Procedura di stima ai minimi quadrati misti. Uno studio sull’interscambio tra aree partitiche in Italia(*)

Agostino Tarsitano
Università degli studi della Calabria
Dipartimento di economia e Statistica
87030 Arcavacata di Rende (Cs)
agotar@unical.it

Riassunto.
Il lavoro discute l’applicazione degli stimatori misti alla determinazione della matrice di transizione fra aree partitiche in Italia. La tecnica degli stimatori misti ingloba le stime già disponibili ed i vincoli analitici come osservazioni aggiuntive proponendosi come valida alternativa ai minimi quadrati generalizzati. I risultati ottenuti nell’ambito del comportamento di voto degli italiani confermano che, in caso di inammissibilità delle stime delle transizioni ottenute con altri metodi, si può fare ricorso agli stimatori misti con risultati discreti.

Keywords: stimatori misti, mobilità elettorale, minimi quadrati generalizzati

(*) Lavoro presentato al convegno nazionale della Associazione Italiana di Sociologia: “Innovazione e regolazione sociale nelle società contemporanee” - Sezione metodologia. Trento, 7-10 ottobre, 1985.

1. lntroduzione

La classificazione delle unità di una popolazione secondo le categorie di due criteri (tabella a doppia entrata) è una situazione ricorrente nell'analisi dei dati. In questo lavoro si studia il caso in cui la classificazione delle unità si ripete regolarmente per ogni occasione *t* ed ogni volta si riescono a conoscere le frequenze assolute marginali di riga e di colonna w_{it} ed ovviamente il totale delle frequenze N_s, ma non è possibile osservare direttamente le frequenze delle celle interne della tabella. Restringiamo l'attenzione al caso un cui il criterio di classificazione di riga sia identico al criterio di classificazione adottato per le colonne. Al fine di poter giungere a congetture ragionevoli sulle frequenze incognite si ipotizza che l'allocazione delle unità nelle celle interne avvenga in base ad un processo markoviano del primo ordine. In pratica, si assume che la frazione ricadente in una data categoria nell'occasione t dipenda linearmente ed esclusivamente dalla ripartizione delle unità nella occasione precedente

$$w_{j,s+1} = \sum_{i=1}^{k} w_{is} r_{ij} \quad j = 1,2,\ldots,k \ ; \quad w_{is} = \frac{n_{is}}{N_s}; \quad r_{ij} \geq 0, \quad \sum_{j=1}^{k} r_{ij} = \sum_{j=1}^{k} w_{ij} = 1 \quad (1)$$

dove *k* è il numero di categorie e n_{is} è il numero di unità collocate nella categoria i-esima per l'occasione s-esima. La quantità r_{ij} rappresenta l'interazione tra le categorie *i* e *j* che è ipotizzata costante rispetto alle varie occasioni. Il problema consiste nel determinare la matrice delle interazioni $\boldsymbol{R}$.

In questi termini possono essere espressi diversi problemi: ad esempio, lo studio della concentrazione industriale (Hart-Prais, 1956); i flussi demografici (Rogers, 1967); i cambiamenti di voto (Miller, 1972); la mobilità della forza lavoro (Pollastri, 1976); la mobilità sociale (Shorroks, 1978).

Il piano del presente lavoro è il seguente: nel prossimo paragrafo si studierà la stima ai minimi quadrati generalizzati della matrice delle interazioni; nel paragrafo successivo se ne studierà la stima ai minimi quadrati misti. Nel paragrafo 4 gli stimatori misti verranno applicati alla verifica di varie ipotesi sulla struttura dell'interscambio tra aree partitiche in Italia.

2. Procura di stima ai minimi quadrati generali

La relazione (1.1) esprime un sistema ideale rigido che ha il suo cardine nella matrice delle interazioni; una volta che questa sia nota il sistema rivela tutte le sue caratteristiche. Le interazioni r_{ij} tuttavia sono incognite e debbono essere stimate in base alle sole manifestazioni osservate delle frazioni ricadenti nelle varie categorie. In questo paragrafo si discuterà della stima di R con il metodo dei minimi quadrati generalizzati.

Nel nostro contesto di studio la matrice $\boldsymbol{R}$ è costante ed ognuna delle n occasioni in cui si possono osservare le quote marginali ne fornisce una misurazione indiretta sebbene viziata da errori.

$$\boldsymbol{w_j} = \boldsymbol{W}\boldsymbol{r}_j + \boldsymbol{e}_j; \qquad j = 1, 2, \ldots k-1$$

$$\boldsymbol{w_j} = \begin{bmatrix} w_{j,1} \\ w_{j,2} \\ \vdots \\ w_{j,n} \end{bmatrix}; \boldsymbol{W} = \begin{bmatrix} w_{1,0} & w_{2,0} & \cdots & w_{k-1,0} \\ w_{1,1} & w_{2,1} & \cdots & w_{k-1,1} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1,n-1} & w_{2,n-1} & \cdots & w_{k-1,n-1} \end{bmatrix}; \boldsymbol{r_j} = \begin{bmatrix} r_{1,j} \\ r_{2,j} \\ \vdots \\ r_{k-1,j} \end{bmatrix}; \boldsymbol{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad (2)$$

La k-esima frazione e la k-esima interazione sono state eliminate in virtù del vincolo di somma unitaria posto nella (1).

Ipotizziamo che la frazioni $w_{j,s}$ provengano da una distribuzione multinomiale con media *qj,s, varianza* $q_{j,s}(1-q_{j,s})/N_s$ e covarianze $-q_{i,s}(1-q_{j,s})/N_s$. Inoltre, sul vettore degli errori $\mathbf{e}_j$ facciamo valere le seguenti relazioni

$$a.\ \boldsymbol{E}(\boldsymbol{e}_j) = \boldsymbol{0}; \qquad b.\ E(\boldsymbol{e}_j \boldsymbol{e}_i^t) = \sigma_i v_{ij}; \qquad c.\ E(\boldsymbol{W}^t \boldsymbol{e}_j) = \boldsymbol{0} \qquad (3)$$

dove la matrice $\boldsymbol{V}=\{v_{ij}\}$ nella b è diagonale con elementi positivi sulla diagonale stessa.

Le relazioni (2,1) possono essere espresse nella forma matriciale aggregata

$$\begin{bmatrix} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \\ \vdots \\ \boldsymbol{w}_{k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{W} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{W} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{W} \end{bmatrix} \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_{k-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_{k-1} \end{bmatrix} \qquad (4)$$

$$\boldsymbol{w} = \boldsymbol{A}\boldsymbol{r} + \boldsymbol{e} \quad con \quad \boldsymbol{A} = (\boldsymbol{I_{k-1}} \otimes \boldsymbol{W})$$

con i vincoli

$$\textbf{\textit{a. r}} \geq \textbf{\textit{0;}} \qquad \textbf{\textit{b. Br}} \leq \textbf{\textit{u}} \quad con \;\; \textbf{\textit{B}} = \left(\textbf{\textit{u}} \otimes \textbf{\textit{I}}_{\textbf{\textit{k-1}}}\right) \tag{5}$$

Il simbolo $\boldsymbol{I}_{k-1}$ denota la matrice iudentità di ordine *(k-1)* ed $\boldsymbol{u}$ è un vettore formato esclusivamente da “1”. Per il vettore degli errori risulta inoltre che

$$E(\boldsymbol{e}) = \boldsymbol{0}; \quad E\left(\boldsymbol{e}\boldsymbol{e}^{\boldsymbol{t}}\right) = \Omega; \quad E\left(\boldsymbol{A}^{\boldsymbol{t}}\boldsymbol{e}\right) = \boldsymbol{0} \tag{6}$$

dove Ω è una matrice quadrata a blocchi *(k-1)xn* in cui ciascun blocco ω_{ij} è una matrice diagonale così costituita

$$\varpi_{ij} = \begin{bmatrix} \dfrac{q_i\left(\delta_{ij} - q_i\right)}{N_1} & 0 & \cdots & 0 \\ 0 & \dfrac{q_i\left(\delta_{ij} - q_i\right)}{N_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dfrac{q_i\left(\delta_{ij} - q_i\right)}{N_n} \end{bmatrix}; \delta_{ij} = \begin{cases} 1 \;\; se \;\; i = j \\ 0 \;\; se \;\; i \neq j \end{cases} \tag{7}$$

In base a questa formulazione possiamo stimare le interazioni $\boldsymbol{r}$ -trascurando per il momento i vincoli (5) con la tecnica dei minimi quadrati generalizzati di Aitken ottenendo

$$\boldsymbol{r}^* = \left[\boldsymbol{A}^{\boldsymbol{t}}\Omega^{-1}\boldsymbol{A}\right]^{-1}\boldsymbol{A}^{\boldsymbol{t}}\Omega^{-1}\boldsymbol{w} \tag{8}$$

la cui matrice di dispersione è data da

$$\boldsymbol{E}\left\{\left[\boldsymbol{r}^* - \boldsymbol{r}\right]\left[\boldsymbol{r}^* - \boldsymbol{r}\right]^t\right\} = \left[\boldsymbol{A}^{\boldsymbol{t}}\Omega^{-1}\boldsymbol{A}\right]^{-1} \tag{9}$$

La matrice Ω è di solito incognita e deve essere stimata. La scelta più semplice è di sostituire alle proporzioni incognite $q_{j,s}$, quelle ottenute dalla procedura dei minimi quadrati ordinari (sperando che nessuna sia nulla) per poi applicare la procedura degli stimatori di Aitken (Lee et al. 1977, pp. 75-80).

Gli stimatori (8) godono delle usuali proprietà degli stimatori di Aitken, ma hanno il problema aggiuntivo del rispetto dei vincoli di ammissibilità. La relazione b della (5) non è molto cogente in quanto possiamo stimare le interazioni omesse imponendo il vincolo di somma unitaria per colonna alla matrice delle interazioni. Un discorso diverso va fatto per la condizione di non negatività. Data la proprietà di consistenza degli stimatori ai minimi quadrati generalizzati, all'aumentare del numero di osservazioni la probabilità di ottenere valori delle interazioni al di fuori dell'intervallo unitario diventa sempre più piccola. Tuttavia, questo non assicura che si abbiano sistematicamente valori ammissibili o che si possano ottenere per la particolare applicazione che si sta effettuando.

Per ottenere stimatori tali da verificare anche la a della (5) Telser (1963) proposto di sostituire ai valori fuori vincolo lo zero se l'interazione è negativa oppure l'uno se essa supera l'unità e di considerare queste assegnazioni come dei vincoli di uguaglianza da esaminare in successione. Altre possibilità sono il ricorso alla procedura dei minimi quadrati vincolati (Judge e Takayama, 1966; Liew, 1976); alla teoria dell'informazione (Theil e Rey, 1966); alla procedura di massima verosimiglianza basata sulla distribuzione multinomiale (Lee et al.,1977); ai modelli loglineari (Bishop et al., 1975), agli stimatori ai minimi quadrati "ridge" (Miller, 1972).

A proposito di tutte queste tecniche bisogna però aggiungere che:
a) Le loro proprietà campionarie ed asintotiche non sono facili a stabilirsi (Faliva, 1973).
b) Producono spesso valori nulli quasi altrettanto inverosimili di quelli negativi, oppure valori delle interazioni diagonali irrisori in applicazioni in cui la dominanza diagonale è un fatto accertato (Rogers, 1967).
c) Lo stimatore vincolato non necessariamente risulta più efficiente di quello ai minimi quadrati ordinari (Lowell e Prescott, 1970), (Liew, 1976), anzi, tende ad essere peggiore (Tarsitano,1984).

3. Procedura di stima ai minimi quadrati misti.

Per superare le stime eventualmente inammissibili dei minimi quadrati generalizzati e le difficoltà teoriche oltreché computazionali e di quelli vincolati la teoria statistica ha sviluppato con Theil-Goldberger (1960), Theil (1963), e Swamy-Metha (1969) la tecnica degli stimatori misti. Lo scopo di tale procedura è di incorporare le informazioni a priori (anche i vincoli analitici) come osservazioni addizionali rispetto a quelle già disponibili sulla dipendente e sui regressori.

Supponiamo che le informazioni a priori disponibili sulle interazioni $\boldsymbol{r}_j$ provengano da uno stimatore corretto e con matrice di dispersione finita

$$\boldsymbol{r}^+ = \boldsymbol{r} + \varepsilon; \qquad dove \quad E(\varepsilon) = \boldsymbol{0}, \quad E\left(\varepsilon\varepsilon^t\right) = \boldsymbol{\Psi} \tag{10}$$

dove $\boldsymbol{r}^+$ è il vettore delle stime a priori costituito perciò da elementi noti e tali da soddisfare le condizioni di ammissibilità e Ψ è una matrice definita positiva. Inoltre, dee valere la condizione $E(\boldsymbol{e}\varepsilon^t)=\boldsymbol{0}$L visto che la procedura mista ha come presupposto l'assenza di correlazione fra le due diverse fonti di informazione.

L'insieme delle informazioni, a priori e correnti, può essere riespresso come segue

$$\begin{bmatrix} \boldsymbol{w} \\ \cdots \\ \boldsymbol{r}^+ \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} \\ \cdots \\ \boldsymbol{I_{k-1}} \end{bmatrix} \boldsymbol{r} + \begin{bmatrix} \boldsymbol{e} \\ \cdots \\ \varepsilon \end{bmatrix}; \qquad \boldsymbol{z} = \boldsymbol{G}\boldsymbol{r} + \boldsymbol{c} \tag{11}$$

Per il vettore degli errori aggregati **c** risultano le seguenti relazioni

$$E(\boldsymbol{c}) = \boldsymbol{0}, \qquad E\left(\boldsymbol{c}\boldsymbol{c}^t\right) = \boldsymbol{D} = \begin{bmatrix} \Omega & 0 \\ 0 & \Psi \end{bmatrix} \tag{12}$$

Al modello (12) può essere applicata la tecnica dei minimi quadrati generalizzati già vista nel paragrafo precedente ottenendo

$$\begin{aligned} \boldsymbol{r}' &= \left[\boldsymbol{G}^t \boldsymbol{D}^{-1} \boldsymbol{G}\right]^{-1} \boldsymbol{G}^t \boldsymbol{D}^{-1} \boldsymbol{z} \\ &= [\boldsymbol{I} + \boldsymbol{H}]^{-1} \boldsymbol{r}^* + \left[\boldsymbol{I} + \boldsymbol{H}^{-1}\right]^{-1} \boldsymbol{r}^+ \quad con \;\; \boldsymbol{H} = \left[\Psi \boldsymbol{A}^t \Omega^{-1} \boldsymbol{A}\right]^{-1} \end{aligned} \tag{13}$$

Vediamo le proprietà del nuovo stimatore. Innanzitutto il (13) è uno stimatore corretto poiché $\mathbf{(I+H)^{-1}=I-(1+H^{-1})^{-1}}$implica che E(r')=r. Inoltre, si ha

$$\boldsymbol{E}\left\{\left[\boldsymbol{r}' - \boldsymbol{r}\right]\left[\boldsymbol{r}' - \boldsymbol{r}\right]^t\right\} = \boldsymbol{M}^t \boldsymbol{\Psi} \boldsymbol{H} \boldsymbol{M} [\boldsymbol{I} + \boldsymbol{H}]^{-1} + [\boldsymbol{I} - \boldsymbol{M}]^t \boldsymbol{\Psi} [\boldsymbol{I} - \boldsymbol{M}] \tag{14}$$

Questa relazione evidenzia che la dispersione dello stimatore misto dipende sia dalla struttura della variabilità nei dati osservati che dalla variabilità delle informazioni a priori.

Effetto delle informazioni a priori
E' ora necessario verificare se l'inserimento dei dati a priori disponibili ha reso più efficiente Io stimatore misto rispetto a quello ordinario. In questo paragrafo dimostreremo un teorema che garantisce, sotto condizioni abbastanza generali, il guadagno di efficienza dello stimatore misto.
Definiamo l'errore quadratico medio dello stimatore **r**' rispetto ad una qualsiasi matrice definita positiva **L** come

$$\boldsymbol{Eqm}(\boldsymbol{r}',\boldsymbol{L}) = \boldsymbol{E}\left\{\left[\boldsymbol{r}' - \boldsymbol{r}\right]\boldsymbol{L}\left[\boldsymbol{r}' - \boldsymbol{r}\right]^t\right\} \tag{15}$$

*Teorema. Se la matrice di varianze-covarianze delle informazioni a priori è della forma $\boldsymbol{\Psi}=\psi\boldsymbol{I}$ con $\psi>0$, allora lo stimatore misto ha errore quadratico medio sempre inferiore a quello ordinario: Eqm(**r'**,**L**)≤Eqm(**r***,**L**) qualunque sia L .*

La relazione *Eqm(**r'**,**L**)≤Eqm(**r***,**L**)* è equivalente alla condizione che la matrice ***H-M*** sia non negativa definita.

$$\begin{aligned} Eqm\left(\boldsymbol{r}^*,\boldsymbol{L}\right) - Eqm\left(\boldsymbol{r}^+,\boldsymbol{L}\right) &= \psi\left[\boldsymbol{H} - \boldsymbol{M}^t\boldsymbol{H}\boldsymbol{M} - \boldsymbol{I} + 2\boldsymbol{M} + \boldsymbol{M}^t\boldsymbol{M}\right] \\ &= \psi\left[\boldsymbol{H} - \boldsymbol{M}\boldsymbol{H}\boldsymbol{M} - \boldsymbol{M}^t\boldsymbol{M}\right] = \psi[H - M] \end{aligned} \tag{16}$$

Dato che ***M*** è positiva definita, perché ***H-M*** sia non negativa definita bisogna e basta (Rao, 1973, p.70) che le radici dell'equazione $|\mathbf{H}-\lambda\mathbf{M}|=0$ ovvero gli autovalori della matrice $\boldsymbol{M}^{-1}\boldsymbol{H} = (\boldsymbol{I}+\boldsymbol{H}^2)$ siano superiori all'unità e poiché $\boldsymbol{H}^2$ è sempre positiva definita il teorema risulta provato (cfr. Theobald, 1974). Il risultato era del resto scontato: l'aggiunta di informazioni incorrelate con quelle già disponibili non può far altro che migliorare (o, al peggio, lasciare invariata) l'efficienza dello stimatore ai minimi quadrati e gli stimatori misti indicano lo schema di aggiornamento di quelli ordinari quando siano disponibili le ulteriori informazioni.

L'ipotesi $\boldsymbol{\Psi}=\psi\boldsymbol{I}$ fornisce inoltre un mezzo per controllare ed orientare l'inserimento delle informazioni a priori, Infatti, ora

$$\boldsymbol{r}'(\psi) = \left[\boldsymbol{I} + \psi^{-1}\left(\boldsymbol{A}^t\Omega^{-1}\boldsymbol{A}\right)^{-1}\right]^{-1}\boldsymbol{r}^* + \left[\boldsymbol{I} + \psi^{-1}\left(\boldsymbol{A}^t\Omega^{-1}\boldsymbol{A}\right)^{-1}\right]^{-1}\boldsymbol{r}^+ \quad (17)$$

E' agevole verificare che

$$\underset{\psi\to\infty}{Lim}\ \boldsymbol{r}(\psi) = \boldsymbol{r}^*; \qquad \underset{\psi\to 0}{Lim}\ \boldsymbol{r}(\psi) = \boldsymbol{r}^+ \quad (18)$$

Se le informazioni a priori sono poco precise (cioè contengono elevata variabilità) lo stimatore misto si tende verso gli stimatori ordinari data la scarsa qualità di ciò che aggiunge ai dati già disponibili. All'aumentare della loro precisione ovvero all'aumentare della fiducia nelle informazioni a priori, gli stimatori misti si allontano dagli stimatori ordinari per avvicinarsi alle valutazioni a priori.

4. Applicazione all'interscambio fra aree partitiche in ltalia.

La teoria sviluppata nel precedente paragrafo viene ora applicata allo studio dei dati rilevati aIl'interscambio tra aree partitiche in Italia. In breve, si sono raggruppate le percentuali di suffragi ottenuti dai partiti in *k* aree partitiche ordinate secondo l'usuale criterio sinistra-destra e si sono posti in relazione i successi elettorali ottenuti in una consultazione con quelli già verificatisi nella consultazione precedente.

$$\boldsymbol{w_t} = \boldsymbol{R}\boldsymbol{w}_{t-1} + \boldsymbol{e}_t \quad (19)$$

La tecnica degli stimatori misti qui applicata si discosta in maniera netta da altri studi elettorali: Brown e Paine (1964), Fabris (1979), Tommasone (1971), Shadee e Corbetta (1964). L'obbiettivo rimane però invariato: stimare **R** in modo da avere risultati attendibili per i coefficienti di transizione. Poiché in questo ambito le osservazioni sono scarse assumono grande importanza le informazioni a priori. In questo paragrafo discuteremo in particolare quattro ipotesi a priori: normalizzazione, mobilità completa, nulla e parziale.

Normalizzazione

La normalizzazione parte dalle stime ottenute dalla procedura dei minimi quadrati generalizzati e ne elimina i valori fuori vincolo:

$$r^{+}_{ij} = \begin{cases} 0 & se \;\; r^{*}_{ij} < 0 \\ r^{*}_{ij} & se \;\; 0 \le r^{*}_{ij} \le 1 \\ 1 & se \;\; r^{*}_{ij} > 1 \end{cases} \tag{20}$$

Alla matrice così ottenuta si applica la procedura di riproporzionamenti iterativi riga-colonna descritta da Friedlander (1961) che ha lo scopo di ricondurre entro fissate somme di riga e di colonna gli elementi di una matrice. Nel nostro caso la somma per righe deve essere uno e quella per colonne uguale alla corrispondente somma delle transizioni stimate con i minimi quadrati.

Ipotesi di mobilità completa

In questo caso si ipotizza che non ci sia alcuna ragione a che l'elettorato di una certa area vi permanga oppure si sposti verso un' area partitica piuttosto che verso un'altra per cui le interazioni sono tutte eguali

$$r^{+}_{ij} = \frac{1}{k-1}, \;\; i,j = 1,2,\ldots,k-1 \tag{21}$$

Ipotesi di assenza di mobilità

L'elettorato si è attestato su posizioni stabili e vi permane

$$r^{+}_{ij} = \begin{cases} 1 & se \;\; i = j \\ 0 & se \;\; i \ne j \end{cases} \tag{22}$$

Ipotesi di mobilità parziale

E' questa una ipotesi articolata che racchiude varie considerazioni:

1) La matrice delle interazioni è a dominanza diagonale ovvero i coefficienti di stabilità, il cosiddetto zoccolo duro, sono maggiori o uguali alla somma delle altre interazioni sulla stessa riga.

2) La interazioni decrescono man mano che le categorie a cui sono associati si allontanano dal coefficiente di stabilità.

Un esempio di definizione della R rispondente a tale ipotesi è

$$r^{+}_{ij} = \begin{cases} 0.5 & se \;\; i = j \\ \dfrac{1}{2a_i\left(1+|i-j|\right)} & se \;\; i \ne j \end{cases}; \quad con \;\; a_i = \sum_{\substack{j=i \\ i \ne j}}^{k-1} \left(\frac{1}{1+|i-j|} \right) \tag{23}$$

Applicazione
Le osservazioni a cui si è applicata la procedura mista sono gli esiti disponibili, a livello nazionale, delle elezioni per la camera dei deputati, dal 1948 al 1987.

Tabella l: voti validi nelle consultazioni politiche (Camera dei Deputati).

Anno	DN [a]	DC+PPST	PLI	PRI	PSDI	PSI	PCI	Altre	Totale
1948	1'257'346	12'836'947	1'004'889	652'477	1'858'356	[b] 2'507'780	[b] 4'461'220	1'689'897	26'268'912
1953	3'438'409	10'987'074	816'627	437'988	1'223'251	3'441'305	6'121'922	626'167	27'092'743
1958	2'844'720	12'657'774	1'046'939	405'574	1'345'750	4'208'111	6'704'706	350'059	30'758'031
1963	2'108'178	11'911'428	2'143'954	420'419	1'876'409	4'257'300	7'768'228	272'115	30'758'031
1968	1'892'217	12'594'507	1'851'060	626'567	[c] 1'383'355	[c] 3'222'477	8'557'404	1'675'666	27'197'421
1972	2'896'762	13'076'034	1'297'105	954'597	1'717'539	3'210'427	9'072'454	1'189'861	33'414'779
1976	2'245'376	14'402'688	478'335	1'314'936	1'237'270	3'542'998	12'622'728	882'942	36'727'273
1979	1'930'639	14'251'189	712'646	1'110'209	1'407'535	3'596'802	11'139'231	2'523'057	36'671'308
1983	2'511'487	12'338'021	1'066'980	1'874'512	1'508'234	4'223'362	11'032'318	2'351'091	36'906'005
1987	2'282'256	13'443'210	810'216	1'429'628	1'140'910	5'505'690	10'254'591	3'725'882	38'592'383

(a) In questa colonna sono stati cumulati i voti del Msi e dei partiti monarchi ci .
(b) I due dati sono stati ottenuti separando i voti del Fr. Dem. Pop. in base alla proporzione che i due partiti hanno riportato nella consultazione del 1953
(c) I due dati sono stati ottenuti separando i voti del PSI-PSDI in base alla proporzione che risulta comemedia geometrica delle proporzioni che i due partiti hanno riportato nel 1963 e nel 1972.

Le aree partitiche che si sono definite sono: Des = (MSI, sigle monarchiche, PLI); Dc = (DC, SVP); Laici = (PSI,PSDI,PRI); Pci; Altre = (Voti validi-totale precedenti).

Tabella 2: proporzioni osservate per le cinque aree partitiche.

Anno	Des	Dc	Laici	Pci	Altre
1948	0.0512	0.5223	0.1430	0.1020	0.1815
1953	0.1299	0.4151	0.0936	0.1300	0.2313
1958	0.0974	0.4333	0.0958	0.1440	0.2295
1963	0.0692	0.3907	0.1457	0.1396	0.2548
1968	0.0628	0.4180	0.1282	0.1070	0.2940
1972	0.0899	0.4058	0.1232	0.0996	0.2815
1976	0.0626	0.4018	0.0645	0.0988	0.3522
1979	0.0565	0.4173	0.0946	0.1053	0.3262
1983	0.0727	0.3571	0.1288	0.1222	0.3193
1987	0.0655	0.3856	0.0970	0.1579	0.2941

Nei dati ci sono, come è facile intuire, molte nebbie, dovute sia alle aggregazioni (sempre effettuate con un certo margine di arbitrarietà) che all'arco di tempo estremamente ampio a cui si riferiscono. Questo rende difficile portare in piena luce le tendenze di fondo, ammesso che queste ci siano.

Nella tabella 3 sono riportati i risultati delle elaborazioni con la stima mista. In particolare, si riportano per ciascuna configurazione a priori, la matrice delle interazioni stimate, i valori del coefficiente di determinazione calcolato in base alla formula:

$$R^2 = 1 - \frac{(y - Gr)^t (y - Gr)}{y^t y} \tag{24}$$

(2.14) ed il valore di ψ che si è reso necessario affinché gli stimatori misti verificassero il vincolo di non negatività. Tali valori sono stati ottenuti fissando inizialmente ψ=1 e procedendo con successivi raddoppi o dimezzamenti in base alla ammissibilità o meno degli stimatori ottenuti con la procedura mista. L'ammissibilità è stata considerata soddisfacente se nessuna delle interazioni con segno negativo era, in valore assoluto, superiore a 0.001.

Tab.3 risultati delle elaborazioni

Procedura	ψ	R^2		Destra	Dc	Laici	Pci	Altre
Ordinari	0	*0.9730*	Destra	0.2063	0.7918	0.1722	0.3251	-0.5006
			Dc	0.3200	0.6100	0.0885	0.2559	-0.2751
			Laici	0.1230	-0.0067	0.4860	-0.2953	0.7427
			Pci	-0.2166	0.2479	0.0043	0.7000	0.2337
			Altre	0.1522	-0.0459	0.5161	0.1392	0.2381
Generalizzati	∞	*0.9725*	Destra	0.0815	0.8632	0.1703	0.3221	-0.4373
			Dc	0.3302	0.5928	0.0888	0.2615	-0.2735
			Laici	0.2105	-0.0731	0.4758	-0.3046	0.0012
			Pci	-0.2949	0.2660	0.0034	0.7017	0.2435
			Altre	0.1342	-0.0020	0.5705	0.1224	0.2547
Normaliz.	Non def.	*0.9495*	Destra	0.0313	0.6858	0.0030	0.1898	0.0000
			Dc	0.1584	0.5884	0.0005	0.1925	0.0000
			Laici	0.1664	0.0000	0.5348	0.0000	0.2986
			Pci	0.0000	0.2929	0.0631	0.5731	0.0708
			Altre	0.1054	0.0000	0.6374	0.1476	0.1093
Mob.Completa	2x10-10	*0.8170*	Destra	0.1801	0.2515	0.2050	0.2264	0.1368
			Dc	0.1145	0.3759	0.2131	0.2921	0.0043
			Laici	0.1550	0.2797	0.2052	0.2412	0.1186
			Pci	0.1074	0.3071	0.1996	0.2624	0.1232
			Altre	0.1873	0.2146	0.2010	0.2080	0.1887
Mob.Nulla	2x10-16	*0.9060*	Destra	0.9997	0.0000	0.0000	0.0000	0.0003
			Dc	0.0000	0.9995	0.0000	0.0002	0.0003
			Laici	0.0000	0.0000	0.9999	0.0000	0.0001
			Pci	0.0000	0.0000	0.0000	0.9999	0.0001
			Altre	0.0000	0.0000	0.0000	0.0000	1.0000
Mob. Parziale	2x10-11	*0.8611*	Destra	0.4882	0.2159	0.1279	0.1118	0.0559
			Dc	0.1180	0.5731	0.1512	0.1566	0.0009
			Laici	0.0876	0.1832	0.4958	0.1725	0.0696
			Pci	0.0397	0.1509	0.1495	0.5315	0.1292
			Altre	0.0736	0.1039	0.1298	0.1995	0.4930

Innanzitutto c'è da notare che ci sono ben pochi cambiamenti nel passare dalle stime ai minimi quadrati ordinari a quelli generalizzati; infatti ne risultano affetti solo il coefficiente di stabilità della "Destra" e l'interazione tra questa e la DC. Per il resto non sembra che il tenere conto della eteroschedasticità abbia portato, in termini di adattamento o di ammissibilità, ad apprezzabili miglioramenti.

La stima ai minimi quadrati generalizzati R presenta parecchi valori negativi segnatamente per le transizioni legate all'area denominata "Altre" e all'area dei cosiddetti "Laici", In questo caso la spiegazione è semplice: le aree partitiche così strutturate sono risultate di troppo eterogenea composizione per consentire una corretta stima dei parametri. Peraltro, il dato anomalo *-0.2949* ottenuto in corrispondenza della probabilità che un elettore passi dal PCI alla destra è quasi certamente dovuto all'effetto di retroazione del coefficiente di stabilità elevatissimo, 0.7017, che porta gli elettori del PCI in posizione remota dall'area di destra,

Tra le varie configurazioni a priori da noi esaminate, la normalizzazione si rivela una procedura molto efficace perché non solo produce valori ammissibili, ma anche l'R^2 non è troppo peggiorato. Rimangono però questionabili gli zeri che essa introduce in qualcuna delle transizioni. Talune corrispondono a situazioni ragionevoli: tra PCI e la Destra, tra le "Altre" e la DC; delle altre invece sono poco plausibili: l'interazione nulla tra l'area "Laici" e la DC e tra gli stessi Laici ed il PCI non sarebbero condivise da nessun esperto o solo attento di questioni politiche italiane. Per le situazioni estreme è netta l'incoerenza della ipotesi di mobilità nulla che infatti produce valori ammissibili solo con un valore di ψ molto prossimo allo zero, segno evidente che nessuna delle aree partitiche ha uno zoccolo duro così denso e pesante da trattenere tutti i propri voti.

L'ipotesi di mobilità completa è quella che dà minore spazio (valore più alto di ψ) alle informazioni esogene. Su tale risultato si riverbera la instabilità delle aggregazioni proposte e la loro mutevolezza nel corso di 40 anni di consultazioni elettorali. E' chiaro che gli assunti presenti nel modello della matrice di transizione e l'arco di tempo trascorso hanno aggiunto dei disturbi alle elaborazioni che rendono problematica l'emersione della struttura profonda delle interazioni tra le aree partitiche. Interessanti in questo senso sono i risultati relativi all'ipotesi di mobilità parziale in cui l'influenza delle stime esogene è solo leggermente superiore alla configurazione di mobilità completa, ma i valori sono molto più rispondente alle attese (anche tenuto conto del peggiore adattamento che ne risulta).

L'aspetto più difforme che il modello fa riscontrare rispetto ai dati campionari è l’uguaglianza tra i coefficienti di stabilità tra le diverse aree partitiche. E’ invece ben noto che quelli della DC e del PCI sono molto più grandi che nelle altre aree come del resto è emerso dai risultati finali di tutte le ipotesi a priori. Tuttavia, non sembra il caso di intervenire sulla definizione della mobilità parziale, piuttosto sarebbe opportuno ridefinire le aree bipartendo magari l’area Laica.

Conclusioni

Nel pur ricco panorama delle tecniche note per la stima della matrice delle interazioni nelle catene markoviane del primo ordine la letteratura aveva trascurato gli stimatori misti. In questo lavoro si è cercato di colmare la lacuna è si è potuto appurare che l’inserimento dei vincoli analitici sotto forma di osservazioni campionarie aggiuntive porta, in generale, ad uno stimatore più efficiente rispetto alla procedura dei minimi quadrati generalizzati. In particolare si è ricondotta la stima della matrice di transizione ad una equilibrata mediazione tra stima ordinaria ed informazioni esogene (vincoli inclusi). Inoltre, la mediazione stessa è governata da un unico parametro i cui valori permettono di valutare l’incidenza delle informazioni esogene sugli stimatori misti.

L’applicazione della tecnica nel contesto delle transizioni tra aree partitiche pur rivelando l’efficacia e la praticità della procedura mista di stima, non ha prodotto risultati realmente validi per le previsioni elettorali. La causa, ovviamente, non è da ricercarsi nella tecnica di stima, ma nelle difficoltà oggettive di definire aree partitiche abbastanza coese da reggere attraverso 40 anni di vita elettorale italiana.

Bibliografia

Bishop Y.M.M., Fienberg S.E., Holland P.W. (1975): Discrete Multivariate Analysis. Theory and practice. The MIT Press, Cambridge, Mass.

Brown P., Payne C. (1984): Forecasting the 1983 British General Election.*The Statistician*. Vol. 33, pp. 217-228

Fabris L. (1979): Metodi statistici per le previsioni elettorali. *Rivista di statistica applicata*. Vol.12, pp. 151-170.

Faliva M. (1973): Stimatori lineari ottimali dei parametri in un modello di regressione lineare con vincoli di disuguaglianza sni parametri. *Statistica*. Vol. 33, pp. 383-394.

Friedlander D. (1961): A Technique for Estimating a Contingency Table Given the Marginal Totals and Some Supplementary Data. *Journal of the Royal Statistical Society, A,* Vol.24, pp.412-420.

Hart P.E., Prais S. (1956): The Analysis of Business Concentration: a Statistical Approach. *Journal of the Royal Statistical Society , Serie A*. VoL.19, pp. 150-175.

Judge G.G.,Takayama T. (1966): Inequality Restrictions in Regression Analysis. *Journal of America Statistical Association*. Vol.61, pp.166-181.

Lee T.C., Judge G.G., Zellner A. (1977): Estimating the Parameters of Markov Probabily Model from Aggregate Time Series Data, North-Holland, Amsterdam.

Liew C.K. (1976): Inequality Constrained Least Squares Estimation. *Journal of American Statistical Association*. Vol.71, pp.746-751.

Lowell M.C., Prescott F. (1970): Mutiple Regression with Inequality Constraints: Pretesting Bias, Hypothesis Testing and Efficiency. *Journal of American Statistical Association*. Vol. 65, pp. 913-925.

Mehta J.S., Swamy P.A. V.B. (1970): The Finite Sample Distribution of Theil's Mixed Regression Estimator and a Related Problem. *Review of lnternational Statistical lnstitute*. Vol. 38, pp. 202-209.

Miller W.L. (1972): Measores of Electoral Cbange using Aggregate Data. *Journal of the Royal Statistical Society, A*. Vol.35, pp.122-142.

Pollastri A.(1976): La stima della matrice di transizione nella analisi della forza lavoro in Italia. *Rivista di Statistica Applcata*. Vol. 9, pp.17-47.

Neudecker H. (1968): The Kronecker Product Matrix and Some of Its Application in Econometrics. *Statistica Neerlandica*. Vol. 22, pp. 69-82.

Rogers A. (1967): Estimating Interregional Population and Migration Operators from Interregional Population Distributions. *Demography*. Vol.

4, pp.515-531.
Shadee H.M.A., Corbetta P.G.(1984): Metodi e modelli di analisi dei dati elettorali, Bologna, Il Mulino.
Shorrocks A.F. (1978): The Measurement of Mobility. *Econometrica*. Vol46, pp. 1013-1024.
Tarsitano A. (1984): Regressioni lineari con vincoli: uno studio di simulazione sugli stimatori misti a confronto con gli stimatori vincolati. In Scritti in onore di Francesco Brambilla. Vol.2, Edizioni Bocconi Comunicazione, Milano.
Theil H. (1963): On the Use of Incomplete Priori in Regression Analysis. *Journal of the American Statistical Association*. Vol 58, pp. 401-414.
Theil H., Goldberger A.S. (1960): On Pure and Mixed Statistical Estimation in Econometrics. *International Statistical Review*. Vol. 2, pp. 65-78.