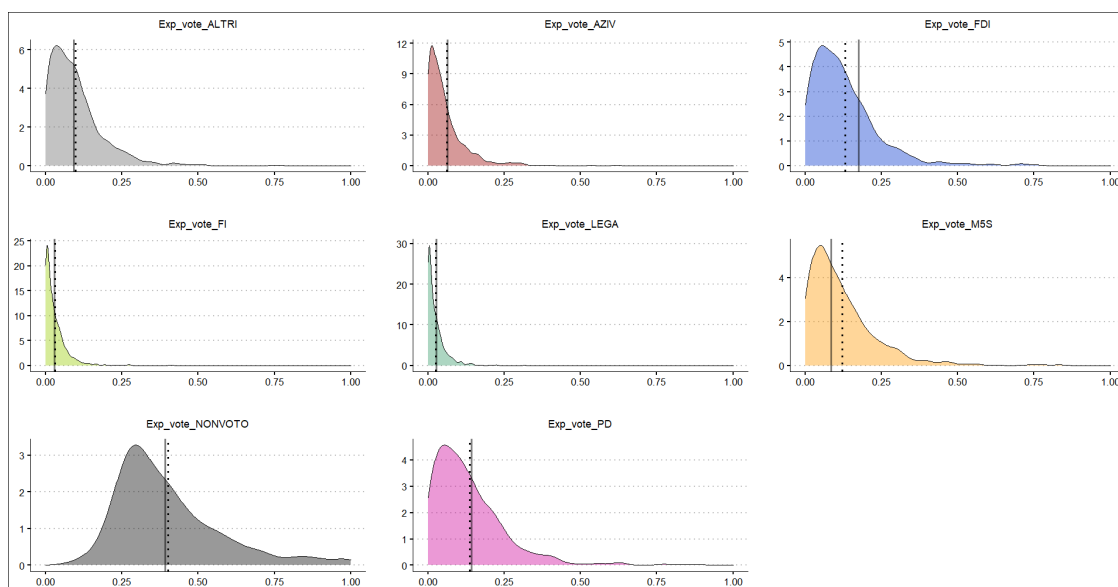
1. Massimizza la funzione di log-verosimiglianza  $l(\theta)$  sotto l'ipotesi di distribuzione  $t$  multivariata (equazione 3.8) e memorizza il vettore di parametri stimati  $\hat{\theta} = \{\hat{\beta}, \hat{\Sigma}, \hat{v}\}$  con la sua matrice di varianza-covarianza  $\hat{V}(\hat{\theta})$ ;
2. Estrai un esito casuale di  $\theta$  da una distribuzione normale multivariata con media  $\hat{\theta}$  e varianza  $\hat{V}(\hat{\theta})$ ;
3. Estrai  $n$  valori casuali di  $\hat{Y}_{ij}$  da una distribuzione  $t$  multivariata con parametri  $\hat{\theta}$  (nel nostro caso  $n = 500$  garantisce risultati "stabili" in termini di variabilità delle stime);
4. Calcola  $n$  simulazioni dei voti predetti  $\hat{V}_{ij}$  in modo deterministico applicando la trasformazione logistica multivariata (equazione 3.4);
5. Calcola il valore medio degli  $n$  voti  $\hat{V}_{ij}$  prodotti dalle simulazioni, così da ottenere una simulazione del voto atteso  $\hat{E}(V_{ij})$ . Il voto atteso della classe " $J$ " usata da riferimento viene calcolato come  $\hat{E}(V_{iJ}) = 1 - \sum_{j=1}^{J-1} \hat{E}(V_{ij})$

Ripetendo  $M$  volte le fasi da 2 a 5 si otterranno  $M$  simulazioni del voto atteso  $\hat{E}(V_{ij})$ , da cui si potrà ricavare la funzione di densità approssimata dall'istogramma delle simulazioni, una stima puntuale dalla media degli  $M$  voti attesi simulati ed il rispettivo errore standard dallo scarto quadratico medio. Con la stessa logica, l'intervallo di confidenza corrisponderà ai valori simulati nei rispettivi percentili di interesse (es: 5% e 95%). Nel caso di studio è stato fissato un valore di  $M$  pari a 1000.

Di seguito viene riportato un esempio concreto della valorizzazione dei voti attesi per tutte le classi politiche a Roma nel 2022 (figura 3.1): il grafico mostra la distribuzione generata dalle 1000 simulazioni utilizzando le stesse covariate osservate nella Capitale. All'interno della distribuzione dei voti attesi di ogni classe compaiono due linee: quella punteggiata mostra la media dei voti attesi ottenuti

dalle simulazioni, mentre quella continua rappresenta la percentuale di voti realmente ottenuti nel 2022. Oltre alla notevole aderenza dei risultati simulati agli esiti reali, le simulazioni ci forniscono anche delle stime della variabilità. La tabella 3.5 contiene i risultati della simulazione con le proporzioni di voto delle singole classi sul totale degli elettori.

Figura 3.1: Grafico delle densità dei voti attesi simulati a Roma - Elezioni 2022.



(Linea continua per la percentuale di voti reali ottenuti dalla classe, linea punteggiata per il voto atteso stimato.)

Passiamo ora alla simulazione dei risultati delle due elezioni per gli otto comuni indicati in precedenza: Roma, Milano, Venezia, Alba, Ancona, Loreto, Cagliari e Vasto.

L'eterogeneità dei comuni selezionati ci permette di verificare se il modello riesce ad approssimare l'ampia varietà dei reali esiti elettorali facendo uso soltanto delle covariate scelte per l'analisi. Le otto città infatti comprendono sia le due maggiormente popolate (Roma e Milano con circa 2,7 e 1,3 milioni di abitanti) ma anche realtà ridotte (come Alba e Loreto, con 31.000 e 13.000 abitanti), collocate in varie regioni italiane, sia lontane che vicine alla costa. Di seguito viene confrontato il

Tabella 3.5: Statistiche descrittive e voti reali a Roma - Elezioni 2022

	Voto reale	Media	Dev. Std.	5% perc.	95% perc.
LEGA	0.0263	0.0249	0.0329	0.0001	0.0869
FI	0.0276	0.0302	0.0406	0.0001	0.1008
FDI	0.1741	0.1294	0.1141	0.0068	0.3308
PD	0.1414	0.1350	0.1211	0.0064	0.357
M5S	0.0848	0.1203	0.1135	0.0039	0.3195
AZIV	0.0632	0.0616	0.0837	0.0004	0.2033
ALTRI	0.0913	0.0981	0.0892	0.0047	0.2672
NON VOTO	0.3914	0.4004	0.1743	0.2001	0.7724

voto atteso simulato con i voti reali del 2018 e del 2022, con i relativi scostamenti (in appendice B.2, tabelle B.5 e B.6).

I dati sugli scostamenti ci danno delle indicazioni importanti. Nonostante quasi la metà degli errori sia contenuto tra un  $\pm 1\%$ , ci sono alcuni risultati attesi lontani da quelli realmente verificati. Ad esempio, per Alba nel 2018 i voti attesi per la Lega sono  $+6,85\%$  rispetto alla proporzione reale. Una motivazione è legata al basso tasso di disoccupazione ( $8,22\%$ ) ed alla quota elevata di stranieri ( $11,46\%$ ) registrata nel comune rispetto alla media nazionale. Infatti, i coefficienti stimati per la Lega nel 2018 per entrambe le covariate sono significativi e di elevata magnitudo anche in confronto ai coefficienti degli altri partiti. In questo modo, il modello tende a simulare un voto atteso più alto a favore della Lega. Altra divergenza da rilevare è il caso di Cagliari nel 2022, in cui il modello simula un voto atteso molto basso per Fratelli d'Italia ( $-5,25\%$  rispetto al reale), rimpiazzato con un  $+5,53\%$  a favore del M5S. Anche in questo caso, la motivazione si riscontra nelle covariate e nei coefficienti stimati. Infatti, oltre ad essere una città dell'area geografica "Isole", presenta alla stesso tempo un tasso di disoccupazione elevato ( $14,49\%$  nel 2021) e una percentuale di residenti con diploma terziario molto alta (la stima dei censimenti Istat è circa  $27\%$  nel 2021). Sono proprio queste due covariate ad avere i

relativi coefficienti determinanti e di segno opposto per le due classi in oggetto ( $\hat{\beta}$  di "Laureati" a -0,74 e +1,41 per FDI e M5S, "Disoccupazione" a -2,88 e +2,27, "Isole" a -0.58 e +0.06).

### Analisi controfattuale degli esiti elettorali

Sempre sfruttando la flessibilità delle simulazioni statistiche, possiamo calcolare la variazione controfattuale dei risultati elettorali in presenza di diversi valori di una covariata. Il metodo impiegato è detto *first difference* e si basa sostanzialmente nella differenza tra due valori attesi:  $\hat{E}(Y_m) - \hat{E}(Y_o)$ , dove  $\hat{E}(Y_o)$  è pari al valore atteso simulato a partire dalle covariate originali (come fatto per Roma nel paragrafo precedente), mentre  $\hat{E}(Y_m)$  è il valore atteso ottenuto dalle simulazioni dopo aver modificato un regressore (a parità di tutti gli altri).<sup>19</sup>

Quindi, il metodo *first difference* non richiede altro se non il calcolo di due valori attesi, utilizzando gli stessi cinque passi nell'elenco descritto sopra per valorizzare il voto atteso  $\hat{E}(V_{ij})$ .

A scopo esplicativo, di seguito viene riportato un esempio (figura 3.2) a partire dai valori delle covariate osservati in media in Italia nel 2018, mostrando come sarebbero cambiate le proporzioni di voto secondo il modello<sup>20</sup> al variare della percentuale di persone con più di sessant'anni (tenendo costanti le altre variabili). Facendo variare la quota di over 60 dal 10% al 60%, si può notare come alcune classi non mostrano differenze sostanziali (Fratelli d'Italia, Forza Italia ed altri partiti), mentre ci siano prospettive a ribasso per Lega e M5S. L'analisi ci suggerisce poi che

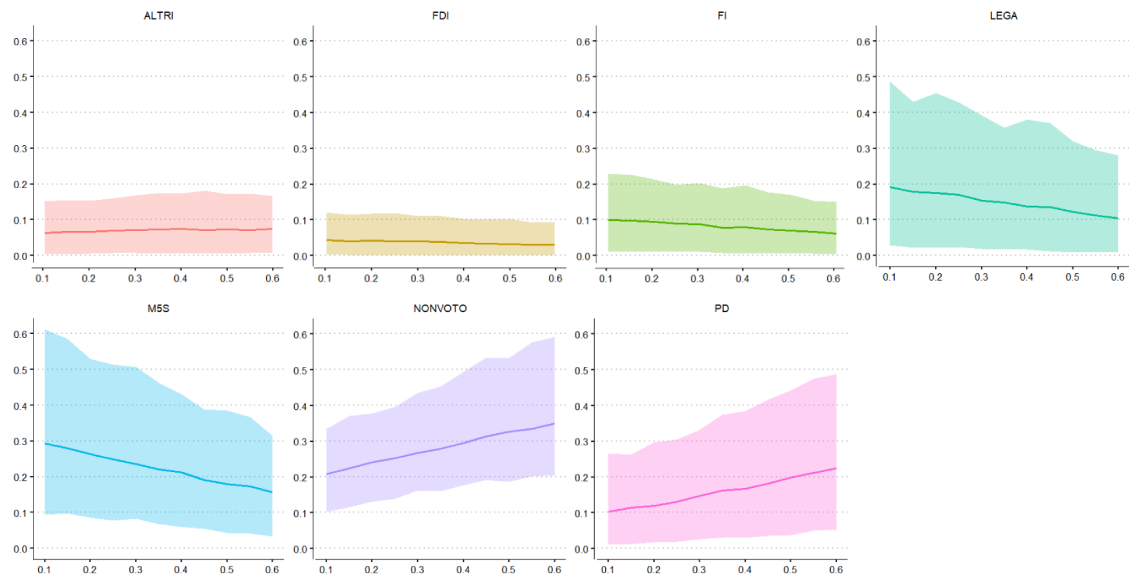
---

<sup>19</sup>Una descrizione dettagliata della tecnica *first difference* in King (2000) [11].

<sup>20</sup>Per garantire il funzionamento della simulazione a partire dai parametri stimati è stato necessario inserire anche dei valori per le variabili dummy. Quindi, nell'esempio è stato ipotizzato che le proporzioni di voto siano relative ad un borgo situato nel nord est, né in montagna e né in zona costiera.

un aumento degli over 60 in Italia nel 2018 avrebbe favorito il Partito Democratico e fatto aumentare la percentuale di non votanti.

Figura 3.2: Proporzioni di voto in media nazionale al variare degli over60 - Elezioni 2018



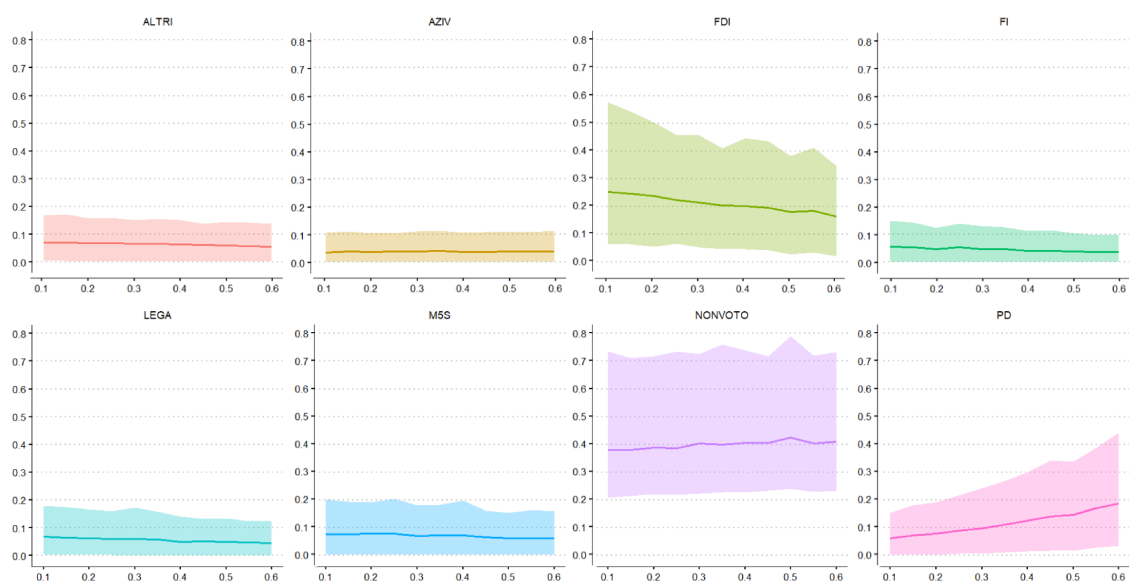
(L'asse  $y$  rappresenta la proporzione di voto della classe in corrispondenza di una diversa percentuale di over60, espressa nell'asse  $x$ .)

Confrontando lo stesso esempio usando i dati del 2022 notiamo, ad esempio, che l'effetto di una variazione degli over60 sui voti attesi di Lega e M5S tende ad appiattirsi.<sup>21</sup> Allo stesso modo, anche la quota di non votanti sembra rimanere più stabile (anche se parte da un valore medio già elevato). Viene poi mantenuta l'aspettativa di crescita per il PD, ma con una maggiore intensità. Il partito che invece sembra trarre svantaggi da una percentuale crescente di over60 è Fratelli d'Italia. Ciò potrebbe confermare quanto visto in precedenza durante l'interpretazione dei

<sup>21</sup>In questo caso la classe Azione - Italia Viva non compromette la confrontabilità dei risultati poiché, se non fosse stata inserita, la sua proporzione di voti ottenuti sarebbe confluita in "Altri partiti".

coefficienti (sottoparagrafo 3.3.1): gli over60 mostrano, in genere, una maggiore propensione a non votare; tuttavia, tra coloro che esprimono un voto, c'è la tendenza a favorire il centro-sinistra.

Figura 3.3: Andamento delle proporzioni di voto in media nazionale al variare degli over60 - Elezioni 2022



(L'asse y rappresenta la proporzione di voto della classe in corrispondenza di una diversa percentuale di over 60, espressa nell'asse x.)

Riprendiamo ora il caso delle otto città integrando un'analisi controfattuale. Proveremo infatti a simulare i possibili esiti elettorali in presenza di lievi modifiche alle covariate.<sup>22</sup>

Ad esempio, se volessimo simulare come sarebbe cambiato il voto atteso di ogni classe politica nel caso di una popolazione residente meno anziana rispetto a quella realmente registrata, basterà calcolare un nuovo voto atteso  $\tilde{E}(V)$  sottraendo per ognuno degli otto comuni un 5% in "Over 60", tenendo costanti le altre variabili. I risultati che otteniamo per il 2018 e per il 2022 sono riassunti nella tabella B.7. In

<sup>22</sup>Il dettaglio dei risultati in appendice B.2, tabelle B.7, B.8, B.9, B.10 e B.11.

essa confrontiamo il voto atteso dal modello in assenza di variazione alle covariate  $\hat{E}(V)$  e dopo una loro modifica  $\tilde{E}(V)$ .

Quello che emerge dalle simulazioni delle elezioni 2018 coincide in larga parte con le simulazioni per il 2022. Infatti, il primo partito ad essere "danneggiato" da una popolazione meno anziana, indipendentemente dalla zona geografica in cui sono collocati gli otto comuni, sarebbe il PD. I partiti che invece ne beneficerebbero variano in base alle città e tra le due elezioni, anche se i maggiori vantaggi sono attesi per il M5S e Lega. Il partito di Giorgia Meloni ottiene dei differenziali positivi principalmente nel 2022 e nei due comuni marchigiani.<sup>23</sup>

Altro aspetto in comune nelle due elezioni è l'aspettativa di una maggiore partecipazione al voto ( $\tilde{E}(V) - \hat{E}(V)$ ) sempre negativo per la classe "non voto", ad eccezione di Vasto nel 2022.

La quota di stranieri è la seconda covariata esaminata. Potremmo chiederci: come si sarebbero potuti modificare i risultati elettorali negli otto comuni, se in ognuno di essi ci fossero stati, ad esempio, il 5% in più di stranieri sul totale dei residenti?

Anche in questo caso, lasciando invariate le altre variabili, possiamo aggiungere il 5% soltanto al dato sugli stranieri per i comuni di interesse, e analizzare i nuovi voti attesi  $\tilde{E}(V)$  (tabella B.8).

Le simulazioni ipotizzano un aumento generalizzato dei voti a favore del PD, sia nel 2018 che nel 2022. Una quota maggiore di stranieri avrebbe potuto favorire anche la Lega nei tre comuni del nord Italia (Milano, Venezia e Alba), soprattutto nel 2018. Nei comuni analizzati, FdI avrebbe potuto beneficiare di maggiori voti quasi

---

<sup>23</sup>Il presidente della regione Marche in carica già da settembre 2020 è Francesco Acquaroli, appartenente a FdI. Come vedremo anche in seguito, i differenziali generati dalle simulazioni più favorevoli a FdI si trovano spesso nei due comuni marchigiani. Anche se tale risultato non è di per sé direttamente collegabile alla presidenza regionale, il tema dei possibili legami tra presidenza regionale ed esiti politici nazionali meriterebbe maggiori riflessioni in futuro.



esclusivamente nel 2022, in particolare nei due paesi marchigiani (Ancona +1,33% e Loreto +1,46%). Il principale partito ad essere sfavorito da questa situazione ipotetica sembra invece il M5S.

Un'altra domanda che possiamo porci è: di quanto sarebbero potuti variare i voti ai partiti in presenza di più persone con diploma terziario? Il diploma terziario si riferisce sia al titolo ottenuto al termine di un corso di laurea, che al certificato conferito alla conclusione di un corso di specializzazione ITS. In ogni caso, per brevità la variabile è stata definita come "laureati", e possono sorgere numerose riflessioni dietro a quei dati. Ad esempio, è ragionevole attendersi (anche dalle analisi demografiche a riguardo) che una persona che abbia proseguito con gli studi abbia minori rischi di rimanere disoccupato, o di finire al di sotto della soglia di povertà. Stessa riflessione per l'accesso a posizioni manageriali all'interno delle organizzazioni. C'è poi il tema dell'inevitabile posticipo dell'ingresso nel mondo del lavoro, con tutti gli impatti che ne conseguono. Da ciò, ci sono anche delle reazioni imprescindibili negli esiti politici. Oltre a quanto detto finora nelle sezioni precedenti in merito alla variabile "laureati", integriamo anche l'analisi controfattuale con i potenziali risultati nei soliti otto comuni in presenza di un ulteriore 5% in più di persone con diploma terziario.

Dalla tabella di sintesi B.9, quello che emerge maggiormente è il ribasso dei voti attesi dei tre partiti di centro-destra in entrambe le elezioni, a favore del PD, del Terzo Polo e della classe con gli altri partiti. Inoltre, un'altra informazione che traspare dalle simulazioni è la riduzione marcata del non voto per il 2018 ma soprattutto per il 2022.

Il reddito medio per lavoratore dipendente è un indice della ricchezza della forza lavoro subordinata. La variabile, espressa in logaritmo, è risultata statisticamente

significativa per ogni classe politica e per entrambe le elezioni, con l'unica eccezione in "Altri partiti" nel 2018. Simuliamo ora i possibili esiti nel caso in cui gli otto comuni presentino un reddito medio di circa il 5% in più rispetto a quello effettivamente osservato (tabella B.10).<sup>24</sup>.

Quello che si nota è l'aspettativa a rialzo per il PD in entrambe le elezioni ed in tutti gli otto comuni analizzati. Anche Azione - Italia Viva mostra un leggero rialzo, ma soltanto nelle tre città più grandi (Roma +0.73%, Milano + 0.42% e Venezia +0.36%). Nel centro-destra non c'è una tendenza chiara data dalla *first-difference*, tuttavia va segnalato che anche per questa variabile le simulazioni prevedono per i due comuni marchigiani un incremento nel voto di FdI nel 2022 (Ancona +1,23% e Loreto +1.32%). Infine, i nuovi voti attesi tendono a ridurre la quota di non votanti: in altre parole, le simulazioni prodotte a partire dal modello ci indicano che, nelle città esaminate, una ricchezza maggiore nelle mani della forza lavoro dipendente avrebbe potuto spingere più elettori a votare; ed il principale partito a beneficiarne (con le dovute distinzioni tra le due elezioni e tra i vari comuni) sarebbe stato il Partito Democratico.

L'ultimo regressore che analizziamo attraverso le simulazioni con il metodo *first difference* è il tasso di disoccupazione. Abbiamo visto durante l'interpretazione dei coefficienti quale fosse il peso di questa covariata all'interno del modello, sempre statisticamente significativa e di elevata magnitudo (sia positiva che negativa tra le varie classi politiche). Inoltre, nella stessa occasione era stata enfatizzata la relazione tra il tasso di disoccupazione ed i voti del M5S, in cui il relativo coefficiente

---

<sup>24</sup>Per precisione, le simulazioni sono state realizzate aumentando il logaritmo del reddito medio di +0.05. La variazione applicata (+0.05) garantisce una buona approssimazione del logaritmo con la variazione percentuale rispetto ai valori assoluti. Per fare un esempio concreto, il reddito medio osservato in Ancona nel 2018 era pari a 20.735€, da cui  $\log(20735) = 9.9396$ . La somma  $9.9396 + 0.05$  è pari a 9.9896, che riportato in valori assoluti grazie dal calcolo esponenziale  $e^{9.9896}$  corrisponde a 21.798€. L'incremento da 21.798 a 20.735 è di 1.063€, pari al 5,12% (il 5% puntuale sarebbe stato 1.037€).

stimato passa da -0.30 a +2.27 nel 2022, diventando anche il primato tra le classi analizzate. Ora riportiamo gli otto esempi concreti, provando a simulare il cambiamento nelle proporzioni di voto in presenza di un tasso di disoccupazione più alto (+5%) rispetto a quello realmente osservato antecedentemente alle due elezioni. Come si può vedere dalla tabella B.11, in entrambe le elezioni e nelle otto città considerate, le simulazioni generano delle aspettative di voto a ribasso per Lega, PD, FdI e per la classe degli altri partiti. Anche Azione - Italia Viva presenterebbe un voto atteso più basso nell'unica elezione in cui ha partecipato tra le due analizzate. Da qui la domanda: dove sarebbero confluiti questi voti in presenza di una maggiore disoccupazione? Una prima parte sarebbe andata a favore di Forza Italia, soprattutto nel 2018 per i motivi già visti in 3.3.1 (slogan storici del leader e reddito di dignità nel programma). Un'altra quota, sostanzialmente più elevata nel 2022, sarebbe stata raccolta dal M5S. L'ultima porzione di voti, invece, non sarebbe stata assegnata a nessuna classe politica. Banalmente, perché dalle simulazioni si sarebbe atteso un incremento significativo anche dei non votanti. I valori dello scostamento in presenza di un +5% nel tasso di disoccupazione osservato ci permette anche di fare confronti tra le due elezioni: ad esempio, possiamo notare come le simulazioni abbiano previsto un +4,11% ed un +2,50% a favore del M5S nei voti di Cagliari e Vasto nel 2022 (contro il +0,28% e +0,68% del 2018). Rilevante anche il +2,57% atteso per il M5S nella Capitale nel 2022 contro il +1,25% di quattro anni prima. Il confronto dei differenziali  $\tilde{E}(V) - \hat{E}(V)$  simulati nelle due elezioni conferma quanto già detto precedentemente in 3.3.1. Una misura di politica attiva come il RdC, entrata realmente in vigore senza restare una mera proposta elettorale, ha suscitato negli elettori un interesse sostanziale nei confronti del partito che l'ha ideata e resa effettiva.

Un dato interessante a riguardo ce lo fornisce l'[Osservatorio Inps sul RdC e PdC](#) attraverso il tasso di inclusione. Esso conteggia il numero di persone coinvolte dalla misura ogni 1.000 abitanti e viene calcolato a livello provinciale su base annuale.

Dal tasso di inclusione riusciamo ad individuare quali sono le province italiane con un maggior numero di beneficiari diretti (intesi come singoli richiedenti della misura) e indiretti (perché parte del nucleo familiare del richiedente); di conseguenza, ci indica anche in quali province circola un maggiore interesse. Prendendo i valori delle sette province in cui sono situati gli otto comuni analizzati, riscontriamo una certa tendenza positiva tra il tasso di inclusione del 2021 e i differenziali  $\tilde{E}(V) - \hat{E}(V)$  per il 2022 dati dalla simulazione con +5% nel tasso di disoccupazione. E ricordiamo, nel modello dal quale provengono le simulazioni non è stato inserito alcun dato specifico sul RdC. Ad esempio, la prima delle sette province per tasso di inclusione più alto è Cagliari, con 86 persone coinvolte ogni 1.000 abitanti, ed è Cagliari il comune ad avere il differenziale maggiore a favore del M5S (+4,11%). Roma è la seconda provincia per tasso di inclusione (66/1000), così come Roma è il comune con il secondo differenziale più alto (M5S +2.57). Segue poi la provincia di Chieti con 49 coinvolti su 1.000 abitanti, nella quale risiede Vasto (differenziale M5S pari a +2,50%). Abbiamo poi Milano con 44/1000 e +1,75% atteso nel comune omonimo, Ancona con 34/1000, dove invece la stessa città capoluogo registra un +2,29% simulato e Loreto un +1,02% a favore del Movimento. Troviamo poi la provincia di Cuneo con tasso di inclusione pari a 25 su 1000, nella quale risiede Alba (+0,74% simulato per il M5S) e infine la provincia di Venezia con 23 su 1000 e +1,35% atteso.

La sintesi comprensiva anche dei dati sul tasso di inclusione è racchiusa nella tabella 3.6.

Tabella 3.6: Tasso di inclusione RdC/PdC provinciale 2021 e differenziali M5S 2022.

Comune	Provincia	Diff. M5S 2022	Tasso inclusione 2021
Cagliari	CA	+4,11%	86/1000
Roma	RM	+2,57%	66/1000
Vasto	CH	+2,50%	49/1000
Milano	MI	+1,75%	44/1000
Ancona	AN	+2,29%	34/1000
Loreto	AN	+1,02%	34/1000
Alba	CN	+0,74%	25/1000
Venezia	VE	+1,35%	23/1000

"Diff. M5S 2022" corrisponde a  $\tilde{E}(V) - \hat{E}(V)$  simulato con +5% in tasso di disoccupazione.

Per concludere, le simulazioni dell'analisi controfattuale, oltre a fornirci degli esempi concreti sui possibili impatti di una variazione di una covariata rispetto al dato osservato, aggiungono anche delle indicazioni in più che la sola analisi dei coefficienti stimati non avrebbe potuto darci: dalle simulazioni possiamo infatti quantificare anche i potenziali scostamenti della classe usata come riferimento, ovvero il non voto. In questo modo, abbiamo potuto verificare in un modo di facile comprensione anche quali sono, tra le covariate considerate all'interno del modello, quelle che potrebbe spingere gli elettori a votare oppure a non farlo.

## Capitolo 4

# Conclusioni

In questo studio, abbiamo cercato di analizzare le elezioni politiche italiane del 2018 e del 2022 con dati a livello comunale in modo multidimensionale. La definizione delle classi politiche ha permesso di tenere in considerazione la frammentarietà dei vari orientamenti. La scelta delle covariate, invece, ha consentito l'integrazione di specifici fattori determinanti nell'analisi elettorale.

Il modello KK, adattato ai due casi italiani, si è dimostrato un valido strumento per stimare i possibili impatti delle covariate sulle variabili dipendenti, ma anche per calcolare e successivamente per esaminare la correlazione dei residui tra le classi politiche.

L'assunzione distributiva di  $t$  multivariata permette una maggiore flessibilità del modello nel trattare casi di comuni molto distanti dalla media nazionale.

Sfruttare la classe del "non voto" come denominatore nel rapporto dei logaritmi che definiscono le variabili dipendenti contribuisce all'interpretabilità dei risultati e aggiunge un patrimonio informativo che, altrimenti, verrebbe ignorato.

Definendo entrambi i modelli a partire dagli stessi regressori, è stato possibile confrontare le stime prodotte per il 2018 e per il 2022.

Guardando ai risultati espressi dai coefficienti e dai rispettivi test di significatività,

possiamo segnalare un'influenza non trascurabile delle covariate sulle varie equazioni del modello; un particolare apporto viene fornito dalla disoccupazione, dalla presenza di persone con oltre sessant'anni e dal reddito dei lavoratori dipendenti. È stato approfondito l'esame del tasso di disoccupazione, comparando gli esiti delle due elezioni: è stato notato un forte spostamento a favore del M5S nel 2022, motivabile principalmente dall'introduzione del RdC.

La matrice delle correlazioni dei residui delle equazioni stimate apporta contribuisce con altre indicazioni: oltre alla chiara vicinanza dei partiti all'interno della coalizione di centro-destra, nel 2022 il M5S sembra presentarsi in modo più chiaro come alternativa ai votanti del PD o degli altri partiti minori, e il Terzo Polo non risulta così "esterno" ai votanti dei partiti tradizionalmente a destra ed a sinistra.

L'analisi predittiva e controfattuale, invece, alleggeriscono l'interpretabilità dei risultati del modello grazie alle simulazioni statistiche. Prendendo i casi concreti di otto comuni italiani, ognuno con una propria peculiarità territoriale o dimensionale, abbiamo ipotizzato altri possibili esiti elettorali al mutare di specifiche covariate. Concentrandoci nuovamente sul tasso di disoccupazione, abbiamo simulato nuovi risultati elettorali e abbiamo confrontato la tendenza positiva tra l'aspettativa di voto espressa dal modello per il M5S con il tasso di inclusione del RdC.

L'analisi, ovviamente, è soltanto una visione parziale e semplificata della realtà. Abbiamo compreso l'importanza dei social network nelle strategie di comunicazione dei vari partiti. Tuttavia, almeno in questa sede, non è stato possibile stimare i possibili impatti negli esiti di voto.

Sfruttando la flessibilità del modello KK, sempre partendo da dati comunali, sarebbe interessante in futuro provare ed integrare altre variabili determinanti nello studio. Oltre al dato sulle inserzioni, meriterebbe un'integrazione con le spese elettorali sostenute dai partiti, oppure l'inserimento di altre covariate legate ad altre misure pubbliche specifiche e ad ampia portata.

# Appendice A

## Statistiche descrittive

### A.1 Mappe dei risultati elettorali

Figura A.1: Proporzione dei voti per comune - Lega

(a) Anno: 2018

(b) Anno: 2022

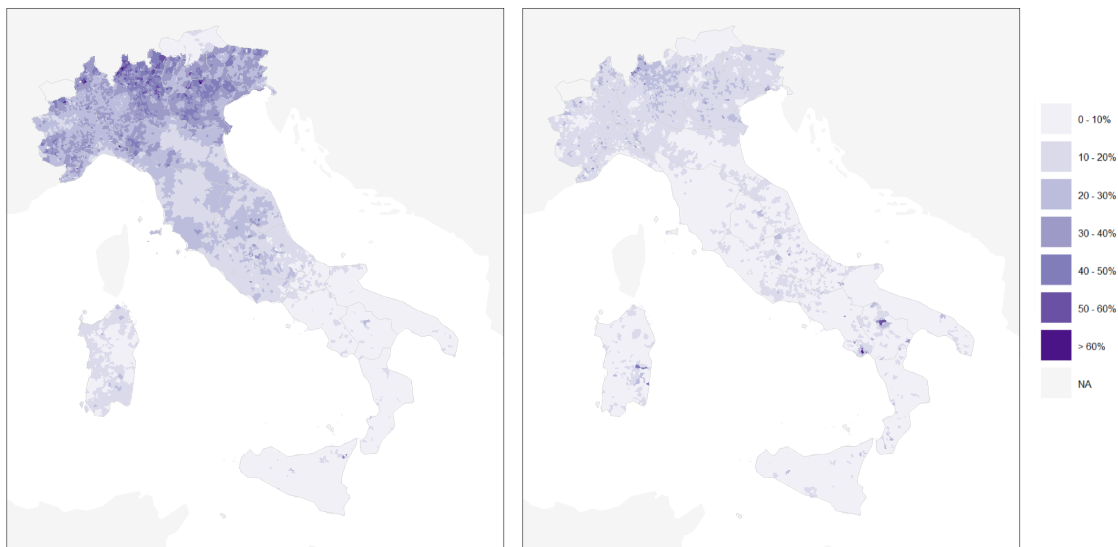




Figura A.2: Proporzion dei voti per comune - Forza Italia (FI)

(a) Anno: 2018

(b) Anno: 2022

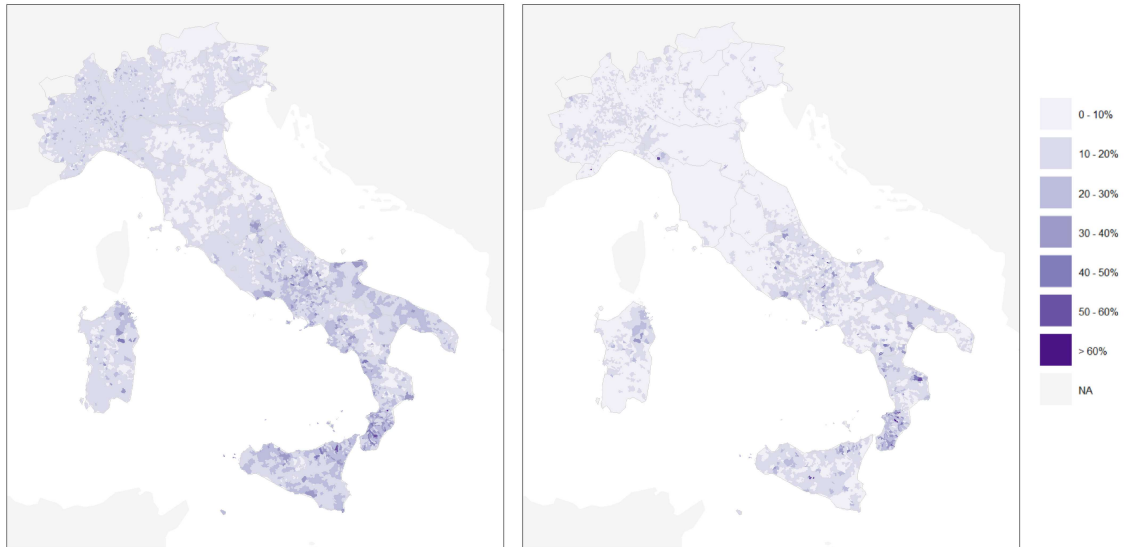


Figura A.3: Proporzion dei voti per comune - Fratelli d'Italia (FDI)

(a) Anno: 2018

(b) Anno: 2022

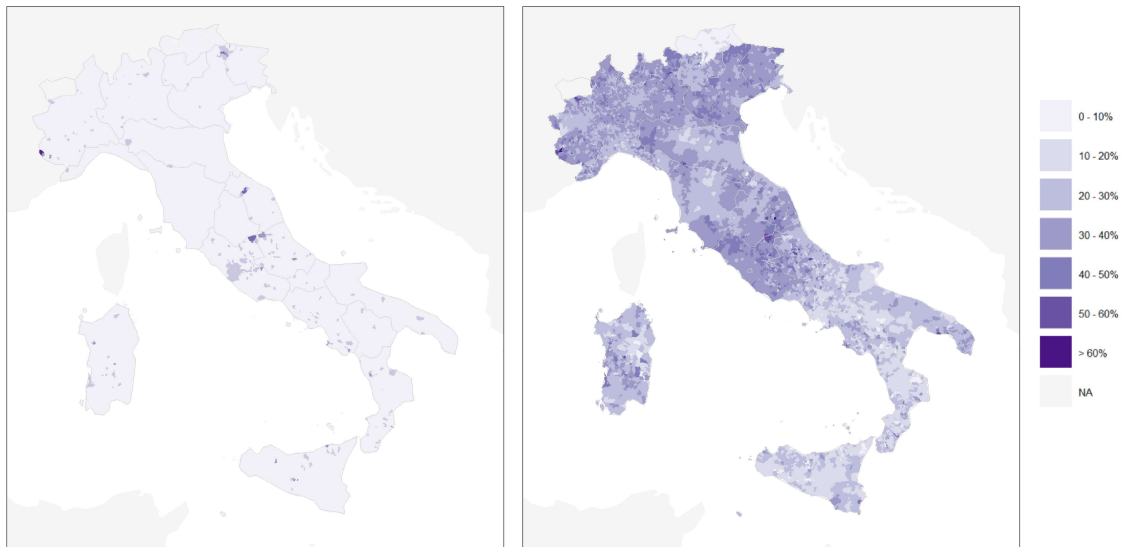


Figura A.4: Proporzioe dei voti per comune - Partito Democratico (PD)

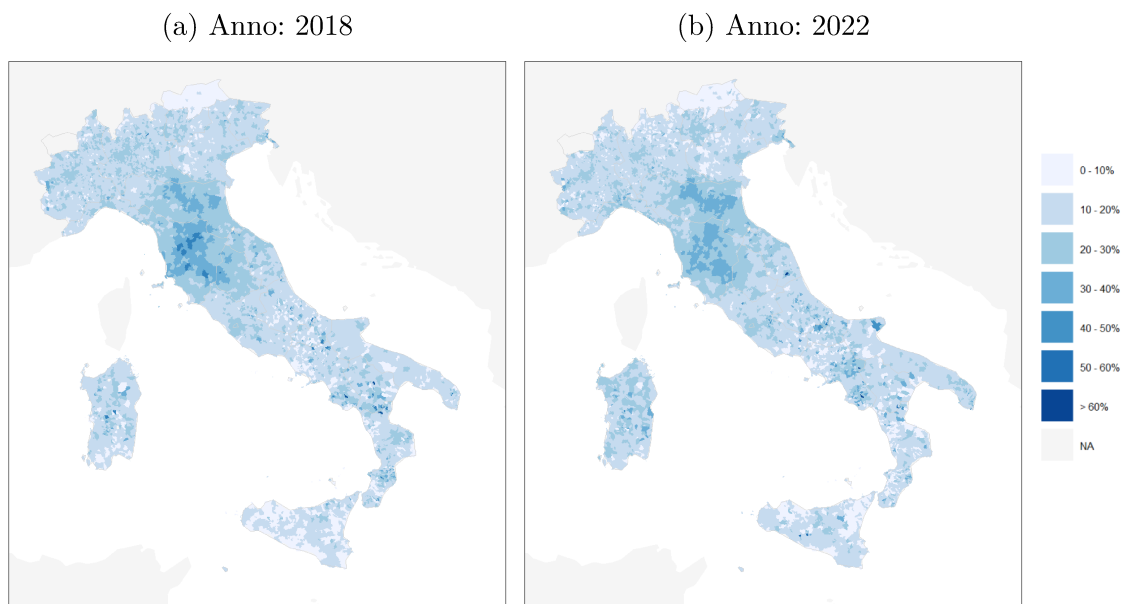


Figura A.5: Proporzioe dei voti per comune - Movimento 5 Stelle (M5S)

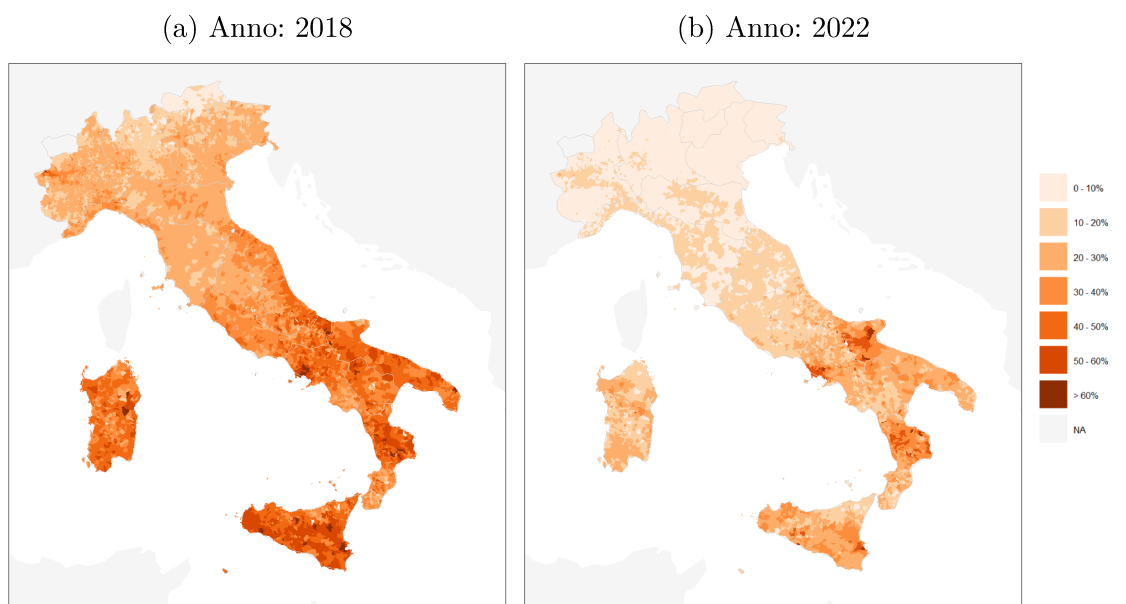


Figura A.6: Proporzioe dei voti per comune - Partiti locali

(a) Anno: 2018

(b) Anno: 2022

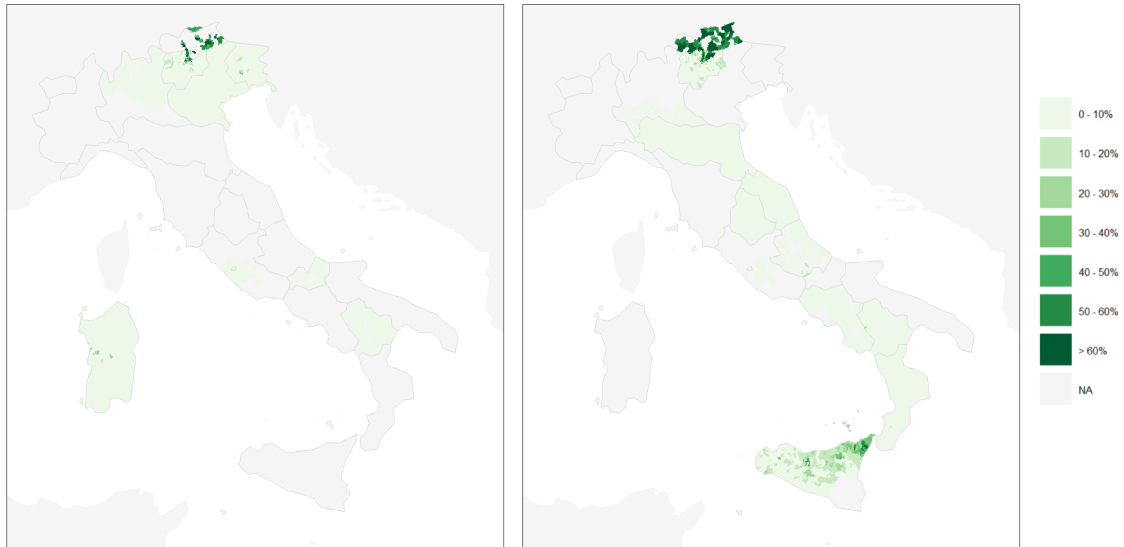


Figura A.7: Proporzioe dei voti per comune - Altri partiti

(a) Anno: 2018

(b) Anno: 2022

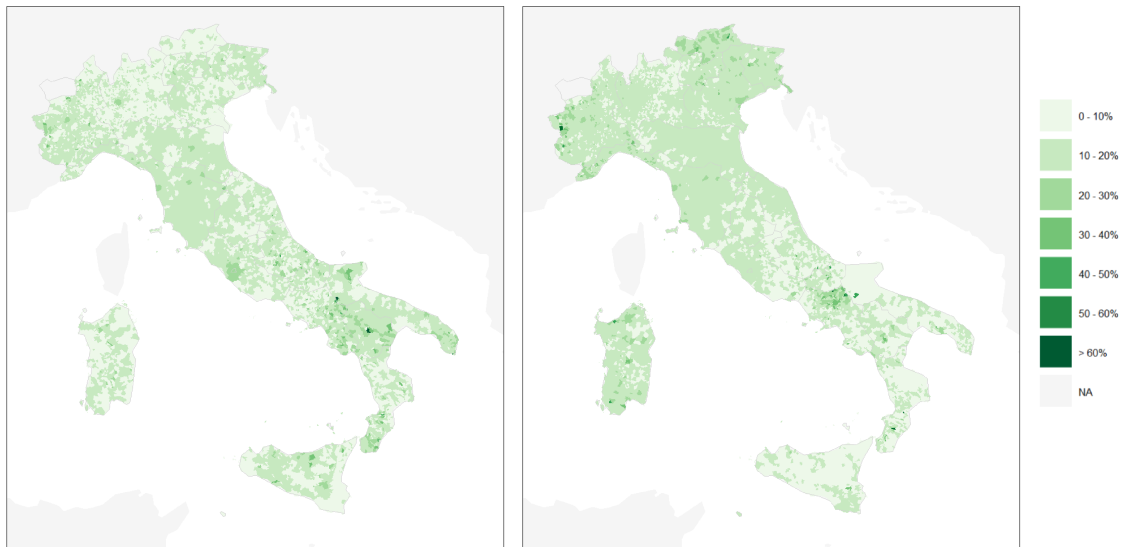
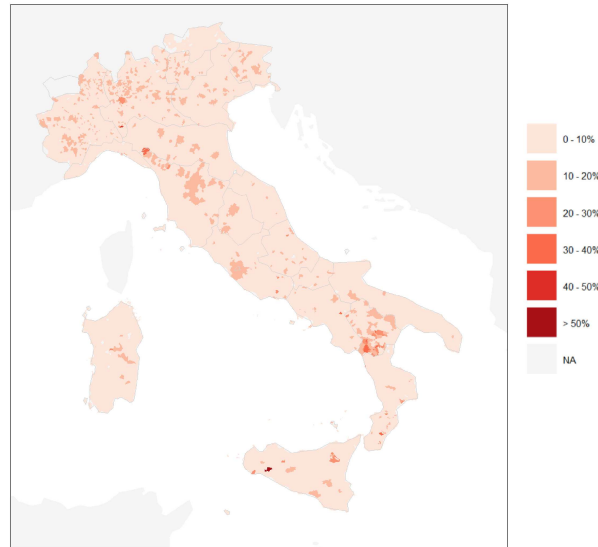


Figura A.8: Proporzioe dei voti per comune - Azione - Italia Viva (AZ-IV)

(a) Anno: 2022



## A.2 Analisi delle corrispondenze

Tabella A.1: Analisi delle corrispondenze (righe) - Elezioni 2018

	Massa(x100)	Inerzia(x100)	Inerzia(%)
Veneto	8,02	0,5280	( 3,14%)
Lombardia	16,12	1,0652	( 6,34%)
Toscana	6,11	0,5171	( 3,08%)
Sardegna	2,94	0,2238	( 1,33%)
Abruzzo	2,25	0,0587	( 0,35%)
Basilicata	1,00	0,0743	( 0,44%)
Sicilia	8,62	0,9272	( 5,52%)
Puglia	7,03	0,4020	( 2,39%)
Piemonte	7,27	0,1376	( 0,82%)
Lazio	9,45	0,0427	( 0,25%)
Campania	9,80	0,9073	( 5,40%)
Calabria	3,31	0,2324	( 1,38%)
Marche	2,55	0,0514	( 0,31%)
Umbria	1,45	0,0589	( 0,35%)
Molise	0,55	0,0318	( 0,19%)
Emilia-Romagna	7,15	0,3849	( 2,29%)
Friuli	2,04	0,0623	( 0,37%)
Liguria	2,64	0,0182	( 0,11%)
Trentino	1,71	11,0672	( 65,91%)

Tabella A.2: Analisi delle corrispondenze (colonne) - Elezioni 2018

	Massa(x100)	Inerzia(x100)	Inerzia(%)
Centro-destra	24,67	1,6598	( 9,88%)
PD	12,66	1,4054	( 8,37%)
M5S	22,01	1,4337	( 8,54%)
Partiti locali	0,38	11,3341	( 67,50%)
Altri partiti	8,19	0,1602	( 0,95%)
Non voto	32,09	0,7977	( 4,75%)

Tabella A.3: Inerzia - Elezioni 2018

Totale Inerzia = 0.16791		
Asse 1	0,1145	(68,16%; cum: 68,16%)
Asse 2	0,0434	(25,84%; cum: 94,00%)
Asse 3	0,0078	(4,66%; cum: 98,66%)
Asse 4	0,0017	(1,03%; cum: 99,96%)
Asse 5	0,0005	(0,31%; cum: 100,00%)

Tabella A.4: Analisi delle corrispondenze (righe) - Elezioni 2022

	Massa(x100)	Inerzia(x100)	Inerzia(%)
Piemonte	7,21	0,1734	( 1,37%)
Lombardia	16,31	1,1083	( 8,70%)
Trentino	1,76	3,5198	( 27,63%)
Veneto	8,10	0,7648	( 6,00%)
Friuli	2,03	0,0964	( 0,76%)
Liguria	2,60	0,0564	( 0,44%)
Emilia-Romagna	7,23	0,6084	( 4,77%)
Toscana	6,11	0,3602	( 2,83%)
Umbria	1,44	0,0244	( 0,19%)
Marche	2,53	0,0267	( 0,21%)
Lazio	9,45	0,0833	( 0,65%)
Abruzzo	2,23	0,0362	( 0,28%)
Molise	0,53	0,0341	( 0,27%)
Campania	9,80	1,5585	( 12,23%)
Puglia	6,99	0,5865	( 4,60%)
Basilicata	0,97	0,0498	( 0,39%)
Calabria	3,25	0,3993	( 3,13%)
Sicilia	8,52	3,0642	( 24,05%)
Sardegna	2,92	0,1889	( 1,48%)

Tabella A.5: Analisi delle corrispondenze (colonne) - Elezioni 2022

	Massa(x100)	Inerzia(x100)	Inerzia(%)
Centro-destra	25,44	1,4320	( 11,24%)
PD	11,14	1,0673	( 8,38%)
AZIV	4,49	0,5458	( 4,28%)
M5S	9,08	2,1761	( 17,08%)
Partiti locali	0,79	6,0582	( 47,55%)
Altri partiti	7,88	0,4283	( 3,36%)
Non voto	41,18	1,0337	( 8,11%)

Tabella A.6: Inerzia - Elezioni 2022

Totale inerzia = 0,1274		
Asse 1	0,0748	(58,70%; cum: 58,70%)
Asse 2	0,0433	(33,95%; cum: 92,66%)
Asse 3	0,0067	(5,23%; cum: 97,80%)
Asse 4	0,0018	(1,39%; cum: 99,28%)
Asse 5	0,0006	(0,44%; cum: 99,71%)
Asse 6	0,0004	(0,29%; cum: 100%)

## A.3 Inserzioni su Meta

Tabella A.7: Distribuzione regionale delle spese per inserzioni su Meta - Elezioni 2022

Valori espressi nell'ordine delle migliaia (x1.000)

Regione	FDI	ALTRI	LEGA	AZIV	PD	M5S	FI	Totale	Elettori
Sicilia	19,0	15,5	11,7	8,5	6,9	7,4	1,2	70,3	3.921
Lazio	18,6	23,7	10,0	13,4	7,4	4,7	0,9	78,8	4.351
Campania	18,3	21,6	13,5	9,5	8,1	8,5	1,3	80,9	4.510
Lombardia	16,3	18,0	14,3	10,1	8,4	1,6	8,2	76,8	7.505
Puglia	12,8	7,9	8,5	6,5	5,6	5,1	0,9	47,2	3.217
Piemonte	8,4	6,1	6,7	5,1	4,1	0,9	0,4	31,6	3.320
Emilia-Romagna	7,5	5,0	5,9	5,1	5,1	0,8	0,4	29,8	3.328
Veneto	7,3	5,0	6,7	4,9	4,9	0,7	0,4	30,0	3.728
Calabria	6,8	3,3	4,3	3,0	2,6	2,6	0,6	23,2	1.496
Toscana	6,7	5,8	6,3	5,7	5,3	0,8	0,4	31,0	2.811
Sardegna	4,7	3,0	2,7	2,1	3,0	1,6	0,3	17,4	1.342
Marche	3,5	1,9	2,0	2,0	1,6	0,2	0,2	11,4	1.165
Liguria	3,2	2,0	2,2	1,8	1,5	0,3	0,2	11,3	1.195
Abruzzo	3,2	1,8	2,1	1,8	1,5	0,8	0,2	11,4	1.026
Umbria	2,5	0,8	1,0	0,9	0,8	0,2	0,1	6,3	662
Friuli	1,1	1,0	0,9	0,9	0,8	0,2	0,1	4,9	936
Basilicata	1,0	0,6	0,4	0,8	0,8	0,5	0,1	4,2	446
Trentino	0,5	4,1	0,6	0,4	0,2	0,1	0,0	6,0	811
Molise	0,0	0,0	0,0	0,0	0,3	0,2	0,0	0,5	243
Italia	141,6	126,9	100,0	82,3	68,8	37,1	16,1	573,0	46.021

Tabella A.8: Spese per inserzioni su Meta ogni mille elettori - Elezioni 2022

Regione	FDI	ALTRI	LEGA	AZIV	PD	M5S	FI	Totale
Sicilia	4,85	3,94	2,99	2,18	1,77	1,88	0,31	17,93
Lazio	4,28	5,44	2,30	3,08	1,71	1,08	0,21	18,11
Campania	4,06	4,79	3,00	2,10	1,80	1,89	0,29	17,94
Lombardia	2,18	2,39	1,91	1,34	1,12	0,21	1,09	10,23
Puglia	3,98	2,44	2,65	2,01	1,74	1,58	0,28	14,67
Piemonte	2,52	1,83	2,02	1,53	1,23	0,26	0,13	9,52
Emilia-Romagna	2,26	1,49	1,78	1,53	1,52	0,23	0,13	8,95
Veneto	1,97	1,35	1,80	1,31	1,32	0,19	0,12	8,04
Calabria	4,52	2,24	2,88	1,99	1,71	1,75	0,40	15,50
Toscana	2,38	2,05	2,26	2,02	1,88	0,27	0,15	11,03
Sardegna	3,48	2,24	2,00	1,57	2,23	1,19	0,24	12,97
Marche	3,02	1,59	1,75	1,68	1,38	0,19	0,16	9,79
Liguria	2,71	1,67	1,83	1,53	1,28	0,26	0,15	9,46
Abruzzo	3,11	1,71	2,06	1,77	1,45	0,82	0,22	11,11
Umbria	3,82	1,28	1,46	1,35	1,14	0,25	0,15	9,52
Friuli	1,18	1,04	0,95	0,94	0,88	0,16	0,09	5,24
Basilicata	2,33	1,40	0,95	1,75	1,72	1,10	0,22	9,42
Trentino	0,65	5,04	0,80	0,45	0,29	0,12	0,00	7,40
Molise	0,00	0,10	0,01	0,06	1,03	0,89	0,08	2,06
Italia	3,08	2,76	2,17	1,79	1,50	0,81	0,35	12,45



Tabella A.9: Distribuzione delle spese per inserioni su Meta in percentuale rispetto al totale di ogni classe - Elezioni 2022

Regione	FDI	ALTRI	LEGA	AZIV	PD	M5S	FI
Sicilia	13,43	12,18	11,73	10,38	10,09	19,86	7,46
Lazio	13,14	18,65	9,99	16,28	10,81	12,70	5,77
Campania	12,93	17,04	13,52	11,50	11,77	23,00	8,06
Lombardia	11,54	14,14	14,31	12,21	12,20	4,20	51,00
Puglia	9,04	6,18	8,53	7,87	8,12	13,71	5,60
Piemonte	5,91	4,79	6,72	6,16	5,92	2,35	2,77
Veneto	5,31	3,91	5,91	6,21	7,34	2,08	2,60
Emilia-Romagna	5,18	3,97	6,69	5,92	7,14	1,94	2,70
Toscana	4,78	2,64	4,31	3,62	3,72	7,06	3,72
Calabria	4,73	4,54	6,34	6,91	7,68	2,06	2,70
Sardegna	3,30	2,37	2,68	2,56	4,36	4,30	1,96
Marche	2,48	1,46	2,04	2,38	2,34	0,60	1,18
Liguria	2,29	1,57	2,19	2,22	2,22	0,84	1,14
Abruzzo	2,25	1,39	2,11	2,21	2,17	2,27	1,42
Umbria	1,79	0,67	0,97	1,09	1,10	0,44	0,61
Friuli	0,78	0,76	0,89	1,07	1,20	0,41	0,55
Basilicata	0,74	0,49	0,42	0,95	1,12	1,33	0,62
Trentino	0,37	3,22	0,65	0,45	0,34	0,25	0,01
Molise	0,00	0,02	0,00	0,02	0,37	0,58	0,13

# Appendice B

## Modello inferenziale

### B.1 Covariate

Tabella B.1: Riepilogo delle covariate usate nel modello

$x$	Descrizione	Dummy
$x_0$	Costante o intercetta	-
$x_1$	Quota di votanti maschi sul totale	-
$x_2$	Quota di stranieri sulla popolazione totale	-
$x_3$	Logaritmo del reddito medio da lavoro dipendente	-
$x_4$	Logaritmo della popolazione residente	-
$x_5$	Logaritmo della popolazione residente al quadrato	-
$x_6$	Quota di persone con più di 60 anni sulla popolazione totale	-
$x_7$	Quota di persone con titolo di studio terziario sulla popolazione totale	-
$x_8$	Tasso di disoccupazione	-
$x_9$	Nord-ovest	$d$
$x_{10}$	Centro	$d$
$x_{11}$	Isole	$d$
$x_{12}$	Sud	$d$
$x_{13}$	Sobborghi o zone a densità intermedia di popolazione	$d$
$x_{14}$	Zone rurali o scarsamente popolate	$d$
$x_{15}$	Comune in montagna	$d$
$x_{16}$	Zona costiera	$d$

Tabella B.2: Statistiche descrittive delle covariate - Elezioni 2018 e 2022

	Anno	Media	Mediana	Dev. Std.	C.V.	Asimmetria	Ecc. Curtosi	5% perc.	95% perc.	Dati mancanti	Dati validi
Maschi	2018	0,5131	0,5092	0,0224	0,043	2,077	10,128	0,4864	0,5519	0 (0 %)	7884 (100%)
	2022	0,5108	0,5069	0,0235	0,046	1,687	7,794	0,4826	0,5520	0 (0 %)	7830 (100%)
Stranieri	2018	0,0655	0,0586	0,0416	0,6355	1,014	1,6392	0,0122	0,1418	75 (1%)	7809 (99%)
	2022	0,0668	0,059	0,0435	0,651	1,106	2,058	0,012	0,1476	29 (0,4%)	7801 (99,6%)
Reddito	2018	18.872	18.894	3.741	0,20	0,41	2,49	12.704	24.489	221 (2,8%)	7663 (97,2%)
	2022	19.579	19.618	3.788	0,19	0,47	2,03	13.597	25.463	0 (0%)	7830 (100%)
log(Reddito)	2018	9,82	9,84	0,20	0,02	-0,42	0,49	9,45	10,10	221 (2,8%)	7663 (97,2%)
	2022	9,86	9,88	0,19	0,02	-0,28	0,31	9,51	10,14	0 (0%)	7830 (100%)
Popolazione	2018	7.629	2.491	42.718	5,60	44,37	2.591	294	24.950	75 (1%)	7809 (99%)
	2022	7.523	2.435	41.697	5,54	44,25	2.586	282	25.031	0 (0%)	7830 (100%)
log( $p$ )	2018	7,85	7,82	1,35	0,17	0,18	0,28	5,68	10,13	75 (1%)	7809 (99%)
	2022	7,83	7,79	1,36	0,17	0,16	0,25	5,64	10,12	0 (0%)	7830 (100%)
log( $p$ ) <sup>2</sup>	2018	63,53	61,16	21,80	0,34	0,77	1,48	32,20	102,51	75 (1%)	7809 (99%)
	2022	63,18	60,8	21,93	0,34	0,76	1,40	31,85	102,57	0 (0%)	7830 (100%)
Over 60	2018	0,318	0,313	0,0596	0,187	0,789	1,8762	0,2316	0,4210	10 (0,1%)	7874 (99,9%)
	2022	0,334	0,328	0,058	0,176	0,746	1,6757	0,2486	0,4371	0 (0%)	7830 (100%)
Laureati	2018	0,1031	0,0996	0,0315	0,3055	1,041	2,8974	0,0592	0,1586	42 (0,5%)	7842 (99,5%)
	2022	0,114	0,1105	0,0343	0,30	0,985	2,6442	0,066	0,1751	23 (0,3%)	7807 (99,7%)
Disoccupazione	2018	0,1397	0,1128	0,0692	0,502	0,8623	-0,003	0,0586	0,2703	14 (0,2%)	7870 (99,8%)
	2022	0,09	0,08	0,0399	0,442	0,7992	0,3593	0,0412	0,1653	3 (0,1%)	7827 (99,9%)

100

Modello inferenziale

Tabella B.3: Matrice delle correlazioni delle covariate - Elezioni 2018

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$	$x_{15}$	$x_{16}$
Maschi	1	-0,07	-0,25	-0,47	-0,45	0,33	-0,26	0,10	-0,04	0,12	0,02	-0,09	-0,33	0,37	0,20	-0,05
Stranieri	-0,07	1	0,22	0,15	0,15	-0,14	0,11	-0,29	0,20	-0,28	-0,29	0,12	0,09	-0,10	-0,17	-0,13
log(Reddito)	-0,25	0,22	1	0,14	0,15	-0,09	0,31	-0,71	0,51	-0,55	-0,31	0,20	0,22	-0,26	-0,09	-0,29
log( $p$ )	-0,47	0,15	0,14	1	0,99	-0,53	0,40	0,09	-0,22	0,04	0,04	0,12	0,53	-0,65	-0,35	0,23
log( $p$ ) <sup>2</sup>	-0,45	0,15	0,15	0,99	1	-0,49	0,42	0,09	-0,21	0,03	0,04	0,12	0,53	-0,66	-0,34	0,24
Over60	0,33	-0,14	-0,09	-0,53	-0,49	1	-0,15	-0,04	0,07	-0,05	0,01	-0,11	-0,38	0,42	0,24	-0,09
Laureati	-0,26	0,11	0,31	0,40	0,42	-0,15	1	-0,03	-0,09	0,11	-0,14	-0,02	0,26	-0,33	-0,13	0,11
Disoccupati	0,10	-0,29	-0,71	0,09	0,09	-0,04	-0,03	1	-0,44	0,52	0,51	-0,39	-0,04	0,01	-0,10	0,37
Nord ovest	-0,04	0,20	0,51	-0,22	-0,21	0,07	-0,09	-0,44	1	-0,42	-0,26	-0,36	0,05	-0,06	0,01	-0,22
Centro	0,12	-0,28	-0,55	0,04	0,03	-0,05	0,11	0,52	-0,42	1	-0,18	-0,25	-0,05	0,03	0,03	0,20
Isole	0,02	-0,29	-0,31	0,04	0,04	0,01	-0,14	0,51	-0,26	-0,18	1	-0,15	-0,02	0,03	-0,10	0,24
Sud	-0,09	0,12	0,20	0,12	0,12	-0,11	-0,02	-0,39	-0,36	-0,25	-0,15	1	0,03	-0,01	0,06	-0,10
Borgo	-0,33	0,09	0,22	0,53	0,53	-0,38	0,26	-0,04	0,05	-0,05	-0,02	0,03	1	-0,93	-0,22	0,13
Rurale	0,37	-0,10	-0,26	-0,65	-0,66	0,42	-0,33	0,01	-0,06	0,03	0,03	-0,01	-0,93	1	0,26	-0,16
Montagna	0,20	-0,17	-0,09	-0,35	-0,34	0,24	-0,13	-0,10	0,01	0,03	-0,10	0,06	-0,22	0,26	1	-0,20
Costa	-0,05	-0,13	-0,29	0,23	0,24	-0,09	0,11	0,37	-0,22	0,20	0,24	-0,10	0,13	-0,16	-0,20	1

Tabella B.4: Matrice delle correlazioni delle covariate - Elezioni 2022

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$	$x_{15}$	$x_{16}$
Maschi	1	-0,08	-0,29	-0,43	-0,41	0,31	-0,23	0,13	-0,10	0,25	-0,08	-0,10	-0,31	0,34	0,20	-0,04
Stranieri	-0,08	1	0,26	0,16	0,16	-0,18	0,09	-0,28	0,22	-0,31	-0,30	0,15	0,09	-0,10	-0,18	-0,14
log(Reddito)	-0,29	0,26	1	0,17	0,17	-0,15	0,32	-0,67	0,51	-0,52	-0,35	0,22	0,23	-0,26	-0,11	-0,31
log( $p$ )	-0,43	0,16	0,17	1	0,99	-0,54	0,40	0,01	-0,21	0,03	0,03	0,13	0,54	-0,65	-0,34	0,23
log( $p$ ) <sup>2</sup>	-0,41	0,16	0,17	0,99	1	-0,51	0,42	0,02	-0,20	0,02	0,03	0,12	0,53	-0,66	-0,34	0,24
Over60	0,31	-0,18	-0,15	-0,54	-0,51	1	-0,15	0,10	0,04	-0,01	0,04	-0,12	-0,38	0,43	0,26	-0,08
Laureati	-0,23	0,09	0,32	0,40	0,42	-0,15	1	-0,03	-0,09	0,11	-0,14	-0,02	0,26	-0,33	-0,14	0,10
Disoccupati	0,13	-0,28	-0,67	0,01	0,02	0,10	-0,03	1	-0,38	0,43	0,52	-0,39	-0,05	0,03	-0,08	0,37
Nord ovest	-0,10	0,22	0,51	-0,21	-0,20	0,04	-0,09	-0,38	1	-0,42	-0,25	-0,36	0,06	-0,07	0,00	-0,22
Centro	0,25	-0,31	-0,52	0,03	0,02	-0,01	0,11	0,43	-0,42	1	-0,18	-0,25	-0,06	0,03	0,04	0,20
Isole	-0,08	-0,30	-0,35	0,03	0,03	0,04	-0,14	0,52	-0,25	-0,18	1	-0,15	-0,02	0,03	-0,10	0,24
Sud	-0,10	0,15	0,22	0,13	0,12	-0,12	-0,02	-0,39	-0,36	-0,25	-0,15	1	0,03	-0,02	0,06	-0,10
Borgo	-0,31	0,09	0,23	0,54	0,53	-0,38	0,26	-0,05	0,06	-0,06	-0,02	0,03	1	-0,93	-0,22	0,13
Rurale	0,34	-0,10	-0,26	-0,65	-0,66	0,43	-0,33	0,03	-0,07	0,03	0,03	-0,02	-0,93	1	0,26	-0,16
Montagna	0,20	-0,18	-0,11	-0,34	-0,34	0,26	-0,14	-0,08	0,00	0,04	-0,10	0,06	-0,22	0,26	1	-0,19
Costa	-0,04	-0,14	-0,31	0,23	0,24	-0,08	0,10	0,37	-0,22	0,20	0,24	-0,10	0,13	-0,16	-0,19	1

## B.2 Analisi controfattuale

Tabella B.5: Voti reali ( $V$ ), attesi ( $\hat{E}(V)$ ) e scostamenti ( $\hat{E}(V) - V$ ) - Elezioni 2018

	Voto (%)	LEGA	FI	FDI	PD	M5S	ALTRI	NON VOTO
Roma	$V$	7,19	7,38	6,49	14,95	20,29	10,98	32,72
	$\hat{E}(V)$	8,28	7,83	4,72	13,22	22,25	10,97	32,74
	$\hat{E}(V) - V$	1,09	0,45	-1,77	-1,73	1,94	-0,01	0,02
Milano	$V$	11,68	10,58	3,04	18,37	12,52	12,46	31,35
	$\hat{E}(V)$	13,02	9,12	4,36	15,83	15,33	12,28	30,06
	$\hat{E}(V) - V$	1,34	-1,46	1,32	-2,54	2,81	-0,18	-1,29
Venezia	$V$	15,77	5,97	2,59	15,01	18,93	10,39	31,34
	$\hat{E}(V)$	15,64	7,71	3,62	14,29	17,38	9,79	31,57
	$\hat{E}(V) - V$	-0,13	1,74	1,03	-0,72	-1,55	-0,60	0,23
Alba	$V$	14,39	8,36	4,10	15,99	15,52	10,19	31,45
	$\hat{E}(V)$	21,24	9,44	3,34	16,16	15,29	8,55	25,99
	$\hat{E}(V) - V$	6,85	1,08	-0,76	0,17	-0,23	-1,66	-5,06
Ancona	$V$	10,28	6,02	3,30	17,87	23,29	9,46	29,79
	$\hat{E}(V)$	9,68	7,56	4,02	19,23	22,48	9,43	27,60
	$\hat{E}(V) - V$	-0,60	1,54	0,72	1,36	-0,81	-0,03	-2,19
Loreto	$V$	13,09	7,14	4,13	12,54	25,95	8,56	28,59
	$\hat{E}(V)$	13,98	8,02	4,06	16,16	23,91	8,24	25,64
	$\hat{E}(V) - V$	0,89	0,88	-0,07	3,62	-2,04	-0,32	-2,95
Cagliari	$V$	6,93	9,52	3,27	11,03	20,86	11,37	37,02
	$\hat{E}(V)$	2,87	7,96	2,69	11,56	27,47	10,52	36,93
	$\hat{E}(V) - V$	-4,06	-1,56	-0,58	0,53	6,61	-0,85	-0,09
Vasto	$V$	7,56	9,37	4,98	10,37	29,11	7,20	31,42
	$\hat{E}(V)$	4,47	10,06	2,84	11,38	29,68	9,28	32,29
	$\hat{E}(V) - V$	-3,09	0,69	-2,14	1,01	0,57	2,08	0,87

Tabella B.6: Voti reali ( $V$ ), attesi ( $\hat{E}(V)$ ) e scostamenti ( $\hat{E}(V) - V$ ) - Elezioni 2022

	Voto (%)	LEGA	FI	FDI	PD	M5S	AZIV	ALTRI	NON VOTO
Roma	$V$	2,63	2,76	17,41	14,13	8,48	6,32	9,13	39,14
	$\hat{E}(V)$	2,49	3,02	12,94	13,50	12,03	6,16	9,81	40,04
	$\hat{E}(V) - V$	-0,14	0,26	-4,47	-0,64	3,55	-0,16	0,68	0,90
Milano	$V$	4,09	3,77	13,25	16,41	4,96	10,07	11,73	35,73
	$\hat{E}(V)$	4,07	3,01	12,76	18,11	7,78	12,28	12,04	29,95
	$\hat{E}(V) - V$	-0,02	-0,76	-0,49	1,70	2,82	2,21	0,31	-5,78
Venezia	$V$	5,97	2,57	15,74	14,13	5,39	5,20	12,81	38,20
	$\hat{E}(V)$	5,46	3,62	14,42	12,92	5,03	5,83	10,46	42,26
	$\hat{E}(V) - V$	-0,50	1,06	-1,33	-1,20	-0,36	0,63	-2,35	4,06
Alba	$V$	5,90	5,87	14,23	13,79	4,83	7,63	9,50	38,25
	$\hat{E}(V)$	7,55	4,68	18,78	13,92	5,11	6,68	8,08	35,21
	$\hat{E}(V) - V$	1,65	-1,19	4,54	0,13	0,27	-0,95	-1,42	-3,04
Ancona	$V$	3,25	2,79	15,79	15,90	9,39	5,66	10,77	36,45
	$\hat{E}(V)$	3,26	3,43	14,54	16,27	9,46	6,14	9,45	37,44
	$\hat{E}(V) - V$	0,01	0,64	-1,25	0,37	0,07	0,47	-1,32	1,00
Loreto	$V$	4,54	4,31	19,88	10,21	9,41	4,91	10,65	36,09
	$\hat{E}(V)$	4,90	4,35	19,05	12,40	8,07	4,74	8,13	38,37
	$\hat{E}(V) - V$	0,35	0,04	-0,83	2,19	-1,34	-0,18	-2,52	2,28
Cagliari	$V$	2,55	3,54	12,20	11,19	9,92	3,85	11,94	44,81
	$\hat{E}(V)$	1,41	3,08	6,95	13,19	15,46	4,33	12,79	42,79
	$\hat{E}(V) - V$	-1,13	-0,46	-5,25	2,00	5,53	0,48	0,85	-2,02
Vasto	$V$	3,78	4,39	15,08	10,36	11,20	4,33	6,21	44,65
	$\hat{E}(V)$	2,61	5,11	10,51	11,03	15,09	3,44	6,61	45,60
	$\hat{E}(V) - V$	-1,16	0,71	-4,56	0,67	3,90	-0,90	0,40	0,94

Tabella B.7: Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con -5% in "Over 60"

	Anno	Voto (%)	Lega	FI	FDI	PD	M5S	AZIV	Altri	Non voto
Roma	2018	$\tilde{E}(V)$	8,34	8,40	4,58	12,41	24,19	-	10,94	31,14
		$\tilde{E}(V) - \hat{E}(V)$	0,07	0,57	-0,15	-0,81	1,95	-	-0,03	-1,59
	2022	$\tilde{E}(V)$	2,46	3,01	13,35	12,64	12,62	6,69	10,26	38,98
		$\tilde{E}(V) - \hat{E}(V)$	-0,03	-0,01	0,40	-0,86	0,58	0,52	0,45	-1,06
Milano	2018	$\tilde{E}(V)$	14,19	9,21	4,48	14,88	16,38	-	12,54	28,32
		$\tilde{E}(V) - \hat{E}(V)$	1,17	0,09	0,12	-0,95	1,05	-	0,25	-1,74
	2022	$\tilde{E}(V)$	4,28	3,97	12,83	16,56	8,03	12,43	12,46	29,44
		$\tilde{E}(V) - \hat{E}(V)$	0,21	0,96	0,07	-1,55	0,24	0,15	0,42	-0,51
Venezia	2018	$\tilde{E}(V)$	16,90	7,51	3,55	13,77	18,24	-	10,08	29,94
		$\tilde{E}(V) - \hat{E}(V)$	1,26	-0,20	-0,06	-0,52	0,87	-	0,29	-1,63
	2022	$\tilde{E}(V)$	5,79	3,92	15,12	11,97	5,43	6,01	10,64	41,13
		$\tilde{E}(V) - \hat{E}(V)$	0,32	0,30	0,70	-0,96	0,40	0,18	0,19	-1,13
Alba	2018	$\tilde{E}(V)$	23,43	9,62	3,49	15,04	15,63	-	8,06	24,73
		$\tilde{E}(V) - \hat{E}(V)$	2,21	0,18	0,15	-1,13	0,34	-	-0,49	-1,26
	2022	$\tilde{E}(V)$	8,27	5,29	19,68	12,61	4,89	6,39	8,26	34,60
		$\tilde{E}(V) - \hat{E}(V)$	0,73	0,60	0,91	-1,30	-0,21	-0,28	0,18	-0,61
Ancona	2018	$\tilde{E}(V)$	10,21	7,77	4,13	17,65	24,54	-	9,38	26,32
		$\tilde{E}(V) - \hat{E}(V)$	0,53	0,21	0,11	-1,58	2,06	-	-0,05	-1,28
	2022	$\tilde{E}(V)$	3,48	3,65	16,31	15,32	10,00	5,93	9,60	35,71
		$\tilde{E}(V) - \hat{E}(V)$	0,22	0,22	1,77	-0,95	0,54	-0,21	0,15	-1,73
Loreto	2018	$\tilde{E}(V)$	15,36	8,51	4,03	15,05	24,59	-	7,95	24,52
		$\tilde{E}(V) - \hat{E}(V)$	1,38	0,49	-0,03	-1,12	0,68	-	-0,29	-1,12
	2022	$\tilde{E}(V)$	4,57	4,28	20,94	11,46	7,67	4,50	8,52	38,06
		$\tilde{E}(V) - \hat{E}(V)$	-0,33	-0,07	1,89	-0,93	-0,40	-0,24	0,39	-0,31
Cagliari	2018	$\tilde{E}(V)$	2,92	8,55	3,10	11,45	28,91	-	10,45	34,63
		$\tilde{E}(V) - \hat{E}(V)$	0,05	0,59	0,41	-0,11	1,44	-	-0,07	-2,30
	2022	$\tilde{E}(V)$	1,50	3,17	6,82	12,24	17,24	4,43	13,31	41,29
		$\tilde{E}(V) - \hat{E}(V)$	0,09	0,09	-0,13	-0,94	1,78	0,09	0,52	-1,50
Vasto	2018	$\tilde{E}(V)$	4,61	10,77	2,93	10,48	31,52	-	9,07	30,63
		$\tilde{E}(V) - \hat{E}(V)$	0,14	0,71	0,09	-0,90	1,84	-	-0,21	-1,67
	2022	$\tilde{E}(V)$	2,68	5,45	10,89	9,70	15,50	3,31	6,76	45,72
		$\tilde{E}(V) - \hat{E}(V)$	0,06	0,34	0,38	-1,33	0,41	-0,13	0,14	0,13



Tabella B.8: Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con +5% in "Stranieri"

	Anno	Voto (%)	Lega	FI	FDI	PD	M5S	AZIV	Altri	Non voto
Roma	2018	$\tilde{E}(V)$	8,29	8,03	4,50	14,12	21,59	-	10,89	32,58
		$\tilde{E}(V) - \hat{E}(V)$	0,02	0,21	-0,23	0,90	-0,66	-	-0,08	-0,15
	2022	$\tilde{E}(V)$	2,41	2,90	13,06	14,94	11,92	6,96	9,64	38,17
		$\tilde{E}(V) - \hat{E}(V)$	-0,08	-0,12	0,12	1,44	-0,11	0,80	-0,17	-1,87
Milano	2018	$\tilde{E}(V)$	13,99	8,71	4,37	16,76	14,51	-	12,38	29,28
		$\tilde{E}(V) - \hat{E}(V)$	0,97	-0,41	0,01	0,93	-0,82	-	0,10	-0,78
	2022	$\tilde{E}(V)$	4,16	3,79	12,44	19,33	7,51	12,77	11,55	28,45
		$\tilde{E}(V) - \hat{E}(V)$	0,09	0,78	-0,31	1,23	-0,27	0,48	-0,49	-1,50
Venezia	2018	$\tilde{E}(V)$	16,71	7,13	3,47	15,53	16,13	-	9,98	31,05
		$\tilde{E}(V) - \hat{E}(V)$	1,07	-0,58	-0,15	1,24	-1,24	-	0,19	-0,52
	2022	$\tilde{E}(V)$	5,67	3,78	14,81	14,14	5,15	6,23	9,97	40,25
		$\tilde{E}(V) - \hat{E}(V)$	0,20	0,16	0,39	1,22	0,11	0,40	-0,48	-2,01
Alba	2018	$\tilde{E}(V)$	23,09	9,13	3,40	17,06	13,79	-	7,97	25,57
		$\tilde{E}(V) - \hat{E}(V)$	1,86	-0,31	0,06	0,90	-1,50	-	-0,58	-0,43
	2022	$\tilde{E}(V)$	8,09	5,09	19,19	14,99	4,62	6,62	7,72	33,68
		$\tilde{E}(V) - \hat{E}(V)$	0,54	0,41	0,41	1,07	-0,48	-0,05	-0,36	-1,53
Ancona	2018	$\tilde{E}(V)$	10,10	7,39	4,04	20,10	21,71	-	9,29	27,37
		$\tilde{E}(V) - \hat{E}(V)$	0,43	-0,17	0,02	0,87	-0,77	-	-0,14	-0,23
	2022	$\tilde{E}(V)$	3,39	3,50	15,87	18,11	9,37	6,13	8,96	34,68
		$\tilde{E}(V) - \hat{E}(V)$	0,13	0,06	1,33	1,83	-0,10	-0,01	-0,50	-2,76
Loreto	2018	$\tilde{E}(V)$	15,31	8,12	3,96	17,15	21,84	-	7,93	25,69
		$\tilde{E}(V) - \hat{E}(V)$	1,33	0,10	-0,10	0,99	-2,06	-	-0,31	0,05
	2022	$\tilde{E}(V)$	4,48	4,13	20,52	13,61	7,25	4,69	7,98	37,33
		$\tilde{E}(V) - \hat{E}(V)$	-0,42	-0,22	1,46	1,22	-0,82	-0,04	-0,14	-1,04
Cagliari	2018	$\tilde{E}(V)$	2,92	8,18	3,06	13,10	25,95	-	10,43	36,36
		$\tilde{E}(V) - \hat{E}(V)$	0,05	0,22	0,37	1,54	-1,52	-	-0,09	-0,57
	2022	$\tilde{E}(V)$	1,47	3,07	6,71	14,49	16,36	4,61	12,54	40,76
		$\tilde{E}(V) - \hat{E}(V)$	0,06	-0,00	-0,24	1,30	0,90	0,27	-0,26	-2,03
Vasto	2018	$\tilde{E}(V)$	4,64	10,37	2,90	12,11	28,34	-	9,13	32,51
		$\tilde{E}(V) - \hat{E}(V)$	0,17	0,30	0,07	0,74	-1,34	-	-0,15	0,22
	2022	$\tilde{E}(V)$	2,64	5,29	10,73	11,52	14,75	3,45	6,39	45,24
		$\tilde{E}(V) - \hat{E}(V)$	0,02	0,18	0,21	0,49	-0,35	0,02	-0,23	-0,35

Tabella B.9: Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con +5% in "Laureati"

	Anno	Voto (%)	Lega	FI	FDI	PD	M5S	AZIV	Altri	Non voto
Roma	2018	$\tilde{E}(V)$	6,78	7,62	4,66	14,02	22,46	-	13,06	31,41
		$\tilde{E}(V) - \hat{E}(V)$	-1,50	-0,21	-0,06	0,80	0,21	-	2,08	-1,33
	2022	$\tilde{E}(V)$	1,95	2,62	11,61	15,29	12,42	7,79	11,74	36,58
		$\tilde{E}(V) - \hat{E}(V)$	-0,55	-0,40	-1,34	1,80	0,38	1,63	1,93	-3,46
Milano	2018	$\tilde{E}(V)$	11,59	8,30	4,56	16,79	15,21	-	14,96	28,58
		$\tilde{E}(V) - \hat{E}(V)$	-1,44	-0,82	0,20	0,96	-0,11	-	2,68	-1,47
	2022	$\tilde{E}(V)$	3,32	3,40	11,01	19,55	7,72	14,09	14,02	26,89
		$\tilde{E}(V) - \hat{E}(V)$	-0,75	0,39	-1,75	1,44	-0,06	1,81	1,98	-3,06
Venezia	2018	$\tilde{E}(V)$	13,84	6,87	3,66	15,74	17,12	-	12,20	30,57
		$\tilde{E}(V) - \hat{E}(V)$	-1,80	-0,84	0,04	1,44	-0,26	-	2,42	-1,00
	2022	$\tilde{E}(V)$	4,59	3,47	13,39	14,63	5,40	7,02	12,35	39,15
		$\tilde{E}(V) - \hat{E}(V)$	-0,87	-0,16	-1,03	1,70	0,37	1,19	1,90	-3,11
Alba	2018	$\tilde{E}(V)$	19,43	8,91	3,63	17,54	14,84	-	9,95	25,70
		$\tilde{E}(V) - \hat{E}(V)$	-1,80	-0,53	0,29	1,38	-0,45	-	1,40	-0,29
	2022	$\tilde{E}(V)$	6,59	4,70	17,44	15,74	4,91	7,58	9,74	33,30
		$\tilde{E}(V) - \hat{E}(V)$	-0,96	0,02	-1,34	1,83	-0,20	0,90	1,66	-1,91
Ancona	2018	$\tilde{E}(V)$	8,28	7,02	4,21	20,03	22,69	-	11,25	26,52
		$\tilde{E}(V) - \hat{E}(V)$	-1,39	-0,54	0,19	0,80	0,21	-	1,82	-1,09
	2022	$\tilde{E}(V)$	2,75	3,18	14,20	18,67	9,83	6,88	11,02	33,48
		$\tilde{E}(V) - \hat{E}(V)$	-0,52	-0,25	-0,34	2,39	0,37	0,74	1,57	-3,96
Loreto	2018	$\tilde{E}(V)$	12,61	7,79	4,17	17,30	23,17	-	9,71	25,25
		$\tilde{E}(V) - \hat{E}(V)$	-1,37	-0,22	0,11	1,13	-0,74	-	1,47	-0,39
	2022	$\tilde{E}(V)$	3,64	3,78	18,57	14,23	7,72	5,37	10,05	36,64
		$\tilde{E}(V) - \hat{E}(V)$	-1,26	-0,57	-0,48	1,84	-0,35	0,63	1,92	-1,73
Cagliari	2018	$\tilde{E}(V)$	2,39	7,68	3,15	12,90	26,75	-	12,46	34,67
		$\tilde{E}(V) - \hat{E}(V)$	-0,48	-0,28	0,46	1,34	-0,72	-	1,94	-2,26
	2022	$\tilde{E}(V)$	1,20	2,76	5,90	14,62	16,83	5,06	15,18	38,46
		$\tilde{E}(V) - \hat{E}(V)$	-0,22	-0,32	-1,04	1,43	1,37	0,73	2,38	-4,33
Vasto	2018	$\tilde{E}(V)$	3,75	9,75	3,00	11,97	29,38	-	10,99	31,16
		$\tilde{E}(V) - \hat{E}(V)$	-0,72	-0,31	0,17	0,59	-0,30	-	1,71	-1,14
	2022	$\tilde{E}(V)$	2,16	4,82	9,67	11,94	15,56	3,90	7,87	44,07
		$\tilde{E}(V) - \hat{E}(V)$	-0,45	-0,28	-0,84	0,92	0,47	0,46	1,25	-1,52

Tabella B.10: Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con circa +5% in "Reddito medio da lavoro dipendente"

	Anno	Voto (%)	Lega	FI	FDI	PD	M5S	AZIV	Altri	Non voto
Roma	2018	$\tilde{E}(V)$	7,91	7,95	4,40	13,52	22,99	-	10,95	32,28
		$\tilde{E}(V) - \hat{E}(V)$	-0,36	0,13	-0,33	0,30	0,74	-	-0,03	-0,46
	2022	$\tilde{E}(V)$	2,37	2,85	12,94	14,31	12,36	6,89	9,89	38,39
		$\tilde{E}(V) - \hat{E}(V)$	-0,12	-0,17	-0,00	0,81	0,33	0,73	0,08	-1,65
Milano	2018	$\tilde{E}(V)$	13,46	8,69	4,30	16,17	15,55	-	12,53	29,30
		$\tilde{E}(V) - \hat{E}(V)$	0,44	-0,43	-0,06	0,34	0,23	-	0,25	-0,76
	2022	$\tilde{E}(V)$	4,10	3,74	12,38	18,59	7,82	12,70	11,91	28,76
		$\tilde{E}(V) - \hat{E}(V)$	0,03	0,72	-0,37	0,48	0,03	0,42	-0,13	-1,18
Venezia	2018	$\tilde{E}(V)$	16,03	7,11	3,42	14,98	17,33	-	10,09	31,04
		$\tilde{E}(V) - \hat{E}(V)$	0,39	-0,61	-0,20	0,69	-0,05	-	0,30	-0,53
	2022	$\tilde{E}(V)$	5,58	3,73	14,72	13,57	5,34	6,19	10,27	40,61
		$\tilde{E}(V) - \hat{E}(V)$	0,12	0,10	0,30	0,65	0,31	0,36	-0,19	-1,65
Alba	2018	$\tilde{E}(V)$	22,28	9,13	3,36	16,50	14,89	-	8,10	25,73
		$\tilde{E}(V) - \hat{E}(V)$	1,05	-0,30	0,02	0,34	-0,40	-	-0,45	-0,26
	2022	$\tilde{E}(V)$	7,98	5,02	19,11	14,40	4,81	6,59	7,97	34,13
		$\tilde{E}(V) - \hat{E}(V)$	0,43	0,34	0,33	0,48	-0,30	-0,08	-0,11	-1,09
Ancona	2018	$\tilde{E}(V)$	9,66	7,34	3,96	19,24	23,22	-	9,37	27,21
		$\tilde{E}(V) - \hat{E}(V)$	-0,01	-0,23	-0,06	0,02	0,74	-	-0,07	-0,40
	2022	$\tilde{E}(V)$	3,34	3,44	15,77	17,36	9,76	6,09	9,22	35,03
		$\tilde{E}(V) - \hat{E}(V)$	0,08	0,01	1,23	1,08	0,29	-0,05	-0,24	-2,41
Loreto	2018	$\tilde{E}(V)$	14,62	8,06	3,88	16,47	23,42	-	7,99	25,56
		$\tilde{E}(V) - \hat{E}(V)$	0,64	0,05	-0,18	0,31	-0,49	-	-0,25	-0,08
	2022	$\tilde{E}(V)$	4,41	4,06	20,37	13,06	7,55	4,66	8,22	37,67
		$\tilde{E}(V) - \hat{E}(V)$	-0,49	-0,29	1,32	0,67	-0,52	-0,08	0,10	-0,70
Cagliari	2018	$\tilde{E}(V)$	2,78	8,06	2,98	12,47	27,46	-	10,44	35,82
		$\tilde{E}(V) - \hat{E}(V)$	-0,09	0,10	0,29	0,91	-0,01	-	-0,09	-1,11
	2022	$\tilde{E}(V)$	1,45	3,01	6,63	13,83	16,91	4,55	12,82	40,79
		$\tilde{E}(V) - \hat{E}(V)$	0,04	-0,07	-0,32	0,64	1,45	0,22	0,03	-2,00
Vasto	2018	$\tilde{E}(V)$	4,39	10,20	2,82	11,50	30,03	-	9,12	31,94
		$\tilde{E}(V) - \hat{E}(V)$	-0,08	0,14	-0,01	0,12	0,34	-	-0,16	-0,36
	2022	$\tilde{E}(V)$	2,59	5,18	10,61	11,01	15,27	3,42	6,54	45,38
		$\tilde{E}(V) - \hat{E}(V)$	-0,02	0,07	0,10	-0,02	0,18	-0,02	-0,07	-0,22

Tabella B.11: Voto atteso ( $\tilde{E}(V)$ ) e variazione ( $\tilde{E}(V) - \hat{E}(V)$ ) con +5% in "Tasso di disoccupazione"

	Anno	Voto (%)	Lega	FI	FDI	PD	M5S	AZIV	Altri	Non voto
Roma	2018	$\tilde{E}(V)$	6,67	8,57	4,32	12,08	23,49	-	10,35	34,52
		$\tilde{E}(V) - \hat{E}(V)$	-1,60	0,74	-0,40	-1,14	1,25	-	-0,62	1,78
	2022	$\tilde{E}(V)$	2,04	2,83	11,66	11,95	14,60	5,66	9,09	42,17
		$\tilde{E}(V) - \hat{E}(V)$	-0,46	-0,19	-1,28	-1,54	2,57	-0,50	-0,72	2,13
Milano	2018	$\tilde{E}(V)$	11,49	9,51	4,27	14,68	16,15	-	11,99	31,90
		$\tilde{E}(V) - \hat{E}(V)$	-1,53	0,39	-0,09	-1,15	0,83	-	-0,29	1,84
	2022	$\tilde{E}(V)$	3,61	3,83	11,65	16,18	9,54	10,83	11,32	33,04
		$\tilde{E}(V) - \hat{E}(V)$	-0,46	0,82	-1,10	-1,93	1,75	-1,45	-0,72	3,09
Venezia	2018	$\tilde{E}(V)$	13,67	7,80	3,39	13,64	18,04	-	9,67	33,79
		$\tilde{E}(V) - \hat{E}(V)$	-1,97	0,08	-0,22	-0,65	0,66	-	-0,12	2,22
	2022	$\tilde{E}(V)$	4,79	3,77	13,56	11,53	6,38	5,23	9,58	45,17
		$\tilde{E}(V) - \hat{E}(V)$	-0,67	0,14	-0,86	-1,39	1,35	-0,60	-0,88	2,91
Alba	2018	$\tilde{E}(V)$	19,21	10,13	3,37	15,09	15,72	-	7,86	28,62
		$\tilde{E}(V) - \hat{E}(V)$	-2,01	0,69	0,03	-1,07	0,42	-	-0,69	2,63
	2022	$\tilde{E}(V)$	6,91	5,10	17,68	12,26	5,85	5,59	7,53	39,08
		$\tilde{E}(V) - \hat{E}(V)$	-0,64	0,42	-1,10	-1,65	0,74	-1,09	-0,55	3,87
Ancona	2018	$\tilde{E}(V)$	8,22	8,01	3,93	17,34	24,07	-	8,94	29,50
		$\tilde{E}(V) - \hat{E}(V)$	-1,46	0,45	-0,09	-1,89	1,60	-	-0,49	1,90
	2022	$\tilde{E}(V)$	2,89	3,47	14,41	14,66	11,76	5,12	8,57	39,12
		$\tilde{E}(V) - \hat{E}(V)$	-0,37	0,04	-0,13	-1,62	2,29	-1,02	-0,88	1,68
Loreto	2018	$\tilde{E}(V)$	12,40	8,85	3,86	14,94	24,43	-	7,67	27,85
		$\tilde{E}(V) - \hat{E}(V)$	-1,58	0,83	-0,20	-1,22	0,52	-	-0,57	2,21
	2022	$\tilde{E}(V)$	3,78	4,08	18,49	11,02	9,09	3,87	7,61	42,06
		$\tilde{E}(V) - \hat{E}(V)$	-1,11	-0,27	-0,56	-1,37	1,02	-0,87	-0,52	3,68
Cagliari	2018	$\tilde{E}(V)$	2,34	8,59	2,89	11,00	27,75	-	9,73	37,71
		$\tilde{E}(V) - \hat{E}(V)$	-0,53	0,63	0,20	-0,56	0,28	-	-0,80	0,78
	2022	$\tilde{E}(V)$	1,25	2,96	5,94	11,46	19,57	3,75	11,56	43,51
		$\tilde{E}(V) - \hat{E}(V)$	-0,17	-0,12	-1,00	-1,73	4,11	-0,58	-1,24	0,72
Vasto	2018	$\tilde{E}(V)$	3,65	10,91	2,74	10,13	30,36	-	8,52	33,69
		$\tilde{E}(V) - \hat{E}(V)$	-0,82	0,85	-0,09	-1,25	0,68	-	-0,76	1,39
	2022	$\tilde{E}(V)$	2,19	5,04	9,39	9,06	17,60	2,79	5,91	48,03
		$\tilde{E}(V) - \hat{E}(V)$	-0,42	-0,07	-1,13	-1,97	2,50	-0,65	-0,71	2,44

# Bibliografia

- [1] John Aitchison. «The statistical analysis of compositional data». In: *Journal of the Royal Statistical Society: Series B (Methodological)* 44.2 (1982), pp. 139–160.
- [2] Yasmine Bekkouche, Julia Cagé e Edgard Dewitte. «The heterogeneous price of a vote: Evidence from multiparty systems, 1993–2017». In: *Journal of Public Economics* 206 (2022), p. 104559.
- [3] John Gibson e Anna Cielecka. «Economic influences on the political support for market reform in post-communist transitions: Some evidence from the 1993 polish parliamentary elections». In: *Europe-Asia Studies* 47.5 (1995), pp. 765–785.
- [4] Michael J Greenacre. «Correspondence analysis». In: *Wiley Interdisciplinary Reviews: Computational Statistics* 2.5 (2010), pp. 613–619.
- [5] James Honacker, Jonathan Katz e Gary King. «A Practical Statistical Model for Multiparty Electoral Data». In: *Political Analysis* 10.1 (2002), pp. 84–100.
- [6] James Honaker, Jonathan N Katz e Gary King. «A fast, easy, and efficient estimator for multiparty electoral data». In: *Political Analysis* 10.1 (2002), pp. 84–100.
- [7] James Honaker, Jonathan N Katz e Gary King. «An Improved Statistical Model for Multiparty Electoral Data». In: (2001).

- [8] James Honaker e Gary King. «What to do about missing values in time-series cross-section data». In: *American journal of political science* 54.2 (2010), pp. 561–581.
- [9] James Honaker, Gary King e Matthew Blackwell. «Amelia II: A program for missing data». In: *Journal of statistical software* 45 (2011), pp. 1–47.
- [10] Jonathan N Katz e Gary King. «A statistical model for multiparty electoral data». In: *American Political Science Review* 93.1 (1999), pp. 15–32.
- [11] Gary King, Michael Tomz e Jason Wittenberg. «Making the most of statistical analyses: Improving interpretation and presentation». In: *American journal of political science* (2000), pp. 347–361.
- [12] Gary King et al. «Analyzing incomplete political science data: An alternative algorithm for multiple imputation». In: *American political science review* 95.1 (2001), pp. 49–69.
- [13] Riccardo Jack Lucchetti. *Basic Econometrics*.
- [14] Michael Tomz, Joshua A Tucker e Jason Wittenberg. «An easy and accurate regression model for multiparty electoral data». In: *Political Analysis* 10.1 (2002), pp. 66–83.