

I DATI PANEL DATA : MODELLI DINAMICI

[D] Variabili esplicative correlate con l'errore

In questa trattazione andremo a modificare le prime due ipotesi di specificazione di tutti i modelli lineari stimati con il metodo OLS: esogenità delle esplicative [Cfr D.1] e modello statico [Cfr D.2]. La trattazione sarà comune in quanto, come vedremo meglio, la presenza della dipendente ritardata fra le covariate del modello panel dinamico è condizione sufficiente per generare correlazione fra esplicative e termini di errore e, quindi, inconsistenza di tutti gli stimatori.

[D.1] L'approccio delle variabili strumentali

Se nel modello classico di regressione lineare viene meno l'ipotesi di incorrelazione delle variabili esplicative con il termine di errore, lo stimatore OLS (o GLS a seconda delle ipotesi di specificazione degli errori del modello) perde le proprietà della correttezza e della consistenza. Due casi di insorgenza di correlazione tra regressori ed errore (errori di misura nelle covariate e simultaneità) sono riscontrabili in modelli di cross-section, di serie storiche e di panel. In quest'ultimo caso, in letteratura esistono stimatori corrispondenti agli OLS pooled, LSDV e GLS da applicare quando le variabili esplicative sono correlate con i termini di errore.

In pratica, dopo avere scelto il modello panel (il grado di poolability) con cui trattare l'eterogeneità (pooled completo, FE, RE), se si sospettano problemi di correlazione fra le esplicative ed i termini di errore, si usano i corrispondenti stimatori IV. Il ricercatore deve aggiungere alle ipotesi di specificazione anche l'elenco delle variabili strumentali. Si ricordi che la condizione (necessaria) di identificazione dei parametri è che il numero di strumenti (esogeni) sia uguale o superiore al numero di variabili endogene esplicative (da strumentare).

Il terzo caso di insorgenza di correlazione tra covariate ed errore è quello di modelli dinamici con errori autocorrelati; pertanto, richiede che i dati abbiano dimensione temporale (serie storiche o panel).

Tuttavia, contrariamente alle serie storiche, nel caso panel la dinamica implica l'insorgere di inconsistenza degli stimatori senza richiedere la contemporanea presenza di errori autocorrelati. Tutto ciò è vero sia nel caso di modelli FE, sia per modelli RE. Questo fatto solleva importanti questioni in quanto le relazioni di comportamento spesso implicano componenti dinamiche. L'argomento è trattato nella seguente Section [D.2]. Con le dovute modifiche, i metodi di stima presentati di seguito per modelli dinamici possono essere applicati anche al caso di modelli statici affetti da problemi di correlazione tra covariate e termine di errore.

[D.2] Modelli dinamici con dati panel

In generale, elementi di dinamica nei modelli panel possono essere introdotti (i) ipotizzando che gli errori siano processi AR (dinamica degli errori), (ii) aggiungendo come regressori alcuni ritardi della variabile dipendente (dinamica dell'equazione). Dato che il caso (i) è ritenuto un approccio meno soddisfacente dal punto di vista metodologico (vedi la discussione preliminare sub 1.2), supponiamo che il modello 'vero' sia dinamico:

$$(7) \quad y_{it} = \rho y_{it-1} + a_i + \varepsilon_{it} \quad \varepsilon_{it} \sim \text{IID}(0, \sigma^2)$$

dove a_i sono gli effetti fissi.

Avvertenza: in ciò che segue, ci riferiremo a modelli con effetti one way (solo individuali), in cui l'unica variabile esplicativa è la dipendente ritardata di un periodo. L'estensione a casi più articolati non modifica la sostanza di quanto enunciato.

In questo modello è possibile assumere:

- x è predeterminata o esogena in senso sequenziale: $E(\varepsilon_{it}|x_{is})=0 \forall t \geq s$ and $E(\varepsilon_{it}|x_{is}) \neq 0 \forall t < s, \forall i=1, \dots, N, \forall t, s=1, \dots, T$. Soltanto le esplicative passate e presenti non sono correlate con gli errori correnti. Infatti, essendo nel modello AR $x_{it} = y_{it-1}$, segue che $E(x_{it}|\varepsilon_{it}) = E(x_{it-1}|\varepsilon_{it}) = 0$, ma $E(x_{it+1}|\varepsilon_{it}) \neq 0$ (e anche $E(x_{it}|\varepsilon_{it-1}) \neq 0$); quindi, la condizione di esogeneità stretta è violata.

La trasformazione within equivale a fare le seguenti elaborazioni (cfr Nickell, 1981, Econometrica).

Sottraggo $y_{it} = \rho y_{it-1}^* + a_i + \varepsilon_{it}$ (con $y_{it-1}^* = T^{-1} \sum_{t=0}^{T-1} y_{it} \neq y_{it-1}$) dalla (7)

$\rightarrow (y_{it} - y_{it-1}) = \rho (y_{it-1} - y_{it-1}^*) + (\varepsilon_{it} - \varepsilon_{it-1})$ in cui:

$Cov[(y_{it-1} - y_{it-1}^*), (\varepsilon_{it} - \varepsilon_{it-1})] = Cov[y_{it-1}, (\varepsilon_{it} - \varepsilon_{it-1})] = -Cov(y_{it-1}, \varepsilon_{it-1}) = -[\sigma_\varepsilon^2 (1-\rho^{t-1})]/[T(1-\rho)] \neq 0$
 [ricorda che $\varepsilon_{it} = (\varepsilon_{i1} + \varepsilon_{i2} + \varepsilon_{i3} + \dots + \varepsilon_{iT})/T$]

\rightarrow La trasformazione within crea una correlazione fra esplicative e residui nel modello che scompare solo asintoticamente ($T \rightarrow \infty$) Ricapitolando:

T grande $\rightarrow Cov = 0 \rightarrow$ incorrelazione asintotica

T piccolo $\rightarrow Cov \neq 0 \rightarrow$ distorsione negativa

Nota che i precedenti risultati valgono nonostante nessuna autocorrelazione sia stata ipotizzata per gli ε_{it}

\rightarrow è necessario “strumentare” y_{it-1} , ma come?

La sostanza non cambia quando il modello dinamico presenta effetti casuali $\mu_i \sim IID(0, \sigma_\mu^2)$:

(8) $y_{it} = a + \rho y_{it-1} + (\mu_i + \varepsilon_{it}) \quad \varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2)$

perché l'esistenza di covarianza non nulla fra le esplicative ed i termini di errore ($\mu_i + \varepsilon_{it}$) la si ottiene per per definizione a causa della presenza dell'effetto individuale μ_i sia in y_{it} sia in y_{it-1} .

[D.2.1] L'approccio IV per panel dinamici: Anderson-Hsiao

La trasformazione in differenze prime delle precedenti equazioni (7) e (8):

(9) $\Delta y_{it} = \rho \Delta y_{it-1} + \Delta \varepsilon_{it}$

elimina gli effetti individuali e presenta:

$\Delta y_{it-1} = (y_{it-1} - y_{it-2})$

$\Delta \varepsilon_{it} = (\varepsilon_{it} - \varepsilon_{it-1})$

\rightarrow esplicative ed errore sono ancora correlati, per definizione, nelle componenti in grassetto.

Tuttavia, qualsiasi variabile datata prima di t-1 è uno strumento valido (se $\varepsilon_{it} \sim IID$).

Anderson-Hsiao (1981) suggeriscono di utilizzare Δy_{it-2} come strumento per Δy_{it-1} e di stimare ρ nella (7) col metodo IV-2SLS.

Arellano (1989) nota che y_{it-2} (in livelli) è, per definizione, uno strumento *rilevante* (non-debole) perché correlato con l'esplicativa Δy_{it-1} e, allo stesso tempo, *esogeno* perché incorrelato con l'errore $\Delta \varepsilon_{it}$. Inoltre, y_{it-2} permette di risparmiare un'osservazione (guadagno in termini di gradi di libertà).

Se il modello di partenza include variabili esplicative time-invariant (è come aggiungere una componente “ $b z_i$ ”), queste vengono azzerate dal calcolo delle differenze prime (al pari degli effetti individuali) e il parametro b non può essere stimato. Per stimare b :

- (i) si stima IV il modello in differenze (quindi escludendo b);
- (ii) si salvano i residui e se ne calcola la media individuale;

(iii) si stima b da una regressione cross-section in cui la dipendente sono le medie individuali dei residui sub (ii) e l'esplicativa è z_i .

Questa stima di b è consistente per $N \rightarrow \infty$, inconsistente se N è fisso e $T \rightarrow \infty$.

[D.2.2] L'approccio GMM-diff per panel dinamici: Arellano-Bond

Lo stimatore IV dell'equazione in differenze (9) ha alcune pregevoli caratteristiche: è molto semplice da implementare e, soprattutto, è consistente. D'altro canto, non è efficiente perché:

- non usa tutte le possibili condizioni di ortogonalità
- non tiene conto della struttura degli errori $\Delta \varepsilon_{it}$

Una risposta ad entrambi i precedenti punti viene fornita dall'approccio del metodo generalizzato dei momenti (GMM-diff) proposto da Arellano-Bond (1991), uno stimatore per panel dinamici quando T è piccolo e N è (molto) grande. Nei casi in cui la dimensione temporale è elevata, il semplice approccio IV (solo consistente) di Anderson-Hsiao [Cfr D.2.1] costituisce una semplice e attraente alternativa.

Arellano-Bond (1991) notano che addizionali strumenti possono essere ottenuti utilizzando le condizioni di ortogonalità fra la dipendente ritardata e i disturbi ε_{it} . Il loro approccio viene illustrato dal semplice modello AR(1):

$$y_{it} = \rho y_{it-1} + v_{it} \quad \text{con: } v_{it} = \mu_i + \varepsilon_{it}$$

dove: $\mu_i \sim \text{IID}(0, \sigma_\mu^2)$ e $\varepsilon_{it} \sim \text{IID}(0, \sigma_\varepsilon^2)$. Allo scopo di ottenere uno stimatore consistente di ρ per T fisso (piccolo) e $N \rightarrow \infty$ si calcolano le differenze prime per eliminare gli effetti individuali μ_i (il risultato è quello già riportato nella precedente (9)):

$$(10) \quad \Delta y_{it} = \rho \Delta y_{it-1} + \Delta \varepsilon_{it}$$

Si noti che l'errore della (10) è un modello MA(1) non invertibile. Se le osservazioni iniziano in $t=1$, il primo periodo in cui la relazione (10) può essere osservata è in $t=3$. La Tabella 3 sintetizza gli strumenti utilizzabili per ogni osservazione.

Tab. 3 – Gli strumenti utilizzabili

t =	equazione (10):	strumenti:
3	$y_{i3} - y_{i2} = \rho (y_{i2} - y_{i1}) + \varepsilon_{i3} - \varepsilon_{i2}$	y_{i1}
4	$y_{i4} - y_{i3} = \rho (y_{i3} - y_{i2}) + \varepsilon_{i4} - \varepsilon_{i3}$	$y_{i1} \quad y_{i2}$
...
T	$y_{iT} - y_{iT-1} = \rho (y_{iT-1} - y_{iT-2}) + \varepsilon_{iT} - \varepsilon_{iT-1}$	$y_{i1} \quad y_{i2} \quad \dots \quad y_{iT-3} \quad y_{iT-2}$

I precedenti strumenti possono essere organizzati nella seguente matrice W_i di dimensione $(T-2) \times [(T-1)(T-2)/2]$:

$$W_i = \begin{bmatrix} y_{i1} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & y_{i1} & y_{i2} & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & y_{i1} & y_{i2} & \dots & y_{iT-3} & y_{iT-2} \end{bmatrix}$$

da cui, sovrapponendo le N matrici W_i , si ottiene W (di $N(T-2)$ righe) che rappresenta l'insieme degli strumenti per tutte le NT osservazioni:

$$W = (W_1', \dots, W_N)'$$

Una procedura IV non tiene conto della particolare struttura della matrice di varianze-covarianze di $\Delta\varepsilon_{it}$:

$$E(\Delta\varepsilon_i \Delta\varepsilon_i') = \sigma_i^2 (I_N \otimes G) = \Phi$$

dove: $\Delta\varepsilon_i' = (\varepsilon_{i3}-\varepsilon_{i2}, \varepsilon_{i4}-\varepsilon_{i3}, \dots, \varepsilon_{iT}-\varepsilon_{i(T-1)})'$, e:

$$G = \begin{bmatrix} 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2 & -1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 2 \end{bmatrix}$$

G è una matrice $(T-2) \times (T-2)$ la cui particolare forma dipende dal fatto che $\Delta\varepsilon_i$ è un processo MA(1) non invertibile.

Le condizioni di ortogonalità (equazioni dei momenti) fra gli strumenti ed i termini di errore che abbiamo descritto sono: $E(W_i' \Delta\varepsilon_i) = 0$ con $i = 1, 2, \dots, N$. Il vec della (10) è premoltiplicato per W' allo scopo di imporre le precedenti condizioni di ortogonalità:

$$(11) \quad W' \Delta y = W' (\Delta y_{-1}) \rho + W' \Delta \varepsilon$$

Per applicare lo stimatore GLS, si definisce la matrice di pesi

$\Omega^{-1} = [W' \Phi W]^{-1}$ (\rightarrow feasible perché la sua struttura è nota) e si ottiene lo stimatore (preliminare e consistente) “one-step” di Arellano-Bond:

$$\begin{aligned} \hat{\rho}_{GLS} &= \{(\Delta y_{-1})' W \Omega^{-1} W' (\Delta y_{-1})\}^{-1} \{(\Delta y_{-1})' W \Omega^{-1} W' \Delta y\} \\ &= \{(\Delta y_{-1})' W [W' (I_N \otimes G) W]^{-1} W' (\Delta y_{-1})\}^{-1} \\ &\quad \times \{(\Delta y_{-1})' W [W' (I_N \otimes G) W]^{-1} W' \Delta y\} \end{aligned}$$

$\hat{\rho}_{GLS}$ non è però ottimale. Lo stimatore ottimale di ρ lo si ottiene nel seguente modo. Dai residui della stima “one-step” $\Delta \hat{\varepsilon}_{GLS}$ è possibile ottenere una stima della matrice delle varianze-covarianze:

$$V_{N, GLS} = \sum_{i=0}^N W_i' (\Delta \hat{\varepsilon}_{GLS, i}) (\Delta \hat{\varepsilon}_{GLS, i})' W_i$$

da utilizzare al posto di Ω per ottenere lo stimatore “two-step” che realizza lo stimatore GMM di Hansen (1982):

$$\hat{\rho}_{GMM} = \{(\Delta y_{-1})' W [V_{N, GLS}]^{-1} W' (\Delta y_{-1})\}^{-1} \{(\Delta y_{-1})' W [V_{N, GLS}]^{-1} W' \Delta y\}$$

I due stimatori $\hat{\rho}_{GLS}$ e $\hat{\rho}_{GMM}$ sono asintoticamente equivalenti se vale la condizione: $\varepsilon_{it} \sim \text{IID}(0, \sigma_i^2)$.

La varianza asintotica è:

$$a \text{Var}(\hat{\rho}_{GMM}) = \{(\Delta y_{-1})' W [V_{N, GMM}]^{-1} W' (\Delta y_{-1})\}^{-1}$$

Lo stimatore GMM non richiede conoscenze a priori su: condizioni iniziali e distribuzione di ε_{it} e μ_i

In uno studio che confronta la performance di alternativi stimatori per modelli panel, Judson-Owen (1999) trovano che:

- quando T è grande, il calcolo della stima GMM è di difficile attuazione;
- lo stimatore GMM “one step” funziona meglio del “two step”;

- riducendo arbitrariamente la dimensione (il numero di ritardi) della matrice degli strumenti W non peggiora sostanzialmente la performance dei GMM.

Infine, Judson-Owen effettuano un confronto, per panel dinamici, degli stimatori di Anderson-Hsiao (AH), di Arellano-Bond (“one-step”, GMM1) e i LSDV la cui distorsione è corretta secondo le indicazioni di Kiviet, 1995 (LSDVC). I loro risultati sono sintetizzati in Tabella 4.

Tab. 4 – I metodi raccomandati dall’esperimento

panel:	$T \leq 10$	$T = 20$	$T = 30$
bilanciato	LSDVC	LSDVC	LSDVC
non bilanciato	GMM1	GMM o AH	LSDV

→ suggerimento: non c’è “il metodo” che funziona al meglio in qualsiasi situazione

[D.2.3] L’approccio GMM-sys per panel dinamici: Blundell-Bond

L’ottimalità della trasformazione in differenze strumentata con ritardi nei livelli viene però meno se l’esplicativa del modello è una variabile integrata. Infatti, in generale, se l’esplicativa x_{it} è $I(1)$, il coefficiente di correlazione fra Δx_{it} e x_{it-1} è quasi nullo perché le differenze prime di una variabile integrata, per definizione, *non possono* essere correlate con i suoi livelli (intuizione: si pensi alla specificazione del modello univariato per effettuare il test di Dickey-Fuller); mentre Δx_{it} e Δx_{it-1} *possono* essere correlate, a meno che x_{it} sia random walk. Infine, si sottolinea che il coefficiente di correlazione fra x_{it} e x_{it-1} sarà sempre quasi unitario per qualsiasi x_{it} integrato.

Per questo motivo, Arellano-Bover (1995) e Blundell-Bond (1998) propongono un approccio GMM-sys in cui, accanto alle equazioni originali in differenze strumentate con livelli (tipiche dell’approccio GMM-diff) aggiungono anche equazioni in livelli strumentate con le differenze. In altri termini, all’interno del sistema “diff”&”lev” sono strumentate con ritardi di “lev”&”diff”.

Pro: nel complesso, GMM-sys si adatta a qualsiasi tipologia statistica
 contro: strumenti deboli, poca identificazione, distorsione in campioni finiti

[E] Non stazionarietà e dinamica nei panel

Negli anni recenti, lo studio dell’interrelazione fra test di radici unitarie e dati panel si è sviluppata in modo considerevole collegando:

- (i) l’area di studi della integrazione e della cointegrazione per serie storiche;
- (ii) l’econometria dei panel.

Scopo dichiarato: combinare l’informazione proveniente dalle serie storiche con quella dalle cross-section, nella speranza che l’inferenza su radici unitarie e cointegrazione possa essere più precisa, grazie alla dimensione individuale.

Gli sforzi lungo questa direzione sono di solito giustificati dalla bassa potenza dei test per serie storiche. Tutto ciò accade specialmente quando le serie storiche sono brevi, ma c’è la possibilità di raccogliere informazioni lungo la dimensione cross-section (paesi, regioni, imprese e settori).

Un altro filone di ricerca (dinamica ed eterogeneità) prende le mosse da Pesaran-Shin (1995) che dimostrano l’inconsistenza degli stimatori di panel eterogenei dinamici con dati aggregati oppure pooled e raccomandano l’uso di stimatori “group-mean”.

[E.1] Test di integrazione

Fin dall'apparizione dei papers di **Levin-Lin** (1992, 1993), d'ora in poi LL, i test panel di radici unitarie sono diventati molto popolari presso gli economisti applicati.

Oramai è dato per acquisito che i test di Dickey-Fuller (DF) e ADF sono poco potenti nel discriminare fra processi integrati ed alternative stazionarie, e che l'uso di dati panel è un modo per incrementare la potenza dei test di radici unitarie per serie storiche.

La potenza del test non è però tutto: si tenga presente che le ipotesi H_0 e H_1 differiscono nei due tipi di test:

ipotesi:	serie storiche	Panel, LL
H_0	$\pi=0$	$\pi_i=0 \forall i$
H_1	$\pi<0$	$\pi_i<0 \forall i$

→ i due test (serie storiche e panel) non rispondono alla stessa domanda e, perciò, non è detto che siano interscambiabili per decidere sulla questione di interesse.

Il primo modello ad affrontare in modo articolato il tema delle radici unitarie nei panel è di LL:

$$(12) \quad \Delta y_{it} = a_i + b_i t + \tau_t + \pi y_{it-1} + \varepsilon_{it}$$

in cui si dà conto dell'esistenza di effetti fissi (a_i); effetti temporali specifici (b_i) e generici (τ_t). Inizialmente LL suppongono che $\varepsilon_{it} \sim \text{IID}(0, \sigma_i^2)$.

A partire dalla (12), LL ricavano un vasto numero di sottomodelli, il cui tratto comune è che sono tutti stimati OLS pooled.

La presenza di effetti fissi ha un ruolo importante nello spiegare l'eterogeneità, visto che il coefficiente π è vincolato ad essere lo stesso per tutti gli individui.

Il test LL verifica: $H_0: \pi = 0$ contro $H_1: \pi < 0$; con ulteriori ipotesi sotto H_0 e H_1 riferite ai coefficienti (a_i b_i τ_t) delle componenti deterministiche (ad esempio, nel modello dove $b_i=\tau_t=0$, si verifica $H_0: \pi=0$ e $a_i=0 \forall i$; contro $H_1: \pi<0$ e a_i non vincolati).

Alcune estensioni in Levin-Lin (1993) permettono una struttura più generale dell'errore (eteroschedasticità e autocorrelzione), ma mantengono l'ipotesi di indipendenza cross-section fra gli individui (ipotesi che talvolta crea problemi ai test, come vedremo).

Il limite principale del test LL è che il parametro π è lo stesso per tutte le osservazioni. Ciò che convince meno sono le sue implicazioni sotto l'alternativa.

Ad esempio, per verificare col test LL l'ipotesi di convergenza della crescita di N paesi, H_0 prevede che nessun paese converga ($\pi=0$ per tutti i paesi), mentre sotto H_1 tutti i paesi convergono alla stessa velocità.

I lavori di LL rivestono una fondamentale importanza perché introducono i principali temi dibattuti nella letteratura sulle radici unitarie, cioè:

- dimostrano che la normalità asintotica degli stimatori è legata ad opportune trasformazioni e correzioni;
- sottolineano la necessità di considerare casi alternativi di tendenza all'infinito di T e N;
- sollevano il tema "omogeneità vs eterogeneità" sotto le ipotesi nulla e alternativa (i loro modelli sono fra i più restrittivi perché richiedono che π sia lo stesso per tutti gli individui del campione);
- fondano le loro analisi nell'ipotesi di indipendenza fra gli individui;

(e) introducono modifiche dei test più elementari per dare conto di residui non sferici ed endogenità dei regressori.

I punti (a)-(b) costituiscono una importante base di partenza per nuovi studi. Gli altri tre temi sono stati ripresi da alcuni autori che hanno proposto importanti estensioni.

Il test di **Im-Pesaran-Shin** (1997) (IPS) è una estensione delle ipotesi relative ai casi (c)-(d) di LL, molto restrittivi. Il punto (e) è importante nel contesto dei test di cointegrazione per panel (si veda la [E.2]).

Il test di IPS estende il test LL per dare conto di eterogeneità dei coefficienti π_i :

$$\Delta y_{it} = a_i + \pi_i y_{it-1} + \varepsilon_{it}$$

Si effettuano N test di significatività di π di tipo DF-ADF individuo per individuo ($t_{i,T}$). Mediante esperimenti di Monte Carlo per panel bilanciati, IPS tabulano i momenti primo $E(t_{i,T})$ e secondo $V(t_{i,T})$ delle statistiche $t_{i,T}$, e dimostrano che:

$$\sqrt{N} \frac{\bar{t}_{N,T} - E(t_{i,T})}{\sqrt{V(t_{i,T})}} \Rightarrow \text{Normal}(0, 1) \text{ con: } \bar{t}_{N,T} = \frac{1}{N} \sum_{i=1}^N t_{i,T}$$

In pratica, il test IPS rappresenta un modo di combinare l'evidenza (eterogenea) di N test di radici unitarie calcolate per N individui. I momenti primo e secondo di $\bar{t}_{i,T}$ sono tabulati ad hoc.

Sotto l'ipotesi nulla, tutte le N serie storiche individuali hanno una radice unitaria:

$$\pi_i = 0 \quad \forall i$$

mentre, sotto l'ipotesi alternativa, gli individui sono raggruppati in due insiemi complementari: N_S di processi non stazionari e N_{1-S} di processi stazionari:

$$\begin{aligned} \pi_i &= 0 & \forall i \in N_S \\ \pi_i &< 0 & \forall i \in N_{1-S} \end{aligned}$$

tali che: $\lim_{N \rightarrow \infty} (N_{1-S} / N) = \delta$, con $0 < \delta \leq 1$

sulla base della quale si possono riformulate le ipotesi:

$$H_0: \delta = 0$$

$$H_1: \delta > 0$$

“The main failing of panel unit root tests is that the null hypothesis (usually that *all* the series in the panel are realisations of unit root processes) will be violated even if a very small number of series in the panel are stationary, and rejection cannot help us calculate the entity of that number.” (Taylor and Sarno, 1998, p. 283).

Il problema dell'eterogeneità dei panel (Cfr Pesaran-Smith, 1995) è forse risolto nella fase di stima, ma non nelle fasi di inferenza.

Alcuni fenomeni (ad esempio le analisi cross-country), caratterizzati da eterogeneità, introducono una asimmetria fra H_0 e H_1 non presente nei modelli per serie storiche: la stessa H_0 è imposta a tutte le i , mentre H_1 può variare lungo la i .

In panel eterogenei con N elevato e $T \approx 15$ è possibile formulare test informativi in “senso medio”: per inferire se H_0 può essere rifiutata per una significativa frazione di individui. Una misura (stima) per δ richiede invece T elevato, ma in tal caso l'esigenza di pooling sarebbe meno sentita.

IPS dimostrano, con un esperimento di Monte Carlo, la migliore performance in campioni finiti del loro test rispetto a quello di LL.

[E.2] Test di cointegrazione

Un filone di analisi di cointegrazione per dati panel è stato sviluppato mediante test che, sotto l'ipotesi nulla, hanno l'assenza di cointegrazione. Tali test sono l'analogo panel del test di Engle-Granger (1987) per serie storiche, perché usano i residui di regressioni statiche per costruire test di radici unitarie le cui distribuzioni sono tabulate ad hoc.

Il test di cointegrazione di **Pedroni** (1999) si fonda su una regressione statica:

$$(13) \quad y_{it} = a_i + b_i x_{it} + c_i t + \varepsilon_{it}$$

che ammette una considerevole eterogeneità nel panel: pendenze eterogenee, effetti fissi e trend individuali sono tutti possibili. Le variabili esplicative possono essere più di una e non serve ipotizzare la loro esogenità. Infine, si impone indipendenza cross-section, a meno di comuni disturbi aggregati.

L'approccio di Pedroni prevede sette modelli alternativi a partire dalla (13) che possono essere classificati in due categorie: (i) quattro modelli sono sviluppati nell'ambito del pooling lungo la dimensione within (i primi tre sono di tipo Phillips-Perron, PP, il quarto di tipo ADF), (ii) tre lungo la dimensione between (due PP e uno ADF).

Se definiamo con ρ_i il coefficiente autoregressivo dei residui OLS dell'i-esimo individuo, la categoria (i) di test, definiti "panel statistic" verifica le ipotesi:

$$H_0: \rho_i = 1 \quad \forall i \quad ; \quad H_1: \rho_i = \rho < 1 \quad \forall i$$

la categoria (ii) di test "group statistic" verifica:

$$H_0: \rho_i = 1 \quad \forall i \quad ; \quad H_1: \rho_i < 1 \quad \forall i$$

L'analogia con i test LL nel caso (i) e IPS nel caso (ii) è evidente, in termini di eterogeneità consentita sotto l'ipotesi alternativa.

Concentriamo l'attenzione sulla procedura di calcolo della "group t-statistic", la terza della categoria (ii), considerata una estensione del test di IPS:

passo 1: stima OLS della (13) per ogni individuo $\rightarrow \hat{\varepsilon}_{OLS,it}$

passo 2: calcolo il test individuale di tipo ADF usando $\hat{\varepsilon}_{OLS,it}$

passo 3: media semplice delle statistiche ADF per individuo

Calcolata una delle sette statistiche test $P_{N,T}$, bisogna standardizzarla mediante: $(P_{N,T} - \mu N^{1/2})/\nu^{1/2} \Rightarrow N(0,1)$

I momenti primo μ e secondo ν di $P_{N,T}$, ottenuti mediante simulazione, sono tabulati ad hoc e variano: (a) per statistica, (b) per numero di esplicative, (c) per nucleo deterministico.

Nota: l'esito dei test di integrazione interagisce con le proprietà di cointegrazione fra individui nel panel (Cfr Banerjee-Marcellino, 2000): la violazione dell'ipotesi di incorrelazione cross-section (cointegrazione fra alcuni i e j) implica size empirici molto maggiori di quelli nominali.

\rightarrow quando vera, la H_0 di radici unitarie è rifiutata troppo spesso.

[E.3] Dinamica, cointegrazione e poolability

Quando la dimensione di T è elevata e dello stesso ordine di grandezza di N, può essere utile concentrare l'attenzione sui legami di lungo periodo e sulla velocità di aggiustamento a tali relazioni.

I parametri dei modelli per dati panel di questo tipo possono essere analizzati mediante 2 procedure (ai due estremi):

- Averaging: stima di equazioni separate per individuo (group) ed esame della distribuzione dei coefficienti fra individui. Lo stimatore “mean group” (MGE) è una stima consistente della media dei parametri (eterogenei).
- Pooling: utilizzo di un modello panel (con effetti fissi o random, FE o RE), dove le pendenze e le varianze dell’errore sono ipotizzate uguali.

Il lavoro di **Pesaran-Shin-Smith** (1999, JASA) considera uno stadio intermedio fra i due precedenti: gli stimatori “pooled mean group” (PMGE) che, come vedremo, implicano averaging e pooling.

Lo stimatore PMGE si applica a modelli che prevedono diverse (da individuo a individuo) intercette, parametri di breve periodo e varianza dell’errore, ma che vincolano i coefficienti di lungo periodo ad essere gli stessi.

Nota: la non imposizione di vincoli sui parametri di breve permette diverse specificazioni dinamiche per individuo.

I vantaggi principali dell’approccio di Pesaran-Shin-Smith sono:

- stima legami comuni di lungo periodo senza imporre implausibili dinamiche comuni;
- si applica al caso di regressori sia stazionari sia non stazionari;
- permette di discriminare fra stimatori MGE e PMGE.

Esempio: Lo studio della domanda di moneta nell’area dell’euro (Golinelli-Pastorello, 2002, EJF). Variabili di interesse m (moneta), y (PIL), r (tasso di interesse a lungo).

Approccio: stima di dati aggregati per l’area e per paese.

L’approccio ARDL(p,p) a livello di area (Pesaran and Shin, 1998):

$$\Delta m_t = \mu + \phi m_{t-1} + \pi x_t + \sum_{i=1}^{p-1} \lambda_i^* \Delta m_{t-i} + \sum_{i=1}^{p-1} \psi_i^* \Delta x_{t-i} + \varepsilon_t$$

dove: μ , ϕ , π , λ e ψ sono parametri; $x_t = (y_t, r_t)'$ e ε_t è l’errore. Sotto ipotesi standard per ε_t e se $\phi < 0 \rightarrow$ esiste una relazione stabile di lungo periodo fra m_t e x_t , e la si definisce con:

$$m_t = \theta' x_t + \eta_t$$

dove $\theta = -\pi/\phi$ e $\eta_t \sim I(0)$.

Our ARDL(2,2) model for m , y and r can be simplified in the final model for the Euro area:

$$\Delta m_t = -0.16 + 0.52 \Delta m_{t-1} - 0.31 \Delta r_{t-1} - 0.11 (m - 1.37 y + 0.68 r)_{t-1}$$

(0.04) (0.08) (0.10) (0.03) (0.06) (0.37)

Main characteristics of the model:

- very parsimonious model
- data congruent
- stable parameter estimates
- forecasting ability 98q1-99q3: p-value 20.1%

Le stesse funzioni di domanda di moneta possono essere studiate per paese (**MGE approach**), specificando un modello ARDL per ogni paese (h) con m variabile dipendente e y e r variabili esplicative.

The results are in Golinelli-Pastorello (2002, Table 6, p. 390). The comparison between the results from aggregate and disaggregated models shows the better properties of the first. The main findings are:

- the regression standard error of the national models is at least double w.r.t the area model;
- the diagnostic tests for national models signal some specification problems;
- a long run relationship is hardly ever detected;
- almost always the estimates of long run elasticity of income are bigger than 1, while the elasticity to the long rate are much more dispersed (they spread from -4 to 5).

This worsening in performance when shifting from the area to the national model can depend on two causes:

- specification problems at country level (our national models are too simplified);
- the disturbing effect of shocks and outliers on the estimation of national money demands (sampling error).

→ Nel secondo dei due casi, è possibile effettuare il “pool” dei parametri di lungo periodo di individui diversi, allo scopo di migliorare la performance del modello.

Seguendo l’approccio di Pesaran, Shin and Smith, è possibile verificare la poolability dei parametri di lungo periodo (long run parameter homogeneity test),

$H_0: \theta_h = \theta \quad \forall h$ ipotesi di poolability

Le stime θ sono ottenute:

- calcolando le medie semplici per paese (MGE, Mean Group Estimator)
- imponendo il vincolo di pooling (PMGE, Pooled Mean Group Estimate).

Sotto la nulla di poolability, entrambi gli stimatori sono consistenti, ma PMGE è anche efficiente.

L’ipotesi nulla può essere verificata mediante due approcci:

- una statistica-test di tipo Hausman (in piccoli campioni non è detto che la differenza fra varianze sia positiva definita);
- un test del rapporto di verosimiglianza, di confronto fra la log-likelihood vincolata (sotto l’ipotesi di pooling) e non vincolata (la somma delle log-likelihood nei modelli per individuo); in piccoli campioni questo test tende a rifiutare in eccesso la nulla.

We made pooled estimates for 4 groups of countries:

1. Euro area without Luxembourg;
2. Core area (Austria, Belgium, France, Germany, Netherlands);
3. Core plus Italy;
4. Core plus Spain.

The results can be summarised as follows:

1. Euro area without Luxembourg:

- long run parameter poolability Hausman test is not rejected; the likelihood ratio test is rejected;
- previous results can be due to the huge dispersion of the unrestricted estimates;
- the fitting of the pooled model is only slightly worse than the unrestricted, and we have no worsening in diagnostic tests;

2. Core area

- long run parameter poolability Hausman test can not be computed; the likelihood ratio test is not 5% rejected;
 - the fitting of the pooled model is broadly the same as the unrestricted;
3. and 4. Core + Italy, and Core + Spain
- the purpose is not to verify the possibility of coexistence of Italy and Spain in a monetary union with the countries of the “core”;
 - the Hausman test can not be computed; the likelihood ratio test 5% rejects the poolability in the core + Italy subset (but does not reject at 1%) and does not 5% reject in the core + Spain subset.

Evidentemente, l’approccio PMGE in parte conferma gli esiti aggregati (confronto solo qualitativo) e in parte fornisce indicazioni sul grado di omogeneità dei singoli paesi/individui (confronto mediante test).

General references

- Arellano, Manuel, *Panel Data Econometrics*, Oxford University Press, 2003.
 Baltagi, Badi H., *Econometric Analysis of Panel Data*, John Wiley & Sons, 2005.
 Hsiao, Cheng, *Analysis of Panel Data*, Cambridge University Press, 2003.
 Manera, Matteo and Galeotti Marzio, *Microeconometria. Metodi ed applicazioni*, Carocci Editore, 2005.
 Sevestre Patrick, *Économétrie des Données de Panel*, Dunod, Paris, 2002.
 Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 2002.

Relevant review-articles:

- Handbook of Econometrics:
 Capitolo 22, *Panel Data*, di Gary Chamberlin.
 Capitolo 29, *Econometric Analysis of Longitudinal Data*, di Heckman e Singer.

Interesting:

- Matyas, L. and P. Sevestre (eds.), *The Econometrics of Panel Data: A Handbook of Theory and Applications*, Kluwer Academic Publishers, 1996.
 Maddala, G.S. (ed.), *The Econometrics of Panel Data*, Vol. I e II, Edward Elgar Publishing, Cheltenham, 1993.
 Baltagi, Badi H. and Baldev Raj (eds.), *Panel Data Analysis*, Physica-Verlag, Heidelberg, 1992.
 Dielman, T.E., *Pooled Cross-Sectional and Time-Series Data Analysis*, Marcel Dekker, 1989.

Good chapters on panel:

- Greene, *Econometric Analysis*, Prentice Hall, 1997, ch. 14.
 Wooldridge, Jeffrey M., *Introductory Econometrics. A modern approach*, Thomson South-Western, 2003, ch. 13, 14.
 Hayashi, Fumio, *Econometrics*, Princeton University Press, 2000, ch. 5.
 Ruud, Paul A., *An Introduction to Classical Econometric Theory*, Oxford University Press, 2000, ch. 24.

Basic:

- Judge, Griffiths, Hill, et. al., *The Theory and Practice of Econometrics*, Wiley, 1985.
 Johnston, Di Nardo, *Econometric Methods*, McGraw-Hill, 1998.

Davidson, Russell and MacKinnon James G., *Estimation and Inference in Econometrics*, Oxford University Press, 1992.

Applications:

Baltagi-Raj (1992): Parte III, e Matyas-Sevestre (1996): Parte III.

Two special issues of the **Journal of Econometrics**: Carraro, Peracchi e Weber (ed.), 1993, Vol. 59, pp. 1-211; and Baltagi (ed.), 1995, Vol. 68, pp. 1-268.