

# istat working papers

N.15  
2016

## **Disaggregazione temporale nei modelli dinamici: recenti avanzamenti nella pratica dei conti economici trimestrali**

*Laura Bisio, Filippo Moauro*



# istat working papers

N.15  
2016

## **Disaggregazione temporale nei modelli dinamici: recenti avanzamenti nella pratica dei conti economici trimestrali**

*Laura Bisio, Filippo Moauro*

### **Comitato scientifico**

Giorgio Alleva  
Tommaso Di Fonzo  
Fabrizio Onida

Emanuele Baldacci  
Andrea Mancini  
Linda Laura Sabbadini

Francesco Billari  
Roberto Monducci  
Antonio Schizzerotto

### **Comitato di redazione**

Alessandro Brunetti  
Romina Fraboni  
Maria Pia Sorvillo

Patrizia Cacioli  
Stefania Rossetti

Marco Fortini  
Daniela Rossi

### **Segreteria tecnica**

Daniela De Luca   Laura Peci   Marinella Pepe   Gilda Sonetti

## **Istat Working Papers**

**Disaggregazione temporale nei modelli dinamici: recenti avanzamenti nella pratica dei conti economici trimestrali.**

N. 15/2016

ISBN 978-88-458-1898-1

© 2016

Istituto nazionale di statistica  
Via Cesare Balbo, 16 – Roma

Salvo diversa indicazione la riproduzione è libera,  
a condizione che venga citata la fonte.

Immagini, loghi (compreso il logo dell'Istat),  
marchi registrati e altri contenuti di proprietà di terzi  
appartengono ai rispettivi proprietari e  
non possono essere riprodotti senza il loro consenso.

# Disaggregazione temporale nei modelli dinamici: recenti avanzamenti nella pratica dei conti economici trimestrali<sup>1</sup>

Laura Bisio e Filippo Moauro<sup>2</sup>

## Sommario

*Il presente lavoro illustra gli avanzamenti nell'ambito delle tecniche di disaggregazione temporale per i conti economici trimestrali che si stanno sperimentando presso l'ISTAT. Essi riguardano l'ampliamento dell'attuale impianto modellistico basato sui metodi di regressione statici alla classe dei modelli dinamici autoregressivi a ritardi distribuiti ADL, il passaggio dall'approccio di regressione a quello spazio negli stati per il trattamento statistico e l'adattamento delle relative procedure informatiche in ambiente Speakeasy. I principali contributi del lavoro sono: lo sviluppo di una procedura integrata per il trattamento statistico della classe di modelli ADL(1,1) e dei metodi statici di regressione; lo sviluppo delle procedure per la disaggregazione temporale nei logaritmi dei dati; l'implementazione di test diagnostici di controllo della stima; la discussione dei risultati di una simulazione comparativa tra metodi su serie reali; la definizione di linee guida per la selezione del modello di disaggregazione più appropriato. L'applicazione mostra che la più ampia classe di modelli ADL(1,1) sia nei livelli che nei logaritmi dei dati può determinare una riduzione consistente degli errori medi assoluti di revisione tra un rilascio e l'altro dei conti economici trimestrali. Limitati risultano i casi di cattivo adattamento dei modelli ADL rispetto ai modelli statici à la Chow-Lin.*

**Parole chiave:** Disaggregazione temporale, Rappresentazione nello spazio degli stati; Filtro di Kalman; Modelli ADL; Conti economici trimestrali

## Abstract

*The paper documents recent developments within the domain of temporal disaggregation techniques, tested in the environment of quarterly national accounts by ISTAT. These developments concern: i) the extension of the temporal disaggregation modeling setup currently in use - based upon static regression models - to the dynamic setup based on the class of autoregressive distributed lag ADL models, ii) the change from a regression-based approach to the state-space framework for statistical treatment and iii) the implementation of the new estimation procedures into a set of computer programs under the Speakeasy environment. Main contributions of the work are: the development of a full-fledged procedure for the statistical treatment of either the dynamic ADL(1,1) model class and the static regression models; the development of temporal disaggregation procedures for models in logarithms; the implementation of a set of diagnostic checking statistical tests; the discussion of a wide empirical application of the techniques here presented, applied to real economic time-series; the provision of some guidelines for the disaggregation model selection. Empirical findings are that the ADL(1,1) model class, either in levels and logarithms, can essentially reduce the mean absolute errors of the quarterly economic accounts revisions occurring between subsequent releases. Few are the cases in which ADL models do not fit well compared to the static models à la Chow-Lin.*

**Keywords:** Temporal disaggregation; State-space form; Kalman filter; ADL models; Quarterly economic accounts.

<sup>1</sup> Il lavoro è frutto della collaborazione tra i due autori, tuttavia il paragrafo 3 e le appendici A e B sono attribuite a Laura Bisio, mentre il paragrafo 2 e le appendici C e D a Filippo Moauro; i paragrafi 1 e 4 sono da attribuire ad entrambi gli autori. Le opinioni espresse in questo working paper sono esclusivamente degli autori e non dell'Istat.

<sup>2</sup> [bisio@istat.it](mailto:bisio@istat.it) e [moauro@istat.it](mailto:moauro@istat.it).

## Indice

|   | Pag. |
|---|------|
| <b>1. Introduzione e obiettivi</b> .....  | 7    |
| <b>2. Modello di riferimento e cenni sul trattamento statistico mediante forme spazio degli stati</b> ..... | 8    |
| 2.1 Da ADL(1,1) al modello AR(1) Chow-Lin .....   | 9    |
| 2.2 Da ADL(1,1) nelle differenze al modello Litterman (1983) e Fernández (1981).....                        | 9    |
| 2.3 Cenni sulla rappresentazione spazio degli stati .....   | 10   |
| 2.4 Il caso dei logaritmi .....   | 11   |
| 2.5 Statistiche test e diagnostica della disaggregazione.....   | 11   |
| <b>3. Analisi empirica</b> .....  | 13   |
| 3.1 Disegno generale dell'esercizio .....   | 13   |
| 3.2 Modelli di trimestralizzazioni a confronto.....   | 14   |
| 3.3 La procedura di selezione del miglior modello .....   | 16   |
| 3.4 Trimestralizzazione delle serie grezze .....  | 18   |
| 3.5 Una valutazione delle disaggregazioni: criteri di scelta e risultati ammissibili .....                  | 20   |
| 3.6 La disaggregazione temporale senza indicatore: procedure a confronto .....                              | 21   |
| <b>4. Conclusioni</b> .....   | 22   |
| <b>Riferimenti bibliografici</b> .....  | 23   |
| <b>Appendice A - Tavole</b> .....   | 25   |
| <b>Appendice B - Figure</b> .....   | 33   |
| <b>Appendice C - Descrizione delle routine Speakeasy</b> .....  | 42   |
| <b>Appendice D - L'output di TRIMEKF</b> .....  | 46   |

## 1. Introduzione e obiettivi

Sin dalla metà degli anni '80, periodo a partire dal quale il rilascio dei conti economici trimestrali assunse un approccio sistematico e tempestivo, i metodi di disaggregazione temporale sono diventati di centrale importanza quale strumento della contabilità trimestrale in ISTAT. Questa fase coincise con una prima informatizzazione dell'istituto che fece tesoro dell'esperienza della Banca d'Italia acquisendo strumenti analoghi di lavoro e adottando molte delle procedure lì sviluppate. Riguardo ai metodi di disaggregazione temporale l'ISTAT cominciò ad utilizzare il programma che implementava il metodo di Chow and Lin (1971) sviluppato da Barbone, Bodo e Visco (1981) in ambiente Speakeasy e il metodo di Denton (1971). Nello stesso ambiente informatico cominciò l'implementazione dei conti trimestrali che ancora oggi vengono elaborati in Speakeasy/Modeleasy.

Un seconda fase innovativa si ebbe a metà degli anni '90 con un'analisi critica dei metodi di disaggregazione temporale offerta dal contributo di Lupi e Parigi (1996), che da un punto di vista tecnico ispirò lo sviluppo all'interno dell'ISTAT di una sofisticata procedura per l'analisi dei risultati della trimestralizzazione. Questa è ancora ampiamente utilizzata nelle fasi operative di produzione dei conti trimestrali, offrendo all'analista un report diagnostico completo della disaggregazione trimestrale.

Successivamente, tra il 2004 e il 2005 fu istituita una commissione di studio sul trattamento dei dati ai fini dell'analisi congiunturale incaricata di formulare proposte relative alle strategie da utilizzare per la disaggregazione temporale nei conti economici trimestrali. I lavori finali della commissione (ISTAT, 2005) documentarono l'impegno dell'ISTAT nell'ammodernamento dei vari strumenti concettuali e tecnici impiegati nella contabilità trimestrale. Le considerazioni finali contenute in Di Fonzo (2005) sono articolate in 9 punti e 4 suggerimenti operativi, la maggior parte dei quali pienamente attuati.

In seguito ai lavori della commissione la gamma di modelli di disaggregazione temporale fu ampliata al metodo di Fernández (1981), correggendo alcuni aspetti della formula di estrapolazione per le stime in corso d'anno.

Infine, il periodo più recente ha visto una serie di sviluppi metodologici relativi alle tecniche di disaggregazione temporale dovuti soprattutto alla spinta propulsiva di Eurostat. Tra gli altri, si vedano i lavori di Frale et al. (2010 e 2011), Grassi et al. (2014) e Moauro (2014). D'altronde, durante gli stessi anni, l'ISTAT ha privilegiato un approccio più pragmatico e orientato agli aspetti operativi di produzione dei dati, attuando i suggerimenti delineati da Di Fonzo (2005) nell'ambito dei modelli univariati. In particolare, lo sforzo è stato rivolto allo sviluppo dei metodi di disaggregazione temporale basati sui modelli autoregressivi a ritardi distribuiti (Proietti, 2005) secondo l'approccio spazio degli stati, nonché alla loro completa integrazione in una nuova procedura informatica in ambiente Speakeasy. Tale estensione, da un lato ha ampliato significativamente la gamma dei possibili modelli utilizzabili nelle fasi di stima dei conti trimestrali, dall'altro, basandosi sul filtro di Kalman (Kalman, 1960), ha apportato dei notevoli vantaggi pratici all'analisi trimestrale, quali il calcolo delle innovazioni, l'elaborazione di misure della qualità previsiva delle stime in corso d'anno e l'impiego di modelli basati sulla trasformazione logaritmica delle serie.

Il presente documento illustra l'insieme degli elementi innovativi recentemente introdotti per la produzione dei conti trimestrali riguardo alle procedure di disaggregazione temporale, alla selezione del modello e al controllo diagnostico dei risultati. Inoltre, la parte applicativa, particolarmente estesa, presenta i risultati di un esercizio di disaggregazione temporale basato su serie storiche ISTAT per un'analisi comparativa della performance delle procedure di disaggregazione temporale e per la definizione di linee guida per la selezione della specificazione più appropriata. Il confronto tra modelli mette in evidenza i casi critici che l'esercizio può determinare in termini di cattiva inizializzazione o eccessiva erraticità della serie disaggregata. Emergono le specificità dei risultati rispetto al modello considerato, alla tipologia di serie considerata in volume o in valore, grezza o destagionalizzata, e rispetto all'efficacia della trasformazione logaritmica e delle possibili estensioni.

Il lavoro è strutturato come di seguito indicato: nella sezione 2 si descrive il modello di riferimento a ritardi distribuiti ADL(1,1) e come questo includa i casi dei modelli statici alla Chow-Lin; nella stessa sezione si illustrano gli elementi principali della rappresentazione a spazio negli stati, l'estensione alla disaggregazione temporale nei logaritmi e i principali test diagnostici implementati; la sezione 3 è dedicata alla parte applicata in cui si illustrano alcuni esempi pratici di disaggregazione basati sull'utilizzo delle nuove subroutine Speakeasy; infine le considerazioni conclusive sono presentate nella sezione 4. Il lavoro include anche 4 appendici: l'appendice A contiene le tavole con i risultati estesi dell'applicazione, la B i principali grafici delle disaggregazioni, la C una descrizione dei principali programmi sviluppati e l'appendice D riporta un esempio di report prodotto dalla nuova procedura di disaggregazione.

## 2. Modello di riferimento e cenni sul trattamento statistico mediante forme spazio degli stati

La metodologia di disaggregazione temporale qui discussa si basa su metodi di regressione dinamici. Questi ipotizzano una semplice relazione lineare univariata tra la variabile dipendente  $y_t$ , il suo valore ritardato  $y_{t-1}$  e un insieme di regressori  $x_t$  nell'intervallo temporale  $t = 1, \dots, T$ . Tuttavia  $y_t$  è disponibile soltanto in forma aggregata in un dato intervallo temporale  $s$  dei valori  $y_t$ . Nel caso di somma tale aggregato denominato qui  $Y_t$  è tale che  $Y_t = y_t + \dots + y_{t-s+1}$ , mentre nel caso di media  $Y_t = (y_t + \dots + y_{t-s+1})/s$ . Perciò  $Y_t$  è osservato soltanto nei periodi  $t=s, 2s, \dots, [T/s]$ , dove  $[T/s]$  è il più grande intero del rapporto  $T/s$ . Invece, le  $k$  covariate  $x_t = (x_{1t}, \dots, x_{kt})$  sono disponibili in tutti i periodi  $t=1, \dots, T$ .

Una rappresentazione generale della relazione tra i due set di variabili è data dai modelli autoregressivi dinamici a ritardi distribuiti ADL(1,1) che vengono formulati alla frequenza d'osservazione più alta e assumono la seguente forma:

$$\Delta^l y_t = \varphi \Delta^l y_{t-1} + m + gt + \Delta^l x'_t \beta_0 + \Delta^l x'_{t-1} \beta_1 + \epsilon_t, \quad \epsilon_t \sim \text{NID}(0, \sigma^2), \quad (1)$$

dove  $\Delta$  è l'operatore alle differenze tale che  $\Delta y_t = y_t - y_{t-1}$ ,  $l$  rappresenta l'ordine della differenziazione che nelle applicazioni che seguono assumerà soltanto valori 0 (cioè nessuna differenziazione dei dati) o 1,  $\varphi$  è il termine autoregressivo tale che  $-1 < \varphi < 1$ ,  $m$  e  $gt$  rappresentano le componenti deterministiche cioè, rispettivamente, una costante e un trend lineare, i vettori  $\beta_0$  e  $\beta_1$  sono i coefficienti di regressione delle covariate  $x_t$  modellate rispettivamente a ritardo 0 e 1;  $\epsilon_t$  è il vettore degli errori stocastici per i quali si assume una distribuzione normale a media nulla e varianza costante nel tempo pari a  $\sigma^2$ .

Il modello ADL(1,1) è stato reso popolare a partire dal lavoro di Hendry e Mizon (1978) dove si sottolineava come il modello rimanesse valido anche se le variabili sottostanti alla regressione determinassero una relazione spuria nei livelli e fossero incorrelate nelle differenze. In seguito allo sviluppo dell'analisi di cointegrazione tra serie si sancì il successo dei modelli a correzione dell'errore (ECM) e fu mostrato come questi rappresentassero una semplice riparametrizzazione dei modelli ADL(1,1) nei livelli dei dati, corrispondenti all'equazione (1) nel caso in cui  $l=0$ .

Nell'ambito della disaggregazione temporale Proietti (2005) propose una metodologia basata sulla parametrizzazione dei modelli ADL(1,1) in una forma spazio degli stati (SSF) e sull'impiego del filtro di Kalman per il trattamento statistico. In particolare il filtro di Kalman è utilizzato per: il calcolo della funzione di log-verosimiglianza, la stima dei parametri incogniti del modello, la distribuzione ad alta frequenza (per esempio nei trimestri) di dati osservati come somma o media in un intervallo temporale più basso (per esempio l'anno) e l'estensione alla disaggregazione non lineare nel caso di dati modellati nei logaritmi.

Riguardo alla stima di massima verosimiglianza dei parametri incogniti del modello (1) la soluzione più adatta è fornita dal metodo dei minimi quadrati generalizzati (GLS). Infatti, tutti i coefficienti di regressione del modello, cioè  $m$ ,  $g$ ,  $\beta_0$ ,  $\beta_1$  e il termine di varianza  $\sigma^2$  dei residui  $\epsilon_t$  possono essere concentrati dalla funzione di verosimiglianza dando luogo ad una verosimiglianza



profilo che dipende dal solo parametro autoregressivo  $\varphi$ . Quindi la sua stima può essere convenientemente impostata come una ricerca a griglia nell'intervallo  $(-1,1)$  dei valori di  $\varphi$  tale da massimizzare la verosimiglianza.

I modelli ADL(1,1) dell'equazione (1) annidano modelli di regressione con residui AR(1), ovvero I(1) o anche ARIMA(1,1,0); nell'ambito della disaggregazione temporale questi tre casi particolari si riconducono rispettivamente ai metodi di Chow-Lin (1971), Fernández (1981) e Litterman (1983).

## 2.1 Da ADL(1,1) al modello AR(1) Chow-Lin

Sotto appropriate assunzioni circa le condizioni iniziali e appropriati vincoli lineari validi per i parametri del regressore  $x_t$ , il modello ADL(1,1) coincide con un modello di regressione lineare i cui residui seguono un processo AR(1). In formule, il modello (1) nei livelli dei dati ( $I=0$ ) e per semplicità senza componenti deterministiche può esprimersi mettendo in evidenza il polinomio a ritardi distribuiti  $(\beta_0 + \beta_1 L)$ , dove  $L$  è l'operatore ritardo tale che  $Lx_t = x_{t-1}$ , dando luogo a:

$$y_t = \varphi y_{t-1} + x'_t(\beta_0 + \beta_1 L) + \epsilon_t$$

che, data la restrizione  $\beta_1 = -\varphi\beta_0$  con  $|\varphi| < 1$ , è possibile esprimere come:

$$y_t(1 - \varphi L) = x'_t\beta_0(1 - \varphi L) + \epsilon_t$$

e, dunque come:

$$\begin{aligned} y_t &= x'_t\beta_0 + \alpha_t \\ \alpha_t &= \varphi\alpha_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{NID}(0, \sigma^2) \end{aligned}$$

dove il residuo  $\alpha_t$  segue un processo autoregressivo stazionario del primo ordine.

## 2.2 Da ADL(1,1) nelle differenze al modello Litterman (1983) e Fernández (1981)

Sotto opportune condizioni iniziali che riflettono la non stazionarietà del modello ADL(1,1) nelle differenze è possibile ottenere da quest'ultimo i modelli alla base dei metodi di disaggregazione temporale Fernández (1981) e Litterman (1983). Nel primo caso, si avrà un modello di regressione lineare i cui residui seguono un processo random walk o I(1), nel secondo caso i residui seguono un processo ARIMA (1,1,0).

In formule, quando  $I=1$  vale:

$$\Delta y_t = \varphi \Delta y_{t-1} + \Delta x'_t \beta_0 + \Delta x'_{t-1} \beta_1 + \epsilon_t, \quad \epsilon_t \sim \text{WN}(0, \sigma^2) \quad (2)$$

il quale sotto l'ipotesi  $\varphi = 0$  e  $\beta_1 = 0$  coincide con il modello di Fernández, cioè:

$$\Delta y_t = \Delta x'_t \beta_0 + \epsilon_t,$$

ovvero:

$$\begin{aligned} y_t &= x'_t \beta_0 + u_t \\ u_t &= u_{t-1} + \epsilon_t \end{aligned}$$

in cui il residuo  $u_t$  segue un processo autoregressivo integrato del primo ordine.

Sotto la condizione  $\beta_1 = -\varphi\beta_0$ , con  $\varphi < 1$  il modello alle differenze (2), restituisce il modello di Litterman, cioè:

$$\Delta y_t = \varphi \Delta y_{t-1} + \Delta x'_t \beta_0 - \Delta x'_{t-1} \varphi \beta_0 + \epsilon_t, \quad \epsilon_t \sim \text{WN}(0, \sigma^2)$$

$$y_t = x_t' \beta_0 + u_t$$

$$\Delta u_t = \varphi \Delta u_{t-1} + \epsilon_t \quad \epsilon_t \sim \text{NID}(0, \sigma^2)$$

in cui il residuo  $u_t$  segue un processo ARIMA(1,1,0).

### 2.3 Cenni sulla rappresentazione spazio degli stati

In generale, nell'ambito delle serie temporali la rappresentazione SSF è caratterizzata da due equazioni: la prima definisce la struttura della serie storica (equazione di misura), la seconda definisce come le componenti strutturali latenti evolvono da un certo stato allo stato successivo (equazione di transizione). La rappresentazione SSF permette di ricorrere alla metodologia del filtro di Kalman che consente di calcolare lo stimatore ottimo del vettore delle variabili di stato al tempo  $t$  per  $t=1, \dots, T$ , date le informazioni note a quello stesso orizzonte. In genere al filtro di Kalman si associa un algoritmo di smoothing che consente la stima ottimale del vettore di stato condizionata a tutto il set informativo.

L'applicazione dell'approccio SSF al problema della disaggregazione temporale è stato introdotto da Harvey e Pierce (1984) e sviluppato in seguito, tra gli altri, da Harvey (1989), Harvey e Chung (2000), Harvey e Koopman (1997) e Moauro e Savio (2002). Le proprietà della rappresentazione SSF nell'ambito delle tecniche di disaggregazione temporale sono state in seguito diffusamente trattate nei contributi di Proietti (2005; 2006) che costituiscono dei fondamentali riferimenti per il presente lavoro. Nel rimandare ai lavori appena citati per gli aspetti più analitici, nel presente contesto mettiamo soltanto in evidenza i principali elementi che caratterizzano l'approccio *spazio degli stati* rispetto a quello tradizionale. Tra i vantaggi: i) la possibilità di trattare opportunamente le condizioni iniziali in presenza di non-stazionarietà delle serie; ii) la disponibilità di statistiche diagnostiche più efficienti per valutare la qualità delle stime di massima verosimiglianza quali le innovazioni; iii) la possibilità di ottenere agevolmente le estrolazioni delle serie in modelli senza covariate. Tra gli svantaggi vi è indubbiamente la maggiore complessità computazionale del filtro di Kalman e la necessità di approntare delle routine di calcolo che in alcuni ambienti informatici possono risultare in una maggiore pesantezza di elaborazione.

Secondo l'approccio suggerito da Harvey (1989, sez. 6.3) il problema di disaggregazione temporale rientra in un problema di stima di "osservazioni mancanti" che viene trattato aggiungendo alla rappresentazione SSF del modello specificato ad alta frequenza una variabile dinamica chiamata "cumulatrice"  $y_t^C$ .  $y_t^C$ , pur essendo definita alle alte frequenze ovvero in tutti gli istanti  $t=1, 2, \dots, T$ , è osservabile soltanto ogni  $s$  periodi sotto forma di aggregato delle  $s$  osservazioni precedenti. In pratica è disponibile a  $t = s, 2s, 3s, \dots$ . Per esempio nel caso di serie trimestrali di flusso (valore aggiunto, consumi, ...)  $y_t^C$  è tale che:

$$\begin{array}{cccc} y_1^C = y_1, & y_2^C = y_1 + y_2, & y_3^C = y_1 + y_2 + y_3, & y_4^C = y_1 + y_2 + y_3 + y_4 \\ y_5^C = y_5, & y_6^C = y_5 + y_6, & y_7^C = y_5 + y_6 + y_7, & y_8^C = y_5 + y_6 + y_7 + y_8 \\ & & \dots & \dots \end{array}$$

ed è disponibile a  $t=4, 8, 12, \dots$

Per il trattamento statistico i passaggi sono i seguenti: la variabile cumulatrice  $y_t^C$  viene aggiunta al vettore di stato della forma SSF del modello espresso alle alte frequenze; si aggiusta di conseguenza l'equazione di misura e il filtro di Kalman per tener conto delle osservazioni mancanti di  $y_t^C$ ; segue il calcolo della funzione di verosimiglianza del modello rappresentato, la sua massimizzazione rispetto al vettore dei parametri incogniti e attraverso l'algoritmo di smoothing la stima delle osservazioni mancanti e quindi della serie disaggregata. Per la trattazione completa della forma SSF e del trattamento statistico si rimanda a Proietti (2005).

## 2.4 Il caso dei logaritmi

La trasformazione logaritmica delle serie storiche è pratica comune nell'econometria delle serie temporali, in special modo quando esse si riferiscono a variabili definite come rapporti tra grandezze di flusso (Proietti, 2005). L'applicazione dei logaritmi alle serie incluse nel modello comporta alcuni vantaggi noti quali la riduzione della volatilità delle serie stesse o la maggior plausibilità delle ipotesi che si assumono valide per il modello di regressione (linearità del modello, omoschedasticità e normalità degli errori).

Nel contesto dei logaritmi, sia il vincolo di aggregazione temporale a cui è sottoposta la serie disaggregata (incognita), sia il vincolo di "osservabilità" rappresentato dalla variabile cumulatrice richiamata in precedenza sono definiti nel dominio non lineare. Ai fini di una maggior chiarezza, si consideri che tra la trasformazione logaritmica della serie disaggregata  $y_t$  e la corrispondente serie aggregata  $Y_\tau$  vale la seguente espressione:

$$Y_\tau = \sum_{j=0}^{s-1} \exp(y_{\tau s-j}), \quad \tau = 1, \dots, \left\lfloor \frac{T}{s} \right\rfloor. \quad (3)$$

Si consideri inoltre che, nel contesto non lineare, la variabile cumulatrice  $Y_t^C$  è funzione sia dei propri ritardi, sia della serie disaggregata nei livelli  $\exp(y_t)$ . Pur rimandando ad altre fonti<sup>3</sup> per una trattazione più analitica di queste specificazioni, si noti che la variabile cumulatrice  $Y_t^C$  non è definita a meno di una stima preliminare della serie disaggregata nei livelli  $\exp(\tilde{y}_t)$  che soddisfi il vincolo (3) a posteriori (*constraint posterior mode estimate*). Una volta ottenuta tale stima preliminare attraverso un'apposita procedura di stima iterativa a partire da un'approssimazione lineare di Taylor di  $\exp(y_t)$ , la variabile  $Y_t^C$  definita come sopra, è inclusa nel sistema delle equazioni di stato della rappresentazione SSF del cosiddetto modello lineare gaussiano approssimato (LGAM) a cui viene applicato il filtro di Kalman, così come nel caso lineare.

## 2.5 Statistiche test e diagnostica della disaggregazione

Una delle caratteristiche della rappresentazione SSF e dell'adozione del filtro di Kalman è la stima degli errori di previsione o innovazioni quali sottoprodotti dell'applicazione dell'algoritmo stesso al sistema di equazioni che definisce la rappresentazione SSF. Le innovazioni sono alla base dei principali test diagnostici implementati nelle nuove procedure dove, in particolare, vengono considerate le innovazioni standardizzate  $\hat{v}_t$  che per i modelli di disaggregazione sono definite soltanto per  $t=s, 2s, \dots, \lfloor T/s \rfloor$ . Data la definizione delle innovazioni  $v_t = Y_t - E(Y_t|Y_{t-1})$  dove  $Y_{t-1}$  è il set informativo al tempo  $t$ , le innovazioni standardizzate si definiscono come:

$$\hat{v}_t = \frac{v_t}{Std(v_t)}.$$

Per una discussione esauriente si rimanda a Harvey (1989, sez.5.4 p.256 e seguenti). In questo contesto mettiamo soltanto in evidenza come per i modelli di disaggregazione statici, tutte le statistiche che si producono attraverso le nuove procedure di calcolo sono coerenti con quelle che si ottengono dalle formule standard della regressione.

Le prime statistiche test prese in considerazione sono il coefficiente di determinazione  $R^2$  e la sua forma corretta  $R_c^2$  definiti rispettivamente come

$$R^2 = 1 - SSR/SST, \quad (4)$$

$$R_c^2 = \frac{1-k}{\lfloor T/s \rfloor - 1} + \frac{\lfloor T/s \rfloor - 1}{\lfloor T/s \rfloor - k} R^2 \quad (5)$$

<sup>3</sup> Harvey (1989) e Proietti(2005).

dove  $SSR$  è la somma dei quadrati dei residui che nell'ambito dei modelli spazio degli stati è semplicemente

$$SSR = [T/s] \cdot \hat{\sigma}^2$$

con  $\hat{\sigma}^2$  dato dalla stima di massima verosimiglianza non corretta di  $\sigma^2$  di cui all'equazione (1) ottenuto dal filtro di Kalman e  $SST$  è la somma dei quadrati degli scostamenti delle differenze prime della serie  $y_t$  dalla sua media. Nell'equazione (5)  $k$  è il numero delle covariate  $x_t$  della regressione (1) inclusi, se presenti, i termini costante e/o il trend lineare.

La disamina in Harvey (1989, p.268-9) mette in forte discussione l' $R^2$  come misura della bontà di adattamento del modello nel contesto dell'analisi delle serie storiche vista la presenza nella maggior parte delle applicazioni di un trend nelle osservazioni della serie  $y_t$  e la conseguente vicinanza del valore di  $R^2$  a uno. Perciò il calcolo di  $SST$  rispetto alle differenze della  $y_t$  è un tentativo di aggiustare la statistica  $R^2$  a serie non stazionarie, rendendola così utile a misurare l'adattamento del modello.

Altre statistiche di bontà di adattamento del modello sono date dall'errore standard di regressione e il valore di massima log-verosimiglianza definite rispettivamente come

$$SER = \sqrt{SSR/([T/s] - k)}, \quad (6)$$

$$\hat{L}_c = -0.5 \cdot \{d_{T+1} + [T/s][\ln \hat{\sigma}^2 + \ln(2\pi) + 1]\} \quad (7)$$

dove  $d_{T+1}$  è la somma dei logaritmi delle varianze delle innovazioni nei periodi  $t=s, 2s, \dots, [T/s]$ .

I criteri informativi AIC e BIC utili per il confronto tra specificazioni alternative sono anch'essi immediatamente ricavabili come semplici formulazioni di  $\hat{\sigma}^2$ , rispettivamente:

$$AIC = 2k/[T/s] + \ln \hat{\sigma}^2 \quad (8)$$

$$BIC = k \cdot \frac{\ln[T/s]}{[T/s]} + \ln \hat{\sigma}^2. \quad (9)$$

Il test di Durbin-Watson (1950, 1951) per valutare al presenza di autocorrelazione del primo ordine nei residui è definito rispetto alle innovazioni standardizzate  $\hat{v}_t$  come:

$$d = \frac{\sum_{t=k+2}^{[T/s]} (\hat{v}_t - \hat{v}_{t-1})^2}{\sum_{t=k+2}^{[T/s]} \hat{v}_t^2}.$$

Il calcolo del test  $N$  di normalità dei residui di Jarque-Bera è derivato direttamente dall'applicazione delle formule riportate da Harvey (1989, p.560 eq.5.4.10 e 5.4.11) rispettivamente per i momenti terzi e quarti dei residui standardizzati  $\hat{v}_t$ . In particolare;

$$\sqrt{b_1} = \hat{\sigma}_*^{-3} \sum (\hat{v}_t - \bar{v})^3 / T^*$$

$$b_2 = \hat{\sigma}_*^{-4} \sum (\hat{v}_t - \bar{v})^4 / T^*$$

dove la relazione tra  $\hat{\sigma}^2$  e  $\hat{\sigma}_*^2$  (varianza dei residui corretta) è tale che  $\hat{\sigma}_*^2/\hat{\sigma}^2 = T/(T-k)$  e  $T^* = T - k$ . La statistica  $N$  è quindi data da:

$$N = \frac{T^*}{6} \cdot b_1 + \frac{T^*}{24} \cdot (b_2 - 3)^2, \quad (10)$$

che sotto l'ipotesi nulla ha una distribuzione chi-quadrato con 2 gradi di libertà per grandi campioni.

Una forma del test di significatività delle prime P autocorrelazioni dei residui è dato dal test Q di Ljung-Box che risulta da:

$$Q = T^*(T^* + 2) \sum_{\tau=1}^P (T^* - \tau)^{-1} r_v^2(\tau) \quad (11)$$

dove  $r_v(\tau)$  sono le autocovarianze campionarie delle innovazioni standardizzate. Per una discussione si rimanda sempre a Harvey (1989) p.259, eq.5.4.7. Nelle applicazioni del contesto dei modelli ADL(1,1) ci è sembrato opportuno confrontare i valori di questa statistica con un chi-quadrato con gradi di libertà pari a  $\sqrt{[T/s]} - 1$ .

L'ultimo test preso in considerazione è il test  $H(h)$  per l'eteroschedasticità dei residui che prende la forma di

$$H(h) = \sum_{t=T-h+1}^T \hat{v}_t^2 / \sum_{t=k+1}^{k+1+h} \hat{v}_t^2. \quad (12)$$

dove h è un intero prossimo a  $T^*/3$ . Per una discussione si rimanda a Harvey (1989) p.259 eq.5.4.9. In questo caso la statistica  $h \cdot H(h)$  è confrontata con in valori di un chi-quadrato con h gradi di libertà.

I test diagnostici discussi sopra meritano alcune considerazioni che derivano dal contesto della disaggregazione temporale da cui muoviamo: in tutti i casi le statistiche implementate sono costruite a partire da residui che possono essere definiti soltanto alle basse frequenze, risultando perciò poco informativi rispetto alla disaggregazione ottenuta. In altri termini queste statistiche offrono delle indicazioni sulla bontà di adattamento dell'indicato rispetto all'indicatore aggregato, rispetto a quanto i residui deviano dalle ipotesi di "bianchezza", assenza di autocorrelazione ed eteroschedasticità alla frequenza più bassa, ma ben poco rispetto a quanto la serie disaggregata sia lontana dall'indicatore di riferimento in termini di dinamica congiunturale. Per questo motivo questi esercizi devono essere sempre accompagnati da un'analisi grafica della serie disaggregata e da statistiche di correlazione tra questa e l'indicatore di riferimento.

Nel paragrafo seguente vedremo come ci possano essere esempi dove la disaggregazione appare ben specificata in termini di statistiche sui residui pur producendo, nel contempo, una serie disaggregata eccessivamente erratica o, al contrario, troppo liscia rispetto all'indicatore.

Altro limite delle statistiche introdotte è ovviamente dato dalla lunghezza delle serie che nella maggior parte degli esercizi di disaggregazione di dati annuali può risultare inferiore a 20 osservazioni. È ben noto infatti che le regioni di significatività dei principali test statistici sono definiti per un numero di osservazioni che tende all'infinito e che quando questi si basano su poche osservazioni perdono di potenza.

### 3. Analisi empirica

#### 3.1 Disegno generale dell'esercizio

Ai fini di un'illustrazione comparativa della performance delle diverse classi di modelli sopra richiamati, presentiamo di seguito i risultati di un esercizio di disaggregazione temporale realizzato utilizzando le subroutine 'TRIMEKF' e 'TRIMEKFL' presentate in dettaglio nell'appendice C. L'esercizio, benché esemplificativo, è volto a riprodurre la pratica corrente di stima dei conti trimestrali, sia per estensione che per natura delle serie trattate. Esso illustra una selezione di trimestralizzazioni basate su serie ISTAT di contabilità nazionale e relativi indicatori congiunturali: l'aggregato annuale è la serie del valore aggiunto relativo alle 17 branche di attività economica di cui si compone il settore dell'industria in senso stretto (sezioni B-E della classificazione NACE

Rev.2) secondo il livello di dettaglio adottato in ambito CET<sup>4</sup>; l'indicatore utilizzato è l'indice di produzione industriale considerato allo stesso livello di dettaglio. Il periodo di riferimento dell'esercizio è l'intervallo temporale 1995 – 2013. Uno schema riassuntivo dei dati utilizzati è presentato nella tavola 1A in Appendice A.

L'esercizio è stato ripetuto considerando sia gli aggregati grezzi, sia gli aggregati destagionalizzati e corretti per i giorni lavorativi. Inoltre, ai fini di una maggiore robustezza, l'esercizio è stato ripetuto sia considerando le serie espresse in valori nominali (prezzi correnti), sia le serie espresse in volume (a valori concatenati con anno di riferimento 2010): cosicché la casistica delle trimestralizzazioni considerate comprende 68 casi distinti.<sup>5</sup> Ciascuna delle 68 disaggregazioni è stata ripetuta considerando l'intera gamma di modelli a disposizione, cioè i modelli di Chow-Lin e Fernández e la classe di modelli ADL nelle specificazioni ADL(1,0) e ADL(1,1).<sup>6</sup> Eccetto che per il modello di Fernández, per ciascuna classe di modelli sono state stimate le specificazioni base, quelle con aggiunta di costante, quelle con aggiunta di costante e trend. Per la classe dei modelli ADL, ciascuna specificazione è stata considerata e stimata sia nei livelli sia nelle differenze prime. Infine, sfruttando l'implementazione della disaggregazione temporale non lineare (Proietti, 2006; Proietti e Moauro, 2006) nella subroutine 'TRIMEKFL', ogni specificazione finora summenzionata è stata trattata anche nei logaritmi, oltre che nei livelli. In totale sono state dunque eseguite 2176 disaggregazioni temporali.

Stimare un così estesa varietà di modelli e di loro specificazioni per ciascuna serie implica senz'altro il rischio di stimare un certo numero di modelli poco o per nulla significativi. Per esempio, è questo il caso in cui si ottengono coefficienti stimati di costante e/o trend prossimi allo zero o che non superano il test di T-student. Ebbene, l'intento dell'esercizio è proprio quello di compiere una valutazione delle prestazioni dei vari modelli simulando la pratica corrente di produzione dei conti trimestrali in cui non è rara l'eventualità in cui non è possibile selezionare la specificazione migliore di modello per ogni esercizio, a causa di carenze di dati, di tempo o di impedimenti organizzativi.

L'esercizio permette di apprezzare i principali vantaggi della stima basata sulla rappresentazione SSF rispetto all'approccio di regressione, come l'ampliamento della gamma dei modelli di trimestralizzazione, la possibilità di adottare la specificazione logaritmica degli stessi modelli e la disponibilità di nuove diagnostiche basate sulla stima delle innovazioni. A tale proposito, vale la pena rimarcare come, essendo i principali modelli tradizionali di disaggregazione riconducibili, sotto i rispettivi set di restrizioni sui parametri, a casi particolari della classe di modelli ADL (si vedano i paragrafi 2.1-2.2), l'entità del contributo innovativo dell'applicazione che presentiamo, in termini di ampliamento di metodi/modelli di disaggregazione, si limita ai casi in cui la stima dei modelli ADL non verifica tali restrizioni. In questi casi i modelli ADL stimati risultano distinti dai modelli tradizionali e possono dunque considerarsi uno strumento supplementare ai modelli statici di regressione.

### 3.2 Modelli di trimestralizzazioni a confronto

Si è deciso di basare l'illustrazione comparativa della performance dei diversi modelli su due indicatori principali che colgono aspetti diversi dell'esercizio: l'errore medio assoluto (MAE) di previsione dei tassi di crescita annuali - cioè la distanza media in valore assoluto tra il tasso di crescita della serie annuale originale e quello della serie prevista annualizzata - e la correlazione tra la variazione congiunturale dell'indicatore e della serie disaggregata. E' bene sottolineare che il MAE, sintetizzando l'informazione trimestrale in informazione a frequenza annuale, fornisce una misura della bontà di adattamento complessiva del modello alla serie storica utile in fase di previsione, mentre il secondo indicatore offre uno strumento per valutare la qualità della

<sup>4</sup> Il livello di dettaglio adottato per la stima dei CET prevede la classificazione delle attività economiche in 44 branche. Per maggiori dettagli si veda ISTAT (2015).

<sup>5</sup> I commenti che seguono si riferiscono unicamente all'esercizio su dati destagionalizzati, avendo constatato la generale coincidenza dei parametri stimati del modello nel caso dei dati grezzi e destagionalizzati.

<sup>6</sup> L'esercizio è stato condotto considerando, tra gli altri, anche il modello Litterman di cui non riportiamo i risultati (disponibili su richiesta) a causa della problematicità della stima di massima verosimiglianza del parametro autoregressivo  $\phi$  che insorge laddove si adotti questo modello. Per una esemplificazione approfondita di questo aspetto si veda Proietti (2005), pp. 104-106.

disaggregazione, utile perciò a discernere quale modello sintetizza meglio l'informazione di breve periodo, a parità di indicatore.

Le tavole 2A-3A in appendice A riportano i valori del MAE ottenuti dalle trimestralizzazioni basate sulle diverse classi di modelli per ciascuna branca per le serie destagionalizzate e corrette per i giorni lavorativi, rispettivamente nel caso delle serie espresse a prezzi correnti e in volume<sup>7</sup>. I valori dell'indicatore variano da branca a branca, evidenziando la problematicità di alcune di esse - si veda ad esempio la branca 7 per cui il MAE medio dei diversi modelli è pari 16.3 nelle serie a prezzi correnti - o l'andamento virtuoso di altri casi - come la branca 6 nelle serie a prezzi correnti in cui la media dei MAE, pari a 1.8 punti percentuali, è il valore minimo tra tutte le branche. Ora, sebbene queste tavole non siano informative circa i dettagli delle specificazioni dei modelli e del relativo output di stima, consentono comunque di apprezzare la prestazione dei modelli della classe ADL i quali, rispetto ai modelli statici di regressione, prevalgono in termini di (minimo) errore di previsione annuale sulle serie a valori nominali in 13 branche su 17, mentre sulle serie in volume soltanto in 6 branche su 17. Da questo primo esame emerge che l'ampliamento delle classi dei modelli comporta comunque un vantaggio apprezzabile in fase di produzione di dati trimestrali, soprattutto in relazione alla stima dei trimestri in corso d'anno, cioè quelli per i quali il valore annuale non è ancora disponibile. Nella stessa appendice A le tavole 6A-7A completano la descrizione dell'esercizio. In esse sono riportati i dettagli della specificazione del miglior modello<sup>8</sup> secondo il criterio del MAE per ciascuna branca, in particolare: il tipo di specificazione, l'eventuale trasformazione logaritmica e/o alle differenze prime, il valore della log-verosimiglianza e i coefficienti stimati dei parametri del modello ( $\varphi$ ,  $\beta_0$ ,  $\beta_1$ , la costante  $c$  e il trend  $g$ ).

I dettagli delle specificazioni dei modelli contenuti nelle due tavole e il confronto tra di loro evidenziano come i modelli definiti nelle differenze prime si adattino meglio alle serie espresse in termini nominali (tavola 6A) piuttosto che alle serie espresse in volume (tavola 7A) coerentemente col fatto che le serie nominali includono la componente inflativa e sono perciò caratterizzate da trend (integrazione) di grado (ordine) più elevato. Riguardo alla trasformazione logaritmica, questa risulta efficace in diversi casi sia per i modelli su serie nominali che in volume.

Le tavole 4A-5A presentano i valori calcolati delle correlazioni prime tra la variazione congiunturale dell'indicatore e della serie disaggregata in ciascuna branca per i (migliori) modelli nelle diverse classi. Si tratta di tavole compilate con la stessa logica delle tavole 2A-3A. Da esse si evince facilmente come i modelli per i quali si ottiene la massima qualità della disaggregazione in termini di accostamento con l'indicatore siano quelli tradizionalmente utilizzati nella produzione dei conti trimestrali cioè i modelli Fernández e Chow-Lin nel caso delle serie nominali: in particolare, il modello di Fernández garantisce la correlazione massima in 12 branche su 17, il modello di Chow-Lin nelle restanti 5 casi su 17. L'esercizio condotto sulle serie in volume mostra ugualmente la netta prevalenza del modello di Fernández, ma evidenzia altresì un prestazione soddisfacente dei modelli ADL(1,1) che prevalgono nelle restanti 5 branche (le numero 7, 11, 13, 14, 15). Un maggior livello di dettaglio è fornito dalle tavole 8A-9A le quali, analogamente alle tavole 6A-7A, mostrano le specificazioni, i valori stimati dei parametri di regressione e la log-verosimiglianza dei modelli per i quali si ottiene la maggior correlazione in ogni branca. Da esse emerge come la trasformazione logaritmica sia nella maggior parte dei casi (26 su un totale di 38 stime tra serie nominali e in volume) opportuna ai fini di ottenere una serie disaggregata altamente correlata con l'indicatore e dunque, da questo punto di vista preferibile alla trattazione delle serie nei livelli.<sup>9</sup> E' bene inoltre ricordare come la trasformazione logaritmica delle serie storiche presenti sia i ben noti vantaggi di riduzione della volatilità o dell'eliminazione dell'eteroschedasticità delle stesse, sia la proprietà di garantire in ambito di disaggregazione

<sup>7</sup> Il valore riportato è relativo alla specificazione del modello per la quale si è ottenuto il valore minimo del MAE nell'ambito di ogni singola classe.

<sup>8</sup> Le specificazioni descritte nelle tavole 6A-7A si riferiscono, ciascuna, al modello per cui si è ottenuto il minor valore del MAE rispetto ai modelli alternativi in ogni branca e per i quali il MAE calcolato è riportato nelle tavole 2A-3A. I grafici delle corrispondenti serie disaggregate stimate sono presentate in appendice B, figure 1B-2B.

<sup>9</sup> Analogamente al caso precedente, presentiamo in appendice B, figure 3B-4B, i grafici delle serie disaggregate ottenute stimando i modelli per cui si è ottenuta la massima correlazione tra indicatore e serie disaggregata di cui alle tavole 8A-9A.

temporale che la serie disaggregata assuma solo valori positivi.<sup>10</sup> In virtù di quanto appena detto, l'ampliamento dei metodi di disaggregazione ai metodi non lineari implementati da 'TRIMEKF' e 'TRIMEKFL' appare indubbiamente un'evoluzione nel segno del perfezionamento della produzione statistica trimestrale.

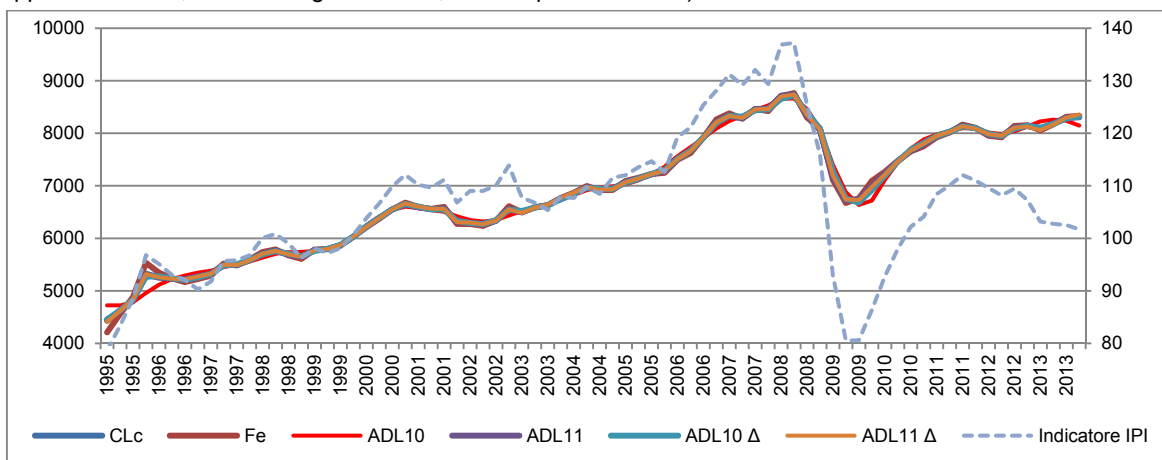
### 3.3 La procedura di selezione del miglior modello

Nel paragrafo precedente abbiamo evidenziato alcuni dei vantaggi dell'estensione della gamma statica di modelli di disaggregazione temporale alla classe dei modelli ADL. Ai fini dell'identificazione di una procedura standard di selezione del modello, in questo paragrafo presentiamo un esercizio di disaggregazione temporale quale esempio del lavoro che l'analista contabile è chiamato a svolgere avendo a disposizione la strumentazione modellistica e diagnostica presentata finora.

L'esempio è basato sulla disaggregazione della serie annuale del valore aggiunto relativa alla branca 14 ("Altri macchinari e apparecchiature") che pesa circa il 2,1% sul valore aggiunto dell'intera economia. L'indicatore trimestrale utilizzato è l'indice di produzione industriale relativo alla medesima branca. L'intervallo temporale considerato è il 1995:q1-2013:q4.

Le figure 1 e 2 presentano l'output della trimestralizzazione secondo i modelli di Chow-Lin (nella specificazione con aggiunta di costante), Fernández, ADL(1,0) e ADL(1,1) nei livelli e nelle differenze (queste ultime contrassegnate col suffisso  $\Delta$  nelle figure per distinguerle dai corrispondenti modelli dei livelli), rispettivamente nel caso di serie destagionalizzate e corrette per i giorni lavorativi, e di serie grezze e non corrette. In entrambi i casi le serie sono valutate a prezzi correnti.

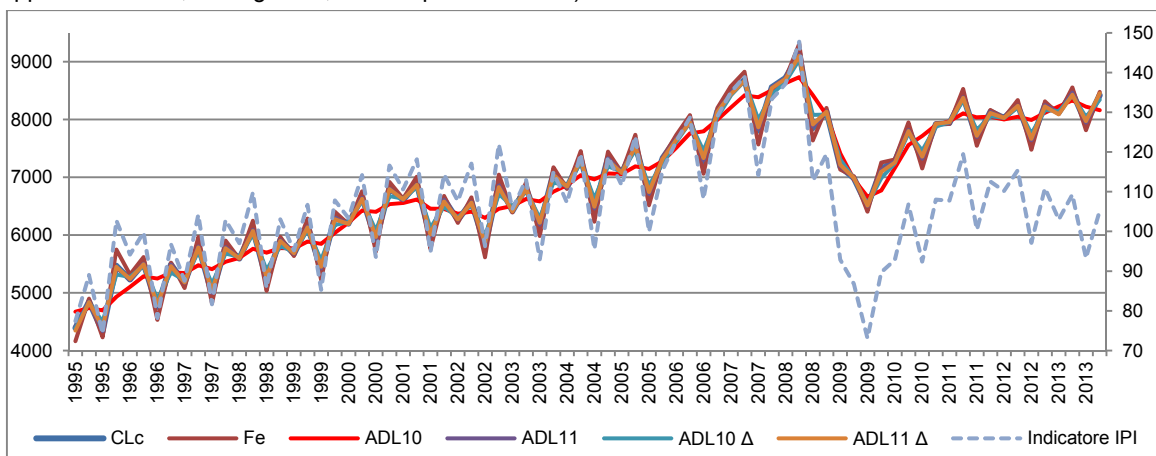
**Figura 1 – Serie trimestralizzate del valore aggiunto e indicatore IPI (Branca "Altri macchinari e apparecchiature"; serie destagionalizzate; valori a prezzi correnti)**



<sup>10</sup> Nel caso specifico delle trimestralizzazioni eseguite per il presente esercizio è stato verificato che nessun modello - con o senza trasformazione logaritmica - abbia prodotto serie contenenti valori negativi.



**Figura 2 - Serie trimestralizzate del valore aggiunto e indicatore IPI (Branca “Altri macchinari e apparecchiature”; serie grezze; valori a prezzi correnti)**



Le figure 1 e 2 mostrano come, nel caso della branca in esame, i diversi modelli producano serie trimestrali aventi andamenti tra loro simili e oscillanti nel medesimo intervallo di valori: le serie sono a tal punto sovrapposte che la semplice ispezione grafica non permette di distinguere la superiorità di un modello sugli altri. Appare tuttavia evidente come il modello ADL(1,0) nei livelli (linea rossa) produca una serie avente andamento decisamente più smussato rispetto alle altre serie, più evidente nel caso dei casi grezzi - la stagionalità appare infatti fortemente ridotta - ma comunque apprezzabile anche nel caso dei dati destagionalizzati dove la linea rossa appare quasi interpolare il “fascio” delle serie prodotte dai modelli alternativi.

Informazioni più dettagliate utili a discernere l’affidabilità dei modelli sono fornite dai tabulati prodotti in fase di stima che offrono le principali statistiche della regressione di ciascun modello e le relative diagnostiche. Ne riportiamo una sintesi nelle tavole 1 e 2, rispettivamente, relative all’esercizio condotto su serie destagionalizzate e grezze.

A riprova del fatto che le serie storiche considerate non presentano particolari anomalie, notiamo innanzitutto come i modelli identificati sui dati grezzi e sui dati destagionalizzati siano tra loro del tutto simili: il parametro autoregressivo e i coefficienti di regressione stimati, nonché le statistiche calcolate (MAE, indici di correlazione, criteri informativi AIC e BIC,  $R^2$  corretto) sono pressoché coincidenti nei due casi. Come vedremo, infatti, il criterio-guida del processo di selezione del modello di disaggregazione sarà univoco e condurrà al medesimo risultato in entrambi i casi.

In secondo luogo rileviamo che la subottimalità del modello ADL(1,0) nei livelli emersa a livello grafico trova conferma nei valori riportati in tavola poiché non solo ad esso è associato il più alto valore del MAE (6.47 nel caso destagionalizzato, 6.39 nel caso grezzo) rispetto agli altri modelli, ma anche i più bassi valori delle correlazioni tra le serie trimestralizzata e indicatore in variazioni prime.

**Tavola 1 - Parametri stimati e statistiche calcolate nell'esercizio di disaggregazione temporale della serie del valore aggiunto (Branca 14, dati destagionalizzati e corretti per i giorni lavorativi; serie a prezzi correnti)**

|           | $\rho$ | $c$     | $\beta_0$ | $\beta_1$ | MAE  | Correlazioni Congiunturali | Correlazioni Tendenziali | Log-lik | AIC  | BIC  | $R^2$ corretto |
|-----------|--------|---------|-----------|-----------|------|----------------------------|--------------------------|---------|------|------|----------------|
| CLc       | 0.996  | 11.49** | 38.47**   | --        | 3.15 | 0.96                       | 0.93                     | -155.35 | 9.8  | 9.9  | 0.996          |
| Fe        | 0      | --      | 51.60**   | --        | 3.30 | 0.93                       | 0.94                     | -158.86 | 10.3 | 10.3 | 0.993          |
| ADL(1,0)  | 0.874  | --      | 8.48**    | --        | 6.47 | 0.75                       | 0.89                     | -167.14 | 11.5 | 11.5 | 0.975          |
| ADL(1,1)  | 0.983  | --      | 37.58**   | -36.32**  | 2.95 | 0.96                       | 0.93                     | -153.58 | 9.7  | 9.8  | 0.996          |
| ADL(1,0)Δ | 0.224  | --      | 31.49**   | --        | 2.84 | 0.80                       | 0.68                     | -152.52 | 9.2  | 9.3  | 0.998          |
| ADL(1,1)Δ | 0.668  | --      | 36.24**   | -22.57**  | 2.51 | 0.80                       | 0.67                     | -151.88 | 8.2  | 8.4  | 0.999          |

**Tavola 2 - Parametri stimati e statistiche calcolate nell'esercizio di disaggregazione temporale della serie del valore aggiunto** (Branca 14, dati grezzi e non corretti per i giorni lavorativi; serie a prezzi correnti)

|                   | $\rho$ | $c$     | $\beta_0$ | $\beta_1$ | MA<br>E | Correlazioni<br>Congiunturali | Correlazioni<br>Tendenziali | Log-lik | AIC  | BIC  | $R^2$ corretto |
|-------------------|--------|---------|-----------|-----------|---------|-------------------------------|-----------------------------|---------|------|------|----------------|
| CLc               | 0.996  | 11.22** | 37.66**   | --        | 2.98    | 0.99                          | 0.93                        | -155.0  | 9.8  | 9.9  | 0.996          |
| Fe                | 0      | --      | 51.48**   | --        | 3.19    | 0.99                          | 0.91                        | -159.2  | 10.3 | 10.4 | 0.992          |
| ADL(1,0)          | 0.876  | --      | 8.33**    | --        | 6.39    | 0.62                          | 0.88                        | -166.6  | 11.4 | 11.5 | 0.976          |
| ADL(1,1)          | 0.982  | --      | 36.62**   | -35.44**  | 2.82    | 0.99                          | 0.93                        | -153.11 | 9.7  | 9.8  | 0.996          |
| ADL(1,0) $\Delta$ | 0.235  | --      | 30.43**   | --        | 2.78    | 0.85                          | 0.76                        | -152.02 | 9.2  | 9.3  | 0.998          |
| ADL(1,1) $\Delta$ | 0.685  | --      | 35.12**   | -22.44**  | 2.45    | 0.87                          | 0.74                        | -151.01 | 8.1  | 8.2  | 0.999          |

Coerenti con tale analisi comparativa risulta anche l'output dei criteri informativi AIC e BIC - il cui valore è inversamente proporzionale alla bontà del modello stimato – entrambi decisamente più alti per il modello ADL(1,0) rispetto a quelli associati ai modelli alternativi, nonché della statistica sulla bontà complessiva di adattamento ( $R^2$  corretto) che risulta la minore tra tutte.

Ad eccezione del modello appena esaminato, la prestazione della classe dei modelli ADL risulta complessivamente all'altezza dello del modello di Chow-Lin. In particolare, il modello che appare distintamente il migliore in termini di bontà complessiva di adattamento alla serie nonché di capacità predittiva è l'ADL(1,1) nelle differenze. Per quanto riguarda il primo aspetto, sia il valore della Log-likelihood, sia il valore dell'  $R^2$  corretto, sia i criteri informativi AIC e BIC – che riportano per questo modello un valore più basso di circa il 20% rispetto alla media dei valori associati agli altri modelli - indicano questo modello come quello che meglio si adatta all'esercizio di disaggregazione in esame. Per quanto riguarda l'aspetto previsivo, il valore del MAE associato all'ADL(1,1) (circa 2.5 nel caso di serie stagionalizzate e grezze) risulta decisamente inferiore agli altri modelli (per i quali si va da un valore minimo di 2.8 a un valore massimo di circa 6.5).

### 3.4 Trimestralizzazione delle serie grezze

La stima dei modelli di disaggregazione temporale è generalmente invariante all'uso di serie annuali corrette per i giorni lavorativi e indicatori stagionalizzati o di serie annuali non corrette e indicatori grezzi e l'esercizio di disaggregazione fin qui presentato ne dà conferma: i coefficienti di regressione, il parametro autoregressivo e la log-verosimiglianza sono pressoché coincidenti nel caso dell'utilizzo dell'uno e dell'altro tipo di dati.<sup>11</sup> Tuttavia, alcune disaggregazioni temporali basate su indicatori non stagionalizzati hanno evidenziato lo specifico problema di mancato trasferimento della dinamica stagionale dall'indicatore alla serie indicata.

L'applicazione alle serie del valore aggiunto fin qui discussa evidenzia tale scollamento tra dinamica stagionale dell'indicatore grezzo e quella della serie disaggregata - la quale risulta caratterizzata da un andamento "smussato" e da una traiettoria che interpola quella delle serie prodotte dai modelli alternativi - quasi sistematicamente in corrispondenza dell'utilizzo dei modelli ADL(1,0) nei livelli.

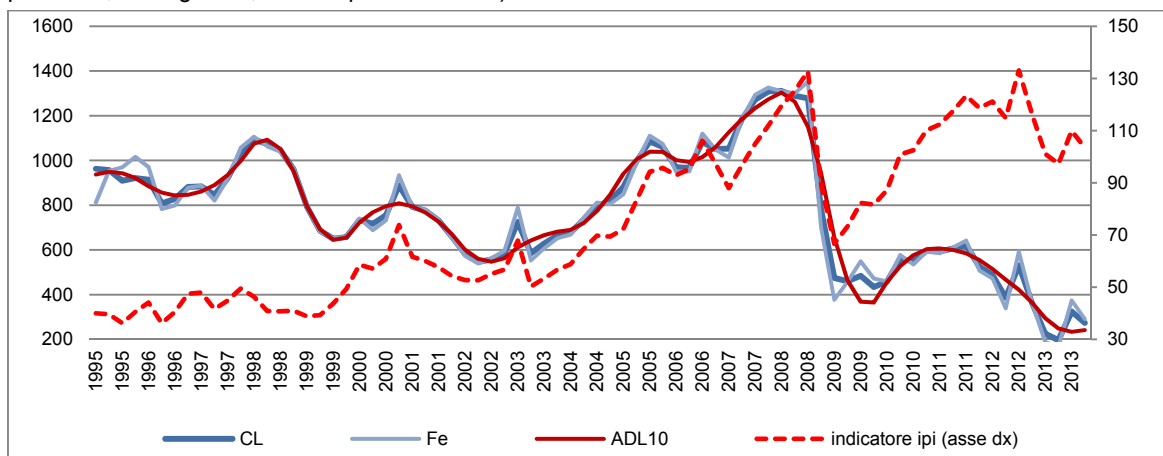
Da un punto di vista grafico, il problema è descritto nelle figure 3 - 4 in cui si riportano a titolo esemplificativo le serie trimestralizzate del valore aggiunto della branca relativa a "Coke e raffinazione di prodotti petroliferi" espresso a valori correnti (figura 3) e a valori concatenati (figura 4), utilizzando alcuni approcci di stima. Alle serie trimestralizzate si accompagna l'indice di produzione trimestrale grezzo riferito alla stessa branca utilizzato come indicatore nella stima. In entrambe le figure si nota facilmente l'assenza di stagionalità della serie associata al modello ADL(1,0) (linea continua rossa) così come il fatto che la sua traiettoria interpola le serie stimate tramite i modelli alternativi Chow-Lin e Fernández. Queste ultime, al contrario, sono caratterizzate da una dinamica stagionale molto aderente dell'indicatore, in entrambi gli esercizi (fig. 3-4).

A complemento dell'analisi grafica, riportiamo nella tavola 10A in Appendice A i valori delle correlazioni tra i tassi di variazione congiunturali e tendenziali tra indicatore grezzo e serie

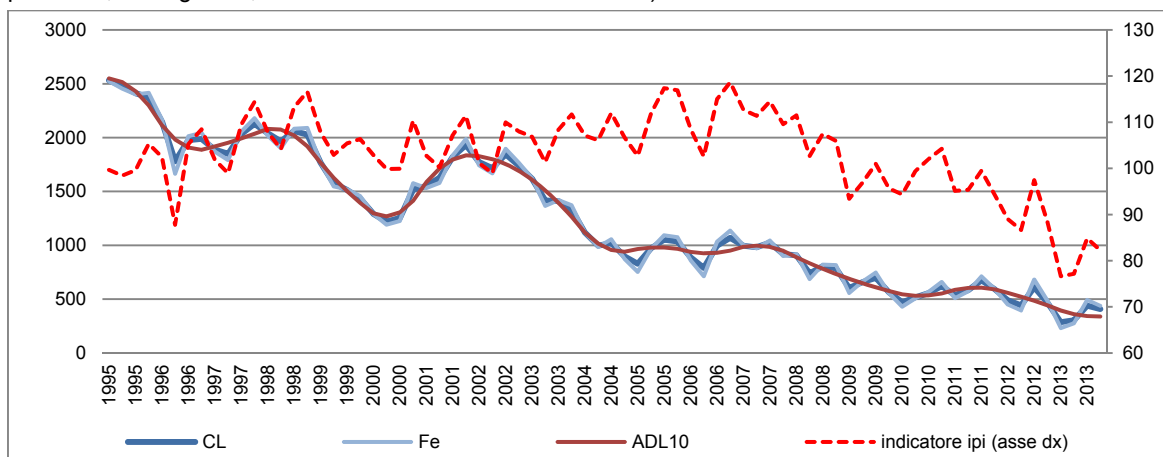
<sup>11</sup> Nel testo e nell'appendice abbiamo riportato finora i tabulati relativi alle stime basate sui soli dati stagionalizzati, per brevità. Quelle basate su dati annuali non corretti/indicatori grezzi sono disponibili su richiesta.

trimestralizzata grezza per ciascuna classe di modello stimata nei livelli.<sup>12</sup> A conferma della tipicità riscontrata per la classe ADL(1,0) ai livelli nel contesto della disaggregazione temporale di serie grezze, i valori delle correlazioni tra tassi congiunturali ( $\Delta q$ ) ad essa associati risultano minori a quelli associati ai modelli alternativi ritenuti ammissibili<sup>13</sup> nell'82% circa dei casi, sia per le serie nominali che in volume. D'altra parte, si noti come nella maggior parte dei casi lo scarto tra le correlazioni delle variazioni prime (dinamica trimestrale) della classe ADL(1,0) rispetto agli altri modelli sia più ampio dello scarto tra le relative correlazioni quarte (dinamica annuale) a riprova del fatto che, a livello di bontà di adattamento complessiva, questa classe di modelli risulta equivalente agli altri.

**Figura 3 – Serie trimestralizzate del valore aggiunto e indicatore IPI (Branca “Coke e raffinazione prodotti petroliferi; serie grezze; valori a prezzi correnti)**



**Figura 4 – Serie trimestralizzate del valore aggiunto e indicatore IPI (Branca “Coke e raffinazione prodotti petroliferi; serie grezze; valori concatenati – anno di rif. 2010)**



<sup>12</sup> Si tratta del valore medio delle correlazioni calcolate per le tre specificazioni possibili - cioè modello senza costante o trend, con costante, con costante e trend. Si noti che le variazioni congiunturali delle serie trimestrali grezze non forniscono un'informazione valutabile dal punto di vista economico per via della presenza della componente stagionale della serie che si sovrappone a quella ciclica. Tuttavia le correlazioni basate sulle variazioni prime riportate in Tavola 10A consentono di valutare la qualità della trimestralizzazione, relativamente al grado di adattamento a frequenza trimestrale della serie obiettivo rispetto all'indicatore.

<sup>13</sup> Ai fini di tale comparazione si escludono i modelli per i quali i valori delle relative correlazioni sono negative o positive ma strettamente inferiori al 0,2, che consideriamo come limite inferiore per l'ammissibilità di un modello.

### 3.5 Una valutazione delle disaggregazioni: criteri di scelta e risultati ammissibili

L'ampliamento dei metodi di trimestralizzazione alla classe dei modelli dinamici presenta il vantaggio di poter scegliere la migliore disaggregazione temporale tra un più alto numero di possibili modelli. Per contro lo stesso ampliamento rende più difficile la scelta. La soluzione a questo trade-off che si riscontra nella letteratura è di adottare un criterio dal generale al particolare, con delle regole determinate che guidano la selezione del modello da una forma ampia e non vincolata a una forma più parsimoniosa e ristretta spesso in maniera automatica. Nell'ambito delle regressioni dinamiche si veda Castle *et al.* (2011).

Ai fini dell'individuazione dei modelli validi, un criterio dal generale al particolare è stato applicato a questo esercizio con riferimento alla presenza o meno della costante e/o del trend nella definizione di tutti i modelli di disaggregazione, ad eccezione del modello di Fernández. Muovendo dai modelli statici a quelli dinamici il criterio adottato si differenzia ulteriormente rispetto all'ordine della differenziazione dei dati e all'ordine del ritardo del modello. Infine la trasformazione logaritmica costituisce un'ulteriore possibile modalità di specificazione di ciascun modello di cui tener conto. Nello specifico del nostro esercizio abbiamo verificato le prestazioni di ciascuna specificazione considerando i principali criteri che guidano la costruzione dei conti trimestrali. In particolare abbiamo verificato che: i) i coefficienti stimati dal modello associati agli indicatori e alle eventuali componenti deterministiche fossero statisticamente significativi (almeno al 5%); ii) il coefficiente stimato del parametro autoregressivo del modello fosse maggiore di zero<sup>14</sup> per evitare l'erraticità delle stime; iii) la correlazione della dinamica congiunturale e tendenziale tra serie indicata e indicatore avesse valore positivo e ragionevolmente elevato iv) la serie trimestralizzata avesse un andamento in linea con l'indicatore in termini grafici.

I risultati di tale analisi sono presentati nella tavola 3. In base ai criteri di cui sopra, la quota di applicazioni valide ottenute dall'intero esercizio (2176 trimestralizzazioni) è pari a circa il 49%. Tale quota risulta leggermente più alta per gli esercizi a prezzi correnti (intorno al 51%) rispetto a quelli in volume (intorno al 48%). Tale risultato può essere un'indicazione che favorisce l'attuale approccio di stima dei conti trimestrali che vede prima la stima degli aggregati nominali, più facili da stimare, e poi di quelli in volume, che necessitano di un maggior lavoro per determinare la stima migliore.

**Tavola 3 - Quote di trimestralizzazioni ritenute valide per ciascuna classe di modello e complessivamente. Valori % sul totale delle branche stimate.**

|                | Aggiustamento stagionale |                        | Valutazione     |                    | Trasformazione      |                       | Quote complessive |
|----------------|--------------------------|------------------------|-----------------|--------------------|---------------------|-----------------------|-------------------|
|                | Dati Grezzi              | Dati Destagionalizzati | Prezzi correnti | Valori concatenati | Modelli nei Livelli | Modelli nei Logaritmi |                   |
| Totale modelli | 49.2                     | 49.6                   | 50.9            | 47.9               | 47.0                | 51.8                  | 49.4              |
| CL             | 71.1                     | 72.1                   | 70.1            | 73.0               | 64.2                | 78.9                  | 71.6              |
| FE             | 100                      | 100                    | 100             | 100                | 100                 | 100                   | 100               |
| ADL(1,0)       | 58.8                     | 63.2                   | 58.8            | 63.2               | 61.3                | 60.8                  | 61.0              |
| ADL(1,1)       | 52.0                     | 52.5                   | 56.9            | 47.5               | 49.0                | 55.4                  | 52.2              |
| ADL(1,0)Δ      | 33.3                     | 32.4                   | 36.3            | 29.4               | 30.4                | 35.3                  | 32.8              |
| ADL(1,1)Δ      | 12.3                     | 10.8                   | 14.7            | 8.3                | 10.3                | 12.7                  | 11.5              |

Con riferimento al tipo di stima, la prestazione dei modelli nei logaritmi garantisce un maggior numero di esercizi ammissibili, corrispondente a circa il 52% dei casi contro il 47% delle stime effettuate nei livelli. La relativa maggior stabilità delle stime non lineari è verificata sia con

<sup>14</sup> Con l'unica eccezione per la classe dei modelli ADL alle differenze per cui sono stati ammesse anche le stime in cui il parametro autoregressivo stimato avesse segno negativo ma valore prossimo allo zero e in particolare, superiore a -0.2. I modelli ADL alle differenze infatti, per valori di  $\rho = 0$  collasano nella specificazione del modello di Fernández, dunque il nostro criterio si basa sull'inclusione delle fattispecie che approssimano quest'ultimo.

riferimento alle trimestralizzazioni ottenute con indicatori non destagionalizzati (51% per i modelli nei logaritmi, 47% per i modelli nei livelli) che per quelle ottenute con indicatori destagionalizzati (53% per i modelli nei logaritmi contro 47% per i modelli nei livelli).

Riguardo alla forma dei modelli, un primo risultato da mettere in evidenza è che nessuna classe fallisce nel fornire disaggregazioni ammissibili. A parte Fernández che non abbiamo diversificato per la presenza di costante e/o trend e che quindi è sempre ammissibile, il modello di Chow-Lin risulta globalmente valido nel 72% dei casi, seguito dai modelli ADL nei livelli, 62% per l'ADL(1,0) e 52% per l'ADL(1,1) e, infine, dai modelli ADL alle differenze per i quali otteniamo minori tassi di trimestralizzazioni ammissibili pari al 33% per gli ADL(1,0) e al 12% per gli ADL(1,1).

L'esercizio evidenzia dunque che la relativa maggior complessità della definizione dei modelli determina una minore probabilità di ottenere coefficienti statisticamente significativi (è questa infatti la causa più frequente di invalidità delle stime): sia nel caso di stime lineari che in quelle non lineari, infatti, i modelli ADL(1,0) assicurano un maggior numero di stime valide (63% e 61% dei casi, rispettivamente) rispetto ai modelli ADL(1,1) (49% e 55% dei casi rispettivamente).

L'ampliamento della gamma di modelli alla specificazione alle differenze ci permette di apprezzare come quest'ultima si adatti relativamente meglio se applicata ad aggregati espressi a valori nominali piuttosto che in volume, il che non ci sorprende poiché nel primo caso la componente inflativa indotta dai prezzi sembra essere opportunamente trattata dalla differenziazione. In particolare, l'applicazione dei modelli ADL(1,0) e ADL(1,1) alle differenze agli aggregati espressi a prezzi correnti risulta valida, rispettivamente, nel circa 36% e 29% dei casi rispettivamente, contro, il 15% e l'8% circa di casi validi quando applicati agli aggregati in volume.

### 3.6 La disaggregazione temporale senza indicatore: procedure a confronto

Tra i vantaggi della rappresentazione SSF e della disaggregazione temporale basata sul filtro di Kalman vi è la possibilità di eseguire in modo efficiente la trimestralizzazione di una serie annuale anche in assenza di una serie ad alta frequenza che guidi l'andamento di breve periodo della serie disaggregata, cioè di un indicatore. Nella prassi della contabilità nazionale, in mancanza di un indicatore trimestrale si usa ricorrere al metodo Denton (1971) secondo l'approccio classico di regressione e all'utilizzo di una serie fittizia - quale ad esempio un trend lineare deterministico - in sostituzione dell'indicatore trimestrale. Tale metodo, come rilevato diffusamente in letteratura (ad esempio si veda Feijoò *et al.* (2003)) non possiede requisiti di ottimalità e induce una forte distorsione nelle prime osservazioni delle serie disaggregate a causa delle condizioni/restrizioni imposte strutturalmente al primo trimestre della serie, come mostrato in Sanz (1981).

A titolo esemplificativo, abbiamo messo a confronto la procedura di trimestralizzazione basata sul metodo di Denton (1971) implementata dalla subroutine 'TRIMEQNA' con quella basata sul modello Chow-Lin (1971) implementata dalla subroutine 'TRIMEKF'. In particolare, nel primo caso abbiamo considerato quale indicatore di riferimento un semplice trend lineare, nel secondo abbiamo la serie disaggregata è stimata senza l'impiego di alcun indicatore. Il confronto è stato effettuato ancora una volta a partire dalle 17 serie storiche che compongono la sezione "Industria" del valore aggiunto, tuttavia nei grafici 5B-6B in Appendice B riportiamo l'esito soltanto delle 4 maggiori branche che compongono l'industria, rispettivamente, per i dati nominali e in volume rispettivamente. Le figure 5B - 6B confermano la problematicità del modello di Denton per quanto riguarda le prime osservazioni trimestrali e la sua subottimalità rispetto alla serie prodotte in assenza di indicatore dalla procedura TRIMEKF, a prescindere dalla natura dei dati utilizzati. In tale circostanza quindi, la subroutine di nuova introduzione garantisce un miglioramento della qualità della pratica di produzione corrente dei conti trimestrali.

#### 4. Conclusioni

Il presente lavoro ha illustrato i recenti sviluppi realizzati dall'ISTAT nell'ambito delle tecniche di disaggregazione temporale nella prospettiva di una loro introduzione nella pratica corrente di produzione dei conti economici trimestrali. In tale ottica, accanto alla descrizione delle nuove metodologie si è proceduto a descrivere i risultati di un'estesa sperimentazione basata su serie storiche ISTAT per una valutazione in senso ampio della loro prestazione.

Le principali innovazioni introdotte riguardano: l'ampliamento della classe dei modelli classici di disaggregazione temporale alla classe dei modelli dinamici ADL, l'adozione dell'approccio spazio degli stati per il trattamento statistico, l'ampliamento del preesistente set diagnostico dei risultati di stima, l'introduzione della disaggregazione temporale non lineare (o nei logaritmi) oltre che nei livelli, la completa integrazione delle nuove tecniche in un insieme di nuove procedure informatiche in ambiente Speakeasy/Modeleasy.

I risultati dell'esercizio empirico, lungi dall'indicare una gerarchia tra i modelli alternativi, hanno messo in evidenza le potenzialità della nuova classe di modelli dinamici. Invero, se da un lato abbiamo osservato una più alta capacità previsiva della classe dei modelli ADL rispetto ai modelli statici, dall'altro questi ultimi producono un più alto numero di casi in cui la serie disaggregata è più aderente all'indicatore in termini di dinamica congiunturale.

Ai fini della selezione del miglior modello l'indicazione che abbiamo dedotto e che dovrebbe guidare la pratica del contabile nazionale è quella di temperare la scelta di un modello che produca una buona estrapolazione della serie annuale (bassa frequenza) a fronte dell'aderenza della serie disaggregata con l'indicatore congiunturale (alta frequenza). Ovviamente quanto più indicato e indicatore sono aderenti tanto meno questo trade-off risulta stringente.

Relativamente all'ampliamento dei metodi di disaggregazione ai metodi non lineari, l'applicazione ha messo in evidenza più volte come questi ultimi risultino preferibili rispetto ai metodi nei livelli. Ne deduciamo che la possibilità di trattare il problema della disaggregazione temporale nei logaritmi costituisca un'evoluzione di assoluto interesse nella direzione di migliorare la qualità dell'informazione statistica.

Un ulteriore interessante aspetto che è emerso nella sperimentazione riguarda l'introduzione di metodi basati su algoritmi in grado di produrre stime in assenza di indicatore. L'esempio presentato ha evidenziato la miglior qualità della serie disaggregata da essa prodotta rispetto alla corrispondente serie prodotta dal modello di Denton.

Tra le criticità emerse dall'impiego dei modelli dinamici si è riscontrato soltanto il mancato trasferimento della dinamica stagionale dell'indicatore alla serie disaggregata per la classe dei modelli ADL(1,0).

In conclusione, l'esercizio ha dimostrato che l'ampliamento alla nuova classe di modelli, degli algoritmi di stima e della specificazione non lineare consente in un numero significativo di casi di migliorare la qualità della trimestralizzazione. Tuttavia i relativi limiti messi in evidenza, seppure ridotti a pochi casi, suggeriscono che la più ampia strumentazione a disposizione si debba accompagnare ad una maggiore capacità critica e una più ampia visione dell'analista.

## Riferimenti bibliografici

- Astolfi, R. e M. Marini. 2005. Procedure di Disaggregazione Temporale utilizzate dall'ISTAT per la Stima dei Conti Economici Trimestrali in: Rapporto finale della Commissione di studio sul trattamento dei dati ai fini dell'analisi congiunturale, ISTAT, Ottobre 2005.
- Barbone L., G. Bodo e I. Visco. 1981. Costi e profitti in senso stretto: un'analisi su serie trimestrali, 1970-1980. *Bollettino della Banca d'Italia*, 36, numero unico.
- Bruno G. e G. Marra. 2005. New features for time series temporal disaggregation in the Modeleasy+ environment, in: Rapporto finale della Commissione di studio sul trattamento dei dati ai fini dell'analisi congiunturale, ISTAT, Ottobre 2005.
- Castle J.L., Doornik J. A., Hendry D. F. (2011) Evaluating automatic model selection, *Journal of Time Series econometrics*, 3(1), DOI: 10.2202/1941-1928.1097.
- Chow, G. C. and A. Lin. 1971. Best linear unbiased interpolation, distribution and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53, 372-5.
- De Jong P. 1991. The Diffuse Kalman Filter. *Annals of statistics*, 19: 1073-1083.
- Denton, F. T. 1971. Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization. *Journal of the American Statistical Association*, 66, 99-102.
- Di Fonzo, T. 2005. Il lavoro svolto e i risultati ottenuti, in: Rapporto finale della Commissione di studio sul trattamento dei dati ai fini dell'analisi congiunturale, ISTAT, Ottobre 2005.
- Durbin, J., e G. S. Watson. 1951a. Testing for Serial Correlation in Least Squares Regression. *Biometrika* 37: 409-428.
- Durbin, J., e G. S. Watson. 1951b. Testing for Serial Correlation in Least Squares Regression. *Biometrika* 38: 159-179.
- Feijó S.R. , Caro A.R. e D.D. Quintana. 2003. Methods for quarterly disaggregation without indicators; a comparative study using simulation. *Computational Statistics & Data Analysis*, 43: 63-78.
- Fernández, R. B. 1981. A methodological note on the estimation of time series. *The Review of Economics and Statistics* 63, 471-6.
- Frale C, Marcellino M, Mazzi G., Proietti T. 2010. Survey data as coincident or leading indicators, *Journal of Forecasting*, 29: 109-131.
- Frale C, Marcellino M, Mazzi G.L., Proietti T. 2011. Euromind: a monthly indicator of the euro area economic conditions. *Journal of the Royal Statistical Society*, A174(2): 439-470.
- Grassi S., Proietti T., Frale C., Marcellino M. and Mazzi G. 2014. EuroMInd-C: a Disaggregate Monthly Indicator of Economic Activity for the Euro Area and member countries. *International Journal of Forecasting*, accepted for publication.
- Harvey, A. C., Pierce R. G. 1984. Estimating missing observations in economic time series. *Journal of the American statistical Association*, 79: 125-131.
- Harvey, A. C. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Harvey A. e C.H. Chung. 2000. Estimating the underlying change in unemployment in the UK. *Journal of the Royal Statistical Society*, A, 163: 303-328.
- Harvey A.C. e S.J. Koopman. 1997. Multivariate Structural Time Series Models (with comments), in C. Heij, J.M. Shumacher, B. Hanzon and C. Praagman (eds.), *System Dynamics in Economic and Financial Models*, Wiley, New York: 269-298.
- Hendry, D.F., e Mizon, G.E. 1978. Serial correlation as a convenient simplification, not a nuisance: A comment on a study of the demand for money by the Bank of England. *Economic Journal*, 88: 549-563.

- ISTAT. 2005. Rapporto finale della Commissione di studio sul trattamento dei dati ai fini dell'analisi congiunturale.
- ISTAT. 2015. I Conti Economici Trimestrali - Principali elementi informativi. Nota Informativa pubblicata il 2 settembre 2015. <http://www.istat.it/it/archivio/167411>
- Jarque, C. M. e A. K. Bera. 1987. A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, 55.: 163–172.
- R.E. Kalman. 1960. A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME. *Journal of Basic Engineering*, 1960, 82, (Series D): 35-45.
- Koopman, S. J. 1997. Exact initial Kalman filtering and smoothing for non-stationary time series models. *Journal of the American Statistical Association*, 9: 1630–8.
- Litterman R.B. 1983. A random walk, Markov model for the distribution of time series. *Journal of Business and Economic Statistics*, 1: 169-173.
- Ljung, G. M. e G.E.P. Box. 1978. On a Measure of a Lack of Fit in Time Series Models. *Biometrika*, 65,2: 297–303.
- Lupi C. e G. Parigi G. 1996. La disaggregazione temporale di serie economiche: un approccio econometrico. *ISTAT, Quaderni di Ricerca*, 3.
- Moauo F. 2014. Monthly Employment Indicators of the Euro Area and Larger Member States: Real-Time Analysis of Indirect Estimates. *Journal of forecasting*, 33, 5: 339-49.
- Moauo F. e G. Savio. 2005. Temporal disaggregation using multivariate structural time series models. *Econometrics Journal*, 8: 214-234.
- Proietti T. 2005. Temporal disaggregation by state space methods: dynamic regression methods revisited, in: “Rapporto finale della Commissione di studio sul trattamento dei dati ai fini dell'analisi congiunturale”, ISTAT, Ottobre 2005.
- Proietti T. 2006. On the Estimation of Nonlinearly Aggregated Mixed Models. *Journal of Computational and Graphical Statistics*, 15,1: 18-38.
- Proietti T. e F. Moauo. 2006. Dynamic factor analysis with non-linear temporal aggregation constraints. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 55,2: 1467-9876.
- Santos Silva J.M.C. e F. N. Cardoso. 2001. The Chow-Lin method using dynamic models. *Economic Modelling*, 18 : 269-280.
- Sanz, R. 1981. Metodos de desagregacion temporal de series economicas. *Banco Espana. Servicio Estudios Economicos*, 22.



## Appendice A – Tavole

**Tavola 1A – Principali caratteristiche delle serie utilizzate nell’esercizio: classificazione delle branche di attività economica, valutazione e periodo di riferimento**

| Branca  | ISIC rev.4<br>NACE rev.2 | Divisioni<br>ATECO | Valutazione              | Periodo   |
|---|--------------------------|--------------------|--------------------------|-----------|
| 3 Attività estrattiva   | B                        | 05 - 09            | Prezzi correnti / Volume | 1995-2013 |
| 4 Industrie alimentari, delle bevande e del tabacco                 | CA                       | 10 - 12            | Prezzi correnti / Volume | 1995-2013 |
| 5 Industrie tessili, confezione e articoli di abbigliamento e pelli | CB                       | 13-15              | Prezzi correnti / Volume | 1995-2013 |
| 6 Legno, carta e stampa   | CC                       | 16-18              | Prezzi correnti / Volume | 1995-2013 |
| 7 Coke e raffinazione di prodotti petroliferi                       | CD                       | 19                 | Prezzi correnti / Volume | 1995-2013 |
| 8 Chimici   | CE                       | 20                 | Prezzi correnti / Volume | 1995-2013 |
| 9 Farmaceutici  | CF                       | 21                 | Prezzi correnti / Volume | 1995-2013 |
| 10 Gomma, plastica e minerali non metalliferi                       | CG                       | 22-23              | Prezzi correnti / Volume | 1995-2013 |
| 11 Metallurgia e prodotti in metallo                                | CH                       | 24-25              | Prezzi correnti / Volume | 1995-2013 |
| 12 Prodotti elettronici, ottici, e di misurazione                   | CI                       | 26                 | Prezzi correnti / Volume | 1995-2013 |
| 13 Apparecchiature elettriche e domestiche                          | CJ                       | 27                 | Prezzi correnti / Volume | 1995-2013 |
| 14 Altri macchinari e apparecchiature                               | CK                       | 28                 | Prezzi correnti / Volume | 1995-2013 |
| 15 Autoveicoli  | CL                       | 29                 | Prezzi correnti / Volume | 1995-2013 |
| 16 Altri mezzi di trasporto   |                          | 30                 | Prezzi correnti / Volume | 1995-2013 |
| 17 Mobili e altre manifatturiere                                    | CM                       | 31-33              | Prezzi correnti / Volume | 1995-2013 |
| 18 Energia elettrica, gas e vapore                                  | D                        | 35                 | Prezzi correnti / Volume | 1995-2013 |
| 19 Acqua e raccolta e trattamento di rifiuti                        | E                        | 36-39              | Prezzi correnti / Volume | 1995-2013 |

**Tavola 2A – Errore medio assoluto di previsione (MAE) dei tassi di crescita annuali (Serie destagionalizzate a prezzi correnti) (a) (b)**

| Modello  | branca<br>3 | branca<br>4 | branca<br>5 | branca<br>6 | branca<br>7  | branca<br>8 | branca<br>9 | branca<br>10 | branca<br>11 | branca<br>12 | branca<br>13 | branca<br>14 | branca<br>15 | branca<br>16 | branca<br>17 | branca<br>18 | branca<br>19 |
|----------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| CL       | 6.64        | 2.44        | <b>2.98</b> | 1.70        | <b>14.61</b> | 3.72        | 2.48        | 2.47         | <b>1.89</b>  | 4.23         | 2.98         | 2.15         | 4.12         | <b>2.19</b>  | 2.17         | 4.63         | 11.65        |
| FE       | 6.98        | 3.18        | 3.27        | 2.20        | 20.35        | 4.46        | 2.40        | 4.18         | 2.71         | 3.36         | 2.30         | 3.30         | 4.84         | 2.69         | 4.13         | 7.45         | 6.44         |
| ADL(1,0) | <b>6.53</b> | <b>1.81</b> | 3.02        | <b>1.47</b> | 15.59        | 4.02        | <b>2.35</b> | 2.45         | 3.25         | 3.39         | <b>2.29</b>  | <b>1.49</b>  | 2.26         | 2.89         | 2.24         | 3.70         | <b>2.99</b>  |
| ADL(1,1) | 7.52        | 1.83        | 3.09        | 1.85        | 14.62        | <b>3.53</b> | 2.40        | <b>2.29</b>  | 2.28         | <b>3.53</b>  | 2.37         | 1.59         | <b>1.88</b>  | 2.53         | <b>1.91</b>  | <b>3.56</b>  | 3.08         |

(a) Per ciascuna branca, si riporta in tabella il valore corrispondente alla specificazione del modello per cui l'errore stimato è il minimo.

(b) Ciascuna classe di modelli (Chow-Lin, Fernández, ADL(1,0), ADL(1,1)) è stata stimata nei livelli e nei logaritmi. Nel caso dei modelli 'Chow-Lin', 'ADL(1,0)' e 'ADL(1,1)', oltre alla specificazione base, è stato stimato anche il modello con aggiunta di costante, e di costante e trend. Nel solo caso dei modelli della classe ADL, tutte le specificazioni sono state stimate anche nelle differenze prime, sia nei livelli che nei logaritmi.

**Tavola 3A – Errore medio assoluto di previsione (MAE) dei tassi di crescita annuali (Serie destagionalizzate in volume) (a) (b)**

| Modello  | branca<br>3 | branca<br>4 | branca<br>5 | branca<br>6 | branca<br>7 | branca<br>8 | branca<br>9 | branca<br>10 | branca<br>11 | branca<br>12 | branca<br>13 | branca<br>14 | branca<br>15 | branca<br>16 | branca<br>17 | branca<br>18 | branca<br>19 |
|----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| CL       | <b>7.58</b> | 2.19        | <b>3.90</b> | <b>1.85</b> | <b>7.24</b> | 5.70        | 3.20        | <b>1.60</b>  | 2.66         | 4.18         | 3.63         | 2.22         | <b>2.90</b>  | 2.55         | 3.41         | <b>3.31</b>  | 4.49         |
| FE       | 8.36        | <b>1.97</b> | 4.65        | 2.22        | 10.17       | 7.67        | 4.27        | 1.86         | <b>2.36</b>  | 3.46         | <b>3.23</b>  | 2.72         | 3.12         | <b>2.03</b>  | 3.98         | 3.61         | 7.54         |
| ADL(1,0) | 8.67        | 1.99        | 4.42        | 2.00        | 7.96        | <b>5.57</b> | <b>3.09</b> | 2.39         | 2.82         | 4.15         | 3.48         | <b>1.71</b>  | 5.39         | 2.36         | 3.18         | 3.54         | <b>3.57</b>  |
| ADL(1,1) | 7.65        | 2.27        | 3.99        | 1.87        | 8.26        | 5.63        | 3.27        | 2.51         | 3.13         | <b>3.45</b>  | 3.60         | 1.73         | 3.13         | 2.56         | <b>3.13</b>  | 3.56         | 4.23         |

(a) Per ciascuna branca, si riporta in tabella il valore corrispondente alla specificazione del modello per cui l'errore stimato è il minimo.

(b) Ciascuna classe di modelli (Chow-Lin, Fernández, ADL(1,0), ADL(1,1)) è stata stimata nei livelli e nei logaritmi. Nel caso dei modelli 'Chow-Lin', 'ADL(1,0)' e 'ADL(1,1)', oltre alla specificazione base, è stato stimato anche il modello con aggiunta di costante, e di costante e trend. Nel solo caso dei modelli della classe ADL, tutte le specificazioni sono state stimate anche nelle differenze prime, sia nei livelli che nei logaritmi.

**Tavola 4A – Correlazione tra la variazione congiunturale dell'indicatore e della serie disaggregata (Serie destagionalizzate a prezzi correnti) (a) (b)**

| Modello  | branca 3    | branca 4    | branca 5    | branca 6    | branca 7    | branca 8    | branca 9    | branca 10   | branca 11   | branca 12   | branca 13   | branca 14   | branca 15   | branca 16   | branca 17   | branca 18   | branca 19   |
|----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| CL       | <b>0.91</b> | 0.85        | 0.93        | 0.80        | 0.91        | 0.88        | 0.97        | 0.82        | <b>0.97</b> | 0.92        | <b>0.92</b> | <b>0.96</b> | 0.96        | 0.94        | <b>0.94</b> | 0.72        | 0.63        |
| FE       | 0.91        | <b>0.85</b> | <b>0.94</b> | <b>0.83</b> | <b>0.92</b> | <b>0.88</b> | <b>0.97</b> | <b>0.84</b> | 0.95        | <b>0.92</b> | 0.91        | 0.93        | <b>0.96</b> | <b>0.94</b> | 0.94        | <b>0.73</b> | <b>0.63</b> |
| ADL(1,0) | 0.81        | 0.45        | 0.89        | 0.71        | 0.81        | 0.62        | 0.91        | 0.76        | 0.97        | 0.67        | 0.92        | 0.95        | 0.90        | 0.91        | 0.83        | 0.18        | 0.10        |
| ADL(1,1) | 0.88        | 0.26        | 0.92        | 0.79        | 0.91        | 0.87        | 0.90        | 0.75        | 0.97        | 0.84        | 0.92        | 0.96        | 0.96        | 0.92        | 0.94        | 0.16        | 0.31        |

(a) In tabella si riporta la specificazione di modello per cui si ottiene la correlazione massima.

(b) Ciascuna classe di modelli (Chow-Lin, Fernández, ADL(1,0), ADL(1,1)) è stata stimata nei livelli e nei logaritmi. Nel caso dei modelli 'Chow-Lin', 'ADL(1,0)' e 'ADL(1,1)', oltre alla specificazione base, è stato stimato anche il modello con aggiunta di costante, e di costante e trend. Nel solo caso dei modelli della classe ADL, tutte le specificazioni sono state stimate anche nelle differenze prime, sia nei livelli che nei logaritmi.

**Tavola 5A – Correlazione tra la variazione congiunturale dell'indicatore e della serie disaggregata (Serie destagionalizzate in volume) (a) (b)**

| Modello  | branca 3    | branca 4    | branca 5    | branca 6    | branca 7    | branca 8    | branca 9    | branca 10   | branca 11   | branca 12   | branca 13   | branca 14   | branca 15   | branca 16   | branca 17   | branca 18   | branca 19   |
|----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| CL       | 0.87        | 0.90        | 0.94        | 0.87        | 0.90        | 0.89        | 0.96        | 0.91        | 0.97        | 0.91        | 0.95        | 0.97        | 0.99        | 0.94        | 0.94        | 0.89        | 0.64        |
| FE       | <b>0.87</b> | <b>0.90</b> | <b>0.94</b> | <b>0.89</b> | 0.85        | <b>0.89</b> | <b>0.96</b> | <b>0.92</b> | 0.97        | <b>0.91</b> | 0.94        | 0.96        | 0.98        | <b>0.94</b> | <b>0.94</b> | <b>0.89</b> | <b>0.65</b> |
| ADL(1,0) | 0.78        | 0.49        | 0.90        | 0.80        | 0.71        | 0.85        | 0.89        | 0.89        | 0.94        | 0.67        | 0.95        | 0.97        | 0.99        | 0.90        | 0.56        | 0.58        | 0.22        |
| ADL(1,1) | 0.77        | 0.77        | 0.90        | 0.88        | <b>0.91</b> | 0.84        | 0.91        | 0.90        | <b>0.97</b> | 0.79        | <b>0.95</b> | <b>0.97</b> | <b>0.99</b> | 0.93        | 0.93        | 0.66        | 0.26        |

(a) In tabella si riporta la specificazione di modello per cui si ottiene la correlazione massima.

(b) Ciascuna classe di modelli (Chow-Lin, Fernández, ADL(1,0), ADL(1,1)) è stata stimata nei livelli e nei logaritmi. Nel caso dei modelli 'Chow-Lin', 'ADL(1,0)' e 'ADL(1,1)', oltre alla specificazione base, è stato stimato anche il modello con aggiunta di costante, e di costante e trend. Nel solo caso dei modelli della classe ADL, tutte le specificazioni sono state stimate anche nelle differenze prime, sia nei livelli che nei logaritmi.

**Tavola 6A – Specificazioni e parametri stimati dei modelli per i quali si ottiene il MAE minimo in ogni branca (Modelli stimati su serie a prezzi correnti) (a)**

| Branca | Modello   | Specifica      | Log-lik | $\phi$ | c                    | g                | $\beta_0$           | $\beta_1$           |
|--------|-----------|----------------|---------|--------|----------------------|------------------|---------------------|---------------------|
| 3      | ADL(1,0)  | $\Delta$       | -142.73 | 0.31   |                      |                  | 8.90<br>(5.52)**    |                     |
| 4      | ADL(1,0)  | --             | -150.69 | 0.98   |                      |                  | 1.70<br>(16.07)**   |                     |
| 5      | Chow-Lin  | --             | -152.76 | 0.81   |                      |                  | 60.15<br>(71.77)**  |                     |
| 6      | ADL(1,0)  | $\Delta$ - log | -140.88 | 0.54   |                      |                  | 0.31<br>(5.59)**    |                     |
| 7      | Chow-Lin  | --             | -150.09 | 0.98   |                      |                  | 9.37<br>(4.52)**    |                     |
| 8      | ADL(1,1)  | --             | -142.01 | 0.96   |                      |                  | 16.00<br>(4.19)**   | -14.85<br>(-3.88)** |
| 9      | ADL(1,0)  | $\Delta$       | -131.93 | 0.51   |                      |                  | 6.24<br>(3.32)**    |                     |
| 10     | ADL(1,1)  | log            | -153.43 | 0.99   |                      |                  | 0.38<br>(3.43)**    | -0.37<br>(-3.26)**  |
| 11     | Chow-Lin  | --             | -156.38 | 0.99   |                      |                  | 53.39<br>(15.35)**  |                     |
| 12     | ADL(1,1)t | log            | -132.67 | 0.93   | 0.05<br>(0.34)       | 0.00<br>(4.41)** | 0.57<br>(3.56)**    | -0.48<br>(-3.04)**  |
| 13     | ADL(1,0)t | $\Delta$ -log  | -132.52 | 0.03   | 0.01<br>(3.30)**     | 0.00<br>(-1.61)  | 0.39<br>(7.51)**    |                     |
| 14     | ADL(1,0)c | $\Delta$       | -147.34 | 0.20   | 0.00<br>(4.04)**     |                  | 0.44<br>(13.30)**   |                     |
| 15     | ADL(1,1)t | --             | -132.46 | 0.37   | 333.74<br>(3.26)**   | 3.59<br>(6.58)** | 34.35<br>(9.85)**   | -22.76<br>(-6.83)** |
| 16     | Chow-Lin  | --             | -136.01 | 0.75   |                      |                  | 16.30<br>(51.104)** |                     |
| 17     | ADL(1,1)  | $\Delta$       | -139.15 | 0.83   |                      |                  | 24.93<br>(10.39)**  | -19.45<br>(-7.57)** |
| 18     | ADL(1,1)c | --             | -152.74 | 0.53   | 1113.98<br>(12.69)** |                  | -35.90<br>(-2.21)** | 54.86<br>(3.38)**   |
| 19     | ADL(1,0)t | --             | -130.59 | 0.81   | 353.58<br>(7.30)**   | 5.42<br>(9.96)** | -1.26<br>(-1.69)*   |                     |

(a) Ciascun coefficiente stimato è corredato dal relativo valore della statistica t-Student tra parentesi; \* p-value  $\leq$  0.001; \*\* pvalue  $\leq$  0.01; \*\*\*pvalue  $\leq$  0.05.

**Tavola 7A – Specificazioni e parametri stimati dei modelli per i quali si ottiene il MAE minimo in ogni branca (Modelli stimati su serie in volume) (a)**

| Branca | Modello    | Specifica      | Log-lik | $\varphi$ | c                 | g                   | $\beta_0$          | $\beta_1$           |
|--------|------------|----------------|---------|-----------|-------------------|---------------------|--------------------|---------------------|
| 3      | Chow-Lin c | --             | -144.90 | 0.89      | 0.54<br>(3.99)*   |                     | 0.51<br>(1.95)*    |                     |
| 4      | Fernández  | --             | -156.41 | 0.00      |                   |                     | 76.91<br>(34.46)** |                     |
| 5      | Chow-Lin   | --             | -158.05 | 0.80      |                   |                     | 58.58<br>(60.05)** |                     |
| 6      | Chow-Lin c | log            | -137.16 | 0.87      | 0.75<br>(16.76)** |                     | 0.55<br>(7.55)**   |                     |
| 7      | Chow-Lin t | log            | -144.12 | 0.70      | 0.39<br>(0.61)    | -0.01<br>(-11.29)** | 1.40<br>(3.05)**   |                     |
| 8      | ADL(1,0)   | --             | -145.41 | 0.25      |                   |                     | 18.68<br>(82.29)** |                     |
| 9      | ADL(1,0)t  | --             | -131.18 | 0.63      | -59.89<br>(-.36)  | 3.23<br>(5.49)**    | 6.53<br>(3.39)**   |                     |
| 10     | Chow-Lin   | --             | -152.81 | 0.999     |                   |                     | 31.67<br>(6.09)**  |                     |
| 11     | Fernández  | --             | -158.95 | 0.00      |                   |                     | 69.94<br>(41.13)** |                     |
| 12     | ADL(1,1)   | --             | -142.49 | 0.98      |                   |                     | 6.84<br>(1.87)*    | -6.60<br>(-1.81)*   |
| 13     | Fernández  | --             | -144.17 | 0.00      |                   |                     | 14.83<br>(24.81)** |                     |
| 14     | ADL(1,0)c  | $\Delta$ - log | -147.63 | 0.00      | 0.00<br>(3.09)**  |                     | 0.64<br>(15.08)**  |                     |
| 15     | Chow-Lin t | log            | -136.92 | 0.91      | 0.34<br>(13.64)** | 0.00<br>(4.43)**    | 0.85<br>(15.22)**  |                     |
| 16     | Fernández  | --             | -144.02 | 0.00      |                   |                     | 21.70<br>(25.78)*  |                     |
| 17     | ADL(1,1)   | --             | -153.12 | 0.94      |                   |                     | 37.55<br>(5.143)** | -34.42<br>(-4.72)** |
| 18     | Chow-Lin   | --             | -159.57 | 0.92      |                   |                     | 66.89<br>(28.26)** |                     |
| 19     | ADL(1,0)   | log            | -145.38 | 0.96      |                   |                     | 0.07<br>(132.33)** |                     |

(a) Ciascun coefficiente stimato è corredato dal relativo valore della statistica t-Student tra parentesi; \* p-value  $\leq$  0.001; \*\* pvalue  $\leq$  0.01; \*\*\*pvalue  $\leq$  0.05 .

**Tavola 8A – Specificazioni e parametri stimati dei modelli per i quali si ottiene la correlazione massima della variazione congiunturale tra serie disaggregata e indicatore, per ciascuna branca (Modelli stimati su serie a prezzi correnti) (a)**

| Branca | Modello    | Specifica | Log-lik | $\varphi$ | c                 | g                 | $\beta_0$          |
|--------|------------|-----------|---------|-----------|-------------------|-------------------|--------------------|
| 3      | Chow-Lin   | log       | -149.00 | 0.94      |                   |                   | 1.57<br>(62.78)**  |
| 4      | Fernàndez  | log       | -163.86 | 0.00      |                   |                   | 2.00<br>(169.21)** |
| 5      | Fernàndez  | log       | -166.48 | 0.00      |                   |                   | 1.89<br>(170.37)** |
| 6      | Fernàndez  | --        | -151.28 | 0.00      |                   |                   | 39.49<br>(25.59)** |
| 7      | Fernàndez  | log       | -152.08 | 0.00      |                   |                   | 1.85<br>(34.35)**  |
| 8      | Fernàndez  | --        | -145.73 | 0.00      |                   |                   | 26.72<br>(23.68)** |
| 9      | Fernàndez  | log       | -141.93 | 0.00      |                   |                   | 1.63<br>(148.65)** |
| 10     | Fernàndez  | --        | -160.44 | 0.00      |                   |                   | 48.21<br>(20.67)** |
| 11     | Chow-Lin t | log       | -150.37 | 0.81      | 1.14<br>(29.63)** | 0.00<br>(6.62)**  | 0.64<br>(14.49)**  |
| 12     | Fernàndez  | log       | -147.42 | 0.00      |                   |                   | 1.44<br>(118.08)** |
| 13     | Chow-Lin t | log       | -131.94 | 0.83      | 0.93<br>(17.78)** | 0.00<br>(12.15)** | 0.43<br>(6.97)**   |
| 14     | Chow-Lin t | log       | -144.04 | 0.68      | 1.92<br>(31.15)** | 0.00<br>(21.21)** | 0.56<br>(13.39)**  |
| 15     | Fernàndez  | log       | -163.64 | 0.00      |                   |                   | 1.67<br>(75.57)**  |
| 16     | Fernàndez  | log       | -143.2  | 0.00      |                   |                   | 1.64<br>(119.58)** |
| 17     | Chow-Lin c | --        | -144.18 | 0.97      | 69.63<br>(5.22)** |                   | 28.23<br>(6.82)**  |
| 18     | Fernàndez  | log       | -181.61 | 0.00      |                   |                   | 2.18<br>(61.05)**  |
| 19     | Fernàndez  | log       | -160.8  | 0.00      |                   |                   | 1.76<br>(68.16)**  |

(a) Ciascun coefficiente stimato è corredato dal relativo valore della statistica t-Student tra parentesi; \* p-value  $\leq$  0.001;\*\* pvalue  $\leq$  0.01; \*\*\*pvalue  $\leq$  0.05 .

**Tavola 9A – Specificazioni e parametri stimati dei modelli per i quali si ottiene la correlazione massima della variazione congiunturale tra serie disaggregata e indicatore, per ciascuna branca (Modelli stimati su serie in volume) (a)**

| Branca | Modello   | Specifica      | Log-lik | $\varphi$ | c                  | g                 | $\beta_0$          | $\beta_1$            |
|--------|-----------|----------------|---------|-----------|--------------------|-------------------|--------------------|----------------------|
| 3      | Fernández | log            | -150.59 | 0.00      |                    |                   | 1.54<br>(83.63)**  |                      |
| 4      | Fernández | log            | -159.15 | 0.00      |                    |                   | 1.98<br>(260.92)** |                      |
| 5      | Fernández | log            | -166.86 | 0.00      |                    |                   | 1.83<br>(184.2)**  |                      |
| 6      | Fernández | --             | -145.66 | 0.00      |                    |                   | 38.46<br>(38.71)** |                      |
| 7      | ADL(1,1)c | log            | -152.33 | 0.99      | 0.12<br>(0.15)     |                   | 2.35<br>(2.68)**   | -2.36<br>(-2.49)**   |
| 8      | Fernández | log            | -152.32 | 0.00      |                    |                   | 1.68<br>(134.43)** |                      |
| 9      | Fernández | log            | -143.71 | 0.00      |                    |                   | 1.63<br>(133.25)** |                      |
| 10     | Fernández | --             | -152.31 | 0.00      |                    |                   | 42.89<br>(35.13)** |                      |
| 11     | ADL(1,1)t | log            | -150.45 | 0.58      | 2.29<br>(19.97)**  | 0.00<br>(12.73)** | 0.81<br>(8.61)**   | -0.50<br>(-5.14)**   |
| 12     | Fernández | log            | -149.26 | 0.00      |                    |                   | 1.50<br>(129.31)** |                      |
| 13     | ADL(1,1)  | $\Delta$ - log | -140.50 | 0.57      |                    |                   | 0.73<br>(7.86)**   | -0.49<br>(-4.73)**   |
| 14     | ADL(1,1)c | $\Delta$ - log | -148.08 | 0.40      | 0.00<br>(2.84)**   |                   | 0.66<br>(10.50)**  | -0.28<br>(-3.77)**   |
| 15     | ADL(1,1)t | --             | -136.01 | 0.89      | 124.41<br>(2.06)** | 0.86<br>(2.69)**  | 23.48<br>(14.61)** | -21.88<br>(-13.38)** |
| 16     | Fernández | log            | -146.54 | 0.00      |                    |                   | 1.65<br>(138.97)** |                      |
| 17     | Fernández | --             | -156.22 | 0.00      |                    |                   | 54.52<br>(29.45)** |                      |
| 18     | Fernández | log            | -168.65 | 0.00      |                    |                   | 1.99<br>(153.7)**  |                      |
| 19     | Fernández | log            | -163.90 | 0.00      |                    |                   | 1.68<br>(85.60)**  |                      |

(a) Ciascun coefficiente stimato è corredato dal relativo valore della statistica t-Student tra parentesi; \* p-value  $\leq$  0.001; \*\* pvalue  $\leq$  0.01; \*\*\*pvalue  $\leq$  0.05.

**Tavola 10A – Correlazioni tra variazioni congiunturali ( $\Delta q$ ) e tendenziali ( $\Delta y$ ) di serie trimestralizzate e indicatore (a)**

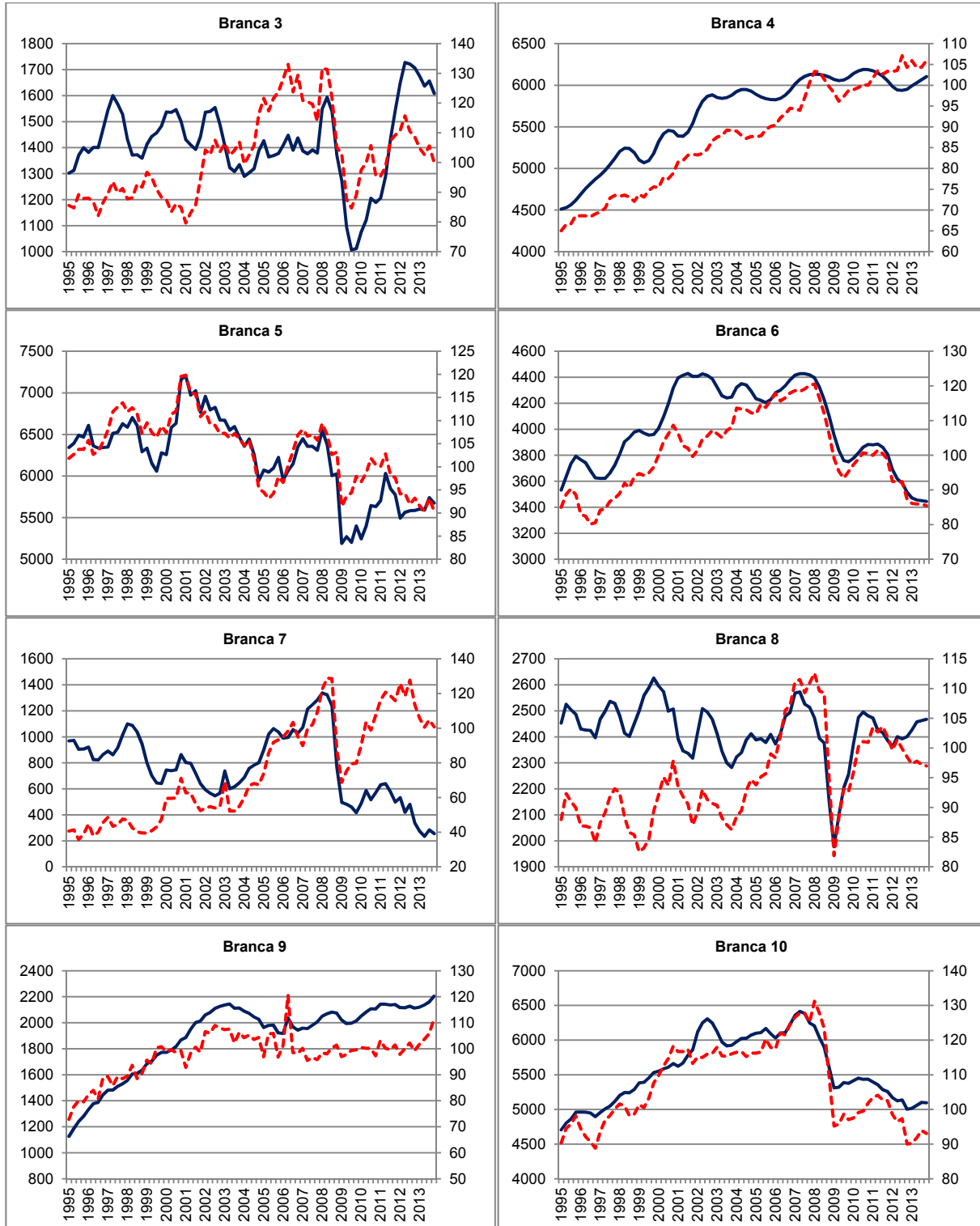
|       |            | Statistiche relative a serie espresse a prezzi correnti |      |             |             |                |                | Statistiche relative a serie espresse in volume |      |             |             |                |                |
|-------|------------|---|------|-------------|-------------|----------------|----------------|---|------|-------------|-------------|----------------|----------------|
|       |            | CL  | Fe   | ADL10       | ADL11       | ADL10 $\Delta$ | ADL11 $\Delta$ | CL  | Fe   | ADL10       | ADL11       | ADL10 $\Delta$ | ADL11 $\Delta$ |
| br 3  | $\Delta q$ | 0.89  | 0.91 | <b>0.39</b> | 0.88        | 0.51           | 0.58           | 0.76  | 0.91 | <b>0.23</b> | 0.37        | 0.76           | 0.83           |
|       | $\Delta y$ | 0.67  | 0.68 | 0.58        | 0.67        | 0.63           | 0.38           | 0.40  | 0.47 | 0.27        | 0.36        | 0.41           | 0.37           |
| br 4  | $\Delta q$ | 0.94  | 0.98 | 0.40        | <b>0.21</b> | 0.53           | -0.82          | 0.77  | 0.99 | <b>0.25</b> | -0.29       | -0.74          | -0.82          |
|       | $\Delta y$ | 0.45  | 0.58 | 0.26        | 0.22        | 0.36           | -0.21          | 0.37  | 0.58 | 0.22        | 0.12        | 0.02           | -0.43          |
| br 5  | $\Delta q$ | 0.99  | 0.99 | <b>0.80</b> | 0.99        | 0.97           | -0.25          | 0.99  | 0.99 | <b>0.80</b> | 0.98        | 0.98           | 0.98           |
|       | $\Delta y$ | 0.81  | 0.82 | 0.70        | 0.79        | 0.78           | 0.41           | 0.75  | 0.75 | 0.66        | 0.73        | 0.76           | 0.77           |
| br 6  | $\Delta q$ | 0.98  | 0.98 | <b>0.64</b> | 0.97        | 0.85           | -0.28          | 0.99  | 0.99 | <b>0.69</b> | 0.99        | 0.94           | -0.68          |
|       | $\Delta y$ | 0.74  | 0.75 | 0.71        | 0.74        | 0.72           | 0.70           | 0.81  | 0.82 | 0.78        | 0.81        | 0.80           | 0.61           |
| br 7  | $\Delta q$ | 0.76  | 0.70 | <b>0.41</b> | 0.76        | 0.77           | 0.76           | 0.48  | 0.82 | -0.01       | <b>0.57</b> | 0.69           | 0.67           |
|       | $\Delta y$ | 0.72  | 0.73 | 0.63        | 0.73        | 0.72           | 0.72           | 0.36  | 0.64 | 0.25        | 0.36        | 0.55           | 0.55           |
| br 8  | $\Delta q$ | 0.59  | 0.97 | <b>0.36</b> | 0.96        | 0.57           | 0.92           | 0.67  | 0.96 | 0.86        | <b>0.76</b> | 0.96           | -0.14          |
|       | $\Delta y$ | 0.54  | 0.72 | 0.49        | 0.70        | 0.56           | 0.67           | 0.57  | 0.70 | 0.69        | 0.59        | 0.72           | 0.65           |
| br 9  | $\Delta q$ | 0.89  | 0.99 | <b>0.70</b> | 0.72        | 0.97           | 0.97           | 0.99  | 0.99 | 0.64        | <b>0.61</b> | 0.96           | 0.95           |
|       | $\Delta y$ | 0.68  | 0.77 | 0.55        | 0.46        | 0.69           | 0.70           | 0.55  | 0.56 | 0.29        | 0.31        | 0.48           | 0.53           |
| br 10 | $\Delta q$ | 0.48  | 0.98 | 0.86        | <b>0.40</b> | 0.84           | 0.84           | 0.83  | 0.99 | <b>0.61</b> | 0.92        | 0.97           | 0.90           |
|       | $\Delta y$ | 0.57  | 0.70 | 0.68        | 0.55        | 0.65           | 0.65           | 0.76  | 0.86 | 0.65        | 0.73        | 0.85           | 0.84           |
| br 11 | $\Delta q$ | 0.99  | 0.99 | <b>0.88</b> | 0.99        | 0.99           | 0.98           | 0.99  | 0.99 | <b>0.71</b> | 0.99        | 0.93           | 0.94           |
|       | $\Delta y$ | 0.95  | 0.94 | 0.92        | 0.95        | 0.94           | 0.94           | 0.94  | 0.94 | 0.88        | 0.94        | 0.92           | 0.91           |
| br 12 | $\Delta q$ | 0.93  | 0.95 | <b>0.20</b> | 0.62        | 0.80           | 0.31           | 0.94  | 0.96 | <b>0.21</b> | 0.95        | 0.85           | -0.69          |
|       | $\Delta y$ | 0.67  | 0.70 | 0.56        | 0.44        | 0.62           | 0.55           | 0.56  | 0.61 | 0.46        | 0.56        | 0.53           | 0.23           |
| br 13 | $\Delta q$ | 0.97  | 0.97 | <b>0.46</b> | 0.97        | 0.96           | 0.82           | 0.98  | 0.98 | <b>0.45</b> | 0.98        | 0.98           | 0.89           |
|       | $\Delta y$ | 0.84  | 0.84 | 0.73        | 0.84        | 0.84           | 0.83           | 0.89  | 0.89 | 0.79        | 0.89        | 0.89           | 0.89           |
| br 14 | $\Delta q$ | 0.99  | 0.99 | <b>0.75</b> | 0.99        | 0.98           | 0.91           | 0.99  | 0.99 | <b>0.29</b> | 0.99        | 0.99           | 0.99           |
|       | $\Delta y$ | 0.93  | 0.91 | 0.90        | 0.93        | 0.94           | 0.93           | 0.95  | 0.95 | 0.65        | 0.95        | 0.95           | 0.95           |
| br 15 | $\Delta q$ | 0.76  | 0.99 | <b>0.64</b> | 0.97        | 0.03           | 0.93           | 0.99  | 0.99 | <b>0.78</b> | 0.99        | 0.98           | 0.99           |
|       | $\Delta y$ | 0.81  | 0.90 | 0.64        | 0.87        | 0.04           | 0.85           | 0.96  | 0.96 | 0.88        | 0.96        | 0.96           | 0.96           |
| br 16 | $\Delta q$ | 0.99  | 0.99 | <b>0.89</b> | 0.99        | 0.98           | 0.08           | 0.99  | 0.99 | <b>0.80</b> | 0.99        | 0.96           | -0.03          |
|       | $\Delta y$ | 0.79  | 0.79 | 0.77        | 0.79        | 0.79           | 0.65           | 0.78  | 0.79 | 0.72        | 0.78        | 0.77           | 0.63           |
| br 17 | $\Delta q$ | 0.99  | 0.99 | <b>0.76</b> | 0.99        | 0.94           | 0.99           | 0.99  | 0.99 | <b>0.62</b> | 0.99        | 0.89           | 0.89           |
|       | $\Delta y$ | 0.79  | 0.78 | 0.68        | 0.79        | 0.73           | 0.77           | 0.76  | 0.77 | 0.65        | 0.76        | 0.58           | 0.55           |
| br 18 | $\Delta q$ | 0.90  | 0.96 | <b>0.45</b> | -0.34       | 0.59           | -0.72          | 0.98  | 0.99 | <b>0.58</b> | 0.91        | 0.97           | -0.12          |
|       | $\Delta y$ | 0.15  | 0.24 | 0.09        | 0.01        | 0.08           | -0.08          | 0.56  | 0.64 | 0.33        | 0.45        | 0.49           | 0.32           |
| br 19 | $\Delta q$ | 0.15  | 0.96 | 0.03        | -0.13       | -0.27          | <b>0.83</b>    | 0.91  | 0.97 | <b>0.41</b> | 0.60        | 0.77           | -0.65          |
|       | $\Delta y$ | -0.06   | 0.17 | -0.06       | -0.07       | -0.06          | 0.02           | 0.25  | 0.29 | 0.17        | 0.18        | 0.19           | 0.12           |

(a) I valori delle correlazioni sono valori medi delle diverse specificazioni (standard, con costante, con costante e trend) di ogni classe di modello stimato nei livelli. L'anno di riferimento degli aggregati espressi in volume (o concatenati) è il 2010.

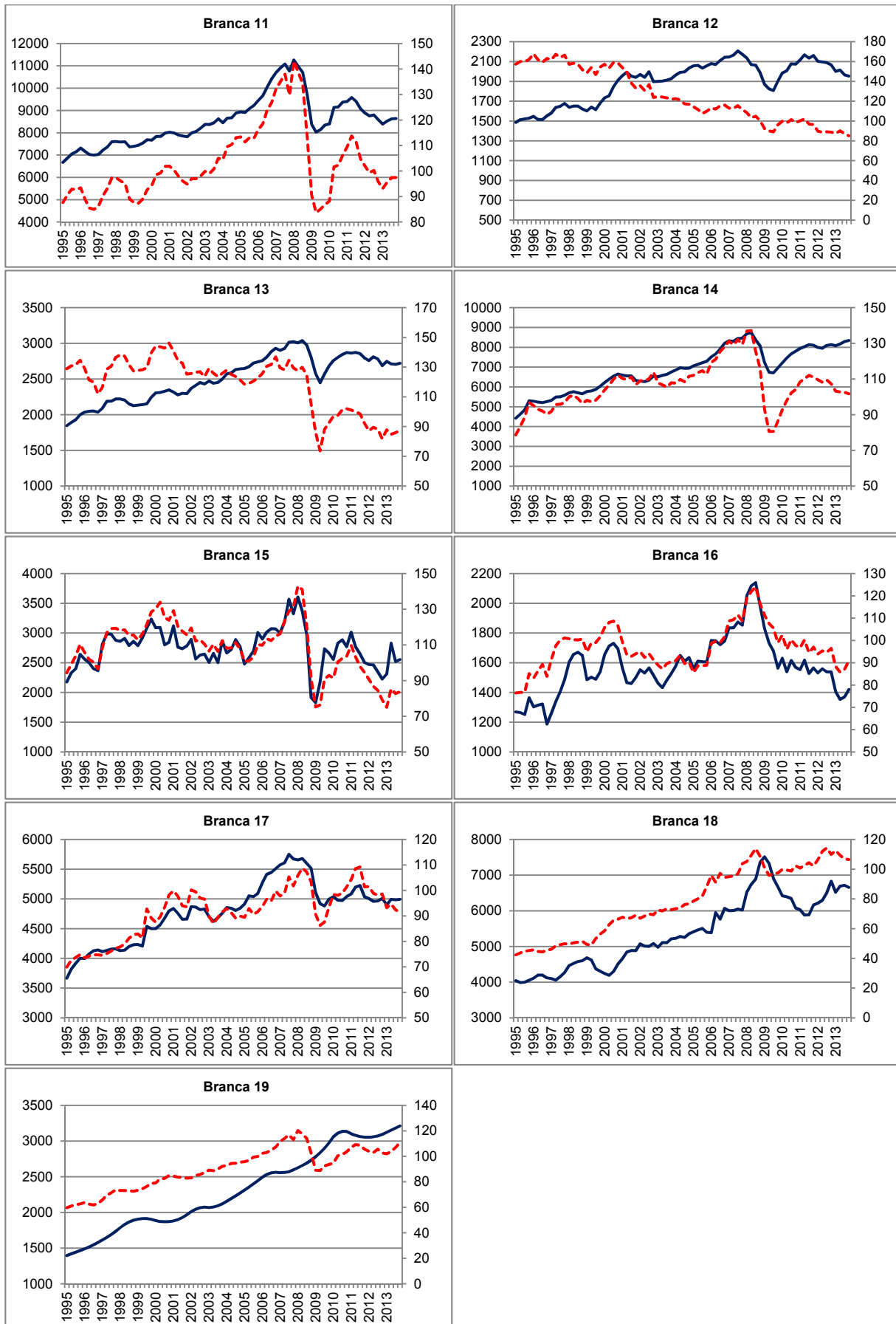


## Appendice B - Figure

**Figura 1B – Serie trimestralizzate in base ai modelli specificati nella tavola 6A (linea continua) e relativo indicatore IPI (linea tratteggiata) nell'intervallo temporale 1995:q1-2013:q4 (Serie a prezzi correnti)**



**Figura 1B (segue) – Serie trimestralizzate in base ai modelli specificati nella tavola 6A (linea continua) e relativo indicatore IPI (linea tratteggiata) nell'intervallo temporale 1995:q1-2013:q4 (Serie a prezzi correnti)**



**Figura 2B – Serie trimestralizzate in base ai modelli specificati nella tavola 7A (linea continua) e relativo indicatore IPI (linea tratteggiata) nell'intervallo temporale 1995:q1-2013:q4 (Serie in volume)**

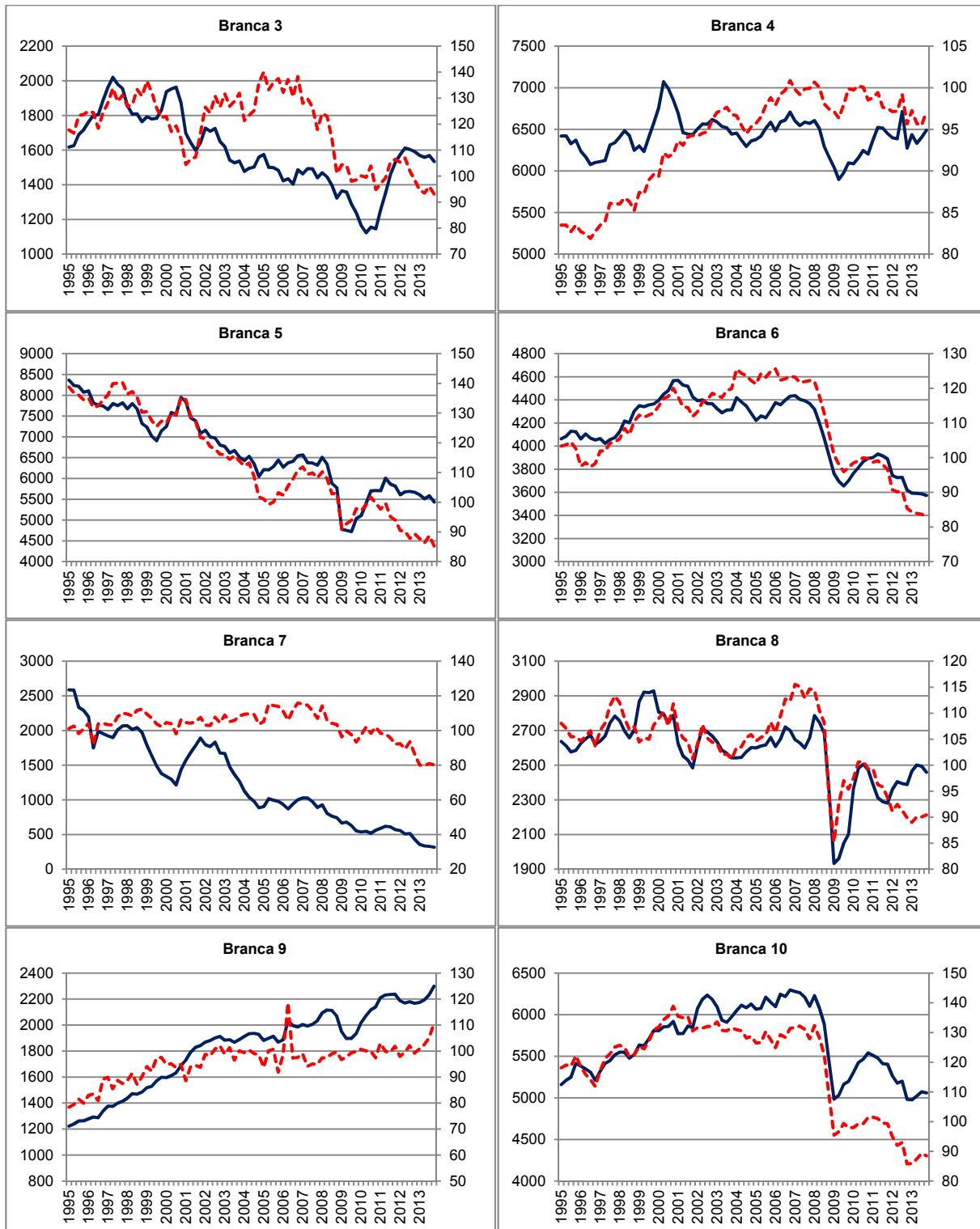
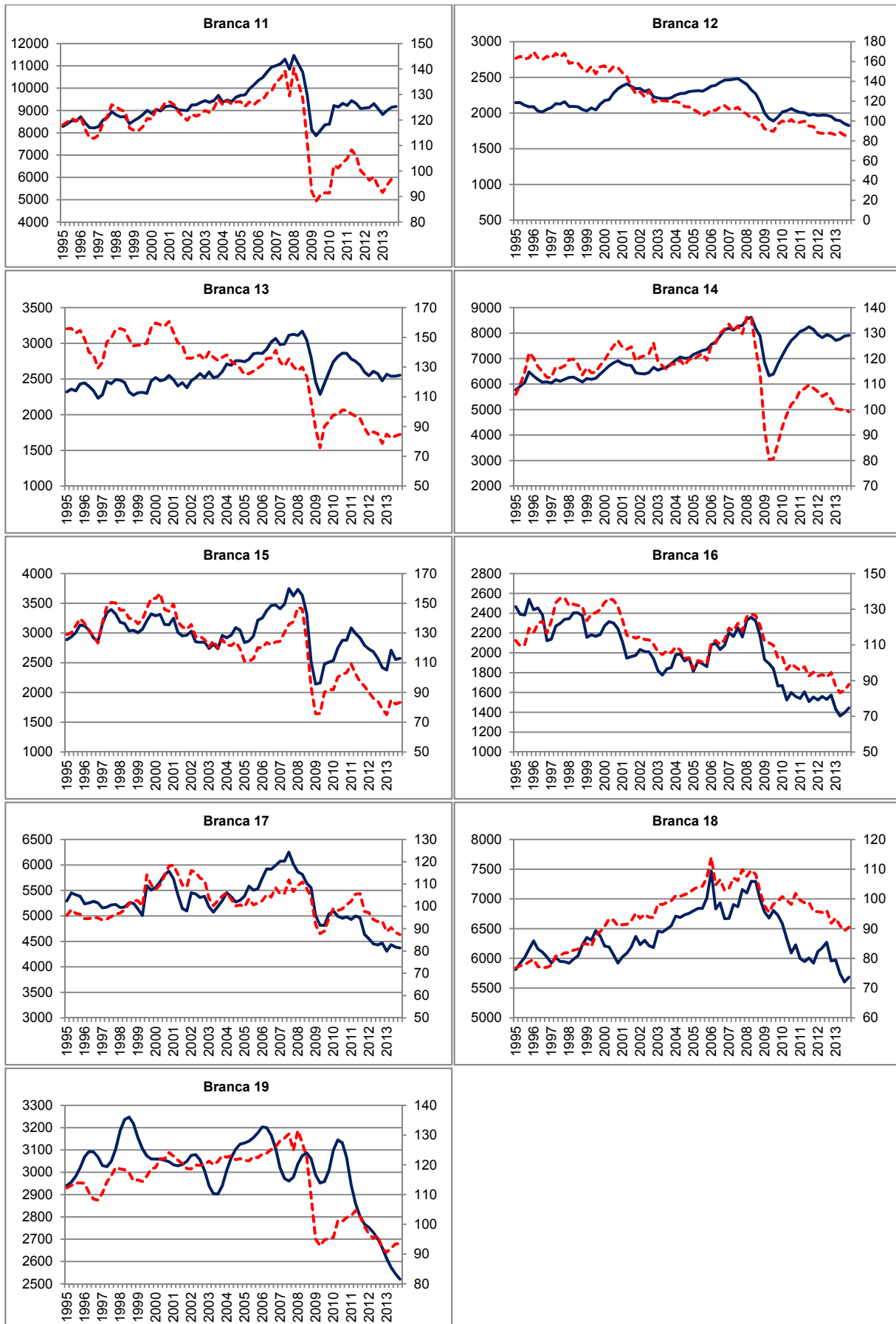
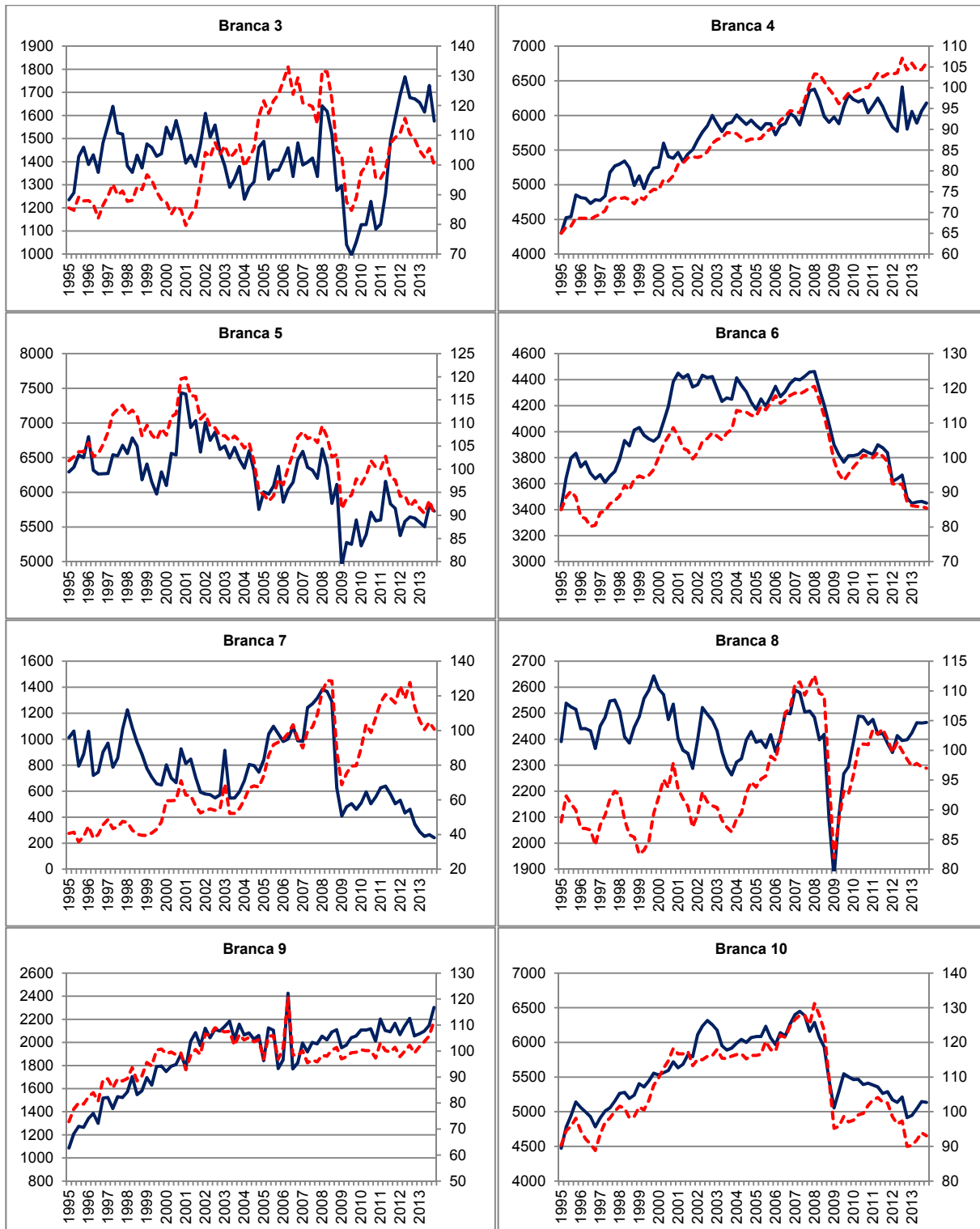


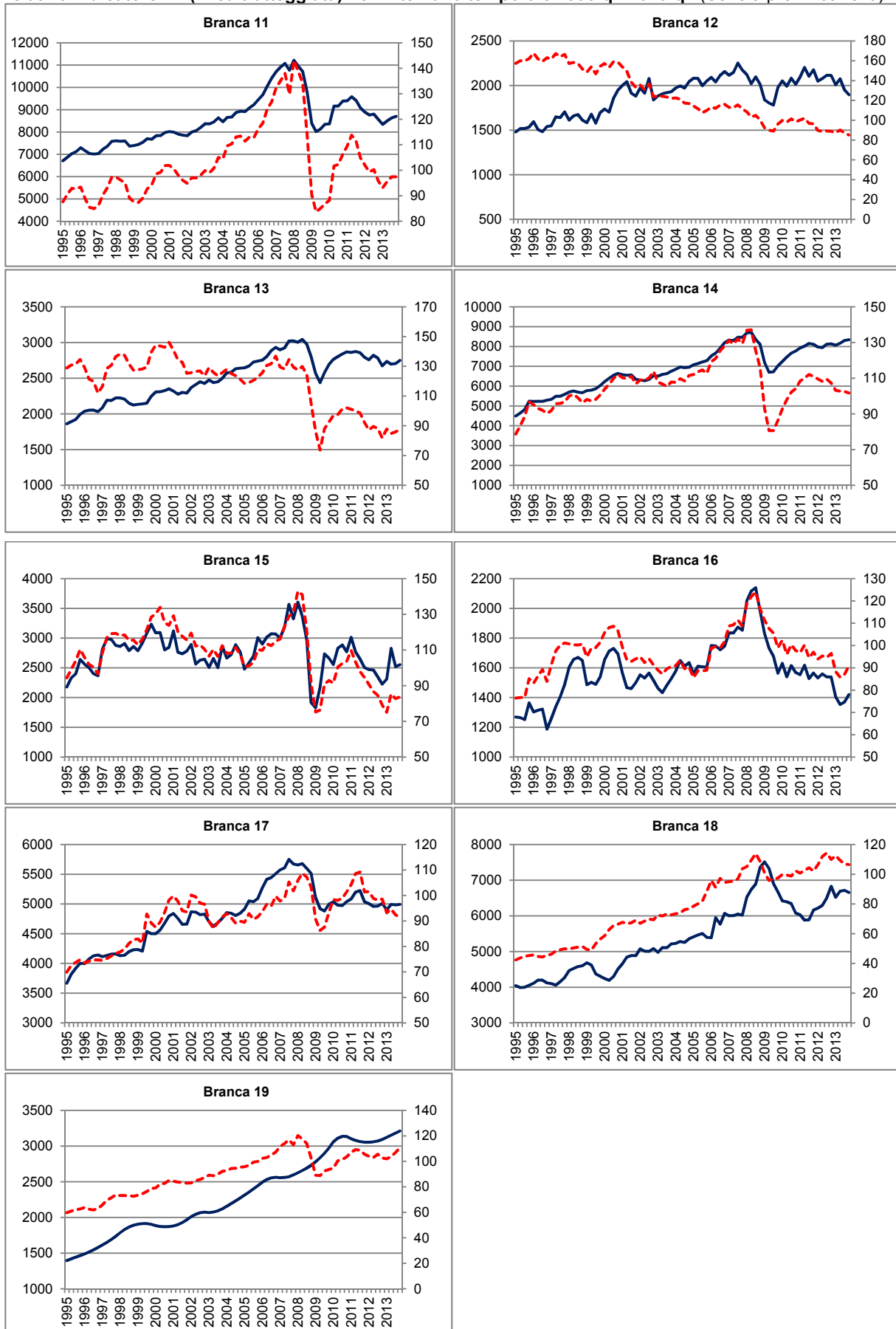
Figura 2B (segue) - Serie trimestralizzate in base ai modelli specificati nella tavola 7A (linea continua) e relativo indicatore IPI (linea tratteggiata) nell'intervallo temporale 1995:q1-2013:q4 (Serie in volume)



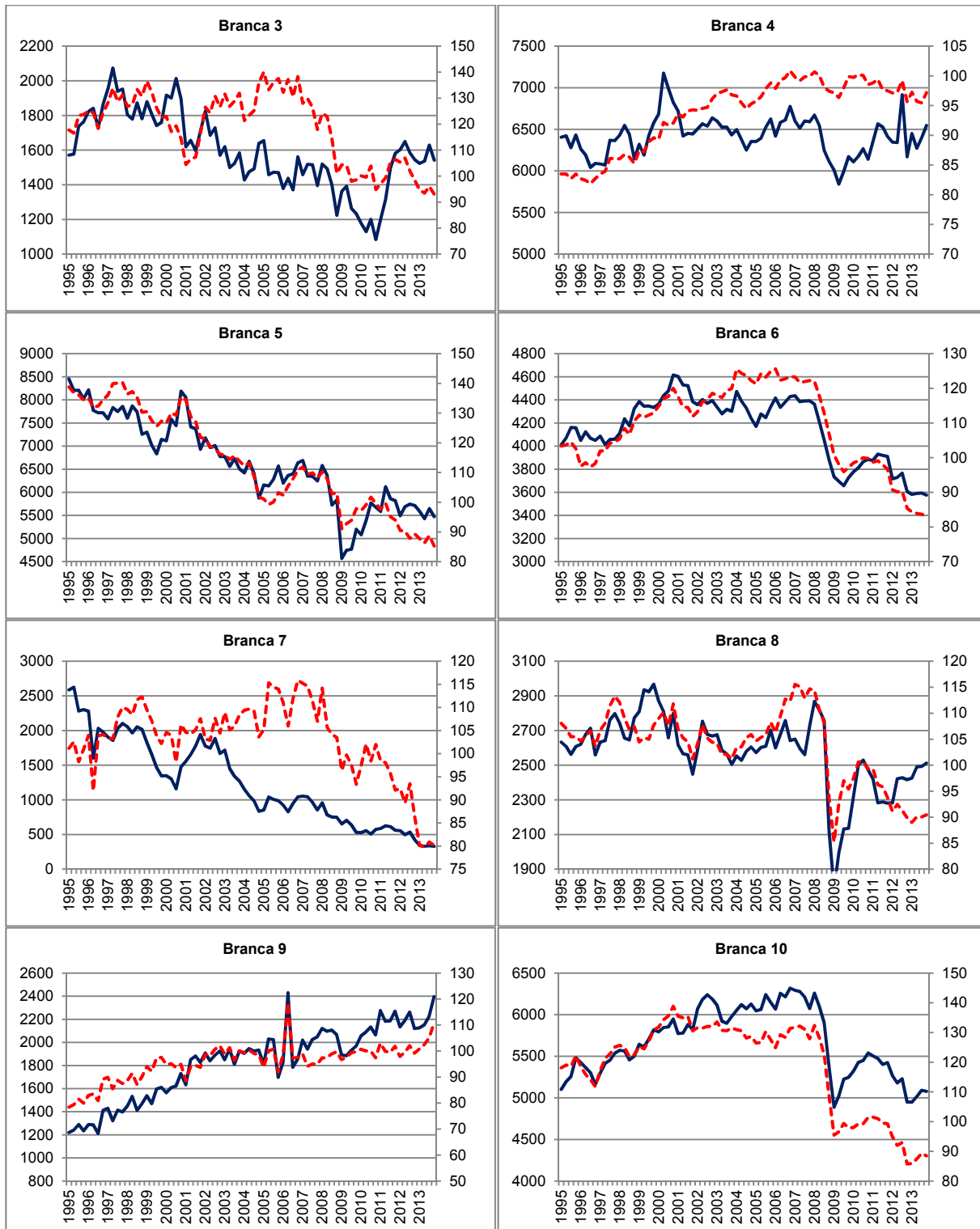
**Figura 3B – Serie trimestralizzate in base ai modelli specificati nella tavola 8A (linea continua) e relativo indicatore IPI (linea tratteggiata) nell'intervallo temporale 1995:q1-2013:q4 (Serie a prezzi correnti)**



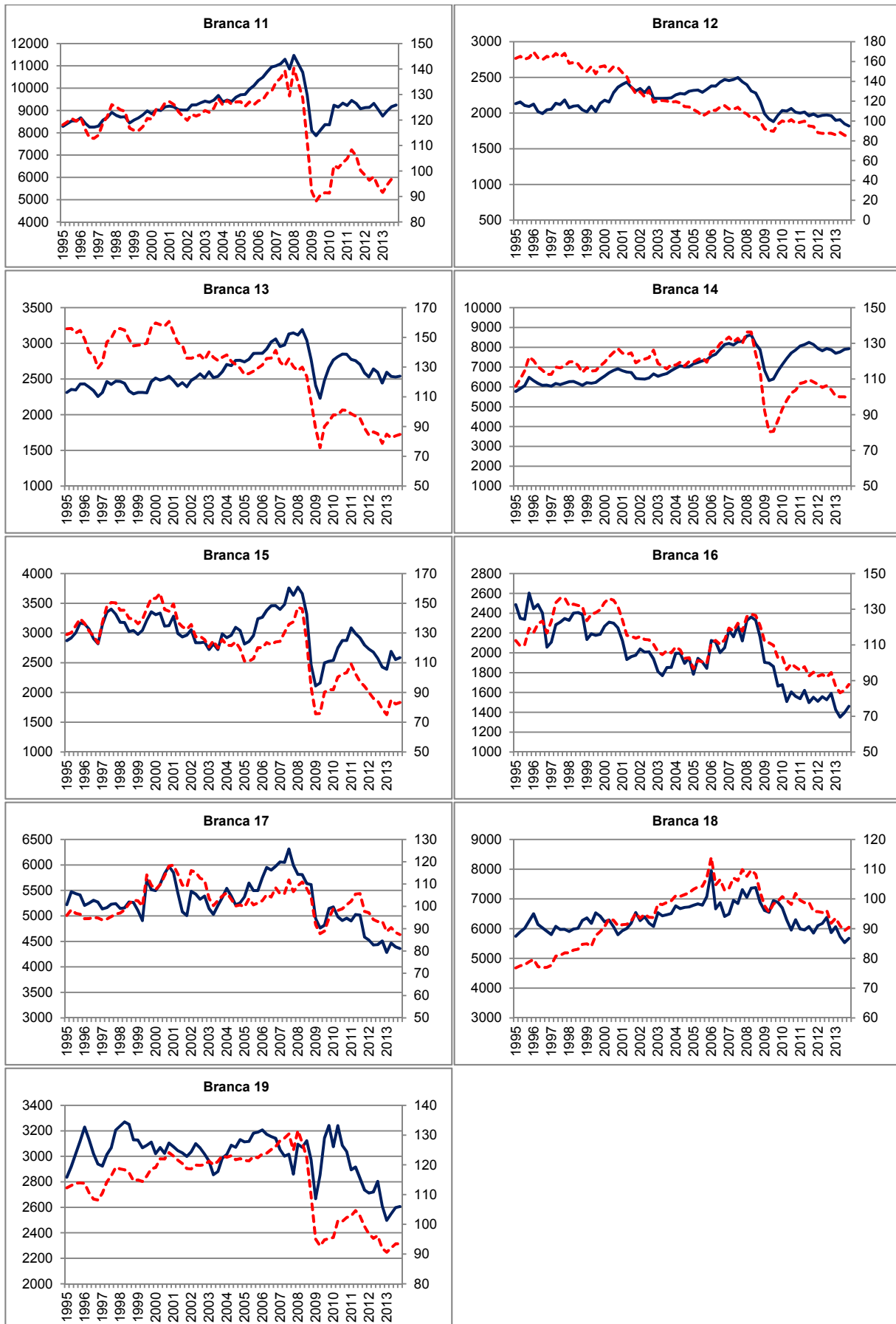
**Figura 3B (segue) – Serie trimestralizzate in base ai modelli specificati nella tavola 8A (linea continua) e relativo indicatore IPI (linea tratteggiata) nell'intervallo temporale 1995:q1-2013:q4 (Serie a prezzi correnti)**



**Figura 4B – Serie trimestralizzate in base ai modelli specificati nella tavola 9A (linea continua) e relativo indicatore IPI (linea tratteggiata) nell'intervallo temporale 1995:q1-2013:q4 (Serie a prezzi correnti)**

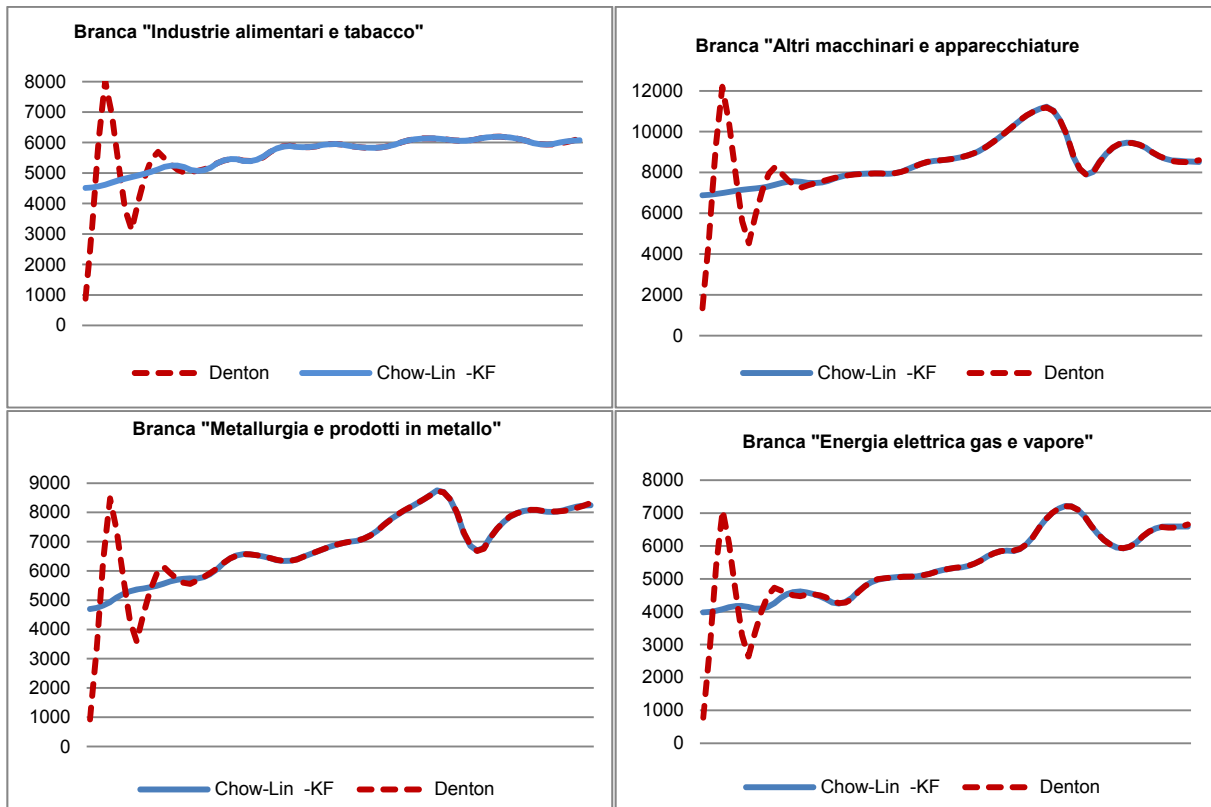


**Figura 4B (segue) – Serie trimestralizzate in base ai modelli specificati nella tavola 9A (linea continua) e relativo indicatore IPI (linea tratteggiata) nell'intervallo temporale 1995:q1-2013:q4 (Serie a prezzi correnti)**

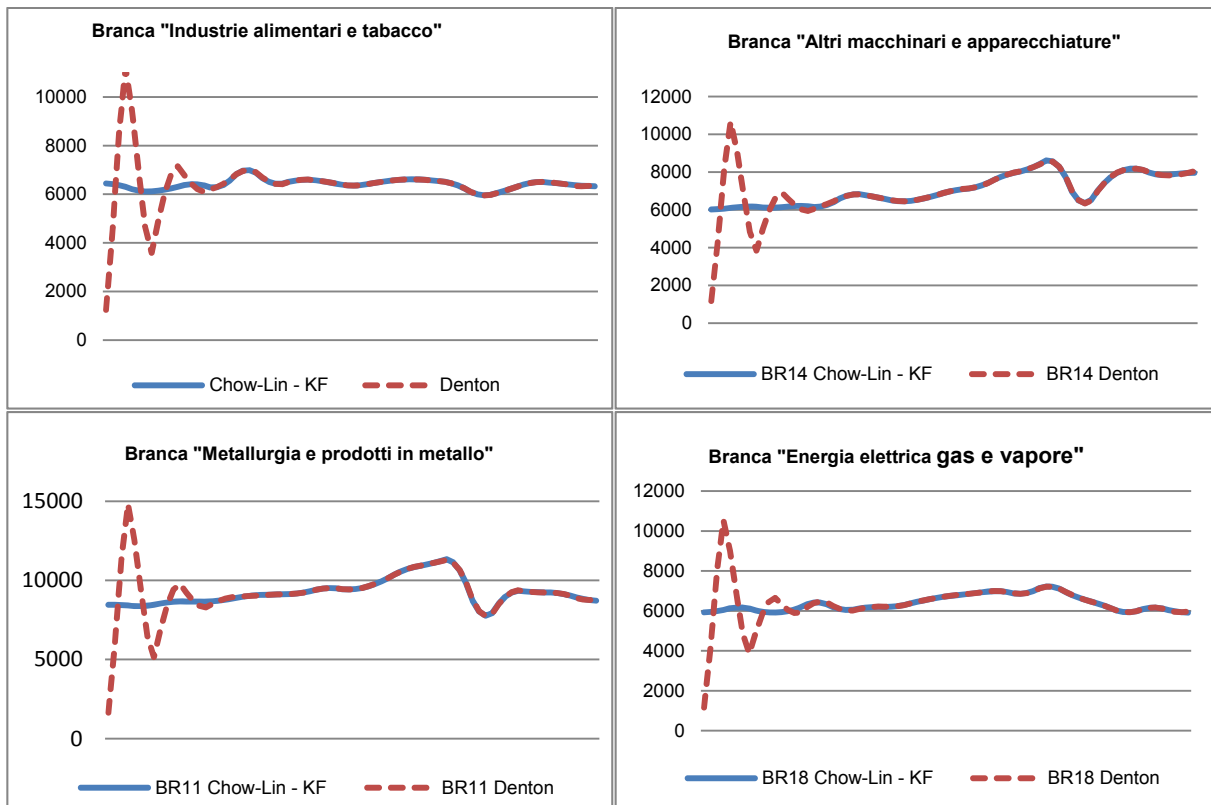




**Figura 5B – Serie trimestrali disaggregate del valore aggiunto in base ai modelli Denton (approccio statico) e Chow-Lin (approccio dinamico, senza indicatore) nell'intervallo 1995:q1-2013:q4 (Serie a prezzi correnti, dati destagionalizzati)**



**Figura 6B – Serie trimestrali disaggregate del valore aggiunto in base ai modelli Denton (approccio statico) e Chow-Lin (approccio dinamico, senza indicatore) nell'intervallo 1995:q1-2013:q4 (Serie in volume, dati destagionalizzati)**



## Appendice C - Descrizione delle subroutine Speakeasy

### Pratica attuale e suggerimenti della commissione di studio

Il trattamento statistico della trimestralizzazione degli aggregati annuali attualmente utilizzato in contabilità nazionale è realizzato dalla subroutine TRIMEQNA in ambiente speakeasy che implementa i metodi di disaggregazione temporale di Chow-Lin (1971) nella versione proposta da Barbone et al. (1981), Fernández (1981) e Denton (1971). TRIMEQNA è la versione attualmente in uso per le attività correnti di produzione dei conti trimestrali, frutto di vari aggiornamenti che si sono succeduti negli anni. Una versione precedente, per molti aspetti superata è descritta in Astolfi e Marini (2005). TRIMEQNA richiama al suo interno la funzione Speakeasy DISAGGR che effettua la disaggregazione temporale in senso stretto e la integra in diversi aspetti, i più importanti dei quali sono: la presentazione dei risultati della disaggregazione (confronti annuali e trimestrali tra valori effettivi e stimati), dei coefficienti di regressione e delle relative statistiche *t* di student, della stima del coefficiente autoregressivo dei residui, delle statistiche sulla bontà di adattamento del modello ( $R^2$  nella versione standard e corretta), della statistica-test *F* sulla significatività congiunta dei coefficienti di regressione del modello, dell'errore standard di regressione, della log-verosimiglianza, dei criteri informativi AIC e BIC, delle statistiche test di Durbin e Watson (1951a, 1951b) e di Jarque e Bera (1987) e delle altre statistiche sul grado di accostamento tra l'indicatore e le stime trimestrali nell'ultima metà del periodo di stima (correlazioni, e distanze medie quadratiche in termini di variazioni prime e quarte).

Inoltre, TRIMEQNA interagisce con altre subroutine per la manipolazione di variabili d'intervento 'impulso', 'a gradino' o 'trend spezzati' spesso utilizzate nelle fasi di stima e si contraddistingue per la rapidità d'esecuzione grazie al richiamo degli eseguibili fortran DISAGGR e DENTON sviluppati da Bruno e Marra (2005). Il report fornito da TRIMEQNA sin dalla sua introduzione è stato apprezzato dagli utilizzatori della contabilità trimestrale in quanto riconosciuto come strumento efficace a fornire un quadro d'insieme della qualità dell'esercizio di stima. Invero, nell'introdurre le modifiche procedurali di tecnica e di calcolo della trimestralizzazione di cui si dirà nel prosieguo del lavoro, si è prestata particolare attenzione nel mantenere quanto più possibile invariata la struttura del report.

Nell'ambito delle indicazioni/suggerimenti dei lavori della commissione di studio sui conti trimestrali, per le quali si rimanda a Di Fonzo (2005), si è proposto l'ampliamento dell'insieme dei metodi di disaggregazione da utilizzare per la stima dei conti trimestrali. Alcuni di questi metodi sono stati sviluppati nel corso dei lavori della commissione stessa, i più interessanti dei quali sfruttano la flessibilità offerta dalla rappresentazione nello spazio degli stati (SSF) e del filtro di Kalman (Kalman, 1960; Harvey, 1989) per il trattamento statistico. In particolare, la classe dei metodi che si basano sui modelli univariati a ritardi distribuiti ADL(1,1) proposta da Proietti (2005) rappresenta la più naturale estensione dei metodi di regressione classici e per questo di più immediato inserimento nel ciclo produttivo dei conti trimestrali.

Dunque, successivamente alla conclusione dei lavori della commissione si è proceduto operativamente all'ampliamento dell'impianto modellistico alla base delle tecniche di disaggregazione preesistenti alla classe dei modelli ADL secondo l'approccio SSF nell'ambiente informatico Speakeasy. È infatti noto come i metodi di disaggregazione temporale classici possono esprimersi come casi particolari della classe di modelli ADL e godere, quindi, anch'essi di una rappresentazione SSF.

### Le nuove procedure

Il passaggio dall'approccio tradizionale a quello SSF ha comportato non solo l'introduzione di nuovi algoritmi di calcolo relativi sia alla fase di trimestralizzazione che alla fase di analisi e diagnostica, ma anche il relativo adattamento computazionale delle diagnostiche dei modelli classici originariamente disponibili in TRIMEQNA. Per consentire tale passaggio in ambiente Speakeasy, si è proceduto alla compilazione di nuove procedure che, pur mantenendo l'impostazione di TRIMEQNA, implementano gli opportuni algoritmi funzionali alla stima della serie disaggregata nello spazio degli stati, il principale dei quali è il filtro di Kalman nella versione di De Jong (1991).

Le nuove subroutine denominate 'TRIMEKF' e 'TRIMEKFL' realizzano la procedura di

trimestralizzazione basata sull'approccio SSF, rispettivamente, nel caso dei modelli esplicitati nei livelli della serie e nei logaritmi. In entrambi i casi le procedure di trimestralizzazione sono definite per diversi modelli alternativi, in particolare: il modello di Chow-Lin (1971), di Fernández (1981), di Litterman (1983) e per la classe dei modelli ADL (1,1). Per ciascun modello (ad eccezione di Litterman) è prevista una duplice specificazione che permette di selezionare l'ipotesi più opportuna circa le condizioni iniziali del processo legate alle proprietà di stazionarietà vs. non stazionarietà della serie. Inoltre, per le specificazioni ADL è possibile la stima sia nei livelli che nelle differenze prime. Entrambe le nuove subroutine producono un report finale di stima la cui struttura è stata mantenuta invariata rispetto a quella del report prodotto da TRIMEQNA poiché particolarmente apprezzato e di uso ormai consolidato in fase di produzione dei conti trimestrali. D'altra parte, la gamma di test diagnostici e di statistiche sulla qualità della stima disponibili nel nuovo report risulta più ampia rispetto al passato. Sono infatti ora disponibili: le stime delle innovazioni e delle previsioni annualizzate per gli ultimi otto anni dell'intervallo di stima, il calcolo del test di Jarque-Bera (1987), del test Q di Ljung e Box (1978), del test "H" sulla eteroschedasticità dei residui, il calcolo di misure medie standard degli errori di previsione semplici (ME), assoluti (MAE) e quadratici (RMSE) rispetto sia ai livelli delle serie, sia alle variazioni congiunturali, sia a quelle tendenziali. A differenza dell'approccio classico di regressione, l'approccio spazio degli stati fa sì che tali diagnostiche siano implementate sfruttando la medesima formulazione valida per tutti i possibili modelli di disaggregazione: il che costituisce un guadagno notevole in termini di efficienza poiché la capacità di controllo sull'esito della stima si amplia a fronte di un onere computazionale minimo. Un esempio di report completo prodotto da TRIMEKF, discusso nella sezione 3, è riportato nell'appendice D.

### Descrizione dei programmi

I programmi Speakeasy *trime\_kf.deck* e *trimekfl.deck* rappresentano i programmi principali che richiamano le subroutine per il trattamento statistico della disaggregazione temporale mediante l'uso di modelli di regressione dinamica.

I modelli implementati sono identificati da un'appropriata stringa per definire gli argomenti di input e di output della subroutine. L'argomento *met0* identifica il modello. Le opzioni sono di seguito elencate:

| <i>Argomento</i> | <i>Modello</i> | <i>Note</i>  |
|------------------|----------------|--|
| CL               | Chow-Lin       | regressione statica con residui AR(1) senza costante e trend |
| CLc              | Chow-Lin       | come sopra con costante                                      |
| CLt              | Chow-Lin       | come sopra con costante e trend                              |
| Fe               | Fernández      | come sopra con residui I(1)                                  |
| L0               | Litterman      | come sopra con residui ARIMA(1,1,0)                          |
| ADL10x           | ADL(1,0)       | regressione dinamica senza costante e trend                  |
| ADL10xc          | ADL(1,0)       | come sopra con costante                                      |
| ADL10xt          | ADL(1,0)       | come sopra con costante e trend                              |
| ADL11x           | ADL(1,1)       | regressione dinamica senza costante e trend                  |
| ADL11xc          | ADL(1,1)       | come sopra con costante                                      |
| ADL11xt          | ADL(1,1)       | come sopra con costante e trend                              |
| ADL10I           | ADL(1,0)       | regressione dinamica senza costante e trend                  |
| ADL10Ic          | ADL(1,0)       | come sopra con costante                                      |
| ADL10It          | ADL(1,0)       | come sopra con costante e trend                              |
| ADL11I           | ADL(1,1)       | regressione dinamica senza costante e trend                  |
| ADL11Ic          | ADL(1,1)       | come sopra con costante                                      |
| ADL11It          | ADL(1,1)       | come sopra con costante e trend                              |

Altri argomenti di input delle subroutine sono:

y0: vettore dei dati da disaggregare

x0: matrice degli indicatori (x0 = ? se senza indicatore)

cs: intero con n. periodi da disaggregare nell'esercizio: 4->annuale-trimestrale; 12->annuale-mensile;

3-> trimestrale-mensile; 6->annuale-bimensile; ecc.  
 rho: scalare con valore di rho. Inserire ? (dato mancante) se si vuole effettuare la stima  
 opt : 1 = somma (flussi); 2 = media (stock)  
 ai: anno iniziale  
 tipovar: nome variabile da trimestralizzare (es. "produzione corrente" oppure "investimenti")  
 nome: titolo sul tabulato  
 out: 0 = senza output; 1 = stampa risultati; 2 = con output grafico; 3 = stampa risultati + output sintetico (solo se senza indicatore)

Tra gli argomenti output abbiamo:

yt: vettore dei dati disaggregati  
 results: matrice di risultati (2 colonne)  
 prime 16 righe stessa sequenza dell'output a video (rho, r2, r2 corretto, etc.), le successive righe contengono i coefficienti stimati e le t-student del modello

Alcune annotazioni:

quando l'argomento  $x_0=?$  (scalare con dato mancante) viene effettuato un esercizio di disaggregazione senza indicatore. Per esercizi di disaggregazione che non siano del tipo annuale-trimestrale inserire out=0. In questi casi le tabulazioni non sono ancora implementate.

Di seguito l'elenco delle subroutine del filtro di Kalman e di quelle accessorie:

### Disaggregazione lineare

*log\_likc.deck*: calcola la log-verosimiglianza profilo per un dato valore di  $\phi$ ;

*grid\_s.deck*: produce la stima di massima verosimiglianza del parametro  $\phi$  attraverso la ricerca a griglia nell'intervallo di valori (-1, 1);

*for\_est.deck*: calcola le previsioni un passo avanti dal modello di disaggregazione selezionato con la relativa varianza;

*smo\_est.deck*: calcola le stime 'smoothed' della disaggregazione dal modello selezionato con la relativa varianza;

*inn\_est.deck*: calcola le innovazioni dal modello di disaggregazione selezionato con la relativa varianza;

*selmodel.deck*: genera un set di parametri necessari alla stima per il modello selezionato;

*indvar.deck*: genera un vettore di valori 0 e 1 utile per la stima della variabile cumulate  $y_t^c$ ;

*settdssf.deck*: costruisce il sistema delle matrici delle equazioni della forma spazio degli stati del modello selezionato.

### Disaggregazione non lineare

*lgam\_inn.deck*: calcola le innovazioni per un dato valore di  $\phi$  relativamente al modello gaussiano approssimato selezionato;

*lgam\_smo.deck*: calcola le stime 'smoothed' della disaggregazione per un dato valore di  $\phi$  relativamente al modello selezionato;

*lgam\_llc.deck*: calcola la log-verosimiglianza profilo per un dato valore di  $\phi$  relativamente al modello selezionato;

*seq\_pm.deck*: calcola la stima finale disaggregata per un dato valore di  $\varphi$  a partire da una serie disaggregata iniziale e iterando fino a convergenza l'algoritmo sequenziale descritto in Proietti (2005) sez.5;

*lgam\_gs.deck*: produce la stima di massima verosimiglianza del parametro  $\varphi$  attraverso la ricerca a griglia nell'intervallo di valori  $(-1, 1)$ ;

## Appendice D - L'output di TRIMEKF

metodo di stima : metodo -- ADL11T  
 tipo di aggregazione : somma  
 equazione :  $Y = C + D * t + B_0 * I_{1,t} + B_1 * I_{1,t-1}$

| Confronto dati annuali veri - previsioni |          |           |          |           |           |           |           |
|--|----------|-----------|----------|-----------|-----------|-----------|-----------|
| ANNIT                                    | liv.veri | liv.prev. | innov.   | var. veri | var.prev. | err% var. | innov.std |
| 2006                                     | 31216.7  | 30500.7   | 716.05   | 7.42      | 4.96      | -2.46     | .04       |
| 2007                                     | 32869.2  | 31533.8   | 1335.38  | 5.29      | 1.02      | -4.28     | .08       |
| 2008                                     | 33226.5  | 32190.7   | 1035.84  | 1.09      | -2.06     | -3.15     | .06       |
| 2009                                     | 26363.2  | 27932.9   | -1569.75 | -20.66    | -15.93    | 4.72      | -.03      |
| 2010                                     | 30134.4  | 30428.4   | -294.06  | 14.30     | 15.42     | 1.12      | -.01      |
| 2011                                     | 32578.3  | 32345.5   | 232.83   | 8.11      | 7.34      | -.77      | .01       |
| 2012                                     | 31562.2  | 32554.9   | -992.74  | -3.12     | -.07      | 3.05      | -.06      |
| 2013                                     | 31753.4  | 31073.7   | 679.70   | .61       | -1.55     | -2.15     | .04       |

| confronto indicatore-serie disaggregata |            |            |              |              |             |              |              |             |  |
|---|------------|------------|--------------|--------------|-------------|--------------|--------------|-------------|--|
| trim                                    | liv indic. | liv disag. | var1% indic. | var1% disag. | discrep v1% | var4% indic. | var4% disag. | discrep v4% |  |
| 201001                                  | 93         | 7160       | 7.33         | 5.19         | -2.14       | .87          | 5.53         | 4.67        |  |
| 201002                                  | 98         | 7445       | 5.28         | 3.97         | -1.31       | 21.87        | 17.63        | -4.24       |  |
| 201003                                  | 102        | 7694       | 4.16         | 3.34         | -.81        | 26.54        | 19.43        | -7.12       |  |
| 201004                                  | 104        | 7836       | 1.54         | 1.84         | .31         | 19.50        | 15.11        | -4.39       |  |
| 201101                                  | 107        | 8093       | 3.67         | 3.29         | -.39        | 15.42        | 13.03        | -2.40       |  |
| 201102                                  | 108        | 8158       | .62          | .80          | .18         | 10.32        | 9.57         | -.74        |  |
| 201103                                  | 110        | 8220       | 1.44         | .77          | -.67        | 7.44         | 6.84         | -.59        |  |
| 201104                                  | 108        | 8108       | -1.21        | -1.37        | -.16        | 4.53         | 3.47         | -1.06       |  |
| 201201                                  | 107        | 7955       | -1.26        | -1.88        | -.63        | -.44         | -1.71        | -1.27       |  |
| 201202                                  | 105        | 7836       | -1.70        | -1.49        | .21         | -2.74        | -3.94        | -1.20       |  |
| 201203                                  | 106        | 7905       | 1.06         | .88          | -.18        | -3.09        | -3.83        | -.74        |  |
| 201204                                  | 104        | 7866       | -2.17        | -.50         | 1.67        | -4.03        | -2.99        | 1.05        |  |
| 201301                                  | 100        | 7821       | -3.49        | -.56         | 2.93        | -6.20        | -1.68        | 4.52        |  |
| 201302                                  | 100        | 7908       | -.45         | 1.11         | 1.56        | -5.01        | .91          | 5.92        |  |
| 201303                                  | 100        | 8001       | .15          | 1.17         | 1.02        | -5.87        | 1.20         | 7.07        |  |
| 201304                                  | 99         | 8024       | -.90         | .29          | 1.20        | -4.65        | 2.01         | 6.66        |  |

## coefficienti stimati del modello (t-student tra parentesi)

---

|          |   |         |            |
|----------|---|---------|------------|
| costante | : | 359.17  | (3.10)**   |
| trend    | : | 4.27    | (10.32)**  |
| b_0      | : | 45.304  | (13.76)**  |
| b_1      | : | -42.524 | (-12.71)** |

## diagnostica del modello (livello prob. tra parantesi)

---

|                            |   |                 |
|----------------------------|---|-----------------|
| rho                        | : | .8856           |
| r2                         | : | .9976           |
| r2 corretto                | : | .9972           |
| test f                     | : | 2116.52 (.00)** |
| standard error regressione | : | 113.14          |
| log-likelihood             | : | -147.15         |
| aic                        | : | 9.64            |
| bic                        | : | 9.84            |
| durbin-watson              | : | 1.99            |
| jarque-bera                | : | .22 (.90)       |
| Q-Ljung-Box                | : | .81 (.85)       |
| H-test                     | : | .49443 (.78068) |

## statistiche sulla disaggregazione (ultima meta' del campione tra parantesi)

---

|                                 |   |                                 |
|---------------------------------|---|---------------------------------|
| correlazione var. prime         | : | .96 (.99)                       |
| correlazione var. quarte        | : | .95 (.98)                       |
| rmse var. prime                 | : | 1.51 (1.89)                     |
| rmse var. quarte                | : | 4.43 (5.28)                     |
| massima discrepanza var. prime  | : | 6.81 [200901] (6.81 [200901])   |
| massima discrepanza var. quarte | : | 13.94 [200903] (13.94 [200903]) |

---

- errori medi di previsione su ultimi 8 anni:

- ME liv.: 142.9 - ME var.: -.492 - MAE liv.: 857.0 - MAE var.: 2.713 - RMSE liv.: 963.1 - RMSE var.: 3.008

---