

Quaderni di Dipartimento

Un Indicatore di Attività Economica per la Lombardia e per le Province di Milano e Pavia

Donatella Baiardi
(Università di Pavia)

Carluccio Bianchi
(Università di Pavia)

130 (11-10)

Dipartimento di economia politica
e metodi quantitativi
Università degli studi di Pavia
Via San Felice, 5
I-27100 Pavia

Novembre 2010

UN INDICATORE DI ATTIVITÀ ECONOMICA PER LA LOMBARDIA E PER LE PROVINCE DI MILANO E PAVIA[†]

Donatella Baiardi[§]

Dipartimento di Economia Politica e Metodi Quantitativi
Università degli Studi di Pavia
Via San Felice, 5
I27100, Pavia
dbaiardi@eco.unipv.it

Carluccio Bianchi

Dipartimento di Economia Politica e Metodi Quantitativi
Università degli Studi di Pavia
Via San Felice, 5
I27100, Pavia
cbianchi@eco.unipv.it

Sommario

Questo lavoro si propone di costruire un indicatore coincidente di attività economica ad alta frequenza per la regione Lombardia e le province di Milano e Pavia, mediante l'utilizzo di un *dynamic factor model* costruito secondo l'approccio di Stock e Watson (1998a e 1998b). La tecnica di analisi delle componenti principali sintetizza le informazioni contenute in un ampio *dataset* di base in un numero limitato di fattori comuni capaci di descrivere le alterne fasi cicliche. L'algoritmo *EM (Expectation Maximization)*, infine, consente di ricavare gli indicatori territoriali desiderati tenendo conto dei dati ufficiali annuali sulla dinamica del PIL regionale o del valore aggiunto provinciale.

JEL: E32, C32, C82.

Parole chiave: Indicatori coincidenti di attività economica, Regioni italiane, Indici di diffusione.

[†] Si ringraziano in particolare Loredana Federico e Mario Alessandro Maggi per i numerosi e preziosi consigli forniti durante la preparazione di questo lavoro. Si ringraziano altresì gli Uffici Studi della Banca d'Italia (sede di Milano), delle Camere di Commercio di Milano e di Pavia, nonché l'Osservatorio RegiosS per avere messo a disposizione le serie storiche regionali e provinciali utilizzate per ricavare gli indicatori proposti.

[§] Corresponding author.

AN ECONOMIC ACTIVITY INDICATOR FOR LOMBARDY AND FOR THE PROVINCES OF MILAN AND PAVIA

Donatella Baiardi[§]

Dipartimento di Economia Politica e Metodi Quantitativi
Università degli Studi di Pavia
Via San Felice, 5
I27100, Pavia
dbaiardi@eco.unipv.it

Carluccio Bianchi

Dipartimento di Economia Politica e Metodi Quantitativi
Università degli Studi di Pavia
Via San Felice, 5
I27100, Pavia
cbianchi@eco.unipv.it

Abstract

This paper aims to construct a high-frequency coincident indicator of economic activity for Lombardy and for the provinces of Milan and Pavia, by using the dynamic factor model approach introduced by Stock e Watson (1998a e 1998b). The principal component technique is first used to summarize the information contained in a large dataset in a limited number of common factors capable of capturing the main features of local business fluctuations. The *EM* (*Expectation Maximization*) algorithm then allows to compute the desired territorial indicators by taking into account the official annual data on regional GDP or provincial value-added growth.

JEL: E32, C32, C82.

Keywords: Coincident Economic Activity Indicators, Italian Regions, Diffusion Indexes.

[§] Corresponding author.

1. Introduzione

Gli indicatori di attività economica sono particolarmente utili e interessanti: forniscono informazioni preziose ed attendibili sulla fase congiunturale in atto, possono essere calcolati con ritardi contenuti e con frequenza maggiore di quella propria delle statistiche ufficiali, ed infine permettono di valutare in tempi rapidi l'opportunità di adottare possibili decisioni anticicliche di politica economica. Oggi gli analisti possono disporre di numerosi indicatori coincidenti ed anticipatori volti a descrivere la situazione corrente e prospettica dell'economia di un Paese. In particolare diverse istituzioni ed enti di ricerca, nazionali ed internazionali, pubblicano indicatori coincidenti e anticipatori riferiti in genere all'intera nazione o ad aree economiche integrate¹. A livello territoriale più ristretto, però, le informazioni statistiche disponibili per l'analisi congiunturale sono molto più limitate e pubblicate con notevoli ritardi. Con riferimento alle regioni, in particolare, le statistiche ufficiali disponibili hanno frequenza solo annuale e un ritardo di circa un anno; a livello provinciale i dati disponibili, pure annuali, sono meno dettagliati, essendo espressi soltanto a prezzi correnti e pubblicati con un ritardo di circa tre anni.

In conseguenza di tali carenze, la costruzione di indicatori di attività economica a livello territoriale ristretto appare oltre modo utile, consentendo di analizzare l'evoluzione congiunturale locale in tempo pressoché reale (superando così i ritardi esistenti nella disponibilità di dati da parte delle fonti ufficiali) e di suggerire eventuali provvedimenti di politica economica in una prospettiva di decentramento decisionale. Ciò appare tanto più importante in un contesto in cui, come evidenziano gli studi sui cicli economici locali (Carlino e Sill, 2001), la dinamica di breve periodo delle varie aree territoriali presenta peculiarità distinte, in conseguenza delle caratteristiche strutturali specifiche del proprio sistema economico (Park e Hewings, 2003)². Tale evidenza costituisce una ulteriore giustificazione dell'utilità di costruire indicatori coincidenti a livello regionale e provinciale, al fine di evidenziare uniformità e discrasie nell'evoluzione ciclica delle aree territoriali di interesse rispetto al contesto macroeconomico di riferimento.

Il presente lavoro si propone pertanto di costruire un indicatore coincidente di attività economica ad alta frequenza per la regione Lombardia e per due province profondamente diverse fra loro, quali Milano e Pavia, seguendo una metodologia statistica proposta alla fine del secolo scorso da Stock e Watson (1998a e 1998b) e più recentemente da Benni e Brasili (2006). Al riguardo, mentre indicatori analoghi, pur con differenze nelle tecniche di costruzione, sono disponibili a livello regionale³, la costruzione di un indicatore coincidente provinciale costituisce una novità assoluta nella letteratura e nella pratica empirica.

¹ Con riferimento all'Italia, si possono citare gli indicatori elaborati da OCSE, ISAE (oggi Istat), RegiosS, nonché da altri centri di ricerca privati; con riferimento all'area dell'euro, poi, l'indicatore coincidente più noto è Eurocoin, pubblicato ed aggiornato mensilmente anche sul sito della Banca d'Italia.

² Tali studi focalizzano inoltre l'attenzione sugli effetti dei cambiamenti inaspettati nella politica monetaria (Carlino e DeFina, 1995) o sull'impatto di shock nazionali o locali su territori maggiormente ristretti come le aree metropolitane (Coulson e Rusher, 1995 e Coulson, 1999).

³ RegiosS pubblica sul suo sito indicatori di attività economica coincidenti per tutte le regioni italiane; si veda al riguardo <http://www.regioss.it/>. Rispetto a tali indicatori, quello proposto nel presente lavoro si differenzia sia per il numero di serie utilizzate (con l'introduzione specifica di serie di origine finanziaria, molto importanti per la regione Lombardia), sia per il numero di fattori comuni considerato, sia infine per la considerazione di ritardi

Il resto del lavoro è organizzato nella maniera seguente: il secondo paragrafo presenta una breve rassegna della letteratura concernente l'identificazione e la misura dei cicli economici attraverso la costruzione di indicatori anticipatori e coincidenti di attività economica. Nel terzo paragrafo si descrivono le caratteristiche essenziali del modello statistico utilizzato per la costruzione dell'indicatore ciclico coincidente e si espongono le variabili che compongono il *dataset* regionale. Nel quarto paragrafo vengono illustrati i risultati dell'analisi empirica svolta per la regione Lombardia e brevemente commentate le caratteristiche essenziali dell'evoluzione congiunturale lombarda, quale emerge dall'andamento temporale dell'indicatore elaborato, anche a confronto con l'esperienza italiana. I due successivi paragrafi descrivono le caratteristiche dei *dataset* provinciali e presentano i risultati dell'esercizio empirico svolto con riferimento alla costruzione di un indicatore coincidente di attività economica per le province di Milano e di Pavia; anche in tal caso vengono brevemente commentate le differenti dinamiche congiunturali delle due province prescelte sulla base dell'andamento degli indicatori costruiti. Alcune considerazioni finali completano il lavoro.

2. L'identificazione e la misura statistica dei cicli economici

La misurazione delle fluttuazioni cicliche è stata oggetto di studio approfondito a partire dagli anni Trenta del secolo scorso, quando Wesley Mitchell svolse una ricerca seminale sui vari aspetti del ciclo economico americano, successivamente completata ed approfondita dal *National Bureau of Economic Research (NBER)*. Secondo tale approccio, il ciclo economico viene definito come il comovimento diffuso e comune delle principali variabili che meglio descrivono l'attività economica stessa (Burns e Mitchell, 1946)⁴. L'idea di base è che i cicli siano fluttuazioni nel livello di attività economica persistenti nel tempo che coinvolgono più settori e variabili. Queste fluttuazioni sono ricorrenti ma non periodiche, con una durata che varia da più di un anno a dieci o vent'anni. Di conseguenza, i cicli sono identificati dai punti di svolta superiore e inferiore (*upturn* e *downturn*) di una pluralità di variabili economiche⁵.

L'analisi dell'*NBER* si basa in primo luogo sull'identificazione delle serie storiche che dovrebbero comporre il ciclo di riferimento. L'ampiezza e la durata del ciclo vengono invece determinate sulla base di un approccio statistico volto a determinare i punti di svolta dei comovimenti dell'attività economica generale (ovvero i massimi ed i minimi che segnano la

o anticipi nelle serie di base, il che rende il modello più intrinsecamente dinamico, in linea con l'approccio di Stock e Watson adottato).

⁴ Tale comovimento viene denominato “ciclo di riferimento”, che si caratterizza per tre elementi fondamentali: generalità, ampiezza e diffusione.

⁵ Secondo una interpretazione comune abbastanza diffusa, soprattutto nella pubblicistica economico-finanziaria, il ciclo economico viene caratterizzato con riferimento alla sola dinamica del PIL reale. Inoltre, in tale contesto, una fase recessiva viene identificata come una caduta del PIL che si protrae per due trimestri consecutivi. Nell'approccio adottato dall'*NBER*, tuttavia, questo punto di vista risulta poco accurato, in quanto da un lato trascura il potenziale esplicativo di altre variabili macroeconomiche (quali, ad esempio, il reddito personale al netto dei trasferimenti, l'occupazione, la produzione industriale, le vendite al dettaglio) e dall'altro non considera natura e dimensione del comovimento mensile dei vari indicatori prescelti (in tal modo, come nel caso della recessione USA del 2001, una fase ciclica negativa può non implicare due trimestri consecutivi di caduta del PIL reale).

fine e l'inizio di un periodo di espansione). In tale contesto, la datazione del ciclo è basata su una definizione classica, che prende in considerazione i livelli assoluti di attività economica⁶.

Sulla base di tale approccio, già a partire dagli anni Cinquanta e Sessanta l'*NBER* ha elaborato una metodologia ampia ed articolata, ancor oggi usata per ottenere misure sintetiche del ciclo economico americano, contribuendo allo sviluppo di diversi indicatori congiunturali (indici anticipatori, coincidenti, ritardati, indici di diffusione, ecc.)⁷. Alla fine degli anni Ottanta Stock e Watson (1989) presentano un progetto di revisione degli indicatori anticipatori e coincidenti sino ad allora utilizzati, usando strumenti econometrici. Dai loro studi emerge la proposta di costruire tre indici sperimentali ad alta frequenza: un indicatore economico coincidente, un indicatore economico anticipatore ed un indice di recessione. Dieci anni più tardi sempre Stock e Watson (1999), nel suggerire una generalizzazione della curva di Phillips, avanzano l'idea di sostituire alla tradizionale variabile tasso di disoccupazione un indice aggregato di attività economica basato su centosessantotto serie storiche, ricavato attraverso l'analisi delle componenti principali. In tale contesto il ciclo economico viene quindi identificato come una variabile latente in un modello stato-spazio. Seguendo tale approccio metodologico viene costruito il *Chicago Fed National Activity Index (CFNAI)*, in cui l'indicatore coincidente è rappresentato dal primo fattore comune standardizzato di un database costituito da ottantacinque serie storiche che descrivono le principali caratteristiche dell'evoluzione congiunturale macroeconomica degli Stati Uniti⁸.

Sulla base di tale metodologia, in Italia Altissimo *et al.* (2000) propongono di costruire un indicatore coincidente ed uno anticipatore, volti rispettivamente a descrivere e ad anticipare l'evoluzione del ciclo economico italiano. Il lavoro segue l'approccio classico dell'*NBER*, ma, al tempo stesso, attinge ai contributi di Stock e Watson (1998a e 1998b), utilizzando tecniche derivate dall'analisi delle serie storiche. L'analisi dell'evoluzione temporale del ciclo economico italiano è stata poi oggetto di approfondimento recente da parte di Brasili e Federico (2006), i quali propongono un indicatore coincidente di attività economica per l'Italia, la Germania e l'Austria, servendosi della stessa metodologia adottata per il *CFNAI*.

Questo tipo di studi ha suscitato anche applicazioni a livello territoriale più ristretto⁹. In tale prospettiva, meritano di essere ricordati ancora il contributo di Carlino e Sill (2001), volto a mettere in luce le differenze tra il ciclo economico nazionale e quelli regionali negli Stati Uniti attraverso l'analisi della dinamica dei redditi pro-capite regionali, e quello di Park e Hewings (2003), il quale dimostra formalmente come le asincronie cicliche tra regioni e Paese nel complesso siano riconducibili alle caratteristiche specifiche della struttura industriale (in

⁶ Per contro, l'approccio del CEPR, alla base della costruzione dell'indicatore Eurocoin, si fonda su una visione del ciclo come deviazione dell'attività economica dal suo trend di crescita di lungo periodo.

⁷ La metodologia NBER è stata introdotta in Italia alla fine degli anni Cinquanta da un gruppo di ricerca all'interno dell'ISCO (poi confluito nell'ISAE).

⁸ Per una sommaria descrizione delle caratteristiche del CFNAI si veda il sito web specifico della Federal Reserve Bank di Chicago: <http://www.chicagofed.org/webpages/publications/cfnai/>. In tale sede sono pure elencate le ottantacinque variabili utilizzate per costruire l'indicatore coincidente in questione. Rispetto all'approccio dell'*NBER*, peraltro, basato come si è detto sulla dinamica del livello assoluto di attività economica, l'indice CFNAI considera piuttosto gli scostamenti rispetto al trend di lungo periodo (come nell'approccio CEPR in precedenza ricordato).

⁹ Come si è ricordato in precedenza, la metodologia in esame è stata pure utilizzata a livello sovranazionale; in particolare Eurocoin rappresenta l'indicatore coincidente per l'Area Euro, sebbene in tal caso la definizione di ciclo di riferimento sia piuttosto quella propria dell'approccio CEPR.

senso lato) di ogni regione. Crone e Clyton-Matthews (2004), invece, elaborano un indicatore coincidente per ciascuno Stato USA ricorrendo alla metodologia proposta da Stock e Watson (1989). In Italia, Chiades *et al.* (2003) presentano un indicatore anticipatore e coincidente per l'economia del Veneto. Infine, l'Osservatorio RegiosS elabora e pubblica periodicamente un indicatore coincidente per ogni regione italiana, sistematicamente aggiornato ogni trimestre circa. Sulla base della metodologia contenuta nello studio di Benni e Brasili (2006), ma con l'introduzione di alcuni affinamenti metodologici¹⁰, questo lavoro propone la costruzione di un indicatore coincidente di attività economica per la regione Lombardia e per le province di Milano e Pavia.

3. Il modello utilizzato e le variabili contenute nel *dataset* regionale

La letteratura presenta numerose tecniche ed applicazioni volte all'estrazione di informazioni sintetiche a partire da un'ampia banca dati originaria. In particolare l'*approximate dynamic factor model* (Stock e Watson, 1998b) si propone di scomporre un ampio insieme di serie storiche effettive X_t in due parti non osservate: una componente comune ed una idiosincratica (calcolata in maniera residuale). Il modello di riferimento parte dalla semplice definizione:

$$X = \Lambda F + e \quad (1)$$

in cui X è la matrice originaria dei dati di dimensione ($N \times T$), dove N è il numero di variabili osservate e T il numero di osservazioni disponibili per ogni variabile, Λ è una matrice di coefficienti ($N \times R$), F è una matrice di “fattori comuni” non osservati ($R \times T$) ed infine e è una matrice ($N \times T$) che rappresenta la componente idiosincratica del modello, cioè la misura dell'errore commesso sostituendo ai dati effettivi X la loro rappresentazione sintetica ΛF . In definitiva, tramite il modello, le N serie originarie vengono sintetizzate in un numero ridotto R di variabili non osservate F , con R abbastanza piccolo da compendiare le molte serie storiche originarie in un numero limitato di fattori comuni¹¹. La matrice dei coefficienti Λ , unica e quindi invariante rispetto al tempo, consente di passare (tramite un processo cosiddetto di *loading*) dai fattori comuni F alle serie originarie X . I fattori comuni F in definitiva corrispondono a delle combinazioni lineari delle serie originarie X , ricavate tramite la metodologia delle componenti principali.

Il modello sinora illustrato è statico, nel senso che tutte le variabili considerate fanno riferimento allo stesso istante temporale. Nulla vieta tuttavia di introdurre nell'analisi possibili considerazioni dinamiche. Il modello di base, in effetti, può essere reso dinamico in tre modi diversi (da qui il termine *dynamic factor model*): a) mediante l'inserimento di una struttura di ritardi dei fattori (per cui la matrice F non contiene soltanto fattori comuni contemporanei

¹⁰ Si veda al riguardo quanto riportato nella precedente nota 3 e nella successiva nota 13.

¹¹ In generale R assume un valore non superiore a 5, il che consente di sintetizzare l'ampio dataset di base in non più di 5 fattori comuni. In taluni casi è addirittura possibile che sia $R=1$, cosicché l'intera banca dati può essere riassunta in un solo fattore comune.

rispetto a \mathbf{X})¹²; b) considerando una matrice \mathbf{X}' estesa, corrispondente alla matrice originaria \mathbf{X} aumentata per l'inclusione di variabili ritardate; c) utilizzando una matrice \mathbf{X} di dimensioni analoghe a quella originaria, ma contenente variabili solo correnti o ritardate o anticipate, a seconda della loro significatività nella spiegazione del livello corrente dell'attività economica^{13,14}. L'approccio proposto è basato su una stima non parametrica del modello (1): il procedimento è quindi del tutto equivalente a quello proprio dell'analisi delle componenti principali (Stock e Watson, 1998b).

Al fine di costruire un indicatore regionale di attività economica, capace di sintetizzare in una sola variabile l'evoluzione ciclica sperimentata, si parte da un *dataset* di base abbastanza ampio, costituito da tutte quelle variabili territoriali in grado di consentire una accurata descrizione delle caratteristiche economiche dell'area di interesse. Con riferimento alla Lombardia, in particolare, le serie storiche disponibili sono in tutto quaranta per il periodo compreso tra gennaio 1993 e dicembre 2010 (ultimo dato utilizzato); esse hanno frequenza mensile e sono derivate da fonti diverse¹⁵. Se si considerano le caratteristiche macroeconomiche delle grandezze disponibili, esse possono essere sintetizzate nelle seguenti categorie:

- produzione industriale, ordinativi e scorte regionali desumibili dalle indagini congiunturali mensili ISAE;
- commercio estero (esportazioni ed importazioni regionali verso l'estero espresse in valore);
- mercato del lavoro (occupazione e forze di lavoro);

¹² Forni *et al.* (2000 e 2001) e Altissimo *et al.* (2001 e 2007), ad esempio, utilizzano delle componenti principali dinamiche.

¹³ I primi due approcci sono quelli suggeriti da Stock e Watson (1998a e 1998b). Nel presente lavoro, invece, si è preferito seguire il terzo metodo, per cui sono state soprattutto opportunamente ritardate le variabili aventi carattere tipicamente anticipatore rispetto al livello corrente dell'attività economica (come ad esempio gli ordini, le tendenze della liquidità, dell'economia, degli ordinativi, dei prezzi e della produzione industriale, il giudizio sulla situazione economica per i successivi dodici mesi, sulla disoccupazione, sulle possibilità di risparmio e sulle intenzioni di acquisto di beni durevoli, gli indici di Borsa ed il tasso di cambio reale). Tale procedura appare giustificata dallo scopo sia di evitare che nella matrice estesa \mathbf{X}' compaiano variabili tanto correnti quanto ritardate, sia di aumentare la significatività delle variabili ritardate e la loro correlazione rispetto ai fattori comuni (il ritardo prescelto è quello che massimizza la correlazione di ciascuna serie del database regionale, misurata con una semplice regressione lineare avente il tasso di crescita della produzione industriale, usata come *proxy* della dinamica dell'attività economica, come variabile dipendente e le serie anticipatrici come regressori; l'errore standard relativo ai coefficienti così stimati è stato utilizzato come guida nella scelta del *lag* ottimale delle serie anticipatrici). Seguendo una procedura del tutto analoga, si sono selezionati *lead* ottimali per le serie aventi carattere *lagging* rispetto alle fluttuazioni cicliche (tipicamente le serie riguardanti il mercato del lavoro e la demografia delle imprese). Al riguardo le tabelle A1 e A2 in appendice riportano i dettagli dei *lead* e dei *lag* utilizzati per tutte le variabili che compaiono nel *dataset* regionale ed in quelli provinciali utilizzati nel presente lavoro.

¹⁴ Nella formulazione originaria di Stock e Watson (1999), peraltro, una ulteriore formulazione dinamica del modello è utilizzata per *prevedere* una sola variabile osservabile al tempo $t+1$ (nel caso specifico il tasso di inflazione π) tramite le variabili incluse in \mathbf{F} derivate dalle serie storiche \mathbf{X} osservate al tempo t . Il modello in tal caso diventa $\boldsymbol{\pi}' = \boldsymbol{\beta}'\mathbf{F} + \boldsymbol{\varepsilon}'$, dove $\boldsymbol{\pi}'$ è il vettore riga ($1 \times T$) delle T osservazioni della variabile da prevedere π_{t+1} , $\boldsymbol{\beta}'$ è il vettore riga ($1 \times R$) dei coefficienti, \mathbf{F} è la matrice ($R \times T$) degli R fattori comuni identificati al tempo t ed infine $\boldsymbol{\varepsilon}'$ è un vettore riga degli errori.

¹⁵ L'elenco completo delle variabili costituenti il *database* regionale, della loro fonte, dei test di radice unitaria applicati e delle trasformazioni apportate alle serie storiche originarie, è presentato in appendice nelle tabelle A1-A2. Le serie storiche originarie utilizzate, prima delle eventuali trasformazioni richieste dalla procedura di calcolo dei fattori comuni, sono tutte destagionalizzate.

- indicatori congiunturali qualitativi desumibili dalle indagini mensili ISAE per ogni regione italiana e per le macroaree territoriali (Nord-Est, Nord-Ovest, Centro e Sud) con riferimento alla fiducia di imprese e consumatori rispettivamente¹⁶;
- indice dei prezzi al consumo nel capoluogo regionale;
- demografia delle imprese;
- immatricolazioni di automobili;
- indicatori nazionali ed internazionali di contesto, quali la produzione industriale italiana, francese e tedesca, ed il tasso di cambio reale effettivo italiano;
- variabili finanziarie, come l'indice di Borsa FTSE-MIB30 e lo spread tra il tasso di interesse sui prestiti e quello sui depositi per l'Italia¹⁷.

Al fine di garantire la stazionarietà delle serie utilizzate¹⁸, i dati mensili originari per le serie nei livelli vengono trasformati nelle rispettive variazioni percentuali tendenziali mentre per le serie espresse come rapporti, a causa dell'inadeguatezza della procedura precedente, si considerano le variazioni dodicesime¹⁹. Da ultimo, al fine di utilizzare serie tra di loro omogenee, come prerequisito utile e consigliato per garantire l'efficacia dell'analisi fattoriale, si è proceduto alla standardizzazione delle variabili²⁰.

4. I risultati dell'analisi empirica svolta per la regione Lombardia

Con riferimento alla regione Lombardia, X è una matrice di dimensioni ($N \times T$), con N pari a quaranta serie storiche e T pari a duecentosedici osservazioni. X è la componente osservata, dove, come si è detto in precedenza, sono state introdotte anche variabili ritardate o anticipate²¹. AF è la componente comune: A è la matrice dei coefficienti (*loading*) avente dimensione ($N \times R$), dove R indica il numero dei fattori comuni non noto a priori. Per tale ragione, R viene inizialmente sostituito da un valore fissato arbitrariamente K , con $K \geq R$

¹⁶ Come è noto, l'indice sintetico di fiducia dei consumatori viene calcolato come media aritmetica semplice delle seguenti variabili riferite ai 12 mesi successivi: cambiamento nella situazione finanziaria delle famiglie, situazione economica generale, disoccupazione, intenzioni di risparmio. In questo lavoro tutte le serie elementari sono considerate separatamente, come per quelle relative alla fiducia delle imprese (si veda la nota precedente).

¹⁷ L'inserimento di queste variabili è motivato dal fatto che la Lombardia risulta essere altamente specializzata nel comparto dei servizi finanziari, che costituiscono circa il 30% del valore aggiunto regionale (il 23% circa del sul totale nazionale). Inoltre, Milano, sede della Borsa italiana, è una delle più importanti piazze d'affari d'Europa.

¹⁸ Come è noto serie non stazionarie possono dar luogo a correlazioni spurie, posto che contengono un elemento di trend comune, generando nel nostro caso fattori comuni non significativi. La stazionarietà delle serie storiche viene verificata tramite il test *Augmented Dickey Fuller (ADF)* (cfr. Said e Dickey, 1984) ed il test *KPSS* (cfr. Kwiatkowski *et al.* 1992), i cui valori sono riportati nelle tabelle A1-A2 in appendice.

¹⁹ Con il termine variazione percentuale tendenziale si intende, come è noto, la variazione percentuale di ogni mese rispetto al corrispondente mese dell'anno precedente. Le differenze dodicesime vengono utilizzate per quelle variabili, incluse nel *dataset*, per cui le variazioni percentuali non avrebbero alcun significato logico, come ad esempio per il tasso di disoccupazione e le serie relative ai differenziali di interesse. Le serie costituite da saldi di risposte alle inchieste condotte presso famiglie e imprese, data la loro natura, non sono state trasformate.

²⁰ Come è noto, l'utilizzo della procedura di standardizzazione consente di evitare, soprattutto quando le variabili considerate hanno unità di misura diversa, che le serie storiche a maggiore varianza abbiano un peso preponderante nella determinazione dei fattori comuni.

²¹ Si veda al riguardo la nota 13.

(nell'esercizio empirico svolto K è posto pari a quindici). Infine, F è la matrice dei fattori comuni con dimensione ($R \times T$), mentre e rappresenta la componente idiosincratica del modello ed è una matrice di dimensioni ($N \times T$)²².

Seguendo l'approccio proposto da Stock e Watson (1998b), l'identificazione del numero ottimale R dei fattori comuni avviene attraverso l'analisi delle componenti principali. In particolare la tabella 1 illustra la quota di varianza singolarmente e cumulativamente spiegata dai primi sei dei dieci fattori comuni considerati. Purtroppo, con riferimento alla Lombardia, i criteri standard di selezione del numero ottimale di fattori ampiamente utilizzati in letteratura nei modelli fattoriali (Bai e Ng, 2002) non consentono di selezionare alcun fattore comune; pragmaticamente si è scelto quindi di utilizzare il numero di fattori per i quali la quota di varianza cumulata supera i 2/3 del totale. Sulla base di tale criterio, il numero di fattori comuni complessivamente utilizzato è pari a quattro²³.

Tab. 1 - I risultati dell'analisi delle componenti principali per la Lombardia

R	Varianza spiegata da ciascun fattore	Varianza cumulata spiegata dai fattori
1	33,49	33,49
2	14,97	48,46
3	10,59	59,05
4	7,65	66,70
5	5,72	72,42
6	3,97	76,39

Come si evince dalla tabella 1, il primo fattore comune, capace di spiegare oltre un terzo della varianza complessiva delle serie storiche disponibili, potrebbe costituire già da solo un indicatore coincidente di attività economica con frequenza mensile²⁴. In effetti considerando

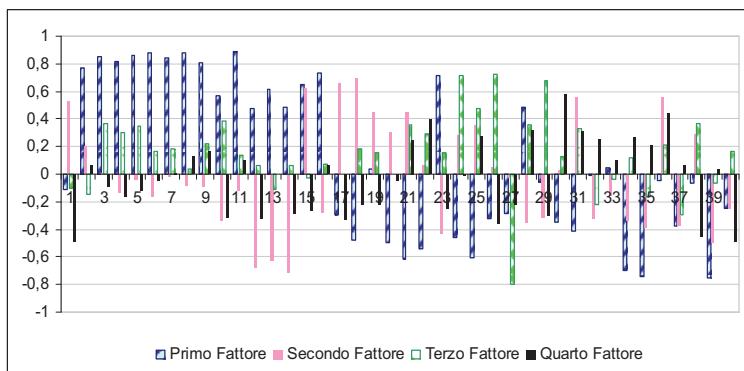
²² La costruzione dell'indicatore coincidente proposto è stata effettuata utilizzando opportuni programmi e codici di Matlab. E-views è stato invece il pacchetto econometrico utilizzato per l'analisi di stazionarietà delle serie storiche e per l'analisi di regressione.

²³ Come si è ricordato nel testo, nella letteratura un metodo ampiamente utilizzato per identificare il numero ottimale di fattori comuni è quello che fa riferimento ai criteri di informazione proposti da Bai e Ng (2002). Nel caso della Lombardia, però, ponendo arbitrariamente un numero massimo di fattori pari a dieci, le due famiglie di criteri di informazione proposti da Bai e Ng non consentono di selezionare alcun fattore comune. I calcoli statistici a tale riguardo sono disponibili su richiesta agli autori. Una regola del pollice alternativa a quella sopra descritta ed utilizzata individua il numero ottimale dei fattori comuni come quello corrispondente al numero totale di fattori i cui autovalori, nella matrice di correlazione, sono maggiori o uguali ad 1. Il limite di questo criterio, noto come Regola di Kaiser, è quello di tendere a selezionare un numero di fattori comuni eccessivo; ad esempio, nel nostro caso, i fattori scelti sarebbero stati in tutto ben nove.

²⁴ In tal caso, la natura dell'indicatore (considerata la procedura di costruzione, per cui si tratta di una variabile standardizzata) sarebbe pressoché analoga a quella propria del *CFNAI* per gli Stati Uniti o dell'indicatore coincidente per l'Italia, la Germania e l'Austria proposto da Brasili e Federico (2006): moltiplicando tale indicatore per il suo scarto quadratico medio si potrebbe ottenere un indicatore di attività reale capace di descrivere la dinamica ciclica come scostamento dalle tendenze di lungo periodo (si veda al riguardo anche la nota 27).

sua struttura, come si nota dalla fig. 1, tale fattore è correlato positivamente con gli ordini interni (0,85), esterni (0,82) e generali (0,86), con la produzione industriale (0,88), con la liquidità tendenziale (0,84), con le tendenze dell'economia (0,88), degli ordini (0,81), della produzione (0,89) e con la liquidità operativa (0,77); esso costituisce quindi un indicatore sintetico del livello di attività reale. Il secondo fattore comune è correlato positivamente con il tasso di occupazione (0,65), il tasso di attività (0,65) e l'occupazione totale (0,69); esso può quindi essere considerato rappresentativo delle condizioni del mercato del lavoro. Il terzo fattore è correlato positivamente con il giudizio delle famiglie sulla situazione economica (0,71), con il bilancio finanziario delle famiglie (0,72) e l'intenzione di acquisto dei beni durevoli (0,68), e negativamente con le possibilità di risparmio per i successivi dodici mesi (0,80); esso sintetizza dunque fondamentalmente lo stato della domanda. Infine il quarto fattore è negativamente correlato con il tasso di cambio reale effettivo (0,44), il differenziale sui tassi di interesse (0,45) e il tasso di inflazione del capoluogo regionale (0,49); esso può quindi cogliere l'influenza della competitività delle imprese e delle variabili monetarie e creditizie.

Fig. 1 - La correlazione di ciascuna serie storica regionale con i quattro fattori comuni



La ricerca di un indicatore regionale coincidente di attività economica potrebbe concludersi a questo punto, sia considerando il primo fattore comune, date le sue caratteristiche sopra descritte, come corrispondente all'obiettivo perseguito, sia, in maniera più appropriata, sfruttando anche l'informazione contenuta negli altri fattori comuni, ottenendo così un indicatore più completo in grado di meglio descrivere la dinamica congiunturale lombarda. Tuttavia, la disponibilità dei valori annuali del PIL regionale a prezzi costanti consente, attraverso l'utilizzo di una appropriata metodologia statistica, di sfruttare in maniera ottimale questa informazione aggiuntiva fondamentale. A tale fine, si è fatto ricorso ad uno strumento ampiamente adottato nell'analisi fattoriale: l'algoritmo *EM* (*Expectation Maximization*)²⁵. Si tratta di una procedura iterativa in cui i dati mancanti nelle serie disponibili ad alta frequenza vengono stimati ricorrendo alla formula (1) in precedenza riportata, in cui i fattori comuni vengono prima calcolati sulla base dei dati esistenti e poi, ad ogni iterazione, ristimati insieme ai dati mancanti, ripetendo la procedura fino al raggiungimento della convergenza. Nel nostro caso, i dati mancanti sono costituiti dalle

²⁵ Per maggiori dettagli sulla natura e sull'utilizzo dell'algoritmo in questione, si veda la successiva nota 26.

osservazioni mensili delle variazioni tendenziali (normalizzate) del PIL lombardo, disponibili con frequenza solo annuale, e con un ritardo di pubblicazione di circa un anno rispetto alle serie mensili disponibili. Nello stimare i dati mancanti si segue fedelmente la procedura elaborata da Stock e Watson (1998b)²⁶. Inoltre, dato che al momento attuale i conti economici regionali sono aggiornati al 2009, l'algoritmo *EM* viene applicato due volte: la prima, per il periodo tra il 1993 e il 2009, per stimare la variazione tendenziale (normalizzata) del PIL mensile attraverso l'informazione statistica fornita dalle serie che costituiscono il *database* e dai fattori comuni selezionati in precedenza; la seconda per completare le osservazioni mancanti dell'indicatore di attività economica per il 2010. La serie così ottenuta, in quanto combinazione lineare di variabili standardizzate, è essa stessa una variabile standardizzata, e quindi andrebbe interpretata, come nell'approccio CFNAI, come deviazione dal trend di lungo periodo. La stessa serie può tuttavia essere trasformata in una più funzionale serie di tassi di crescita tendenziali mensili del PIL utilizzando la definizione di variabile normalizzata.²⁷ Quest'ultima variabile, in definitiva, risulta essere l'indicatore coincidente di attività economica ricercato, il cui andamento è riportato nella figura 2.

Vale la pena di osservare, al riguardo, come l'indicatore qui proposto presenti peculiarità specifiche che lo differenziano nettamente da quelli presenti in letteratura, ed in particolare da quello calcolato dall'Osservatorio RegiosS. In primo luogo, coerentemente con le caratteristiche della struttura produttiva della Lombardia, nel *database* si sono inserite alcune variabili finanziarie rilevanti (come il differenziale tra i tassi di interesse sui prestiti e sugli impieghi bancari e la dinamica dell'indice di Borsa FTSE-MIB30). In secondo luogo, alcune

²⁶ Nel nostro caso, in particolare, la variazione percentuale media annua del PIL regionale, pubblicata dall'Istat, viene utilizzata come termine di riferimento per la stima delle variazioni percentuali tendenziali mensili del PIL non osservate. Come è ben noto, infatti, la media delle variazioni tendenziali mensili del PIL costituisce una approssimazione accettabile (e generalmente utilizzata) della crescita media annua. Più precisamente, la serie delle variazioni percentuali annue standardizzate del PIL regionale non viene mai inserita nel *dataset*, ma si impone il vincolo che la media dei dati mensili non osservati sia pari al valore annuo disponibile all'interno di un processo iterativo di stima dei dati mensili mancanti (si veda anche Stock e Watson, 2002, caso C, pag. 156). La serie storica mensile del PIL non osservata (nella forma sopra ricordata) viene quindi inserita nel *database* originario di serie mensili come variabile aggiuntiva N+1.esima, per cui $X_{N+1,t} = Y_t$), tenuto conto del vincolo annuo per cui:

$$Y_t^a = \frac{1}{12} (Y_{t-11} + Y_{t-10} + \dots + Y_t) \quad \text{per } t = 12, 24, 36, \dots \quad (2)$$

dove Y_t^a è un valore mancante per tutti gli altri valori di t .

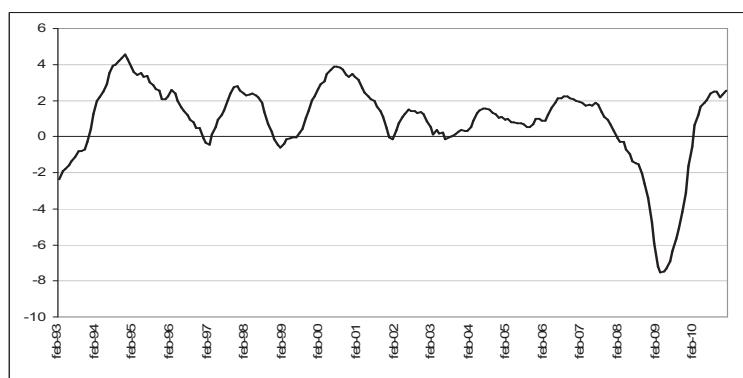
In questa procedura, dunque, l'algoritmo *EM* viene applicato per omogeneizzare le difformi frequenze delle serie esistenti, interpolando il valore della variazione tendenziale del PIL annuale con quello delle variabili e dei fattori comuni selezionati in precedenza, disponibili a cadenza mensile; il risultato è la serie storica delle variazioni tendenziali del PIL regionale ad alta frequenza (mensile) computabile sino al dicembre dell'anno di pubblicazione dei dati ufficiali ISTAT sul PIL regionale. I dati relativi ai mesi successivi alla fine dell'anno per cui si dispone dell'ultima variazione di contabilità nazionale vengono calcolati, sempre utilizzando l'algoritmo *EM*, senza fare ricorso al vincolo definito dalla (2). Per ulteriori dettagli sull'uso dell'algoritmo EM si vedano anche Benni e Brasili (2006) e Canova (2007).

²⁷ Poiché, come è noto, la variazione percentuale standardizzata del PIL Y_t è definita come $Y_t = \frac{V_t - M_V}{\sigma_V}$, dove

V_t è la variazione percentuale tendenziale mensile, mentre M_V e σ_V sono rispettivamente la media e lo scarto quadratico medio delle stesse variazioni percentuali nell'intervallo di osservazione, la variazione percentuale tendenziale mensile del PIL può essere ricavata come $V_t = \sigma_V Y_t + M_V$. Eliminando l'ultimo addendo si otterebbe invece la variazione percentuale tendenziale del PIL al netto del trend (come nell'approccio CFNAI).

serie storiche sono state ritardate (o anticipate)²⁸ in maniera opportuna al fine di aumentare la loro significatività nella determinazione dei fattori comuni; in tal modo, il modello qui proposto costituisce una variante dinamica alternativa non solo rispetto alla specificazione propriamente statica di Benni e Brasili (2006), e quindi dell’Osservatorio RegiosS, ma anche dei precedenti lavori di Stock e Watson (1998a, 1998b)²⁹. Infine, mentre i fattori comuni selezionati dall’Osservatorio RegiosS sono in tutto tre, l’adozione della regola pragmatica adottata in questo lavoro di spiegare almeno i 2/3 della varianza cumulata indica come quattro fattori comuni costituiscano per il nostro *database* una scelta più ragionevole, in grado di consentire una rappresentazione più realistica dell’evoluzione congiunturale lombarda.

Fig. 2 - L’indicatore di attività economica coincidente della Lombardia nel periodo 1993-2010



Nel periodo considerato dalla nostra analisi (1993-2010) l’indicatore coincidente di attività economica della regione Lombardia presenta quattro grandi fasi cicliche, di cui due negli anni Novanta e due nel primo decennio del nuovo secolo. Il primo ciclo (quello connesso alla crisi valutaria dell'estate del 1992 e ai provvedimenti di politica economica restrittiva per l'ingresso nell'UME, di cui qui si coglie peraltro solo la seconda espansiva post-svalutazione della lira) è caratterizzato da una fase espansiva che parte dall'inizio del 1993 e raggiunge il picco nel novembre del 1994; ad esso segue una fase di recessione, connessa ai sopra citati provvedimenti di politica fiscale restrittiva necessari per l'ingresso nell'Unione Monetaria Europea, che termina nel mese di febbraio del 1997. Un secondo ciclo (breve, collegato alla crisi dei paesi del Sud-Est asiatico di fine anni '90) vede, dopo la fase di recupero seguita alla recessione precedente, che raggiunge il massimo nel novembre del 1997, un nuovo periodo di calo produttivo che raggiunge un minimo storico nel gennaio 1999. In seguito, prende avvio una nuova fase espansiva (collegata al periodo di massima euforia della *new economy* negli Stati Uniti e alla successiva esplosione della bolla *dot.com*) che perviene al suo picco nel luglio del 2000, per terminare nel dicembre 2001. A partire dal 2002, analogamente a quanto evidenziato dalle statistiche ufficiali annuali, l’indicatore regionale evidenzia un andamento tendenzialmente stazionario, caratterizzato da modeste e brevi oscillazioni sino all'inizio del 2006, in un contesto generale di modesta crescita produttiva

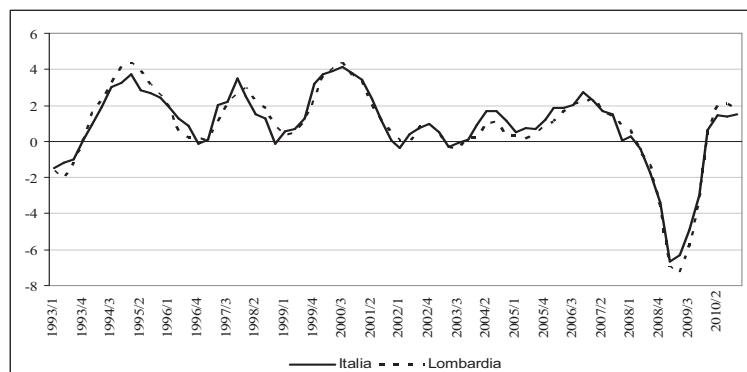
²⁸ Al riguardo si veda quanto specificato nella precedente nota 13. Per maggiori dettagli sui *lead* e *lag* utilizzati si consultino le tabelle A1 e A2 in appendice.

²⁹ Forni *et al.* (2000, 2001) propongono una specificazione dinamica alternativa e più complessa.

(circa l'1% l'anno). In concomitanza con la grave crisi finanziaria mondiale che ha colpito tutti i paesi industrializzati a partire dalla seconda metà del 2007, l'indicatore lombardo segnala un progressivo rallentamento del ritmo di crescita, che si trasforma in una grave recessione, già intensa nel corso del 2008 e poi drammatica nella prima metà del 2009. I dati disponibili sembrano indicare nel marzo del 2009 il punto di minimo di tale fase ciclica negativa, con una variazione tendenziale del PIL regionale (-7,8%) decisamente più consistente di quella media nazionale. La ripresa successiva risulta più sostenuta in Lombardia, soprattutto nel corso del 2010, nonostante la crescita tendenziale del PIL si affievolisca dopo il mese di agosto.

L'evoluzione congiunturale della regione Lombardia, appena illustrata, può essere utilmente comparata con quella italiana, utilizzando per il confronto le medie trimestrali dell'indice mensile regionale qui proposto con le variazioni tendenziali del PIL italiano quali risultano dai conti economici trimestrali dell'Istat. Al riguardo la figura 3 mostra come nel periodo considerato (1993-2010), la dinamica ciclica lombarda e quella del nostro Paese siano nel complesso molto simili. In particolare, tuttavia, mentre i tassi di crescita medi annui del PIL sono praticamente uguali³⁰, la Lombardia si caratterizza per fluttuazioni cicliche più accentuate, come testimoniato sia dalla più elevata deviazione standard dei tassi di crescita (circa 0,2 punti percentuali in più della media nazionale nel periodo considerato) sia dalla maggiore ampiezza dei cicli regionali³¹.

Fig. 3 – La dinamica congiunturale dell'Italia e della Lombardia a confronto



Tale risultato è del resto perfettamente coerente con le caratteristiche strutturali dell'economia lombarda, in cui dal lato della domanda si riscontra una maggiore incidenza degli investimenti fissi lordi e delle esportazioni (grandezze tipicamente più sensibili alle fluttuazioni cicliche rispetto alle altre componenti della domanda aggregata), mentre dal lato

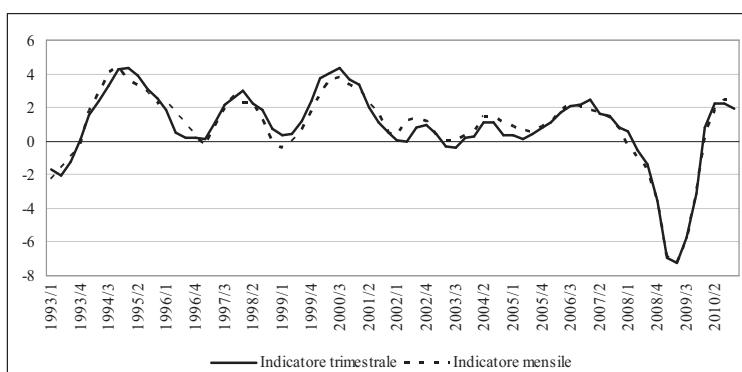
³⁰ Nel periodo in esame il tasso di crescita medio annuo italiano e lombardo sono stati pari entrambi allo 0,9%.

³¹ Una semplice regressione lineare tra le variazioni tendenziali dell'indicatore lombardo e quelle del PIL italiano mostra come la variabilità ciclica lombarda sia mediamente superiore a quella italiana di circa il 5 per cento.

dell'offerta si registra un maggior peso dell'industria in senso stretto (anch'essa più sensibile alle alterne fasi congiunturali)^{32,33}.

La notevole correlazione ciclica tra la Lombardia e l'Italia appena emersa suggerisce di sfruttare tale informazione aggiuntiva nella costruzione dell'indicatore coincidente di attività economica regionale, calcolando un nuovo indicatore, questa volta a frequenza trimestrale, in cui tra le variabili contenute nel *dataset* di base compaia altresì la dinamica del PIL italiano (operativamente la sua variazione percentuale tendenziale standardizzata). Tale nuovo indicatore può essere costruito utilizzando la stessa metodologia già adottata in precedenza, con un numero di serie storiche disponibili ora pari a quarantuno, ed un numero di osservazioni di conseguenza ridotto a settantadue.

Fig. 4 – Un confronto tra l'indicatore di attività economica regionale con frequenza trimestrale e quello con frequenza mensile



Il risultato delle stime effettuate è riportato nella figura 4, in cui si mettono a confronto le dinamiche relative dei due indicatori regionali ottenibili. Come si può osservare, essi mostrano andamenti molto simili, con qualche lieve differenza soltanto nel periodo iniziale compreso tra il 1994 e il primo trimestre del 1999 e nei primi anni dello scorso decennio. Tale evidenza costituisce una prova di robustezza dell'indicatore mensile in precedenza costruito, dato che esso esclude per definizione dalle informazioni di base una variabile rilevante come il PIL nazionale; ciò vale a maggior ragione per il periodo più recente, per il quale non sono disponibili, come a livello nazionale, dati statistici ufficiali sulla congiuntura regionale.

In definitiva, la disponibilità di un indicatore coincidente di attività economica a livello regionale, quale quello proposto in questo lavoro, appare molto utile: da un lato esso consente

³² Con riferimento ai Paesi stranieri, il particolare ruolo del settore manifatturiero nell'evoluzione ciclica regionale è stato studiato da Carlino e Sill (2001), Park e Hewings (2003), e da Coulson (1999) con riferimento alle aree metropolitane.

³³ Confrontando l'evidenza empirica della Lombardia con quella di altre regioni italiane, secondo quanto risulta dalle elaborazioni prodotte dall'Osservatorio RegiosS, si nota inoltre come anche per altre regioni del Nord, quali Piemonte, Veneto, Emilia Romagna e Toscana, la dinamica congiunturale sia simile a quella media italiana. Le regioni del Centro e del Sud, invece, mostrano invece un'evoluzione congiunturale più erratica e talvolta molto diversa rispetto a quella media nazionale (con difformità particolarmente accentuate in alcuni anni).

di descrivere l’evoluzione economica della Lombardia in tempo pressoché reale e ad alta frequenza, rendendo inoltre possibile un confronto con la corrispondente dinamica ciclica nazionale ed internazionale; dall’altro esso permette altresì di quantificare oggettivamente la crescita economica registrata. In effetti, dal punto di vista della possibile implementazione di misure di politica economica anticicliche, questo sembra essere il vero problema regionale: i dati sulle grandezze macroeconomiche aggregate sono disponibili con notevoli ritardi (dell’ordine di circa un anno), non hanno una frequenza sufficientemente elevata (essendo disponibili solo per un arco temporale annuale), presentano numerosi *caveat* e sono spesso oggetto di ampie revisioni contabili.

5. Il dataset disponibile per l’analisi provinciale, con riferimento a Milano e Pavia

Come si è proceduto in precedenza a livello regionale, è possibile ricavare, utilizzando lo stesso approccio metodologico, un indicatore coincidente di attività economica anche a livello provinciale. La costruzione di un indicatore coincidente su una scala territoriale così ristretta costituisce una novità assoluta, permettendo non solo di evidenziare le peculiarità cicliche locali, ma anche di supplire a numerosi problemi presenti nell’analisi congiunturale a livello locale dovuti alla carenza di statistiche ufficiali sintetiche ed aggiornate in tempi rapidi. Per svolgere tale esercizio all’interno della realtà lombarda, si sono scelte due province profondamente diverse tra loro, in termini sia di importanza economica, sia di differenze nella struttura produttiva di base: Milano e Pavia. La provincia di Milano costituisce in effetti, per l’alta specializzazione nell’intermediazione finanziaria (e più in generale nei servizi) e per la sua rilevanza quantitativa (circa la metà della regione e quindi un decimo del valore aggiunto italiano), un termine di riferimento di notevole importanza. Pavia, invece, è una delle province minori della Lombardia e si caratterizza principalmente per una maggiore incidenza della produzione agricola (e dei collegati prodotti agro-industriali), nonché dei servizi legati all’istruzione, alla sanità e ai servizi ricettivi ad essi collegati.

La costruzione di un *dataset* appropriato a livello provinciale incontra tuttavia maggiori ostacoli rispetto al caso regionale, a causa della carenza di informazioni dal punto di vista sia spaziale (poche osservazioni specifiche disponibili) sia soprattutto temporale. In termini aggregati, il valore aggiunto provinciale a prezzi solo correnti è l’unico indicatore economico sintetico esistente; esso presenta frequenza annuale ed i suoi aggiornamenti da parte dell’Istat sono forniti con un ritardo di circa tre anni³⁴. Da un punto di vista più disaggregato, vi è una notevole carenza di dati ad alta frequenza: l’unica eccezione è rappresentata dalle serie storiche a frequenza trimestrale connesse alle indagini svolte dalle Camere di Commercio di ciascuna provincia. Nel complesso, con riferimento alla realtà milanese, è possibile costruire un *dataset* composto da trentanove serie storiche con frequenza trimestrale ($N = 39$) e cinquantadue osservazioni ($T = 52$), per un intervallo temporale compreso tra il primo

³⁴ Come è noto il PIL corrisponde al valore aggiunto più le imposte indirette nette sulla produzione e sulle importazioni, pari a circa il 10% dello stesso PIL. La mancata disponibilità di dati in termini reali costituisce tuttavia una carenza più grave, in quanto non consente di distinguere, data la dinamica del valore aggiunto a prezzi correnti, la componente reale da quella puramente monetaria, soprattutto in contesti in cui l’inflazione risulta variabile.

trimestre del 1998 e l'ultimo trimestre del 2010; per la provincia di Pavia, invece, nello stesso periodo ($T = 52$), si riscontra una maggiore esiguità delle serie storiche disponibili ($N = 36$).

Come già fatto a livello regionale, se si considerano le caratteristiche macroeconomiche delle variabili utilizzate, esse possono essere suddivise nelle seguenti categorie sintetiche:^{35,36}

- produzione industriale provinciale³⁷ e principali indicatori ad essa collegati (ordini, fatturato, tasso di utilizzo degli impianti³⁸);
- commercio estero (esportazioni ed importazioni provinciali verso l'estero espresse in valore);
- mercato del lavoro provinciale (variazione del numero degli addetti nel trimestre e occupazione attesa per i successivi tre mesi);
- indici di fiducia delle imprese per la Lombardia e dei consumatori per il Nord Ovest, nonché altri indicatori congiunturali qualitativi specifici desumibili dalle indagini mensili ISAE, con la stessa disaggregazione territoriale;
- indice dei prezzi al consumo nelle città di Milano e di Pavia;
- indicatori regionali e nazionali di riferimento, come la produzione industriale lombarda ed il tasso di cambio reale effettivo italiano;
- variabili finanziarie, come l'indice di Borsa FTSE-MIB30, lo spread tra i tassi di interesse sui prestiti e sui depositi, i valori degli impieghi e dei depositi bancari per provincia.

Come si può constatare, poiché le serie storiche provinciali specifiche sono limitate, si è dovuto fare ricorso anche ad alcune variabili di “contesto”, ovvero legate sia alle macroaree in cui è inserita la provincia (la regione o il Nord-Ovest) sia alla realtà nazionale. I dati relativi alle immatricolazioni di automobili in provincia non sono pubblicamente disponibili, pur potendo essere ottenuti su richiesta all’Anfia; essi hanno però frequenza solo annuale per gli anni compresi tra il 2001 ed il 2004 (mentre diventano mensili successivamente); per tale motivo, considerata la limitatezza dell’intervallo temporale per cui i dati a frequenza elevata sono disponibili, la serie storica in questione non è stata inserita nel *dataset* di base. Per motivi analoghi, anche la serie sui consumi elettrici per usi produttivi, disponibile a livello regionale e locale da Terna con frequenza solo annuale, è stata omessa dal *database*. Con riferimento al mercato del lavoro le uniche variabili presenti nel *database* provinciale sono quelle relative ai giudizi delle imprese sull’occupazione effettiva e quella prevista per il trimestre successivo, quali risultano dalle rilevazioni trimestrali delle Camere di Commercio di ciascuna provincia, all’interno delle indagini congiunturali svolte³⁹.

³⁵ L’elenco completo delle variabili disponibili a livello provinciale e della loro fonte è riportato in appendice, nella tabella A2.

³⁶ Si noti che nel *database* provinciale vengono utilizzate le stesse serie ritardate o anticipate contenute nel *dataset* regionale, con ritardi tuttavia commisurati alla differente scansione temporale di riferimento (che ora è il trimestre e non più il mese). Per maggiori dettagli si rimanda al contenuto della nota 13.

³⁷ I dati relativi alla produzione industriale provinciale sono disponibili pubblicamente presso il sito internet della Camera di Commercio di ciascuna provincia; a differenza dei dati regionali, diffusi dall’ISAE sotto forma di saldi delle risposte delle imprese, le serie provinciali sono nei livelli.

³⁸ Quest’ultima serie, tuttavia, è disponibile solo per la provincia di Milano.

³⁹ Con riferimento alla provincia di Pavia, le serie dell’occupazione e della produzione attesa per i successivi tre mesi sono disponibili solo per il periodo 2000-2010. Per questa ragione, esse sono state omesse dall’analisi.

6. La costruzione di un indicatore provinciale coincidente

La metodologia adottata per la costruzione di un indicatore coincidente di attività economica provinciale è del tutto analoga a quella illustrata nel paragrafo terzo del presente lavoro. A partire dalle N serie disponibili, in primo luogo si sceglie il numero ottimale R dei fattori comuni tramite l'analisi delle componenti principali. Nel contesto provinciale, però, il numero ottimale di fattori può essere scelto utilizzando i criteri di informazione di Bai e Ng (2002), ampiamente utilizzati in questo tipo di letteratura⁴⁰. In particolare, ponendo un numero massimo di fattori pari a dieci ($K = 10$), il criterio di informazione di Bai e Ng denominato Icp_2 mostra il suo valore minimo in corrispondenza di cinque (minimo relativo) e otto (minimo assoluto) fattori comuni per il *database* di Milano (si veda la tabella A3 in appendice), mentre esso indirizza la scelta verso tre fattori comuni nel caso di Pavia (tab. A4 in appendice), a cui corrisponde il 67% della varianza cumulata spiegata totale. La scelta migliore e più parsimoniosa nel caso di Milano, considerando le indicazioni fornite dai criteri di Bai e Ng, risulta quindi pari a cinque fattori comuni, che insieme spiegano circa l'85% della varianza cumulata complessiva (tab. 2a).

Tab. 2a - I risultati dell'analisi delle componenti principali per la provincia di Milano

<i>R</i>	Varianza spiegata da ciascun fattore	Varianza cumulata spiegata dai fattori
1	38,14	38,14
2	15,41	53,55
3	10,41	68,96
4	5,74	79,37
5	4,72	85,11
6	3,69	89,83

Tab. 2b. - I risultati dell'analisi delle componenti principali per la provincia di Pavia

<i>R</i>	Varianza spiegata da ciascun fattore	Varianza cumulata spiegata dai fattori
1	36,58	36,58
2	19,53	56,11
3	10,81	66,92
4	5,88	72,81
5	4,76	77,56
6	4,62	82,18

Come in precedenza, le figure 5a e 5b presentano le correlazioni esistenti tra ciascuna serie storica inclusa nei *dataset* provinciali e i fattori comuni considerati, le quali confermano i legami attesi tra le variabili contenute nei *database* ed i fattori comuni individuati. In particolare, il primo fattore è correlato positivamente con la produzione industriale locale

⁴⁰ Come si è accennato in precedenza, nel caso della Lombardia, i criteri di informazione di Bai e Ng (2002) non consentono di selezionare alcun fattore comune, dato che, posto un numero massimo di fattori pari a dieci, non si raggiunge mai alcun valor minimo. Contrariamente al caso regionale, nelle applicazioni provinciali, invece, i test di Bai e Ng risultano indicativi. Per questa ragione, mentre per la Lombardia è stata utilizzata la regola pragmatica di scegliere un numero di fattori comuni in grado di spiegare almeno i 2/3 di varianza cumulata, nei due casi provinciali si sono seguite le indicazioni statistiche più robuste offerte dal test di Bai e Ng, presentate nelle tabelle A3 e A4 in appendice. È peraltro utile sottolineare come anche in questo caso i risultati del test di Bai e Ng (2002) selezionano un numero di fattori tale da soddisfare anche la regola pragmatica dei 2/3 della varianza cumulata spiegata utilizzata per la Lombardia (si veda al riguardo la tabella 2).

(0,80 e 0,78)⁴¹, con il fatturato totale (0,79 e 0,68), con gli ordini interni (0,68 e 0,59), con quelli esterni (0,80 e 0,20), con le tendenze dell'economia (0,83 e 0,85), della liquidità (0,92 e 0,80), degli ordini (0,78 e 0,76) e della produzione (0,88 e 0,85), con le importazioni (0,72 e 0,62) e le esportazioni provinciali (0,79 e 0,52) e regionali (0,74 per entrambe le province). Il primo fattore è quindi particolarmente rappresentativo delle condizioni generali del livello di attività economica reale, cosicché, come per la regione, esso potrebbe essere già propriamente considerato come l'intero indicatore coincidente che si vuole costruire.

Fig. 5a – La correlazione di ciascuna serie storica con i fattori comuni selezionati per la provincia di Milano

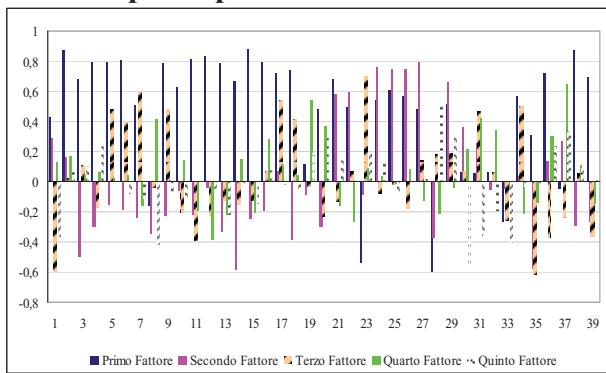
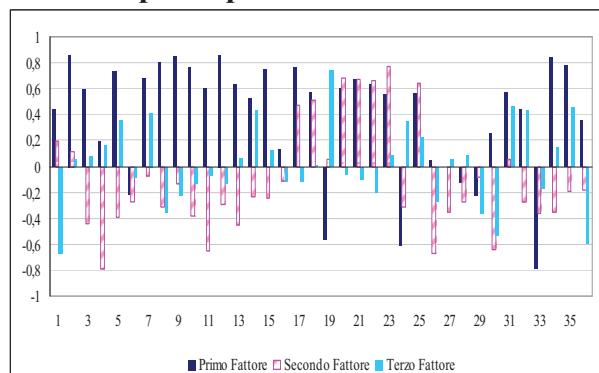


Fig. 5b. - La correlazione di ciascuna serie storica con i fattori comuni selezionati per la provincia di Pavia



Il secondo fattore è correlato positivamente con i giudizi (0,76 e 0,69) e la previsione delle famiglie sulla situazione economica generale (0,76 e 0,67), con il bilancio delle famiglie (0,75 e 0,66), con le possibilità di risparmio (0,79 e 0,78) e con le intenzioni di acquisto di beni durevoli (0,68 e 0,64). Come si è già accennato, il secondo fattore è quindi sostanzialmente rappresentativo delle condizioni della domanda delle famiglie.

Il terzo fattore comune, invece, è correlato negativamente con le giacenze di magazzino (0,59 e 0,67), e positivamente con il grado di utilizzo degli impianti (0,59)⁴², con le importazioni (0,54 e 0,42), le attese sulla disoccupazione per i successivi dodici mesi (0,69 e 0,74) e l'indice di Borsa FTSE-MIB30 (0,50 e 0,53).

Infine, nel caso di Milano (dove i fattori comuni selezionati sono cinque), il quarto fattore comune è correlato positivamente con l'occupazione corrente e quella attesa (0,54 e 0,37 rispettivamente), con la quota di fatturato estero sul totale (0,42), con le imprese iscritte (0,41) e cessate (0,34), con gli impieghi (0,30) e con i depositi (0,65). Il quinto fattore è negativamente correlato con le imprese attive (0,54) ed il tasso di interesse reale (0,39).

Una volta selezionato il numero ottimale di fattori comuni, l'indicatore di attività economica coincidente a livello provinciale viene ottenuto, in maniera analoga a quanto fatto per quello regionale, tramite l'applicazione dell'algoritmo *EM*. In questo caso si tratta di

⁴¹ Per semplicità, il primo valore indicato nelle parentesi fa riferimento alla provincia di Milano, mentre il secondo alla provincia di Pavia.

⁴² Come si è già detto, tale serie è disponibile solo per la provincia di Milano nell'intero intervallo temporale oggetto di indagine.

includere nel *dataset* di origine le informazioni relative alla variazione percentuale tendenziale del valore aggiunto provinciale a prezzi costanti⁴³. Poiché i dati pubblicati dall’Istat relativamente a tale grandezza sono limitati al 2008, anche nel contesto provinciale l’algoritmo in questione viene utilizzato due volte: una prima per interpolare dati con diversa frequenza (e convertire i dati annuali disponibili nei corrispondenti dati trimestrali desiderati); una seconda per completare le serie storiche del valore aggiunto provinciale mancanti (nello specifico per tutti i trimestri degli anni 2009 e 2010).

7. L’indicatore di attività economica di Milano e Pavia a prezzi costanti

Il valore aggiunto, seppure lievemente diverso dal PIL, è comunque un aggregato che consente di valutare in maniera significativa l’evoluzione congiunturale di ogni sistema economico, essendo pari alla differenza tra il valore della produzione di beni e servizi delle varie branche produttive ed il valore dei beni e servizi intermedi impiegati⁴⁴. Le rilevazioni del valore aggiunto ai prezzi base (in euro correnti) al lordo e al netto dei Servizi di Intermediazione Finanziaria Indirettamente Misurati (SIFIM) sono rese disponibili dall’Istat al momento attuale per il periodo tra il 1995 e il 2008.

Come accennato in precedenza, l’indicatore coincidente di attività economica per le province di Milano e di Pavia proposto in questo lavoro è stato calcolato includendo tra le serie di base anche il valore aggiunto a prezzi costanti, ricostruito tramite una metodologia *ad hoc*. Infatti, considerata la mancanza di dati ufficiali sul valore aggiunto a prezzi costanti, quest’ultimo può essere ricavato elaborando una serie di opportuni deflatori. In particolare il deflatore del valore aggiunto provinciale può essere ottenuto, a partire da un certo anno base (tipicamente il 2000, come per le statistiche ufficiali nazionali e regionali), ipotizzando che la dinamica di prezzo provinciale di ciascuna grande branca produttiva disponibile⁴⁵ sia la stessa della regione, ponderando quindi ciascuna variazione di prezzo settoriale con l’incidenza di ciascuna branca sul valore aggiunto complessivo della provincia nell’anno base⁴⁶.

⁴³ Come è noto, l’ISTAT pubblica soltanto dati relativi al valore aggiunto provinciale a prezzi correnti. Qualora si desideri disporre di una valutazione del valore aggiunto a prezzi costanti, al fine di costruire un indicatore di attività economica reale, tale grandezza deve essere ricostruita *ad hoc*. La metodologia seguita in questo lavoro per calcolare serie annuali del valore aggiunto provinciale a prezzi costanti è descritta più avanti nel testo.

⁴⁴ Dal punto di vista del reddito generato, esso corrisponde alla somma delle retribuzioni di tutti i fattori produttivi impiegati nella produzione e degli ammortamenti. Il valore aggiunto può essere calcolato ai prezzi base o ai prezzi di mercato: il primo è pari al saldo tra la produzione ed i consumi intermedi, valutando la produzione ai prezzi base, ossia al netto delle imposte sui prodotti e al lordo dei contributi. Il secondo, invece, è calcolato sulla base dei prezzi di vendita sul mercato dei beni e dei servizi prodotti (e per questa ragione viene appunto definito valore aggiunto ai prezzi di mercato).

⁴⁵ Le branche produttive per le quali sono disponibili deflatori disaggregati sono in totale sei: una per l’agricoltura, due per l’industria (industria in senso stretto e costruzioni), tre per i servizi (commercio, riparazioni, alberghi e ristoranti, trasporti e comunicazioni; intermediazione monetaria e finanziaria, attività immobiliari e imprenditoriali; altre attività di servizi).

⁴⁶ In tal modo il deflatore provinciale differisce da quello regionale per la diversa incidenza delle varie branche produttive sul valore aggiunto. Con riferimento alla provincia di Milano, in particolare, rispetto alla media regionale, si registra una minore importanza dell’industria in senso stretto (4,7 punti percentuali in meno) e dell’agricoltura (-1,3%) ed una maggiore rilevanza di commercio, alberghi, trasporti e comunicazioni (+3,1%) e delle attività di intermediazione finanziaria e imprenditoriale (+3,6%), come ipotizzabile a priori. Nel caso di Pavia, invece, si registra una maggiore incidenza dell’agricoltura (+2,5 punti percentuali in più rispetto alla

Ovviamente, come per il PIL regionale, anche i dati sul valore aggiunto provinciale sono disponibili con frequenza solo annuale; ai fini del presente esercizio, pertanto, le serie annuali ricostruite sono state successivamente trimestralizzate, sempre ricorrendo all'algoritmo EM.

Ciò premesso, le figure 6a e 6b riportano gli indicatori di attività economica provinciali coincidenti ottenuti secondo la metodologia descritta in precedenza. Nel periodo storico considerato, gli indici calcolati presentano significative peculiarità e differenze. Con riferimento anzitutto alla provincia di Milano (fig. 6a), l'indicatore mette in luce l'evidenza di un periodo di apprezzabile e continua crescita economica tra la fine del 1998 e la prima metà del 2001 (connessa alla grande espansione della finanza sino allo scoppio della bolla Internet sui mercati mobiliari). La successiva contrazione del livello di attività è ovviamente più forte nella provincia milanese; inoltre la crisi del settore della finanza e dell'intermediazione creditizia e finanziaria più in generale nell'intero periodo 2002-2006 determinano una sostanziale stazionarietà dell'attività economica, assai più accentuata di quella media regionale (circa 0,6 punti percentuali di minor crescita annua). Nel corso del 2006 un accenno di svolta ciclica si dimostra di breve durata, posto che già a partire dall'inizio del 2007, e quindi in anticipo rispetto alla corrispondente esperienza regionale e nazionale, si registra una marcata flessione dell'attività produttiva in seguito al manifestarsi dei prodromi della recente gravissima crisi finanziaria globale. La contrazione dell'attività economica milanese risulta, come attendibile a priori, più intensa di quella media lombarda ed italiana, con una caduta tendenziale del livello di attività che supera il 9% nel secondo trimestre del 2009. La successiva ripresa è più lenta e meno consistente di quella regionale, con variazioni tendenziali del livello di attività reale appena sopra lo zero, e dunque nettamente al di sotto di quelli regionali (circa 2 punti percentuali di differenza), a testimonianza delle persistenti difficoltà strutturali di un sistema economico in cui il settore dell'intermediazione finanziaria e creditizia riveste un ruolo preponderante.

Con riferimento alla provincia di Pavia (fig. 6b), invece, la dinamica economica sperimentata nel periodo oggetto di indagine risulta completamente diversa. Dopo una recessione più intensa nel biennio 1998-99, già dal quarto trimestre di tale anno si registra l'avvio una forte ripresa, che consente di ottenere ritmi di crescita reale molto elevati per quasi tutto il decennio successivo, sino al 2008. In particolare la recessione successiva allo scoppio della bolla di Wall Street è appena avvertita in provincia, e nell'intero periodo 2002-2006 la crescita media reale (3,4%) è più che doppia di quella lombarda (con uno scarto di 3,7 punti percentuali rispetto a Milano). Anche la recente crisi economica globale sembra colpire meno intensamente la provincia di Pavia, e con ritardo rispetto alle altre realtà territoriali in precedenza commentate. Sebbene la flessione produttiva tra il punto di massimo e quello di minimo ciclico sfiori gli 8 punti percentuali nei tassi di variazione tendenziale (la corrispondente cifra è pari al 10,4% per Milano e al 9,4% per la Lombardia), tuttavia la caduta del livello di attività si commisura attorno ai 4 punti percentuali negativi. Piuttosto, nel caso di Pavia, la successiva fase di ripresa congiunturale appare piuttosto modesta, con tassi di crescita tendenziali ancora in territorio negativo per tutto il 2010, probabilmente come conseguenza delle mutate caratteristiche della domanda dei consumatori, in risposta al

media regionale) e delle cosiddette altre attività di servizi (+3,6%), mentre industria in senso stretto (-5%) e commercio, alberghi, trasporti e comunicazioni (-2,2%) sono sotto la media regionale.

protrarsi della crisi, che hanno influenzato negativamente la specializzazione produttiva pavese⁴⁷.

Fig. 6a - L'indicatore di attività economica milanese

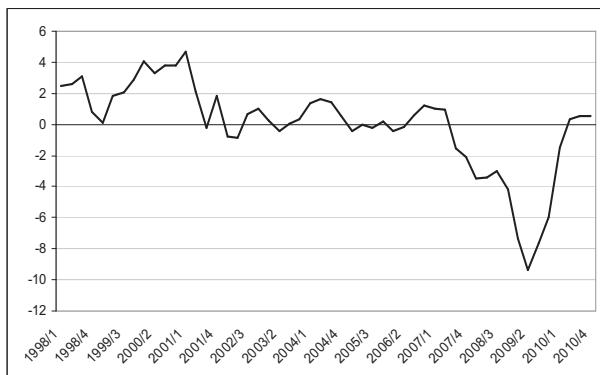
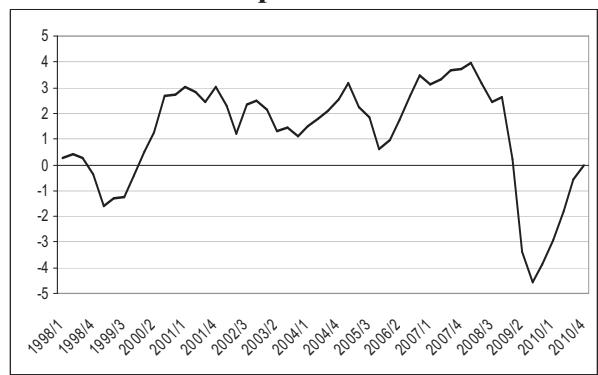


Fig. 6b - L'indicatore di attività economica pavese



8. Conclusioni

Il processo di decentramento decisionale in atto negli ultimi anni nel nostro Paese, unito alle caratteristiche strutturali proprie delle diverse aree territoriali, ha contribuito ad accentuare le differenze registrate nell’evoluzione economica congiunturale a livello locale. In tale contesto la disponibilità di indicatori di attività economica coincidenti, distinti per area territoriale, può risultare molto utile, anche in una prospettiva di politica economica: tali indici permettono infatti di descrivere meglio e in tempi rapidi le dinamiche locali specifiche e di designare eventuali opportuni interventi anticyclici. Ciò appare particolarmente utile soprattutto con riferimento alle regioni, le quali possono disporre di risorse finanziarie significative per adottare misure di sostegno all’economia in caso di necessità.

I conti territoriali ufficialmente pubblicati dall’Istat presentano peraltro numerosi *caveat*: sono disponibili con ritardo, hanno frequenza solo annuale e sono oggetto di revisioni spesso marcate. In particolare, la grandezza che esprime in modo sintetico la situazione dell’economia, ovvero il PIL a livello regionale ed il valore aggiunto a livello provinciale, è reso disponibile con circa un anno di ritardo per le regioni e tre anni per le province, e la pubblicazione di ogni nuovo dato è spesso accompagnata da revisioni significative di quelli precedenti. Nel caso delle province, poi, i dati pubblicati sono solo a prezzi correnti, il che, in un contesto di inflazione variabile, può ostacolare una corretta interpretazione dell’evoluzione economica reale sottostante.

Le difficoltà nel reperire fonti statistiche ampie e dettagliate rappresentano un ostacolo significativo e non trascurabile nei confronti di qualsiasi analisi congiunturale a livello locale.

⁴⁷ Tale esperienza appare perfettamente in linea con i lavori di Quah (1996) e Park e Hewings (2003) citati in precedenza, in cui si suggerisce che le differenti esperienze congiunturali delle varie aree territoriali sono correlate alla struttura produttiva prevalente.

In questo lavoro si propone quindi la costruzione di un indicatore coincidente di attività economica in grado di cogliere con un ritardo limitato (due o tre mesi circa) e con un notevole grado di significatività la dinamica produttiva dell'area di interesse, sintetizzando il contenuto informativo delle variabili che costituiscono il *dataset* disponibile. Tale indicatore, inoltre, può essere immediatamente rapportato al tasso di crescita del PIL o del valore aggiunto, permettendo quindi un'analisi completa e sufficientemente accurata dell'evoluzione congiunturale in atto a livello territoriale. Questa tipologia di analisi appare altresì in grado di segnalare con buon anticipo eventuali problemi nell'interpretazione delle dinamiche territoriali (in presenza di evoluzioni discordanti nelle grandezze elementari di riferimento), tramite il confronto tra i valori dell'indicatore e le variazioni percentuali annuali del PIL regionale o del valore aggiunto provinciale desumibili dalle fonti ufficiali.

Bibliografia

- Altissimo F., Bassinetti A., Cristadoro R., Forni M., Lippi M., Reichlin L., Veronese G. (2001), A Real Time Indicator of the Euro Area Business Cycle, Roma: Banca d'Italia, *Tema di Discussione della Banca d'Italia*, n. 436.
- Altissimo F., Cristadoro R., Forni M., Lippi M., Veronese G. (2007), New Eurocoin: Tracking Economic Growth in Real Time, Roma: Banca d'Italia, *Tema di Discussione della Banca d'Italia*, n. 631.
- Altissimo F., Marchetti D. J., Oneto G. P. (2000), The Italian Business Cycle: Coincident and Leading Indicators and Some Stilized Facts, Roma: Banca d'Italia, *Temi di Discussione del Servizio Studi*, n. 377.
- Bai J., Ng S. (2002), Determining the Number of Factors in Approximate Factors Models, *Econometrica*, 70, 1: 191-221.
- Benni F., Brasili A. (2006), Un indicatore sintetico di attività economica per le regioni italiane, *Rivista di Economia e Statistica del Territorio* 2: 5-24.
- Brasili A., Federico L. (2006), Using Factors Models to Construct New Indicators of the Economic Activity in Austria, Germany and Italy, Roma: ISAE, *Working Paper presented at the 28th CIRET Conference*.
- Burns A. F., Mitchell W.C. (1946), *Measuring Business Cycles*, NBER, New York, Columbia University Press.
- Canova F. (2007), *Methods for Applied Macroeconomic Research*, Oxford: Princeton University Press.
- Carlino G. A., DeFina, R.F. (1995). Regional Income Dynamics, *Journal of Urban Economics*, 37: 88-106.
- Carlino G. A., Sill K. (2001), Regional Income Fluctuations: Common Trends and Common Cycles, *The Review of Economics and Statistics*, 83, 3: 446-456.
- Chiades P., Gallo M., Venturini A. (2003), L'utilizzo degli indicatori compositi nell'analisi congiunturale territoriale: un'applicazione all'economia del Veneto, Roma: Banca d'Italia, *Tema di Discussione della Banca d'Italia*, n. 485.
- Coulson, N. E. (1999). Sectoral Sources of Metropolitan Growth, *Regional Science and Urban Economics*, 29: 723-743.
- Coulson, N.E., Rusher, S.F. (1995). Sources of Fluctuations in the Boston Economy, *Journal of Urban Economics*, 38: 74-93.
- Crone T., Clayton-Matthews A. (2004), Consistent Economic Indexes for the 50 States, Philadelphia: Federal Reserve Bank of Philadelphia, *Working Paper n. 04-9*.
- Forni M., Hallin M., Lippi M., Reichlin L. (2000), The Generalized Dynamic Factor Model: Identification and Estimation, *The Review of Economics and Statistics*, 82: 4, 540-554.
- Forni M., Hallin M., Lippi M., Reichlin L. (2001), Coincident and Leading Indicators for the EURO area, *The Economic Journal*, 101, 471: 62-85.
- Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Y. Shin (1992), Testing the Null of Stationarity Against the Alternative of a Unit Root: How Sure Are we the Economic Time Series Have a Unit Root, *Journal of Econometrics*, 54, 1: 159-178.

- Park, Y., Hewings, G. (2003). Does Industry Mix Matter in Regional Business Cycles?, REAL Discussion Paper 03-T-29.
- Quah D., (1996). Aggregate and Regional Disaggregate Fluctuations, *Empirical Economics*, 21, 137-159.
- Said E., Dickey D.A. (1984), Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order, *Biometrika*, 71: 599–607.
- Stock J.H., Watson M.W. (1989), *New Indexes of Coincident and Leading Economic Indicators*, NBER Macroeconomic Annual, 4: 351-394.
- Stock, J.H., Watson M.W. (1998a), Business Cycle Fluctuations in U.S. Macroeconomic Time Series, Handbook of Macroeconomics, 1, Part A: 3-64.
- Stock J. H., Watson M.W. (1998b), Diffusion Indexes, New York: NBER, *Working Paper*, n. 6702.
- Stock J.H., Watson M. W. (1999), Forecasting Inflation, *Journal of Monetary Economics*, 44, 2: 293-335.
- Stock J.H., Watson M.W. (2002), Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business & Economic Statistics*, 20, 2: 147-162.

APPENDICE

Tab. A1 - La composizione del dataset per la regione Lombardia

Numero	Variabile	Fonte	ADF	KPSS	Trasformazione
1	Giacenze dei prodotti finiti	ISAE	-3,72 (0,00) [2]	0,339	Nessuna
2	Stato della liquidità	ISAE	-3,00 (0,04) [1]	0,406	Nessuna
3	Ordini interni	ISAE	-3,55 (0,00) [4]	0,113	Nessuna Ritardo=3
4	Ordini esterni	ISAE	-2,81 (0,05) [1]	0,126	Nessuna Ritardo=3
5	Ordini generali	ISAE	-3,42 (0,01) [4]	0,118	Nessuna Ritardo=3
6	Produzione industriale	ISAE	-2,88 (0,04) [1]	0,163	Nessuna
7	Tendenza alla liquidità	ISAE	-3,85 (0,00) [1]	0,174	Nessuna Ritardo=3
8	Tendenza dell'economia	ISAE	-3,72 (0,00) [0]	0,153	Nessuna Ritardo=3
9	Tendenza degli ordini	ISAE	-5,09 (0,00) [0]	0,183	Nessuna Ritardo=3
10	Tendenza dei prezzi	ISAE	-3,14 (0,02) [1]	0,092	Nessuna Ritardo=3
11	Tendenza della produzione	ISAE	-2,97 (0,03) [1]	0,174	Nessuna Ritardo=3
12	Esportazioni	ISTAT (Coeweb)	-1,49 (0,54) [2]	1,808	Var. tendenziale
13	Importazioni	ISTAT (Coeweb)	-1,46 (0,55) [9]	1,823	Var. tendenziale
14	Esportazioni Nord Ovest	ISTAT (Coeweb)	-1,54 (0,51) [2]	1,788	Var. tendenziale
15	Tasso di occupazione*	ISTAT	-1,14 (0,92) [0]	0,272	Diff. dodicesime
16	Tasso di disoccupazione*	ISTAT	-1,41 (0,57) [0]	0,860	Diff. dodicesime
17	Tasso di attività*	ISTAT	-0,84 (0,96) [6]	0,226	Diff. dodicesime
18	Occupazione totale*	ISTAT	-2,47 (0,34) [0]	0,228	Var. tendenziale
19	Occupazione nell'industria*	ISTAT	-2,17 (0,50) [0]	0,222	Var. tendenziale
20	Occupazione nel settore dei	ISTAT	-0,25 (0,93) [0]	1,875	Var. tendenziale

	<u>servizi</u>				
21	Giudizio sulla situazione economica per i successivi dodici mesi	ISAE	-2,52 (0,11) [0]	0,298	Nessuna Ritardo=3
22	Previsione sulla situazione economica per i successivi dodici mesi	ISAE	-4,14 (0,00) [0]	0,181	Nessuna Ritardo=3
23	Disoccupazione prevista per i successivi dodici mesi	ISAE	-2,62 (0,08) [0]	0,382	Nessuna Ritardo=3
24	Giudizio sulla situazione personale delle famiglie	ISAE	-2,07 (0,26) [1]	0,625	Nessuna
25	Previsione sulla situazione economica delle famiglie per i successivi dodici mesi	ISAE	-2,97 (0,04) [1]	0,382	Nessuna Ritardo=3
26	Bilancio finanziario delle famiglie	ISAE	-2,69 (0,07) [2]	1,530	Nessuna
27	Possibilità di risparmio per i successivi dodici mesi*	ISAE	-2,53 (0,10) [4]	1,419	Nessuna Ritardo=3
28	Convenienza al risparmio	ISAE	-3,00 (0,13) [1]	0,421	Nessuna
29	Intenzione di acquisto di beni durevoli nei i successivi dodici mesi	ISAE	-2,89 (0,04) [1]	0,696	Nessuna Ritardo=3
30	Immatricolazioni di automobili	Anfia, ACI	-2,98 (0,13) [0]	0,202**	Var. tendenziale
31	Numero di imprese attive	Movimprese, Infocamere	-1,63 (0,99) [0]	1,910	Var. tendenziale
32	Numero di imprese iscritte	Movimprese, Infocamere	-3,51 (0,00) [0]	0,344	Var. tendenziale Lead=3
33	Numero di imprese cessate	Movimprese, Infocamere	-3,45 (0,01) [0]	0,379**	Var. tendenziale Lead=3
34	Produzione industriale tedesca	Eurostat	-1,52 (0,51) [3]	1,550	Var. tendenziale
35	Produzione industriale francese	Eurostat	-1,70 (0,43) [3]	0,985	Var. tendenziale
36	Tasso di cambio reale effettivo	Banca d'Italia	-0,84 (0,35) [4]	0,420	Var. tendenziale, Ritardo=6
37	Indice di Borsa FTSE-MIB30	Eurostat	-1,63 (0,46) [1]	0,719	Var. tendenziale, Ritardo=3
38	Differenziale tra tasso di interesse sui prestiti e sui depositi bancari	Banca d'Italia	-1,45 (0,55) [3]	1,729	Diff. dodicesime, Ritardo=3
39	Produzione industriale italiana	ISTAT	-2,05 (0,26) [3]	0,400	Var. tendenziale
40	Indice generale dei prezzi di Milano*	ISTAT	-1,77 (0,72) [1]	0,372	Var. tendenziale

Test *Augmenting Dickey Fuller* (ADF) e *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS) per testare la presenza di radici unitarie nelle serie storiche del *database* nel periodo compreso tra il 1992 e il 2010. L'ipotesi nulla del test ADF è la presenza di radice unitaria nella serie di interesse. *Test statistic*, p-value nelle parentesi tonde e *lag length*, selezionato automaticamente secondo lo *Schwarz Information Criterion*, nelle parentesi quadrate.

L'ipotesi nulla del test KPSS è l'assenza di radice unitaria (o stazionarietà) della serie di interesse. Il valore critico asintotico al 5% per il test KPSS è 0,463. L'ipotesi nulla viene rifiutata ogni volta il *test statistic* calcolato è superiore al valore critico asintotico (0,463).

Salvo diversa indicazione, tutti i test sono computati includendo la costante nella test equation.

* Nella *test equation* è stata inserita sia la costante, sia il linear trend. Il valore critico asintotico al 5% in questo caso per il test KPSS è 0,146.

** Significativo al 10% (il valore critico asintotico al 10% per il test KPSS è 0,347).

Var. tendenziale indica la variazione tendenziale. Diff. dodicesime le differenze dodicesime.

Tutte le serie storiche sono state infine standardizzate.

Tab. A2 - La composizione del dataset per le province di Milano e Pavia

N	Variabile	Fonte	ADF MI	ADF PV	KPSS MI	KPSS PV	Trasformazione
1	Giacenze dei prodotti finiti	ISAE	-3,23 (0,02) [0]	-3,23 (0,02) [0]	0,498	0,498	Nessuna
2	Stato della liquidità	ISAE	-1,65 (0,45) [0]	-1,65 (0,45) [0]	0,679	0,679	Nessuna
3	Ordini interni	Camera di Commercio provinciale	-1,56 (0,49) [0]	-2,87 (0,05) [1]	0,703	0,172	Var. tendenziale Ritardo=1
4	Ordini esterni	Camera di Commercio provinciale	-0,928 (0,77) [0]	-1,18 (0,68) [2]	0,918	0,479**	Var. Tend. Ritardo=1
5	Produzione industriale	Camera di Commercio provinciale	-1,48 (0,53) [1]	0,240 (0,97) [10]	0,815	0,479	Var. tendenziale
6	Produzione delle piccole imprese	Camera di Commercio provinciale	-1,49 (0,53) [1]	-	0,680	-	Var. tendenziale
7	Tasso di utilizzo degli impianti	Camera di Commercio provinciale	-2,22 (0,20) [0]	-	0,721	-	Diff. Quarte
8	Quota di fatturato estero sul fatturato totale	Camera di Commercio provinciale	-5,42 (0,00) [0]	-4,25 (0,00) [0]	0,422	0,250	Nessuna
9	Fatturato totale	Camera di Commercio provinciale	-0,60 (0,86) [0]	-0,43 (0,89) [1]	0,845	0,763	Var. tendenziale
10	Attese della produzione per il trimestre successivo	Camera di Commercio di Milano	-2,59 (0,10) [4]	-	0,376	-	Nessuna
11	Tendenza alla liquidità	ISAE	-2,86 (0,05) [0]	-2,86 (0,05) [0]	0,279	0,279	Nessuna Ritardo=1
12	Tendenza dell'economia*	ISAE	-3,60 (0,00) [1]	-3,60 (0,00) [1]	0,312	0,312	Nessuna Ritardo=1
13	Tendenza degli ordini	ISAE	-3,54 (0,01) [0]	-3,54 (0,01) [0]	0,241	0,241	Nessuna Ritardo=1

14	Tendenza dei prezzi	ISAE	-2,37 (0,15) [0]	-2,37 (0,15) [0]	0,116	0,116	Nessuna Ritardo=1
15	Tendenza della produzione	ISAE	-3,84 (0,00) [1]	-3,84 (0,00) [1]	0,371	0,371	Nessuna Ritardo=1
16	Esportazioni provinciali	ISTAT	-1,98 (0,29) [4]	-1,25 (0,64) [0]	0,611	0,869	Var. tendenziale
17	Importazioni provinciali	ISTAT	-2,04 (0,27) [4]	-0,81 (0,80) [0]	0,627	1,020	Var. tendenziale
18	Esportazioni della Lombardia	ISTAT	-1,26 (0,64) [5]		0,874		Var. tendenziale
19	Esportazioni di Milano	ISTAT	-	-1,98 (0,29) [4]	-	0,611	Var. tendenziale
20	Occupazione	Camera di Commercio provinciale	-7,55 (0,00) [0]	-5,58 (0,00) [0]	0,109	0,229	Nessuna Lead=1
21	Occupazione attesa per il trimestre successivo	Camera di Commercio provinciale	-3,88 (0,00) [0]	-	0,101	-	Nessuna Lead=1
22	Giudizio sulla situazione economica per i successivi dodici mesi	ISAE	-2,18 (0,22) [1]	-2,18 (0,22) [1]	0,551	0,551	Nessuna Ritardo=1
23	Previsione sulla situazione economica per i successivi dodici mesi	ISAE	-3,60 (0,00) [0]	-3,60 (0,00) [0]	0,416	0,416	Nessuna Ritardo=1
24	Disoccupazione prevista per i successivi dodici mesi	ISAE	-2,19 (0,21) [0]	-2,19 (0,21) [0]	0,334	0,334	Nessuna Ritardo=1
25	Giudizio sulla situazione personale delle famiglie	ISAE	-1,18 (0,67) [0]	-1,18 (0,67) [0]	0,607	0,607	Nessuna
26	Previsione sulla situazione economica delle famiglie per i successivi dodici mesi	ISAE	-1,72 (0,42) [0]	-1,72 (0,42) [0]	0,656	0,656	Nessuna Ritardo=1
27	Bilancio finanziario delle famiglie	ISAE	-1,52 (0,51) [0]	-1,52 (0,51) [0]	0,798	0,798	Nessuna
28	Possibilità di risparmio per i successivi dodici mesi	ISAE	-1,63 (0,46) [0]	-1,63 (0,46) [0]	0,675	0,675	Nessuna Ritardo=1
29	Convenienza al	ISAE	-0,82	-0,82	0,650	0,650	Nessuna

	risparmio		(0,80) [1]	(0,80) [1]			
30	Intenzione di acquisto di beni durevoli nei i successivi dodici mesi	ISAE	-2,57 (0,10) [0]	-2,57 (0,10) [0]	0,416	0,416	Nessuna Ritardo=1
31	Imprese attive*	Movimprese, Infocamere	-1,55 (0,79) [0]	-1,11 (0,74) [4]	0,222	0,811	Var. tendenziale
32	Imprese iscritte*	Movimprese, Infocamere	-1,43 (0,84) [3]	-1,05 (0,26) [3]	0,174	0,152	Var. tendenziale, lead=1
33	Imprese cessate	Movimprese, Infocamere	-0,16 (0,62) [2]	-0,44 (0,45) [2]	0,513	0,184	Var. tendenziale, lead=1
34	Tasso di cambio reale effettivo	Banca d'Italia	-1,25 (0,64) [0]	-1,25 (0,64) [0]	0,514	0,514	Var. tendenziale, Ritardo=2
35	Indice di Borsa FTSE-MIB30	Eurostat	-2,15 (0,51) [1]	-2,15 (0,51) [1]	0,115	0,115	Var. tendenziale, ritardo=1
36	Differenziale tra tasso di interesse sui prestiti e sui depositi bancari	Banca d'Italia	-2,05 (0,26) [1]	-2,05 (0,26) [1]	0,836	0,836	Diff. Quarte, ritardo=1
37	Impieghi bancari della provinciali	Banca d'Italia	-1,83 (0,36) [4]	-1,28 (0,88) [0]	0,874	0,242	Var. tendenziale
38	Depositi bancari della provinciali	Banca d'Italia	-0,85 (0,79) [0]	-0,38 (0,90) [1]	0,880	0,594	Var. tendenziale
39	Produzione industriale della Lombardia	ISAE	-3,02 (0,04) [1]	-3,02 (0,04) [1]	0,254	0,254	Nessuna
40	Produzione industriale di Milano	Camera di Commercio di Milano	-	-1,48 (0,53) [1]	-	0,815	Var. tendenziale
41	Indice dei prezzi al consumo (MI-PV)	ISTAT	-1,09 (0,71) [2]	-3,02 (0,04) [1]	0,904	0,155	Var. tendenziale

Test *Augmenting Dickey Fuller* (ADF) e *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS) per testare la presenza di radici unitarie nelle serie storiche del *database* nel periodo compreso tra il 1998 e il 2010. L'ipotesi nulla del test ADF è la presenza di radice unitaria nella serie di interesse. *Test statistic*, p-value nelle parentesi tonde e *lag length*, selezionato automaticamente secondo lo *Schwarz Information Criterion*, nelle parentesi quadrate. L'ipotesi nulla del test KPSS è la assenza di radice unitaria (o stazionarietà) della serie di interesse. Il valore critico asintotico al 5% per il test KPSS è 0,463. L'ipotesi nulla viene rifiutata ogni qual volta il *test statistic* calcolato è superiore al valore critico asintotico (0,463).

Salvo diversa indicazione, tutti i test sono computati includendo la costante nella *test equation*.

* Nella *test equation* è stata inserita sia la costante, sia il linear trend. Il valore critico asintotico al 5% in questo caso per il test KPSS è 0,146.

Var. tendenziale indica la variazione tendenziale. Diff. quarte le differenze quarte.

Tutte le serie storiche sono state infine standardizzate.

Tab. A3 - I criteri di informazione di Bai e Ng (2002) per il dataset di Milano

R	Icp ₁	Icp ₂	Icp ₃	Pcp ₁	Pcp ₂	Pcp ₃
1	-0,36041	-0,3353	-0,40575	0,613668	0,614921	0,611405
2	-0,50765	-0,45743	-0,59834	0,469467	0,471973	0,464942
3	-0,62213	-0,54679	-0,75815	0,374317	0,378077	0,36753
4	-0,65637	-0,55593	-0,83774	0,324959	0,329971	0,315909
5	-0,68643	-0,56088	-0,91314	0,285618	0,291883	0,274305
6	-0,70311	-0,55245	-0,97517	0,256341	0,263859	0,242766
7	-0,73268	-0,5569	-1,05007	0,229945	0,238716	0,214107
8	-0,76889	-0,56801	-1,13163	0,207714	0,217738	0,189613
9	-0,78579	-0,55979	-1,19387	0,19267	0,203947	0,172306
10	-0,79424	-0,54313	-1,24766	0,18175	0,19428	0,159124

Tab. A4 - I criteri di informazione di Bai e Ng (2002) per il dataset di Pavia

R	Icp ₁	Icp ₂	Icp ₃	Pcp ₁	Pcp ₂	Pcp ₃
1	-0,3155	-0,29076	-0,35968	0,637859	0,638906	0,635989
2	-0,51094	-0,46148	-0,5993	0,462221	0,464314	0,45848
3	-0,60771	-0,53351	-0,74025	0,372101	0,375242	0,366491
4	-0,62891	-0,52999	-0,80565	0,324382	0,328569	0,316901
5	-0,6542	-0,53055	-0,87512	0,283807	0,289042	0,274457
6	-0,70747	-0,55909	-0,97257	0,244583	0,250864	0,233362
7	-0,75819	-0,58508	-1,06747	0,213897	0,221225	0,200806
8	-0,80836	-0,61052	-1,16183	0,189786	0,19816	0,174824
9	-0,8363	-0,61372	-1,23395	0,17361	0,183031	0,156779
10	-0,8816	-0,63429	-1,32343	0,159222	0,16969	0,140521