

**STIMA PRECOCE DEI CONSUMI PRIVATI DI  
CONTABILITA' NAZIONALE MEDIANTE  
MODELLI STRUTTURALI DINAMICI**

Fabio Corradi\* Giuseppina Guagnano\*\*

Rapporto di ricerca n.20

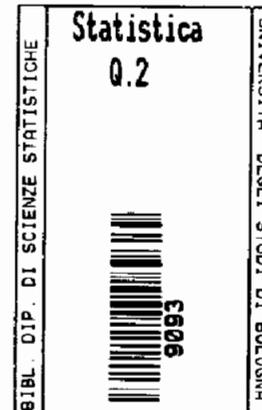
CON PRI - La misura dei consumi privati

I lavori raccolti in questa collana hanno avuto origine nell'ambito del progetto di ricerca dell'ISTAT «Le statistiche dei consumi privati nel sistema statistico nazionale» e del progetto di ricerca MURST 40% «La misura dei consumi privati: uno studio sull'accuratezza, coerenza e qualità dei dati». Al progetto di ricerca hanno partecipato i ricercatori dell'ISTAT e dei seguenti Dipartimenti e Istituti universitari:

- Dipartimento di Scienze Statistiche, Bologna
- Dipartimento di Contabilità Nazionale, Roma
- Dipartimento Statistico, Firenze
- Istituto di Statistica e Matematica, Istituto Universitario Navale, Napoli
- Dipartimento di Scienze Statistiche, Perugia
- Istituto di Statistica, Messina.

\* Istituto di Statistica - Università di Torino

\*\* Dipartimento di Scienze Statistiche - Università di Padova



Dipartimento di Scienze Statistiche "Paolo Fortunati"  
dell' Università degli Studi di Bologna

Ottobre 1993



## INDICE

|  |       |
|--|-------|
| 1. Premessa  | p. 5  |
| 2. Modelli strutturali dinamici per previsioni <i>k</i> passi avanti   | p. 6  |
| 2.1 Previsioni condizionate  | p. 9  |
| 3. Impiego di Modelli Lineari Dinamici univariati e bivariati per la previsione di alcune serie trimestrali dei consumi di Contabilità Nazionale | p. 12 |
| 3.1 Specificazione dei modelli   | p. 13 |
| 3.2 Le serie utilizzate  | p. 13 |
| 3.3 Analisi dei risultati  | p. 14 |
| 4. Sintesi conclusiva  | p. 16 |
| <i>Note</i>  | p. 17 |
| <i>Riferimenti bibliografici</i>   | p. 18 |
| <i>Appendice</i>   | p. 20 |
| A1. Varianza dello stato <i>k</i> passi in avanti nel modello univariato qualora si adotti il <i>discount approach</i>                           | p. 20 |
| A2. Media dello stato <i>k</i> passi avanti nel modello condizionato   | p. 21 |
| A3. Varianza dello stato <i>k</i> passi in avanti nel modello condizionato   | p. 22 |
| A4. Tabelle e grafici  | p. 24 |

## 1. Premessa

Nell'ambito del progetto CON.PRI<sup>1</sup>, rilievo è stato riservato alle discrepanze fra le stime dei consumi di Contabilità Nazionale (CN) e quelle desumibili dall'indagine sui Bilanci di Famiglia (BF). Varie cause di tali divergenze sono state individuate (Mantegazza e Tassinari, 1991), essenzialmente riconducibili a differenti definizioni degli aggregati di consumo, differenze nella popolazione di riferimento e nei metodi di rilevazione e stima.

Il punto di divergenza più rilevante risulta quello relativo ai metodi di rilevazione e stima, essendo le differenze nelle popolazioni di riferimento e nelle definizioni controllabili, almeno in linea di principio. La metodologia adottata nell'indagine campionaria BF è, infatti, propriamente statistica e fondata sul metodo della spesa, mentre il criterio di computo dei consumi di CN utilizza una pluralità di fonti e di indicatori di riferimento, fra cui gli stessi dati BF, e si basa sul metodo della disponibilità, almeno per la maggior parte dei capitoli di spesa.

Dalle differenze metodologiche appena richiamate discende anche la diversa cadenza temporale delle due serie: i dati di BF sono infatti immediatamente disponibili ogni trimestre, mentre i dati trimestrali di CN vengono calcolati *ex-post*, sulla base del dato annuale (Monducci, 1991). Si manifesta così l'esigenza di disporre di stime trimestrali contemporanee di CN. Tale obiettivo può essere perseguito modellando le serie CN mediante modelli strutturali dinamici. Il problema si configura come una previsione in cui si voglia utilizzare, se conveniente, il contenuto informativo di BF ai fini della stima di CN.

Nel § 2 verranno presentati tre tipi di modelli per la previsione dei dati CN: il primo fa riferimento a modelli dinamici lineari univariati, chiusi all'informazione esterna dei dati BF; gli altri due, invece, a modelli dinamici aperti a tale informazione, che viene inclusa o come regressore o come serie condizionante. In questo modo è possibile valutare se e come l'inclusione di BF migliori la capacità previsiva su CN. Nel § 3 verranno discussi i criteri operativi seguiti nello studio delle serie relative ai consumi alimentari e verranno presentati alcuni risultati. Nel § 4, infine, verranno tratte alcune principali conclusioni.

## 2. Modelli strutturali dinamici per previsioni $k$ passi avanti

Nel caso univariato il modello dinamico lineare (DLM) proposto da Harrison e Stevens (1976), è:

$$y_t = F_t' \theta_t + v_t, \quad v_t \sim N(0, V) \quad (1)$$

$$\theta_t = G_t \theta_{t-1} + w_t, \quad w_t \sim N(0, W_t) \quad (2)$$

ove  $\theta_t$  è un vettore di  $p \times 1$  parametri di stato,  $F$  è un vettore  $p \times 1$  contenente valori noti di variabili indipendenti,  $G$  è la matrice  $p \times p$  di evoluzione dello stato, e valgono le usuali assunzioni di indipendenza seriale e reciproca per le componenti aleatorie  $v_t$  e  $w_t$ .

In un modello univariato, chiuso ad informazioni esterne alla serie considerata, il vettore  $F_t$  è costituito da variabili indicatore che attivano o meno le corrispondenti componenti del vettore di stato. In un modello lineare dinamico con regressori, il vettore  $F_t$  comprende anche le osservazioni al tempo  $t$  delle variabili di regressione; i corrispondenti elementi di  $\theta_t$  sono i coefficienti di regressione ad esse relativi.

Sulla base dell'ultima informazione disponibile  $D_{t-1}$ , le distribuzioni  $k$  passi in avanti per lo stato e per le previsioni risultano,  $\forall t$  e per  $k \geq 1$ :

$$\text{distribuzione dello stato: } (\theta_{t-1+k} | D_{t-1}, V) \sim N(a_{t-1}(k), VR_{t-1}(k)), \quad (3)$$

$$\text{distribuzione delle previsioni: } (y_{t-1+k} | D_{t-1}, V) \sim N(f_{t-1}(k), VQ_{t-1}(k)), \quad (4)$$

con momenti definiti ricorsivamente da:

$$f_{t-1}(k) = F_{t-1+k}' a_{t-1}(k) \quad (5)$$

$$Q_{t-1}(k) = F_{t-1+k}' R_{t-1}(k) F_{t-1+k} + 1, \quad (6)$$

$$\text{dove: } a_{t-1}(k) = G_{t-1+k} a_{t-1}(k-1), \quad (7)$$

$$R_{t-1}(k) = G_{t-1+k} R_{t-1}(k-1) G_{t-1+k}' + W_{t-1+k}, \quad (8)$$

con valori iniziali  $a_{t-1}(0)$  e  $R_{t-1}(0)$  eguali ai corrispondenti momenti,  $m_{t-1}$  e  $C_{t-1}$ , della distribuzione di  $\theta$  aggiornata al tempo  $t-1$ . Se la matrice  $G$  è costante nel tempo, come avviene nel caso di modelli strutturali, la (7) può essere riscritta come

$$a_{t-1}(k) = G^k a_{t-1}(0) \quad (7 \text{ bis})$$

e quindi il valore atteso delle previsioni diventa:

$$f_{t-1}(k) = F_{t-1+k}' G^k m_{t-1}, \quad (9)$$

dove emerge più chiaramente come la previsione venga effettuata iterando l'equazione di stato senza operare, ad ogni passo, l'usuale aggiornamento di  $\theta$  condizionatamente alle osservazioni. In questo modo viene implicitamente assunta l'ipotesi di stabilità delle componenti del vettore di stato, la cui ultima stima aggiornata è supposta una buona approssimazione dello stato anche per il periodo di previsione. L'incertezza aumenta al crescere dell'intervallo previsivo e se ne tiene conto nel modello aggiungendo la componente aleatoria  $w$ , di varianza  $W$ , ad ogni iterazione della (8).

Un'ulteriore specificazione della (8) può essere data, seguendo il *discount approach* (Ameen e Harrison, 1985), esprimendo la varianza  $W_t$  come una frazione costante di  $C_{t-1}$ . Questo approccio suggerisce di impiegare diversi fattori di sconto ( $d$ , con  $0 \leq d \leq 1$ ) per le differenti componenti dello stato. Posto  $\delta = (d^j - 1)$ , i fattori  $\delta$  esprimono la frazione di incremento di variabilità, o di perdita di informazione, relativamente alle singole componenti dello stato fra due tempi contigui. Essi vengono applicati alle sottomatrici di varianze e covarianze delle corrispondenti singole componenti dello stato. Generalmente, si ipotizzano valori più piccoli di  $\delta$  per le componenti di regressione e di stagionalità che per quelle di trend e di livello, considerate meno stabili nel tempo.

Considerando, ad esempio, un modello connotato da trend lineare (costituito, quindi, da due componenti: livello e crescita), un regressore e stagionalità trimestrale (rappresentata in forma di Fourier e quindi da tre coefficienti), sarà utile definire la matrice  $\Delta = \text{blockdiag}[\delta_T U_2, \delta_R U_1, \delta_S U_3]$ , essendo  $U_i$  una matrice di 1 di dimensione  $i \times i$ . Esprimendo  $W_t = \Delta * C_{t-1}$ , dove l'operatore  $*$  indica il prodotto di Hadamard, la (8) può essere riscritta (vedi l'Appendice, A1) nel modo seguente:

$$R_{t-1}(k) = \sum_{j=0}^k \binom{k}{j} \Delta^{[j]} * G^{k-j} C_{t-1} G^{j, k-j}, \quad (8 \text{ bis})$$

con  $\Delta^{[j]}$  indicante il prodotto di Hadamard di  $j$  termini  $\Delta$ , per cui:  $\Delta^{[n]} = \text{blockdiag}[\delta_T^n U_2, \delta_R^n U_1, \delta_S^n U_3]$  e  $\Delta^{[0]}$  è una matrice di 1 diagonale a blocchi.

Si noti che la varianza dello stato  $k$  passi in avanti è completamente predeterminata e risulta crescente all'aumentare di  $k$ , noto il valore dell'ultimo aggiornamento  $C_{t-1}$  e fissato il valore dei coefficienti di sconto in  $\Delta$ . La varianza di previsione  $Q_{t-1}(k)$  definita dalla (6) risulta pertanto non decrescente nel caso di modelli chiusi, mentre, in presenza di regressori nel vettore  $F$ , i suoi valori dipenderanno anche da questi ultimi e quindi non sarà più possibile stabilire, in generale, se il suo andamento sia crescente o meno nel tempo.

Un metodo alternativo di previsione, che sfrutti la conoscenza relativa ad altre serie simili rispetto a quella da prevedere, può essere proposto nell'ambito dei modelli dinamici lineari a componenti comuni (DLMCC) (Quintana, 1985, 1987). Essi sono definiti come:

$$y'_t = F'\Theta'_t + v'_t \quad v_t \sim N(0, V) \quad (10)$$

$$\Theta'_t = G\Theta'_{t-1} + \Omega'_t, \quad \Omega_t \sim {}_m N(0, V \otimes W_t) \quad (11)$$

dove  $\Theta_t$  è la matrice  $q \times p$  composta di  $q$  vettori di stato relativi a ciascuna delle serie osservate, mentre  $F$  è un vettore  $p \times 1$ , comune a tutte le serie; la matrice  $\Omega_t$   $q \times p$  è distribuita secondo una *normal matrix-variate* (Dawid, 1981). Oltre alle assunzioni relative agli errori, richiamate nel modello definito dalle (1) e (2), si assume che i vettori riga costituenti  $\Theta_t$  siano distribuiti secondo una normale multivariata con matrice di varianza e covarianza,  $\sigma_{ii}R_i$ , identica a meno del fattore  $\sigma_{ii}$  che esprime la varianza delle serie osservate.

Le principali conseguenze risiedono nell'esigenza di operare con serie per le quali una modellizzazione in termini delle stesse componenti non sia in contraddizione con le conoscenze disponibili sulle serie medesime. Ad esempio, non sarebbe possibile operare con un DLMCC su serie aventi differente cadenza stagionale. Inoltre, la circostanza che  $R_i$  sia comune a tutte le serie fa sì che il filtro di Kalman, diverso per le singole componenti dello stato, sia anch'esso comune alle diverse serie.

Sia nel modello univariato, che nel DLMCC, si è finora considerata nota la varianza  $V$  delle osservazioni, condizionatamente alla quale sono state derivate le distribuzioni dello stato,  $k$  passi in avanti, le quali, date le assunzioni sugli errori, risultano normali. Spesso la varianza  $V$  deve essere stimata, in quanto incognita. Per i modelli definiti dalle (1) e (2) e dalle (10) e (11), è possibile effettuare un'analisi

coniugata esatta per la distribuzione della varianza delle osservazioni (Quintana e West, 1987). Questa possibilità costituisce il principale vantaggio dell'impiego del DLMCC in un contesto multivariato, sebbene le condizioni da esso imposte siano abbastanza restrittive. La varianza delle osservazioni viene modellata tramite una *gamma inversa*, ovvero una *Wishart inversa* nel DLMCC, la quale, risultando coniugata alla *Normale*, è aggiornata assieme al vettore di stato. Da risultati sviluppati in ambito Bayesiano (Dawid, 1981), le distribuzioni marginali dello stato e delle osservazioni risultano delle  $t$ , rimanendo tuttavia immutate le equazioni di aggiornamento dello stato da cui si ricavano le (3) e (4) e la (12), nelle quali alla varianza  $V$ , finora supposta nota, viene sostituita la sua stima aggiornata. Nel caso multivariato, l'aggiornamento della  $V$  riguarda anche le covarianze fra le serie che assumono particolare rilievo nel modello condizionato. L'aggiornamento si basa sugli errori di previsione registrati per le singole serie, per cui, a partire da un *a priori* ( $S_0, n_0$ ), l'aggiornamento risulta essere (West e Harrison, 1989; Quintana, 1985, 1987):

$$S_t = n_t^{-1}(n_{t-1}S_{t-1} + e_t e_t' / Q_t). \quad (18)$$

### 2.1. Previsioni condizionate

Con riferimento al problema previsivo, è possibile utilizzare i risultati ottenuti da Corradi e Guagnano (1992), relativamente alla stima di dati mancanti nell'ambito dei DLMCC<sup>2</sup>.

Per il caso bivariato, in cui la serie  $y_2$  non viene osservata al tempo  $t$  e quindi  $D_t = \{D_{t-1}, y_{1,t}\}$ , si ha che, date

$$p(\theta_{1,t}, \theta_{2,t} | D_{t-1}, V) = N(a_{1,t}, a_{2,t}; V \otimes R_t) \text{ e } V = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}, \quad (12)$$

l'aggiornamento sui parametri relativi a  $y_2$  risulta:

$$p(\theta_{2,t} | D_t, V) = N(m_{2,t}^* = a_{2,t} + b(m_{1,t} - a_{1,t}); C_{2,t}^* = V^* + b^2 \sigma_{11} C_t), \quad (13)$$

e la distribuzione del dato mancante:

$$p(y_{2,t} | D_t, V) = N(f_{2,t}^* = F' m_{2,t}^*; F' C_{2,t}^* F + \sigma_{22}) \quad (14)$$

dove:

$$b = \frac{\sigma_{12}}{\sigma_{11}}, \quad V^* = \sigma_{22}R_t - \sigma_{21}R_t\sigma_{11}^{-1}R_t^{-1}R_t\sigma_{12} = R_t(\sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}) = R_t\sigma_{22.1}. \quad (15)$$

La caratteristica fondamentale di tale approccio è che esso fornisce una metodologia per l'aggiornamento del vettore di stato  $\theta_{2,t}$  anche in assenza di osservazioni. Considerando la (13) si può notare che tale aggiornamento dipende da quello effettuato sul vettore di stato  $\theta_{1,t}$  in base all'osservazione  $y_{1,t}$  e dal legame lineare fatto registrare fra gli errori di previsione delle serie, espresso da  $b = \sigma_{12}/\sigma_{11}$ . In particolare, essendo  $m_{1,t} - a_{1,t} = K_t e_{1,t}$  con  $K_t = R_t F_t / Q_t$  comune alle due serie,  $m_{2,t}^*$  della (13) si può esprimere come:  $m_{2,t}^* = a_{2,t} + K_t b e_{1,t}$ . Rispetto all'aggiornamento standard, la differenza è nel sostituire l'errore di previsione  $e_{2,t}$ , non osservabile, con la stima ottenibile dalla regressione rispetto ad  $e_{1,t}$ , filtrando poi, come di consueto, tramite il filtro di Kalman  $K$ . In questo modo, l'errore commesso nel prevedere la serie  $y_1$  viene trasferito separatamente sulle componenti dello stato.

Iterando il procedimento di aggiornamento proposto nella (13), è possibile determinare per induzione i momenti delle distribuzioni dello stato,  $k$  passi in avanti (vedi Appendice). Avremo pertanto:

$$a_{2,t-l}^*(k) = G^k a_{2,t-l}(0) + b \sum_{j=1}^{k-1} G^j (m_{1,t-l+k-j} - a_{1,t-l+k-j}) \quad (16)$$

$$m_{2,t-l}^*(k) = G^k a_{2,t-l}(0) + b \sum_{j=0}^{k-1} G^j (m_{1,t-l+k-j} - a_{1,t-l+k-j}) \quad (16 \text{ bis})$$

$$R_{2,t-l}^*(k) = \sigma_{22} \sum_{j=0}^k \binom{k}{j} \Delta^{[j]} G^{k-j} C_{t-l} G^{j-k} + \quad (17)$$

$$- \frac{\sigma_{12}^2}{\sigma_{11}} \sum_{x=1}^{k-1} \sum_{j=0}^{k-x} \binom{k-x}{j} \Delta^{[j]} G^{k-x-j} (R_{t+x-l} - C_{t+x-l}) G^{j-k-x}$$

$$C_{2,t-l}^*(k) = \sigma_{22} \sum_{j=0}^k \binom{k}{j} \Delta^{[j]} G^{k-j} C_{t-l} G^{j-k} + \quad (17 \text{ bis})$$

$$- \frac{\sigma_{12}^2}{\sigma_{11}} \sum_{x=1}^k \sum_{j=0}^{k-x} \binom{k-x}{j} \Delta^{[j]} G^{k-x-j} (R_{t+x-l} - C_{t+x-l}) G^{j-k-x}.$$

Si osservi come entrambi i momenti, primo e secondo, dello stato siano costituiti di due parti. La prima di queste è uguale alla stima del corrispondente momento del modello univariato chiuso (cfr. (7 bis) e (8 bis)), sempre che la parametrizzazione risulti la medesima<sup>3</sup>. La seconda componente, che per i momenti a priori si attiva soltanto<sup>4</sup> per  $k > 1$ , rappresenta l'effetto del condizionamento operato rispetto all'altra serie e quindi esprime il contributo informativo fornito dalla procedura proposta. Nella stima della media, tale informazione è costituita dalla somma dei guadagni valutati per la media di  $\theta_1$   $\forall t$ , proiettati in avanti tramite opportune potenze della matrice  $G$  di transizione e trasferiti alla media di  $\theta_{2,t}$  tramite il coefficiente  $b$ . Nella stima della varianza, l'informazione aggiuntiva, indotta dal condizionamento a  $y_1$ , produce una riduzione, dipendente dalle riduzioni di variabilità registrate per  $\theta_1$  ad ogni tempo; anche queste vengono proiettate al tempo  $t+k$  tramite la matrice  $G$  ed opportune potenze della matrice  $\Delta$ , per essere trasferite infine alla varianza di  $\theta_2$  mediante il coefficiente  $\sigma_{12}^2/\sigma_{11}$ .

Si può concludere, pertanto, che il modello condizionato produce stime derivate dal modello univariato chiuso, dove l'elemento di differenziazione è analiticamente individuabile, dipendendo esclusivamente dall'evoluzione del modello stimato per la serie condizionante e dal legame accertato in precedenza fra quest'ultimo e il modello condizionato.

Se si considera il modello di regressione, con regressore uguale alla serie condizionante del precedente modello, confronti analitici tra i corrispondenti momenti dello stato non sono più possibili, essendo differente la struttura parametrica: in questo modello, infatti, il vettore di stato  $\theta_2$  include anche il coefficiente di regressione. Confronti sono possibili unicamente per quanto riguarda le prestazioni dei modelli previsivi. Affinché tali confronti siano coerenti, i due modelli aperti dovranno utilizzare lo stesso insieme di informazioni; pertanto, essendo le previsioni  $y_{2,t}(k)$  del modello di regressione funzione delle osservazioni  $y_{1,t+k}$  nel modello condizionato, a differenza del modello chiuso, le previsioni non andranno determinate sulla base dei momenti a priori (eqq. (16) e (17)), ma in funzione dei momenti a posteriori (eqq. (16bis) e (17bis)).

### 3. Impiego di Modelli Lineari Dinamici univariati e bivariati per la previsione di alcune serie trimestrali dei consumi di Contabilità Nazionale

Ricordando le caratteristiche delle serie di CN, l'analisi successiva consiste nella ricerca del "miglior" modello previsivo che fornisca, all'inizio di ogni anno, stime precoci per le osservazioni dei quattro trimestri, disponibili usualmente soltanto alla fine dell'anno.

Il problema può essere schematizzato nel seguente modo. Si indichi con  $t$  il quarto trimestre di un dato anno, in corrispondenza del quale si rendano disponibili le quattro osservazioni della serie CN per l'anno stesso; l'insieme di informazione diviene allora:

$D_t = \{y_{1,t}, y_{2,t}\}$ , comprendente, cioè, tutte le osservazioni disponibili rispettivamente per le serie BF e CN fino al tempo  $t$ . Ciò è vero  $\forall t+4n$ ,  $n \in N^+$ , mentre in generale  $D_{t+k} = \{y_{1,t+k}, y_{2,t}\}$ , con  $k=1,2,3,4$ .

Per ottenere stime precoci per i valori  $y_{2,t+k}$  si possono effettuare, per ogni anno, previsioni *k-passi-avanti*, sulla base delle informazioni  $D_{t+k}$ .

Durante i quattro passi previsivi di ogni anno, non è possibile effettuare gli aggiornamenti sul vettore di stato, salvo che, indirettamente, nel modello condizionato. Alla fine del quarto trimestre, invece, essendo disponibili le quattro osservazioni dell'anno, si possono rideterminare, a partire dal tempo  $t$ , in corrispondenza del quale è stato effettuato l'ultimo aggiornamento, i momenti a posteriori dello stato fino al tempo  $t+4$ , mediante i due modelli *1 passo avanti*, rispettivamente, univariato chiuso e con regressore. Al primo dei due, in particolare, fanno riferimento gli aggiornamenti all'istante  $t+4$  per il modello condizionato e per quello chiuso *4 passi avanti*; mentre, al secondo, gli aggiornamenti per il modello di regressione previsivo. A questo punto è possibile procedere ad un nuovo ciclo previsivo per i quattro trimestri dell'anno successivo.

Una volta completato il procedimento previsivo, si può effettuare un confronto tra le prestazioni dei tre modelli *k passi avanti* (chiuso, con regressore e condizionato) ed il DLM chiuso *1 passo avanti* preso a riferimento. Anzitutto, si può procedere alla valutazione degli errori di previsione in valore assoluto, in rapporto ai dati osservati.

Successivamente, si può calcolare la MSE sull'intera serie di errori relativi (Tab. 1, in Appendice). In tal modo è possibile confrontare le prestazioni dei modelli per differenti serie di consumi. Per valutare, infine, in che misura le informazioni esterne possano migliorare le prestazioni dei modelli chiusi, sono stati considerati i seguenti indici di miglioramento:  $(MSE_A - MSE_B)/MSE_B$ , dove  $MSE_A$  è relativa, alternativamente, ad uno dei due modelli univariati chiusi, mentre  $MSE_B$  ad uno dei due modelli aperti (cfr. Tab. 2, Appendice A4).

#### 3.1. Specificazione dei modelli

Per consentire la massima confrontabilità dei risultati ottenibili tramite i diversi modelli considerati, sono stati seguiti criteri uniformi.

Si è tenuto conto della natura stagionale dei consumi e della circostanza che i dati sono espressi a prezzi correnti e quindi presentano un andamento evolutivo; le componenti strutturali dei modelli includono, pertanto, trend lineare e stagionalità, ai quali andrà eventualmente aggiunta una componente regressiva.

Le due matrici di varianza e covarianza  $V$  e  $W_t$ , incognite a priori, sono stimate, rispettivamente, secondo un'analisi coniugata esatta e secondo il *discount approach*, descritti nel § 2. Per i coefficienti  $\delta$  corrispondenti alle diverse componenti dello stato, sono stati assunti i seguenti valori (cfr. West e Harrison, 1989), uguali per tutti i modelli<sup>6</sup>:  $\delta_t=0.11$  per le componenti livello e crescita del trend,  $\delta_s=0.04$  per la componente stagionale,  $\delta_r=0.02$  per la componente regressiva. La scelta operata può aver limitato la capacità previsiva di alcuni modelli, per i quali sarebbe stata opportuna una diversa aderenza degli aggiornamenti delle stime ai dati, rispetto a quella determinata dai fissati fattori di sconto.

Per la specificazione dei *priors*, infine, si è deciso di effettuare una *Reference Analysis* (Pole e West, 1989), preliminare alla stima dei modelli.

#### 3.2. Le serie utilizzate

Le serie dei consumi prese in esame sono quelle trimestrali della classificazione a 50 voci della Contabilità Nazionale, per il periodo 1970:1 - 1990:4. Prevalentemente, esse fanno riferimento al comparto alimentare, per il quale è più facilmente ipotizzabile una situazione di interdipendenza tra le serie BF e CN (Monducci, 1991); tuttavia, sono

state analizzate anche alcune voci di consumo non alimentare<sup>7</sup>.

Da un'analisi descrittiva delle serie sono emerse alcune caratteristiche che possono essere così sintetizzate:

- i dati di CN sono, in generale, più elevati di quelli di BF, ad eccezione delle serie relative ai capitoli *Pane, Olii, Caffè, Bevande alcoliche e Bevande analcoliche* per il comparto alimentare, e *Abitazioni e Trasporti* per quello non alimentare;
- le due fonti forniscono serie di dati meno differenti per i capitoli *Patate e Frutta*, tra i beni alimentari, e *Vestiario e Calzature*, tra quelli non-alimentari;
- la riduzione delle differenze tra le corrispondenti serie dei due gruppi a partire dal primo trimestre 1985 può essere stata causata dall'introduzione, in tale anno, di alcune modifiche nell'indagine BF (Mantegazza e Tassinari, 1991).

La stima del DLMCC *1 passo avanti* (eqq. (10) e (11)), per ogni voce di consumo, per ciascuna coppia di serie CN e BF e la stima delle due varianze di osservazione, ottenuta mediante analisi coniugata, permettono di calcolare stime dei coefficienti di correlazione lineare. L'esame dell'andamento nel tempo di questi coefficienti (Fig. 1) consente di distinguere almeno due gruppi di serie: il primo presenta valori dei coefficienti di correlazione relativamente più elevati (in genere superiori a 0.8), più stabili nel tempo e generalmente convergenti; il secondo, invece, presenta coefficienti di correlazione relativamente meno elevati, più instabili nel tempo e convergenti, quando ciò avvenga, più lentamente.

### 3.3 Analisi dei risultati

Dall'esame delle MSE degli errori relativi di previsione (Tab. 1), il modello univariato chiuso *1 passo avanti* risulta mostrare una capacità di adattamento ai dati molto diversa a seconda delle serie analizzate. Le migliori previsioni si hanno per le serie: *Carne, Bevande alcoliche, Pane, Latte e Zucchero*; le peggiori si registrano invece per: *Alberghi, Patate, Trasporti, Mobili e Combustibili*. I DLM, quindi, risultano spiegare meglio le serie alimentari rispetto alle altre<sup>8</sup>, le quali presentano andamenti più irregolari (salti di livello, cambiamenti di variabilità, instabilità della componente stagionale). Fatta eccezione per *Alberghi*, il livello degli errori relativi risulta quasi sempre inferiore al 2%.

Se si considerano i modelli univariati chiusi *k passi avanti*, si registra, ovviamente, un peggioramento delle MSE, ma le prestazioni

relative di tali modelli per i vari capitoli di spesa rimangono sostanzialmente invariate rispetto ai DLM *1 passo avanti*.

Se si introduce l'informazione BF e si considerano, quindi, i due modelli aperti *k passi avanti*, si ottiene un miglioramento delle previsioni per la quasi totalità delle serie (Tab. 2, colonne 2 e 4). L'effetto positivo si realizza sia per il modello condizionato che per quello con regressore per le serie: *Carne, Patate, Combustibili, Vestiario, Mobili, Bevande Alcoliche*. Per queste serie il modello di regressione produce stime più accurate del modello condizionato, il quale a sua volta è preferibile al modello chiuso *1 passo avanti*. Ciò deriva dalla circostanza che, in presenza di una struttura dinamica stabile, la perdita di informazione dovuta al mancato aggiornamento sulla serie CN è circoscritta ed è più che compensata dal guadagno di informazione derivante dall'osservazione dei dati BF contemporanei alle stime da effettuare per CN. Si può quindi concludere che le serie omologhe CN e BF considerate, pur mostrando uno stretto legame, non vengono rappresentate completamente dal DLMCC nella loro struttura dinamica, come invece avviene per le serie *Pane, Olii, Zucchero*, per le quali risulta preferibile il modello condizionato. Ad ulteriore conferma, per tali serie, le MSE del modello *1 passo avanti* risultano relativamente basse.

Per quanto riguarda la relazione fra i due modelli aperti, la principale causa del loro diverso grado di accuratezza deriva dal diverso impiego della medesima informazione aggiuntiva. Infatti, mentre nel modello di regressione si specifica una relazione lineare fra le serie BF e CN direttamente per le osservazioni, nel modello condizionato la relazione lineare fra le componenti dello stato per le due serie è il risultato del condizionamento effettuato utilizzando distribuzioni normali. Ciò significa che nel modello di regressione il dato BF agisce direttamente sulla previsione CN, mentre nel modello condizionato esso viene prima filtrato dal modello *1 passo avanti* specificato per la serie BF e poi trasferito, in termini di errore di previsione, al modello per CN. Nel modello condizionato si è fatto a meno dell'introduzione della componente regressiva, che implica la specificazione del tipo di legame fra le serie. Se il modello regressivo produce previsioni più accurate del modello condizionato, si può concludere che il legame esistente fra le serie è stato colto dalla specificazione adottata. Viceversa, si può concludere che, pur esistendo un legame fra le serie, questo non è del tipo specificato dal modello di regressione e può essere meglio individuato mediante il legame fra le componenti dei vettori di stato

delle due serie.

Infine, per quanto riguarda le serie *Latte* e *Bevande analcoliche*, l'informazione BF risulta fuorviante in entrambi i modelli aperti<sup>9</sup>. Ciò sembra suggerire che la derivazione *ex-post* della serie trimestrale di CN abbia prodotto una rappresentazione del fenomeno esaminato sostanzialmente diversa da quella fornita dai dati BF.

A livello di varianza delle previsioni (cfr. Figg. 2d-4d), infine, c'è un divario netto tra il modello di regressione e quello condizionato: mentre il primo presenta varianze instabili nel tempo, con ampie oscillazioni, per il secondo le varianze mostrano un andamento abbastanza regolare nel tempo con livelli talvolta assai inferiori. Inoltre, per il modello condizionato l'andamento è quasi coincidente con quello del modello chiuso *1 passo avanti*. Nelle Figg. 2, 3 e 4 vengono illustrati esemplificativamente i casi di *Carne*, *Pane*, *Latte*. Per alcune serie non alimentari<sup>(10)</sup> e per le serie *Patate*, invece, le varianze delle previsioni condizionate e di quelle *1 passo avanti* risultano decisamente superiori a quelle del modello di regressione.

La somiglianza tra il modello condizionato ed il modello chiuso *1 passo avanti* risulta particolarmente evidente nel confronto tra gli andamenti delle componenti di stato stimate. Il confronto non è possibile con le componenti relative al modello di regressione, che ne comprende una in più. Al riguardo, si considerino di nuovo le Figg. 2, 3 e 4, nelle quali sono rappresentati gli andamenti delle componenti di livello, crescita, regressione, stimate con tutti e quattro i modelli considerati.

#### 4. Sintesi conclusiva

I risultati ottenuti confermano, per molte delle serie esaminate, l'esistenza di uno stretto legame tra le stime dei consumi CN e BF. Queste ultime mostrano un elevato potere esplicativo rispetto alle prime. Infatti, l'introduzione dell'informazione BF nel DLM di regressione o nel modello condizionato (entrambi *k passi avanti*), determina un miglioramento nelle previsioni dei dati CN non solo rispetto al DLM chiuso *k passi avanti*, ma spesso anche rispetto al DLM chiuso *1 passo avanti*. Fanno eccezione le serie *Latte*, *Bevande analcoliche* e, in misura minore, *Alberghi*, per le quali non risulta presente né la relazione di tipo lineare fra CN e BF misurata dal modello di regressione, né la più generale relazione espressa dal modello

condizionato. Infine, per le serie *Oli*, *Pane*, *Zucchero*, sembra emergere una relazione fra i dati BF e CN esprimibile in termini degli errori di previsione e quindi delle componenti dello stato.

#### Note

1 Programma di ricerca su: "La misura dei consumi privati: uno studio sull'accuratezza, coerenza e qualità dei dati", finanziato dal MURST sui fondi 40%.

2 Si ricordi che, in generale, nei DLM la procedura di previsione coincide con la procedura di stima per dati mancanti.

3 Nella (8 bis), in realtà, la varianza dello stato va ulteriormente moltiplicata per la varianza  $\sigma_{\epsilon t}$  delle osservazioni (cfr. la (3)), che è invece già presente nelle  $C^*$  e  $R^*$  per il modello condizionato.

4 Per  $k=1$ , i momenti a priori risultano, invece, identici per i due modelli.

5 Qualora si impieghi un DLM univariato chiuso, va considerato soltanto il sottosistema di  $D_{t+k}$  relativo a  $y_t$ .

6 Ciò in quanto non si ritiene di possedere a priori una conoscenza specifica sul grado di stabilità delle diverse componenti dello stato, per ogni serie analizzata.

7 Le serie considerate sono:

*Pane e cereali, Latte formaggi e uova, Carne, Pesce, Olii e grassi, Frutta e ortaggi, Patate, Zucchero, Caffè the e cacao, Bevande alcoliche, Bevande analcoliche*, per il comparto alimentare;

*Vestiaro, Calzature, Abitazione, Combustibili ed energia elettrica, Mobili ed accessori, Acquisto di mezzi di trasporto, Spese per alberghi e pubblici esercizi*, per il comparto non alimentare.

I consumi non alimentari qui considerati, ad eccezione di *Vestiaro* e *Calzature*, sono gli stessi analizzati da Daddi e Viviani (1992), scelti in base alla rilevanza economica ed al grado di utilizzo dei relativi dati BF per la produzione delle stime CN. Oltre che per i beni alimentari, infatti, le stime CN fanno riferimento ai dati BF anche per il capitolo *Abitazione*, mentre se ne prescinde completamente per *Combustibili ed energia elettrica*, e per *Mobili ed accessori*; per il capitolo *Spese per alberghi e pubblici esercizi* è addirittura diversa la popolazione di riferimento.

8 Almeno per le voci di consumo non alimentare qui considerate.

9 Ciò, nonostante la serie CN per il *Latte* presenti stime dei coefficienti di correlazione con la serie BF abbastanza elevate, anche se instabili nel tempo.

10 Si tratta di *Alberghi, Calzature, Vestiaro, Combustibili, Mobili*.

## Riferimenti bibliografici

- AMEEN J.R.M., HARRISON P.J. (1985): *Normal discount Bayesian models*, in "Bayesian statistics 2", J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith (Eds.). North-Holland, Amsterdam, and Valencia University Press.
- BARBOSA E.P., HARRISON P.J. (1989): *Variance Estimation for Multivariate DLM'S*, Research Report, 160, Dep. of Statistics, University of Warwick.
- CORRADI F., GUAGNANO G. (1992): *Dati mancanti in serie storiche multivariate: l'utilizzo del modello lineare dinamico a componenti comuni*, in "Atti della XXXVI Riunione Scientifica SIS", vol. 2, p. 119.
- DADDI P., VIVIANI A. (1992): *Modificazioni strutturali ed analisi temporale delle serie trimestrali dei consumi in Italia*, in "Atti della XXXVI Riunione Scientifica SIS", vol. 2, p.35.
- DAWID A.P. (1981): *Some matrix-variate distribution Theory: Notational considerations and a Bayesian application*, in "Biometrika", 68.
- DE WAAL D.J.(1985): *Matrix valued distribution*, in "Encyclopedia of Statistical Sciences", 5,p.327, Wiley NY.
- DICKEY J.M. (1967): *Matricivariate generalization of the multivariate t distribution and the inverted multivariate t distribution*, in "Ann. Math. Statist.", 38.
- HARRISON P.J., STEVENS C. (1976): *Bayesian forecasting (with discussion)*, in "The Statistician", 36.
- HARVEY A.C. (1986): *Analysis and generalisation of a multivariate exponential smoothing model*, in "Management Science", 32.
- MANTEGAZZA S., TASSINARI G. (1991): *Alcune osservazioni in tema di confronto fra i dati sui consumi privati della cn e dell'indagine sui Consumi delle Famiglie*, Rapporto di ricerca, Dpt. Scienze Statistiche, Univ. di Bologna.
- MONDUCCI R. (1991): *Utilizzazione dei bilanci di famiglia nelle trimestralizzazioni dei consumi privati*, Nota interna ISTAT, Roma.
- POLE A., West, M., (1989): *Reference analysis of the DLM*, in "J. Time Series Analysis".
- PRESS (1973): *Applied Multivariate Analysis*. NY: Holt, Rinehart & Winston.
- QUINTANA J.M. (1985): *A dynamic linear matrix-variate regression model*, Research Report, 83, Dep. of Statistics, University of Warwick.
- QUINTANA J.M. (1987): *Multivariate Bayesian forecasting models*, Unpublished Ph.D. Thesis, University of Warwick.
- QUINTANA J.M., WEST M. (1987): *Multivariate time series analysis*, in "The Statistician", 36.
- QUINTANA J.M., WEST M. (1988): *Time series analysis of compositional data*, in "Bayesian Statistics 3", J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith (Eds.). Oxford Univ. Press.
- RIZZI, (1988): *Il linguaggio delle matrici*. NIS, Roma.
- WEST M., HARRISON P.J. (1989): *Bayesian Forecasting and Dynamic Models*. New York: Springer-Verlag.

**A1. - Varianza dello stato  $k$  passi in avanti nel modello univariato, qualora si adotti il *discount approach* (equazione 8 bis)**

Sia  $y_{t-1}$  l'ultima informazione disponibile sulla serie; a partire dal tempo  $t$  la varianza a posteriori  $C_t$  dello stato risulterà uguale a quella a priori  $R_t$ , non essendoci riduzione di variabilità dovuta all'aggiornamento.

Avendo posto  $W_t = \Delta * C_{t-1}$ , ove l'operatore  $*$  indica il prodotto di Hadamard, la  $R_t$  può essere espressa come:

$$R_{t-1}(1) = GC_{t-1}G' + \Delta * C_{t-1} = C_{t-1}(1).$$

Iterando, i termini della successione risulteranno:

$$\begin{aligned} R_{t-1}(2) &= GC_{t-1}(1)G' + \Delta * C_{t-1}(1) \\ &= G^2 C_{t-1} G'^2 + G(\Delta * C_{t-1})G' + \Delta * (GC_{t-1}G') + \Delta^{[2]} * C_{t-1}, \end{aligned}$$

dove  $\Delta^{[j]}$  indica il prodotto di Hadamard di  $j$   $\Delta$  consecutivi.

Essendo  $G(\Delta * C_{t-1})G' = \Delta * (GC_{t-1}G')$  si ha

$$R_{t-1}(2) = G^2 C_{t-1} G'^2 + 2\Delta * GC_{t-1}G' + \Delta^{[2]} * C_{t-1}.$$

$$\begin{aligned} R_{t-1}(3) &= GC_{t-1}(2)G' + \Delta * C_{t-1}(2) \\ &= G^3 C_{t-1} G'^3 + 2G(\Delta * GC_{t-1}G')G' + G(\Delta^{[2]} * C_{t-1})G' + \\ &\quad + \Delta * (G^2 C_{t-1} G'^2) + 2\Delta^{[2]} * (GC_{t-1}G') + \Delta^{[3]} * C_{t-1} \\ &= G^3 C_{t-1} G'^3 + 3\Delta * G^2 C_{t-1} G'^2 + 3\Delta^{[2]} * GC_{t-1}G' + \Delta^{[3]} * C_{t-1} \\ &= \sum_{j=0}^3 \binom{3}{j} \Delta^{[j]} * G^{3-j} C_{t-1} G'^{3-j}, \end{aligned}$$

da cui, per induzione, segue:

$$R_{t-1}(k) = \sum_{j=0}^k \binom{k}{j} \Delta^{[j]} * G^{k-j} C_{t-1} G'^{k-j}.$$

**A2. - Media dello stato  $k$  passi in avanti nel modello condizionato (equazione 16)**

Sia  $y_{2,t-1}$  l'ultima osservazione disponibile sulla serie da prevedere; a partire dal momento primo della più recente distribuzione a posteriori stimata per la serie,  $m_{2,t-1}$ , posto  $a_{2,t-1}(0) = m_{2,t-1}$  ed utilizzando successivamente la (13), si ha:

$$a_{2,t-1}(1) = Ga_{2,t-1}(0)$$

$$m_{2,t-1}(1) = a_{2,t-1}(1) + b(m_{1,t} - a_{1,t}) = Ga_{2,t-1}(0) + b(m_{1,t} - a_{1,t})$$

$$a_{2,t-1}(2) = Gm_{2,t-1}(1) = G^2 a_{2,t-1}(0) + bG(m_{1,t} - a_{1,t})$$

$$m_{2,t-1}(2) = a_{2,t-1}(2) + b(m_{1,t+1} - a_{1,t+1}) =$$

$$= G^2 a_{2,t-1}(0) + bG(m_{1,t} - a_{1,t}) + b(m_{1,t+1} - a_{1,t+1})$$

$$a_{2,t-1}(3) = Gm_{2,t-1}(2) = G^3 a_{2,t-1}(0) +$$

$$+ bG^2(m_{1,t} - a_{1,t}) + bG(m_{1,t+1} - a_{1,t+1})$$

$$m_{2,t-1}(3) = a_{2,t-1}(3) + b(m_{1,t+2} - a_{1,t+2})$$

$$= G^3 a_{2,t-1}(0) + bG^2(m_{1,t} - a_{1,t}) + bG(m_{1,t+1} - a_{1,t+1}) +$$

$$+ b(m_{1,t+2} - a_{1,t+2})$$

$$= G^3 a_{2,t-1}(0) + b \sum_{j=0}^2 G^j (m_{1,t+2-j} - a_{1,t+2-j})$$

$$a_{2,t-1}(4) = Gm_{2,t-1}(3) = G^4 a_{2,t-1}(0) +$$

$$b \sum_{j=0}^3 G^{j+1} (m_{1,t+2-j} - a_{1,t+2-j}).$$

In generale, quindi:

$$a_{2,t-1}(k) = G^k a_{2,t-1}(0) + b \sum_{j=1}^{k-1} G^j (m_{1,t-1+k-j} - a_{1,t-1+k-j}),$$

$$m_{2,t-1}(k) = G^k a_{2,t-1}(0) + b \sum_{j=0}^{k-1} G^j (m_{1,t-1+k-j} - a_{1,t-1+k-j}),$$

nelle quali il termine  $a_{2,t-1}(k)$  costituisce la base per le previsioni  $k$  passi avanti, mentre il termine  $m_{2,t-1}(k)$  rappresenta l'ultimo aggiornamento dello stato al tempo  $t+k$ , dal quale si ottengono le previsioni per il tempo  $t+k+1$ .

### A3. - Varianza dello stato $k$ passi avanti nel modello condizionato (equazione 17)

Con riferimento alla (13) e alla (15) e ricordando che fino al tempo  $t-1$  le matrici  $R$  e  $C$  sono comuni alle due serie, si ha:

$$R_{2,t-1}(1) = R_{1,t-1}(1) = GC_{t-1}G' + \Delta * C_{t-1}$$

$$C_{2,t-1}^*(1) = \sigma_{22} R_{2,t-1}(1) - \frac{\sigma_{12}^2}{\sigma_{11}} (R_{1,t} - C_{1,t})$$

$$R_{2,t-1}^*(2) = GC_{2,t-1}^*(1)G' + \Delta * C_{2,t-1}^*(1)$$

$$= \sigma_{22} GR_{2,t-1}(1)G' - \frac{\sigma_{12}^2}{\sigma_{11}} G(R_{1,t} - C_{1,t})G' +$$

$$+ \sigma_{22} \Delta * R_{2,t-1}(1) - \frac{\sigma_{12}^2}{\sigma_{11}} \Delta * (R_{1,t} - C_{1,t})$$

$$= \sigma_{22} G^2 C_{t-1} G'^2 + \sigma_{22} G(\Delta * C_{t-1})G' - \frac{\sigma_{12}^2}{\sigma_{11}} G(R_{1,t} - C_{1,t})G' +$$

$$+ \sigma_{22} \Delta * (GC_{t-1}G') + \sigma_{22} \Delta^{[2]} * C_{t-1} - (\sigma_{12}^2 / \sigma_{11}) \Delta * (R_{1,t} - C_{1,t})$$

$$= \sigma_{22} (G^2 C_{t-1} G'^2 + 2\Delta * (GC_{t-1}G') + \Delta^{[2]} * C_{t-1}) +$$

$$- (\sigma_{12}^2 / \sigma_{11}) (G(R_{1,t} - C_{1,t})G' + \Delta * (R_{1,t} - C_{1,t}))$$

$$C_{2,t-1}^*(2) = R_{2,t-1}^*(2) - \frac{\sigma_{12}^2}{\sigma_{11}} (R_{1,t+1} - C_{1,t+1})$$

$$R_{2,t-1}^*(3) = GC_{2,t-1}^*(2)G' + \Delta * C_{2,t-1}^*(2)$$

$$= \sigma_{22} (G^3 C_{t-1} G'^3 + 2\Delta * G^2 C_{t-1} G'^2 + G\Delta^{[2]} * C_{t-1} G') +$$

$$- \frac{\sigma_{12}^2}{\sigma_{11}} (G^2 (R_{1,t} - C_{1,t})G'^2 + G\Delta * (R_{1,t} - C_{1,t})G' +$$

$$+ G(R_{1,t+1} - C_{1,t+1})G') + \sigma_{22} (\Delta * G^2 C_{t-1} G'^2 + 2\Delta^{[2]} * (GC_{t-1}G') +$$

$$+ \Delta^{[3]} * C_{t-1}) - \frac{\sigma_{12}^2}{\sigma_{11}} (\Delta * G(R_{1,t} - C_{1,t})G' + \Delta^{[2]} * (R_{1,t} - C_{1,t}) +$$

$$+ \Delta * (R_{1,t+1} - C_{1,t+1}))$$

$$= \sigma_{22} (G^3 C_{t-1} G'^3 + 3\Delta * G^2 C_{t-1} G'^2 + 3\Delta^{[2]} * GC_{t-1}G' + \Delta^{[3]} * C_{t-1}) +$$

$$- \frac{\sigma_{12}^2}{\sigma_{11}} (G^2 (R_{1,t} - C_{1,t})G'^2 + 2\Delta * G(R_{1,t} - C_{1,t})G' +$$

$$+ \Delta^{[2]} * (R_{1,t} - C_{1,t}) + G(R_{1,t+1} - C_{1,t+1})G' + \Delta * (R_{1,t+1} - C_{1,t+1}))$$

$$= \sigma_{22} \sum_{j=0}^3 \binom{3}{j} \Delta^{[j]} * G^{3-j} C_{t-1} G'^{3-j} +$$

$$- \frac{\sigma_{12}^2}{\sigma_{11}} \sum_{x=1}^2 \sum_{j=0}^{3-x} \binom{3-x}{j} \Delta^{[j]} * G^{3-x-j} (R_{1,t+x} - C_{1,t+x}) G'^{3-x-j}$$

In generale:

$$R_{2,t-1}^*(k) = \sigma_{22} \sum_{j=0}^k \binom{k}{j} \Delta^{[j]} * G^{k-j} C_{t-1} G'^{k-j} +$$

$$- \frac{\sigma_{12}^2}{\sigma_{11}} \sum_{x=1}^{k-1} \sum_{j=0}^{k-x} \binom{k-x}{j} \Delta^{[j]} * G^{k-x-j} (R_{1,t+x} - C_{1,t+x}) G'^{k-x-j}$$

$$C_{2,t-1}^*(k) = \sigma_{22} \sum_{j=0}^k \binom{k}{j} \Delta^{[j]} * G^{k-j} C_{t-1} G'^{k-j} +$$

$$- \frac{\sigma_{12}^2}{\sigma_{11}} \sum_{x=1}^k \sum_{j=0}^{k-x} \binom{k-x}{j} \Delta^{[j]} * G^{k-x-j} (R_{1,t+x} - C_{1,t+x}) G'^{k-x-j}$$

#### A4. - Tabelle e grafici

**Tab.1** - MSE\* dei modelli previsivi (valori x 100)

| Serie            | DLM chiuso     | DLM chiuso     | DLM          | DLM con    |
|------------------|----------------|----------------|--------------|------------|
|                  | 1 passo avanti | k passi avanti | condizionato | regressore |
| Carne            | 0.30           | 0.49           | 0.21         | 0.10       |
| Bevande alcol.   | 0.30           | 0.50           | 0.28         | 0.22       |
| Pane             | 0.36           | 0.59           | 0.21         | 0.31       |
| Latte            | 0.39           | 0.60           | 0.78         | 5.78       |
| Zucchero         | 0.42           | 0.65           | 0.39         | 2.08       |
| Abitazione       | 0.45           | 0.73           | 0.84         | 0.36       |
| Frutta           | 0.54           | 0.91           | 0.58         | 0.40       |
| Oli e grassi     | 0.58           | 0.90           | 0.49         | 1.24       |
| Pesce            | 0.58           | 0.93           | 0.61         | 0.56       |
| Vestiaro         | 0.70           | 1.04           | 0.58         | 0.36       |
| Calzature        | 0.71           | 1.10           | 0.82         | 0.76       |
| Bevande analcol. | 0.80           | 1.20           | 25.96        | 1.53       |
| Caffè            | 1.14           | 1.71           | 1.53         | 0.90       |
| Combustibili     | 1.34           | 2.00           | 1.06         | 0.66       |
| Mobili           | 1.70           | 2.17           | 1.41         | 0.88       |
| Trasporti        | 2.02           | 2.35           | 2.23         | 1.91       |
| Patate           | 4.13           | 6.43           | 3.21         | 1.80       |
| Alberghi         | 16.30          | 18.24          | 22.65        | 16.39      |

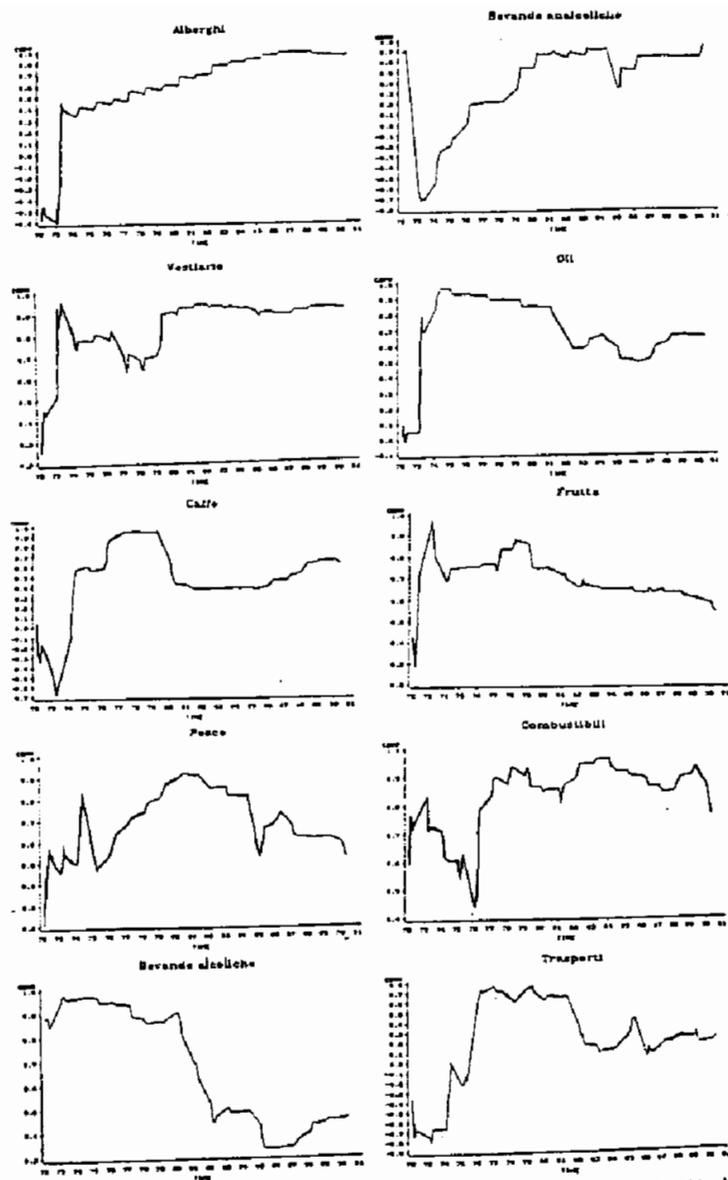
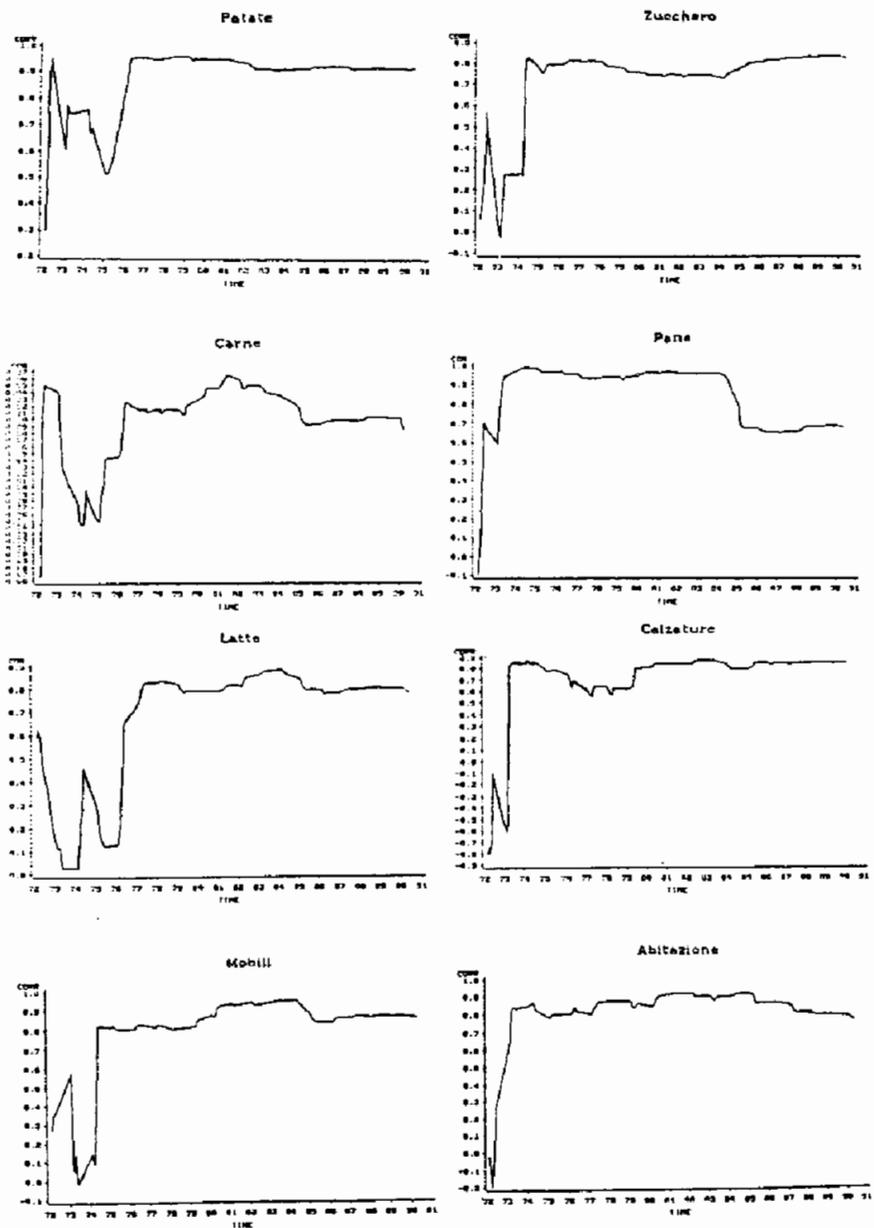
(\*) Calcolati sugli errori relativi di previsione.

**Tab.2** - Indici di miglioramento dei DLM aperti rispetto ai DLM chiusi (valori percentuali)

| Serie       | $\frac{MSE_{Rk}-MSE_{U1}}{MSE_{U1}}$ | $\frac{MSE_{Rk}-MSE_{Uk}}{MSE_{Uk}}$ | $\frac{MSE_{Ck}-MSE_{U1}}{MSE_{U1}}$ | $\frac{MSE_{Ck}-MSE_{Uk}}{MSE_{Uk}}$ |
|-------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|
|             | Carne                                | 66.7                                 | 79.6                                 | 30.0                                 |
| Patate      | 56.4                                 | 72.0                                 | 22.3                                 | 50.1                                 |
| Combustib.  | 50.7                                 | 67.0                                 | 20.9                                 | 47.0                                 |
| Vestiaro    | 48.6                                 | 65.4                                 | 17.1                                 | 44.2                                 |
| Mobili      | 48.2                                 | 59.4                                 | 17.1                                 | 35.0                                 |
| Bev. alcol. | 26.7                                 | 56.0                                 | 6.7                                  | 44.0                                 |
| Frutta      | 25.9                                 | 56.0                                 | -7.4                                 | 36.3                                 |
| Caffè       | 21.1                                 | 47.4                                 | -34.2                                | 10.6                                 |
| Abitazione  | 20.0                                 | 50.7                                 | -86.7                                | -15.1                                |
| Pane        | 13.9                                 | 47.4                                 | 41.7                                 | 64.4                                 |
| Trasporti   | 5.4                                  | 18.7                                 | -10.4                                | 5.1                                  |
| Pesce       | 3.4                                  | 39.8                                 | -5.2                                 | 34.4                                 |
| Alberghi    | -0.6                                 | 10.1                                 | -39.0                                | -24.2                                |
| Calzature   | -7.0                                 | 30.9                                 | -15.5                                | 25.5                                 |
| Bev. analc. | -91.2                                | 27.5                                 | -3145.0                              | -2063.3                              |
| Oli-grassi  | -113.8                               | -37.8                                | 15.5                                 | 45.6                                 |
| Zucchero    | -395.2                               | -220.0                               | 7.1                                  | 40.0                                 |
| Latte       | -1382.1                              | -817.5                               | -100.0                               | -23.8                                |

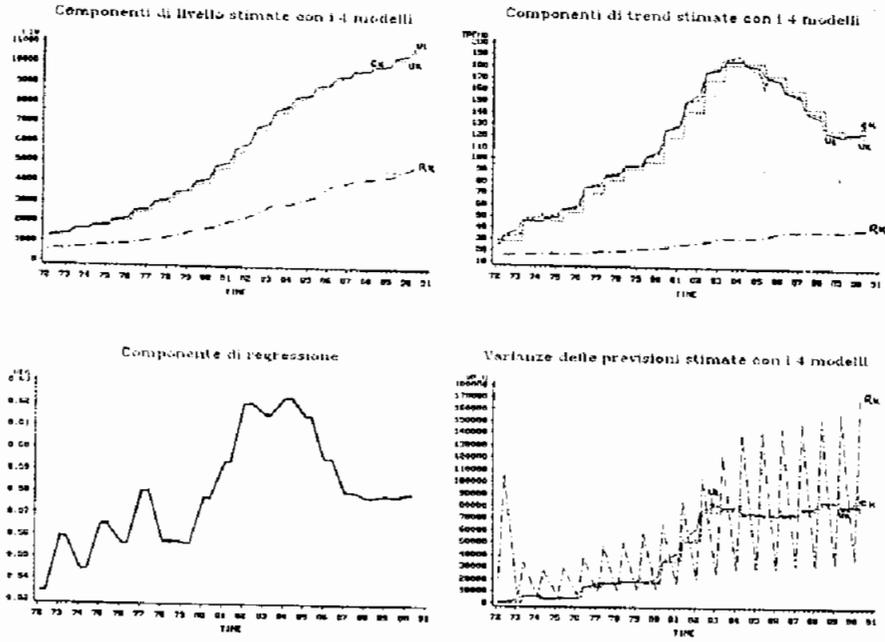
$MSE_{U1}$ =MSE del DLM univariato chiuso 1 passo avanti;  $MSE_{Rk}$ =MSE del DLM con regressore k passi avanti;  $MSE_{Uk}$ =MSE del DLM univariato chiuso k passi avanti;  $MSE_{Ck}$ =MSE del DLM condizionato k passi avanti.

Fig. 1 - Stime dei coefficienti di correlazione lineare fra le serie CN e BF nell'ambito dei DLMCC / passo avanti



(\*) Il salto nella stima del coefficiente di correlazione lineare, tra il terzo trimestre del 1984 ed il secondo del 1985, è dovuto al contemporaneo ed improvviso incremento delle stime BF del consumo di

Fig. 2 - Andamenti delle componenti di stato stimate per la serie Carne



U=componente stimata con il modello univariato chiuso *t* passi avanti  
 Uk=componente stimata con il modello univariato chiuso *k* passi avanti  
 Ck=componente stimata con il modello condizionato *k* passi avanti  
 Rk=componente stimata con il modello di regressione *k* passi avanti

Fig. 3 - Andamenti delle componenti di stato stimate per la serie Pane

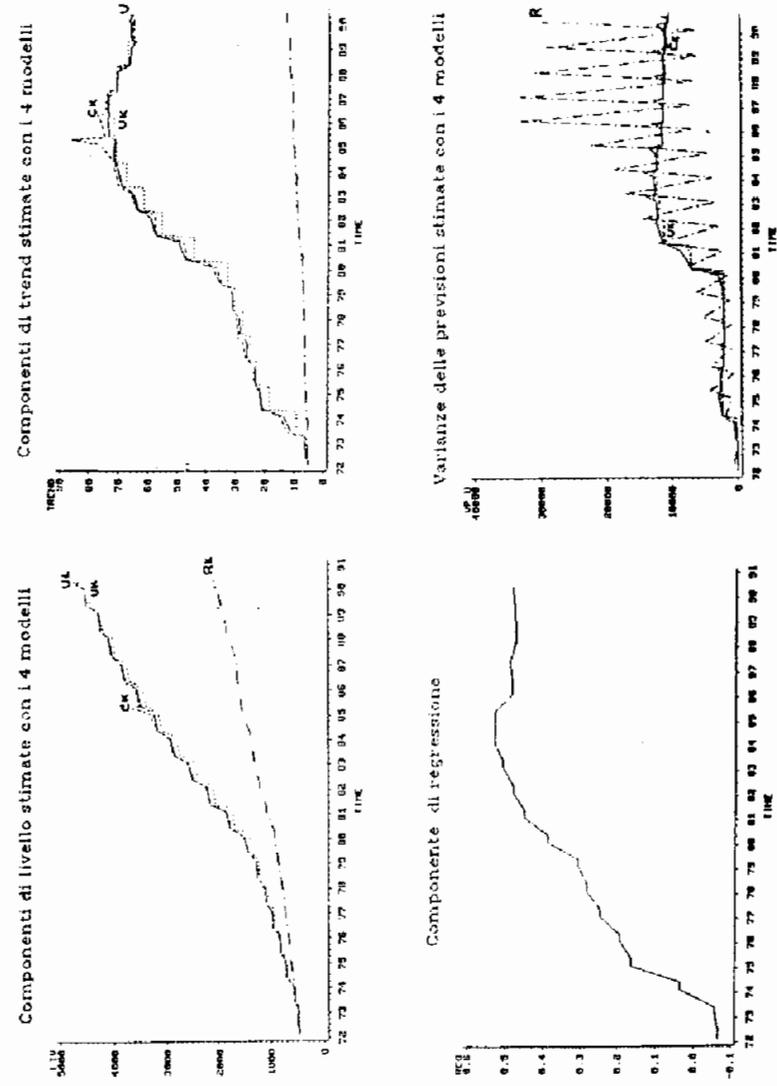


Fig. 4 - Andamenti delle componenti di stato stimate per la serie Lallu

