



UNIVERSITÁ DEGLI STUDI DI PADOVA

Dipartimento di scienze economiche ed aziendali “Marco Fanno”

Corso di laurea in economia

PROVA FINALE

Analisi Della Serie Storica Degli Investimenti

Relatore:

Ch.mo. Prof. Nunzio Cappuccio

Laureanda: **Laura Gantz**

Matricola No. **1135997**

ANNO ACCADEMICO 2018 – 2019

Il nostro obiettivo è quello di studiare e analizzare la serie storica degli investimenti fissi lordi. La principale finalità è quella di studiarne le caratteristiche per stimare il modello che più si adatti ai nostri dati, così da poter fare inferenza sulla serie stessa. *Inferenza* significa poter riuscire a prevedere, in modo abbastanza accurato, i valori futuri di questa nostra variabile, che rappresenta una componente importante dell'economia di un paese, e una delle più mutevoli. Inoltre, è importante verificare se la serie presenta *radice unitaria* (ovvero non stazionarietà), la quale è una caratteristica delle serie macroeconomiche, che indica un comportamento deterministico della serie, con conseguenze economiche importanti.

La prima parte dell'elaborato introduce le nozioni statistiche alla base dell'analisi delle serie storiche, come i processi stocastici e stazionari, fino ad arrivare ad illustrare l'operatore ritardo L , che viene definito con riferimento alla generica serie storica. L'operatore ritardo viene successivamente usato per definire i processi stocastici *MA* (Moving Average), *AR* (AutoRegressive) e *ARMA* (la loro generalizzazione). La seconda parte riguarda l'analisi dei dati con la definizione di un modello statistico adatto, analizzando i grafici e applicando dei test di stazionarietà, chiamati test di radice unitaria.

La parte terza concerne l'interpretazione e la conclusione rispetto ai dati e alle conseguenze macroeconomiche.

Lista delle figure

<i>Figura 2.1: Grafico della serie storica annuale dei logaritmi degli investimenti</i>	14
<i>Figura 2.2: Grafico della serie storica trimestrale dei logaritmi degli investimenti</i>	14
<i>Figura 2.3: Correlogramma della serie storica annuale dei logaritmi degli investimenti</i>	15
<i>Figura 2.4: Correlogramma della serie storica trimestrale dei logaritmi degli investimenti</i>	15
<i>Figura 2.5: Grafico del tasso di crescita annuale degli investimenti</i>	16
<i>Figura 2.6: Grafico del tasso di crescita annuale degli investimenti, a cadenza trimestrale</i>	16
<i>Figura 2.7: Correlogramma del tasso di crescita annuale degli investimenti</i>	17
<i>Figura 2.8: Correlogramma del tasso di crescita annuale degli investimenti, a cadenza trimestrale</i>	17
<i>Figura 2.9: Grafico della serie storica trimestrale dei logaritmi degli investimenti confrontato con retta dei minimi quadrati</i>	18
<i>Figura 2.10: Grafico del tasso di crescita annuale, a cadenza trimestrale degli investimenti confrontato con retta dei minimi quadrati</i>	18
<i>Figura 2.11: Grafico dell'andamento del valore della statistica test, riferito al test di Zivot-Andrews</i>	30
<i>Figura 2.12: Grafico dei residui della regressione rispetto al tempo</i>	36
<i>Figura 2.13: Correlogramma dei residui della regressione</i>	36

Lista delle tabelle

<i>Tabella 2.1: Risultati dei test di radice unitaria su campioni trimestrali</i>	27
<i>Tabella 2.2: Risultati del test Zivot-Andrews su campioni trimestrali</i>	30
<i>Tabella 2.3: Risultati dei test di radice unitaria su campioni annuali</i>	31

Indice

Introduzione	1
<i>1.1 Processi stocastici</i>	2
<i>1.2 La stazionarietà</i>	4
<i>1.3 Processi ergodici</i>	4
<i>1.4 L'operatore ritardo</i>	5
<i>1.5 Processi MA</i>	6
<i>1.6 Processi AR</i>	7
<i>1.7 Processi ARMA</i>	8
Analisi della serie storica degli investimenti fissi lordi italiani	10
<i>2.1 Dati</i>	10
<i>2.2 Processi a radice unitaria</i>	20
<i>2.3 Risultati</i>	26
Conclusioni	34
Appendice A	I
Appendice B	IV
Appendice C	VI
Bibliografia	VIII

Introduzione

Un *processo stocastico* è una famiglia di variabili aleatorie ordinate rispetto a un parametro. Una *serie storica* è un insieme finito di osservazioni relative ad un certo fenomeno, ordinate nel tempo, e può essere vista, quindi, come una realizzazione finita di un processo stocastico. La cadenza temporale delle osservazioni può fare riferimento ad una successione di intervalli temporali o ad istanti temporali: nel primo caso si parla di *serie di stato o posizionali*, nel secondo di *serie di flusso*.

Una serie storica si può intendere composta di una componente sistematica ed una accidentale. Secondo l'approccio più tradizionale, denominato *approccio classico*, nella componente sistematica si possono rilevare trend, ciclo e stagionalità¹:

- Il *trend* è il comportamento di lungo periodo e rappresenta una componente generale sistematica lineare o non lineare, che mette in evidenza una evoluzione strutturale del fenomeno dovuta a cause che agiscono in modo sistematico sullo stesso.
- Il *ciclo* è originato dal presentarsi di condizioni più o meno favorevoli, di espansione o contrazione, del contesto economico nel quale si colloca il fenomeno in esame.
- La *stagionalità* è la componente che si ripete ad intervalli sistematici nel tempo; essa determina cicli di periodo inferiore ad un anno, se si tratta di serie economiche.

La *componente casuale* o componente di disturbo comprende le variazioni residue non spiegate dalle precedenti componenti.

Lo studio riferito alle serie storiche si è poi evoluto, progredendo verso un *approccio moderno*. Secondo questa analisi, l'obiettivo è quello di usare la serie storica osservata per individuare il processo stocastico stazionario che si adatti maggiormente ai dati reali, da utilizzare per calcolare previsioni future della serie stessa. I processi stocastici stazionari godono di particolari

¹ Warren Milton Persons, *Indices of General Business Conditions*. Harvard Univ. Committee on Economics research, 1919.

proprietà probabilistiche che risultano utili per modellare molte serie storiche che si incontrano nella realtà.

I due approcci, pur ponendo quesiti diversi, sono matematicamente equivalenti, come dimostrato da *Wiener e Khintchine*², essendo semplicemente metafore diverse per concepire e trattare i dati.

1.1 Processi stocastici

Un *processo stocastico*³ è una famiglia y_t , $t \in T$ di variabili aleatorie definite su uno stesso spazio di probabilità (Ω, F, P) , dove Ω indica uno spazio campionario, F è una σ -algebra di sottoinsiemi di Ω e P è una misura di probabilità. L'insieme T è detto *insieme dei tempi*. Un processo stocastico è, quindi, un'applicazione che associa un numero reale a una coppia (ω, t) costituita, rispettivamente, da una eventualità e da un istante temporale.

Dato il processo stocastico $y(\omega, t): t \in T$, per ogni $\omega \in \Omega$, con $S \equiv (\Omega, F)$, la mappa

$$y(., t): t \rightarrow S$$

è detta *funzione campionaria* o realizzazione.

Una serie storica è un insieme definito di osservazioni ordinate nel tempo e solitamente misurate ad intervalli di tempo equi spazati (mensile, trimestrale, annuale): è, quindi, definita come il prodotto della funzione campionaria.

Una serie storica $y_t : t \in T_0$ risulta, dunque, essere una successione finita di valori, ognuno dei quali è la realizzazione di una variabile aleatoria che fa parte di un processo stocastico $y_t : t \in T$, con $T_0 \subseteq T$.

Ci chiediamo, allora, dato un processo stocastico y_t , quali grandezze numeriche ad esso associate siano di interesse per la risoluzione di problemi concreti. Da un numero ragionevole di dati sperimentali si possono ottenere buone approssimazioni di certi valori medi, che serviranno per lo studio di questi processi.

Introduciamo, allora, la funzione *valore medio*

$$\mu_t = E[y_t],$$

la *funzione varianza*

² Gottman, J.M., *Time series analysis. A comprehensive introduction for social scientists*, Cambridge 1981

³ Brockwell Peter and Davis Richard. *Time Series: Theory and Methods, Second Edition*, Springer-Verlag 1991

$$\sigma^2_t = E [(y_t - E(y_t))^2],$$

E la *funzione di autocovarianza*

$$Cov_{t,s} = E [(y_t - \mu_t)(y_s - \mu_s)].$$

Se il processo ha una distribuzione normale multivariata, allora i primi due momenti (media e varianza) definiscono completamente la funzione di ripartizione del processo.

Essendo la covarianza una funzione che dipende dall'unità di misura e dalla grandezza dei dati stessi, si usa introdurre il *coefficiente di correlazione*, il quale è adimensionale

$$R_{t,s} = \rho_{t,s} = \frac{cov(x,y)}{\sqrt{var(x)var(y)}}$$

Per ogni s , ρ_s rappresenta il coefficiente di correlazione tra due variabili dello stesso processo, le quali risultano separate da un intervallo di tempo di lunghezza h .

A livello di tali processi si usa introdurre la funzione $\rho_{t,s} = \rho(y_t, y_s)$. L'autocorrelazione esprime quindi la *dipendenza lineare* che esiste tra il processo al tempo t e sé stesso al tempo $t + h$.

Per confrontare tra loro funzioni di autocorrelazione di diversi processi, si introduce la normalizzazione

$$\frac{R_t}{R_0}$$

Per $t = 0$, vale sempre 1. R_0 è la varianza del processo.

Purtroppo, non è assicurato che si possano ottenere le funzioni media e varianza al variare di t , poiché in molti casi concreti si conosce solo una singola realizzazione del processo stocastico in esame, e per calcolare suddette funzioni avremmo bisogno di un certo numero di valori sperimentali, non solo di uno.

È utile, quindi, cercare di ricondurre questa serie storica a un processo con determinate caratteristiche indipendenti dal tempo, così da poterla studiare e analizzare.

1.2 La stazionarietà

L'analisi delle serie storiche ha come obiettivo principale quello di individuare un opportuno processo stocastico che abbia traiettorie che si adattino ai dati, per poter poi formulare previsioni. Per fare questo, risulta necessario restringere la nostra attenzione ad una classe di processi stocastici che ci consenta di:

1. Identificare univocamente il processo, ossia il modello;
2. Fare inferenza sui momenti del processo stesso, cioè di ottenere delle stime corrette e consistenti dei momenti del processo.

Per queste ragioni diventa determinante il concetto di processo stocastico stazionario.

Un processo stocastico $y_t, t \in Z$ è detto stazionario se soddisfa le seguenti condizioni:

- Stazionarietà della media

$$E(y_t) = \mu_t = \mu \text{ per } t = 1, \dots, T$$

- Stazionarietà della varianza

$$Var(y_t) = E \left[(y_t - \mu_t)^2 \right] = \sigma_t^2 = \sigma^2$$

- Stazionarietà della covarianza

$$Cov(y_t, y_{t+s}) = E[(y_t - \mu_t)(y_{t+s} - \mu_{t+s})] = \gamma(s)$$

ossia se ha momenti finiti e costanti e il valore dell'autocovarianza dipende solo dall'intervallo di tempo e non dallo specifico istante temporale.

Un processo è detto debolmente stazionario se presenta stazionarietà nella media e nella covarianza. La quantità s viene chiamata *lag* o passo.

Ponendo $s = 0$ nella funzione della covarianza, si ha

$$\gamma(0) = Var(X_t) = E[(X_t - \mu_t)(X_t - \mu_t)]$$

1.3 Processi ergodici

Affinché il calcolo dei momenti campionari risulti valido per determinare i veri momenti del processo asintoticamente è necessario applicare l'*ipotesi di ergodicità*. L'idea di ergodicità ci dice che da ogni traiettoria è possibile calcolare qualsiasi proprietà del processo.

Se un processo stocastico stazionario y_t è ergodico ed $x_i, i = 1, 2, \dots, t$ è una sua realizzazione, valgono le seguenti equazioni:

$$\hat{\mu} = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^t x_i$$

$$\hat{\gamma}_0 = \sigma^2 = R_0 = \lim_{t \rightarrow \infty} \sum_{i=1}^t (x_i - \hat{\mu})^2$$

$$\hat{\gamma}_\tau = R_\tau = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^{t-\tau} (x_i - \hat{\mu}) (x_{i+\tau} - \hat{\mu}) \text{ per } \tau = 1, \dots, t-1.$$

Dove $\hat{\mu}$, $\hat{\gamma}_0$ e $\hat{\gamma}_\tau$ sono gli stimatori consistenti rispettivamente della media, della varianza e dell'autocovarianza.

Qualora il processo non dimostri un comportamento stazionario, questo tipo di inferenza “standard” non è più applicabile con correttezza.

Se si vogliono applicare a delle serie storiche tecniche di previsioni valide per processi stazionari, servono dei metodi utili per ricondursi a una serie stazionaria, e qui entrano in gioco i processi *AR*, *MA* e *ARMA*⁴.

1.4 L'operatore ritardo

L'operatore ritardo, indicato generalmente con L (*Lag*) è un operatore che trasforma una serie storica x_t in un'altra y_t , tale che la nuova sequenza abbia gli stessi valori di x_t , ma differenti di un periodo.

In formule,

$$y_t = Lx_t = x_{t-1}.$$

Per convenzione si pone $L^0 = 1$.

L'operatore ritardo è lineare, ovvero *additivo e omogeneo* (di grado 1):

$$L(x_t + w_t) = Lx_t + Lw_t \quad L(ax_t) = aLx_t$$

L'operatore ritardo può essere applicato più volte, viene usata la notazione L^k e quindi si ha $L^k x_t = x_{t-k}$.

⁴ Peter J. Brockwell Richard A. Davis, Introduction to Time Series and Forecasting, Springer, 1996

1.5 Processi MA, a media mobile

Un processo *a media mobile di ordine q* è una successione di variabili aleatorie del tipo:

$$y_t = \theta \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} = \sum_{i=0}^q \theta_i \varepsilon_{t-i}$$

e si indica con $MA(q)$.

L'operatore ritardo consente una definizione più sintetica:

$$y_t = C(L)\varepsilon_t$$

dove $C(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$ è un polinomio di grado q nell'operatore ritardo, $\varepsilon_t \sim WN$, *white noise*.

Analizziamo ora i suoi momenti. Per quanto riguarda il valore atteso si ha

$$E(y_t) = E\left[\sum_{i=0}^q \theta_i \varepsilon_{t-i}\right] = \sum_{i=0}^q \theta_i E(\varepsilon_{t-i}) = 0$$

Quindi un processo MA ha media 0.

Per quanto riguarda la varianza, considerando che il primo momento è nullo,

$$V(y_t) = E(y_t^2) = E\left[\left(\sum_{i=0}^q \theta_i \varepsilon_{t-i}\right)^2\right]$$

Sviluppando il quadrato, possiamo scomporre:

$$\left(\sum_{i=0}^q \theta_i \varepsilon_{t-i}\right)^2 = \left(\sum_{i=0}^q \theta_i^2 \varepsilon_{t-i}^2\right) + \sum_{i=0}^q \sum_{j=i}^q \theta_i \theta_j \varepsilon_{t-i} \varepsilon_{t-j}$$

Dato che, dalla proprietà del *white noise*, il valore atteso della seconda sommatoria è nullo, diviene

$$E(y_t^2) = E\left[\sum_{i=0}^q \theta_i^2 \varepsilon_{t-i}^2\right] = \sum_{i=0}^q \theta_i^2 E(\varepsilon_{t-i}^2) = \sum_{i=0}^q \theta_i^2 \sigma^2 = \sigma^2 \sum_{i=0}^q \theta_i^2$$

Che ha valore finito quando q ha valore finito.

Infine, l'autocovarianza di ordine k è data da

$$E(y_t y_{t+k}) = E\left[\left(\sum_{i=0}^q \theta_i \varepsilon_{t-i}\right)\left(\sum_{j=0}^q \theta_j \varepsilon_{t-j+k}\right)\right] = \sum_{i=0}^q \theta_i \left(\sum_{j=0}^q \theta_j E(\varepsilon_{t-i} \varepsilon_{t-j+k})\right)$$

Considerando ancora le proprietà del white noise, si ha che $E(\varepsilon_{t-i}\varepsilon_{t-j+k}) = \sigma^2$ per $j = i + k$ e 0 negli altri casi, cosicché l'espressione diventa

$$\gamma_k = E(y_t y_{t+k}) = \sigma^2 \sum_{i=0}^q \theta_i \theta_{i+k}$$

Dove $\theta_i = 0$ per $i > q$.

Si noti che per $k > q$ le autocovarianze sono nulle.

Un processo $MA(q)$ è un processo, quindi, ottenuto come combinazione di diversi elementi di uno stesso *white noise*, che presenta delle caratteristiche di persistenza tanto più forti quanto più alto è il suo ordine.

Un processo a media mobile finita $MA(q)$, $q < \infty$, è stazionario ed ergodico.

1.6 Processi AR, autoregressivi

Un processo autoregressivo è una successione di variabili aleatorie ciascuna delle quali è funzione delle precedenti:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t$$

e si indica con $AR(p)$.

Il nome deriva dal fatto che un modello AR assomiglia molto a un modello di regressione in cui le variabili esplicative sono i valori passati della variabile dipendente.

Anche in questo caso l'operatore ritardo consente una definizione più sintetica:

$$A(L)y_t = \varepsilon_t$$

dove $A(L)$ è un polinomio di grado p nell'operatore ritardo, con $A(0)=1$.

Per derivare i momenti di questo processo consideriamo prima una serie $AR(1)$.

Per calcolare i suoi momenti possiamo procedere in modo intuitivo, ipotizzando la stazionarietà del processo.

Si suppone quindi che il processo abbia media costante μ , questo implica:

$$\mu = E(y_t) = \varphi E(y_{t-1}) + E(\varepsilon_t) = \varphi \mu$$

Questa espressione può essere vera in due casi: quando $\mu=0$, e quindi vera per qualunque valore di φ ; oppure nel caso $\varphi=1$ e la media è indeterminata. In questo secondo caso si dice che il processo presenta una radice unitaria perché vi è una soluzione dell'equazione per il valore $\varphi=1$.

Per quanto riguarda i secondi momenti, supponiamo come sopra che il *white noise* ε_t abbia varianza pari a σ^2 . Supponendo che la varianza $V(y_t)$ esista e sia costante nel tempo

$$V(y_t) = E(y_t^2) = E[(\varphi y_{t-1} + \varepsilon_t)^2] = \varphi^2 V + \sigma^2 + 2\varphi E(y_{t-1}\varepsilon_t)$$

L'ultimo elemento vale zero poiché è una combinazione lineare di autocovarianze di un white noise.

Se ne deduce che $V = \frac{\sigma^2}{1-\varphi^2}$

Secondo questa equazione si può parlare di varianza costante nel tempo solo se $|\varphi| < 1$ (per $|\varphi| > 1$ si parla di processi a radice esplosiva).

V è la varianza non condizionale di y_t , mentre σ^2 è la varianza condizionale $y_t | \zeta_{t-1}$.

Tanto maggiore è φ , ovvero tanto maggiore è persistente il processo, tanto più la varianza condizionale al proprio passato sarà minore della sua varianza non condizionale.

L'autocovarianza di ordine 0 è uguale alla varianza, l'autocovarianza di ordine 1 è

$$\gamma_1 = E(y_t, y_{t-1}) = E[(\varphi y_{t-1} + \varepsilon_t)y_{t-1}] = \varphi V$$

L'autocovarianza generica di ordine k è

$$\gamma_k = E(y_t, y_{t-k}) = E[(\varphi y_{t-1} + \varepsilon_t)y_{t-k}] = \varphi \gamma_{k-1}$$

Se ne deduce che $\gamma_k = \varphi^k \frac{\sigma^2}{1-\varphi^2}$

Le autocorrelazioni, che sono un indice di memoria del processo, sono tanto più grandi quanto più grande è φ , confermando l'interpretazione di questo parametro come caratteristica di persistenza.

1.7 Processi ARMA

La classe dei processi ARMA (autoregressivi a media mobile) comprende sia i processi AR che i processi MA come casi particolari. Un processo ARMA(p, q) è definito da:

$$y_t + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} = \theta \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

$$A(L)y_t = C(L)\varepsilon_t$$

dove p è l'ordine del polinomio $A(L)$ e q è l'ordine del polinomio $C(L)$.

Dato che qualsiasi processo $MA(q)$ è stazionario ed ergodico per q finito, le proprietà di un modello $ARMA(p, q)$ dipendono solo dalla sua componente autoregressiva.

Un modello $ARMA$ viene costruito facendo un'ipotesi a priori sui gradi dei polinomi $A(L)$ e $C(L)$, una volta poi stimati i coefficienti dei polinomi, si analizzano le autocorrelazioni campionarie della serie corrisponde ai residui ε_t^* . Se queste non sono troppo grandi, si possono approssimare i ε_t^* come white noise.

Analisi della serie storica degli investimenti fissi lordi italiani

2.1 Dati

Gli Investimenti Fissi Lordi, come definiti dal Sistema Europeo dei Conti⁵, sono costituiti dalle acquisizioni, al netto delle cessioni, di capitale fisso effettuate dai produttori residenti a cui si aggiungono gli incrementi di valore dei beni materiali non prodotti. Il capitale fisso consiste di beni materiali e immateriali prodotti destinati ad essere utilizzati nei processi produttivi per un periodo superiore ad un anno.

Gli investimenti fissi lordi comprendono valori sia positivi sia negativi: sono valori positivi l'acquisizione di beni nuovi e usati, sono valori negativi la cessione di beni usati.

Le serie trimestrali e annuali utilizzate per questo studio sono state realizzate dall'Istat⁶. La serie trimestrale ha come intervallo temporale il periodo dal 1991:1 al 2014:2. La serie annuale ha come intervallo temporale il periodo dal 1995 al 2018.

Per tutto l'elaborato vengono utilizzati i logaritmi dei dati stessi, l'analisi viene svolta, quindi, sui tassi di crescita della serie storica. L'anno base è stato fissato al 2010: il concatenamento fornisce una misura dell'aggregato economico di interesse in termini di volume, ossia al netto della dinamica dei prezzi ad esso sottostanti. Inoltre, questi dati sono dati destagionalizzati, ovvero sono al netto delle fluttuazioni stagionali. I dati destagionalizzati si intendono anche corretti per gli effetti di calendario, laddove questi risultano significativi. La correzione per gli effetti di calendario e la destagionalizzazione sono effettuate utilizzando la procedura Tramo-Seats, sviluppata da Gómez e Maravall⁷.

⁵ *Sistema europeo nazionale dei conti SEC*, Istat, Istituto nazionale di statistica, www.istat.it

⁶ *Conti e aggregati economici trimestrali e annuali*, Banca dati I.Stat, <http://dati.istat.it>

⁷ Gómez Victor e Maravall Agustin, *Estimation, Prediction and Interpolation for Nonstationary Series With the Kalman Filter*, 1996.

Come si vede dalle *figure 2.1 e 2.2* l'andamento di questa serie storica è soggetto a numerose variazioni.

Inoltre, come mostrato dalla *figura 2.9*, la serie presenta una maggiore variabilità nei tempi di crisi, dove ha raggiunto tassi di crescita negativi fino al 15%. Tra il 2007 e il 2015 gli investimenti in Italia sono scesi di ben 109,7 miliardi di euro, pari, in termini percentuali, a una diminuzione di 29,8 punti.

Le forti diminuzioni avvengono, quindi, in concomitanza di eventi macroeconomici importanti. In primis, nel 1992 abbiamo avuto un calo degli investimenti fissi e contemporaneamente un calo del PIL, dovuti alla svalutazione della lira, che in questo periodo perse in tre mesi il 40% del suo valore, rispetto al marco tedesco.

Negli anni precedenti il 1992 c'erano stati degli accordi più solidi fra gli stati, che avevano l'obiettivo di stabilizzare i cambi fra le valute europee aderenti allo SME. L'Italia vi aderì, ma in mancanza di politiche di contenimento della spesa pubblica e pur partendo da un cambio che garantiva inizialmente competitività al nostro sistema industriale, dopo pochi anni la gran parte delle aziende italiane perse la propria competitività nel mercato estero. Si arrivò a tassi di inflazione annui a due cifre, perdita di fiducia dell'Italia nei mercati internazionali, bilancia commerciale ancora più in deficit, causa l'aumento dei costi delle importazioni. Tutto ciò rese inevitabile la svalutazione della lira.

Da questo momento la serie denota una crescita costante, anche se sono presenti fluttuazioni di diversa entità. Dal terzo trimestre del 2008 il valore degli investimenti è diminuito in modo drastico, e ancora adesso non ha recuperato il valore che aveva prima di questa crisi, siamo infatti agli stessi valori dei primi anni '90.

Come nel 1992, nel 2011 l'Italia vive una grave crisi, questa volta di carattere finanziario, con serie ripercussioni sul sistema economico e sulle condizioni di vita della popolazione.

Nel corso degli anni gli interventi della BCE si sono fatti sempre più incisivi. Le due *Long Term Refinancing Operations (LTROs)* hanno avuto effetti positivi parziali e sono servite solo ad arginare la crisi. L'OMT del settembre 2012, con l'acquisto illimitato di titoli governativi, ha rappresentato un forte segnale da parte della BCE. Nel settembre 2014 la BCE ha assunto due misure straordinarie, il *Targeted Long Term Refinancing Operation (TLTRO)* e l'acquisto di titoli *ABS e Covered Bond*, misure attraverso le quali la BCE e gli organismi internazionali hanno cercato di arginare la crisi.

Le condizioni di mercato eccezionali generate dalla crisi del debito sovrano hanno ostacolato, tuttavia, la trasmissione all'economia reale delle azioni di stimolo adottate dalla Banca Centrale Europea. In particolare, la frammentazione del mercato finanziario unico ha impedito alle misure di politica monetaria di generare un miglioramento delle condizioni di accesso al

finanziamento per le banche dei paesi più esposti alla crisi. Inoltre, le tensioni che le banche hanno sperimentato sul fronte della redditività (in un contesto di bassi tassi di interesse) hanno ridotto la propensione degli istituti bancari a ridurre il costo dei prestiti a famiglie e imprese.⁸

Si può notare come le due serie, annuali e trimestrali, presentino lo stesso andamento, ma la serie annuale è molto più smussata: l'andamento di lungo periodo è lo stesso, ma le variazioni di breve periodo o inferiori a un anno, si possono ben notare solo nella serie trimestrale, che denota una maggiore variabilità rispetto alla prima serie. Come si può osservare, anche solo da un'analisi grafica sommaria, questa serie non è stazionaria, poiché presenta alta variabilità e non presenta un andamento preciso, crescente o decrescente.

Inoltre, è utile una breve osservazione del correlogramma per avviare una analisi più approfondita e precisa della stessa. Il correlogramma ci mostra il grado di autocorrelazione tra la variabile al tempo t , e la stessa osservata, in questo caso, fino a 12 ritardi. La funzione di autocorrelazione, ACF, ci mostra come il valore osservato degli investimenti fissi lordi, della serie annuale, sia correlato in modo significativo, e in modo decrescente, alla variabile fino un ritardo di 3 anni. Invece, la serie trimestrale presenta molta più persistenza: i valori passati di questa serie storica sono ancora significati dopo i 12 ritardi. La funzione di autocorrelazione parziale, PACF, invece misura quale sia la correlazione tra due osservazioni al netto della correlazione di queste osservazioni con i ritardi intermedi. Potrebbe essere, quindi, che la correlazione tra il valore attuale e il valore corrispondente al terzo ritardo sia dovuta alle correlazione con i dati intermedi, piuttosto che a una correlazione effettiva tra questi due specifici valori.

Infatti, nel correlogramma parziale il terzo ritardo non è significativo. Togliendo l'effetto di correlazione intermedia, la funzione di autocorrelazione parziale indica i primi due ritardi come realmente significativi, sia nella serie annuale che nella serie trimestrale.

Una possibile alternativa di rappresentazione della serie può essere di considerare la serie Δy_t . Visto che stiamo lavorando in logaritmi, la serie può essere interpretata come la serie storica dei tassi di crescita annuali degli investimenti fissi lordi, a cadenza trimestrale.

In questo caso, le due serie storiche sono molto più differenti. L'andamento è sufficientemente lo stesso, ma la serie trimestrale è molto più variabile: l'ampiezza delle sue oscillazioni infatti è maggiore. Si può inoltre notare come questa serie risponda gravemente agli shock economici. Durante i periodi di crisi ha avuto tassi di crescita negativi fino al 15%, nei periodi di stabilità economica invece la serie presenta un andamento quasi aleatorio, dove a tassi di crescita di leggere entità positivi, si contrappongono tassi di crescita negativi sempre della stessa entità.

⁸ <http://www.consob.it/web/investor-education/crisi-debito-sovrano-2010-2011>

Osservando il correlogramma riferito alla serie dei tassi di crescita annuali, si nota come solo il primo valore, sia nella ACF sia nella PACF, sia significativo e influenzi il comportamento della serie storica. Nella PACF, addirittura, molti valori sono quasi nulli.

Invece, il correlogramma della serie storica trimestrale è molto più interessante. Nella ACF i ritardi fino al quarto sono significativi e tendono ad annullarsi con il tempo, mentre nella PACF solo i primi due ritardi sono significativi, ed è appunto nulla dopo il secondo lag.

Questo comportamento è tipico di un processo autoregressivo. Probabilmente l'andamento della serie di lungo periodo può essere rappresentata da, appunto, un processo autoregressivo, ma dati gli ultimi dati recenti, di una crisi con conseguenze mai viste prima dal nostro sistema economico, la media della serie si è spostata, quindi è utile includere o considerare un processo autoregressivo a media mobile.



Figura 2.1: Grafico della serie storica annuale dei logaritmi degli investimenti

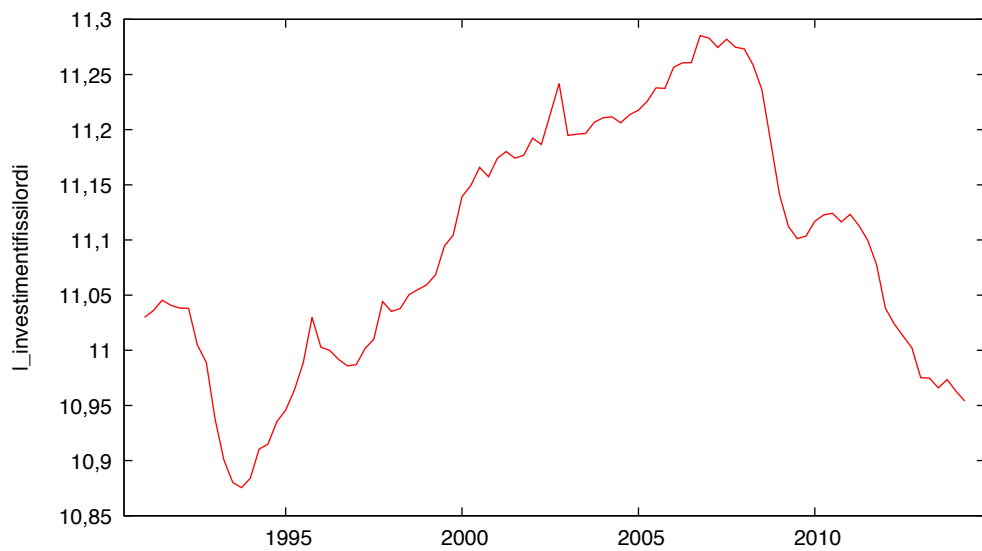


Figura 2.2: Grafico della serie storica trimestrale dei logaritmi degli investimenti

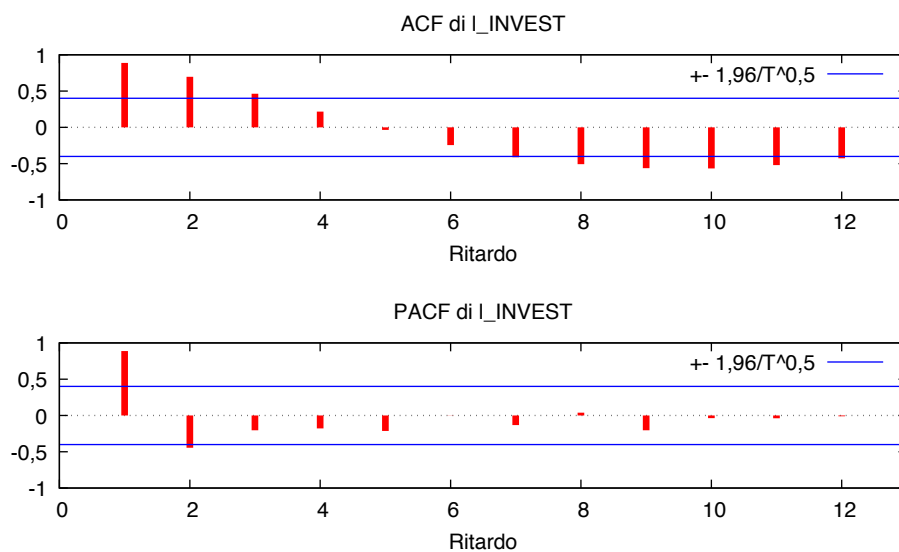


Figura 2.3: Correlogramma della serie storica annuale dei logaritmi degli investimenti

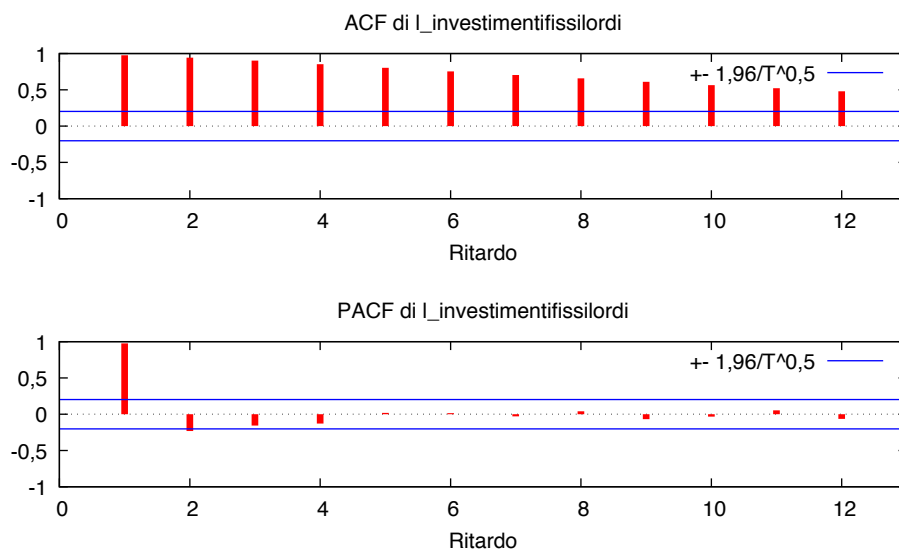


Figura 2.4: Correlogramma della serie storica trimestrale dei logaritmi degli investimenti



Figura 2.5: Grafico del tasso di crescita annuale degli investimenti

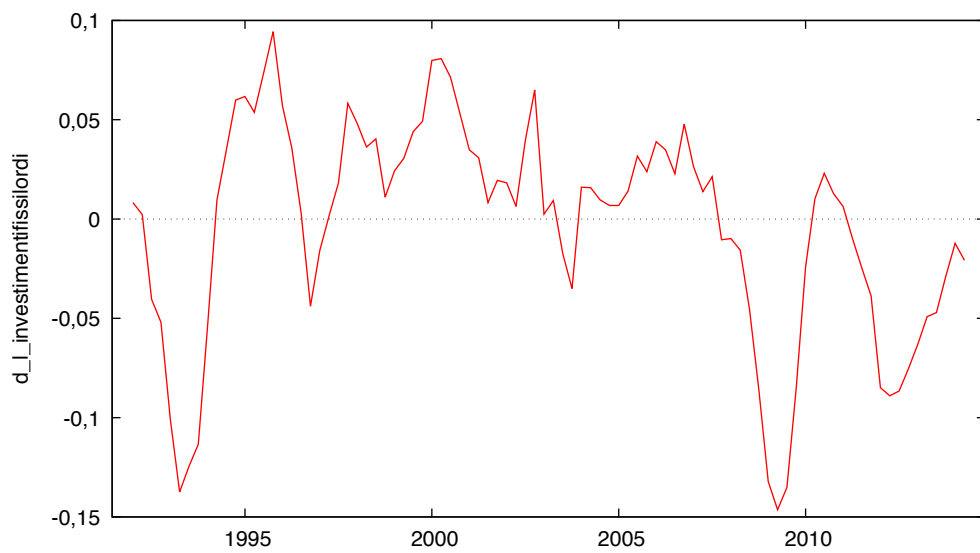


Figura 2.6: Grafico del tasso di crescita annuale degli investimenti, a cadenza trimestrale

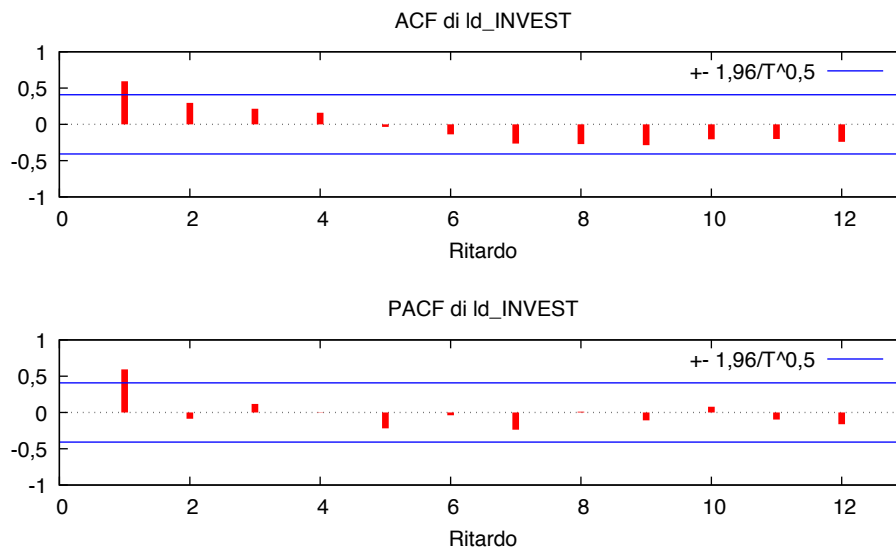


Figura 2.7: Correlogramma del tasso di crescita annuale degli investimenti

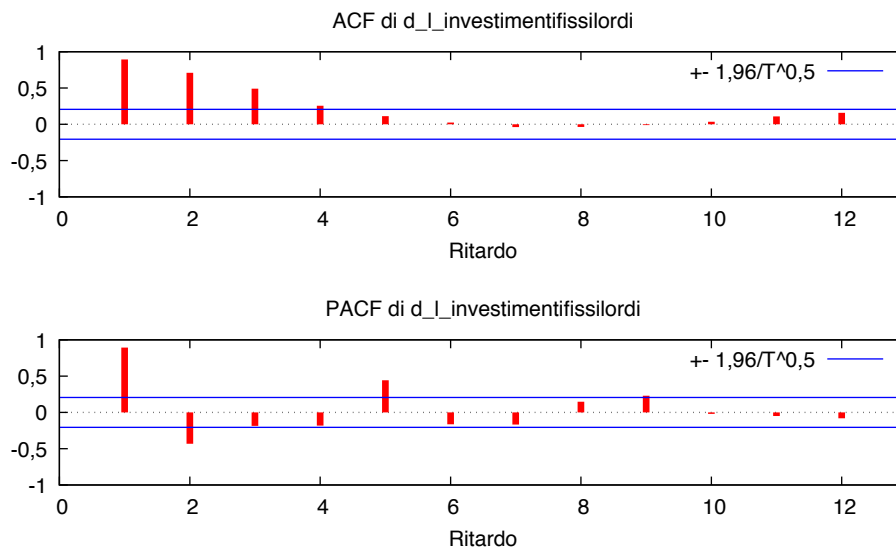


Figura 2.8: Correlogramma del tasso di crescita annuale degli investimenti, a cadenza trimestrale

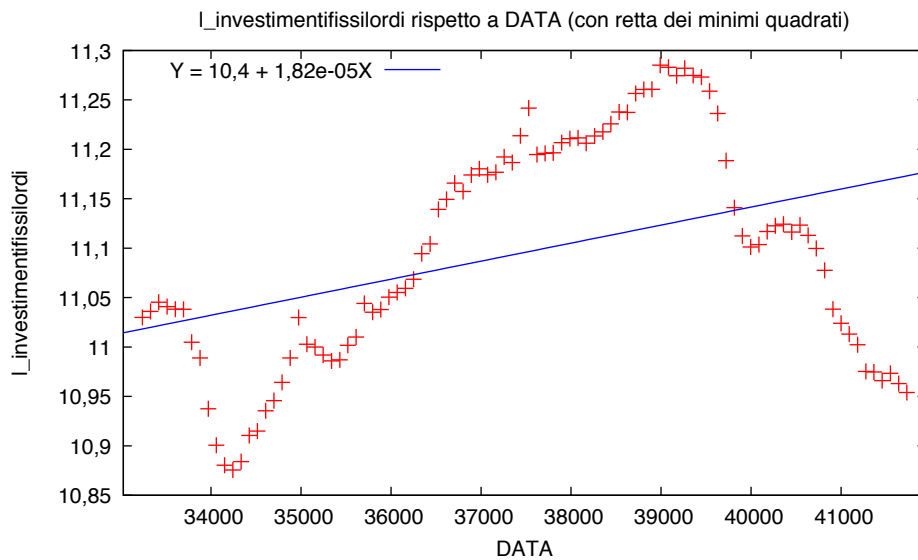


Figura 2.9: Grafico della serie storica trimestrale dei logaritmi degli investimenti confrontata con retta dei minimi quadrati

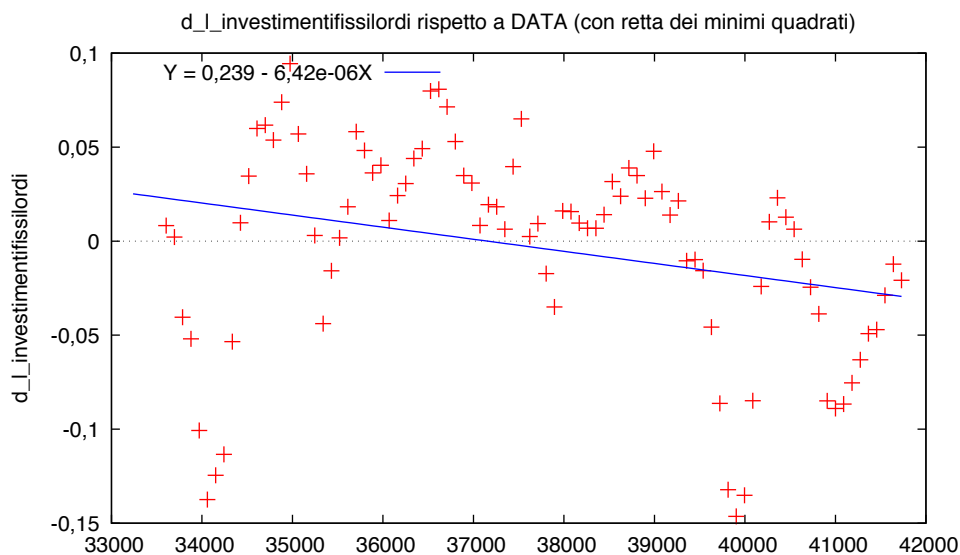


Figura 2.10: Grafico del tasso di crescita annuale degli investimenti, a cadenza trimestrale, confrontato con retta dei minimi quadrati

La precedente analisi sommaria della serie storica ha evidenziato come l'osservazione degli investimenti al generico tempo t mostri un qualche tipo di correlazione con le precedenti. Inoltre, come si osserva dal grafico 2.10, la media del processo sembra diminuire nel tempo. Non è dunque scorretto ipotizzare una modellazione autoregressiva a media mobile (ARMA). Per scegliere il numero dei coefficienti utilizziamo due criteri per metodo di confronto: il Criterio di Schwarz e il Criterio di Akaike.

	Criterio di Akaike	Criterio di Schwarz
ARMA(2,4)	-462,0016	-442,0032
ARMA(2,3)	-464,0008	-446,5021
ARMA(2,2)	-444,2801	-429,2813

Il modello che minimizza entrambi i criteri è il modello ARMA (2,3), rappresentabile come:

$$y_t = -\phi_1 y_{t-1} - \phi_2 y_{t-2} + \theta \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

Modello 1: ARMA, usando le osservazioni 1992:1-2014:2 (T = 90)

Variabile dipendente: d_1_investmentifissilordi

Errori standard basati sull'Hessiana

	<i>Coefficiente</i>	<i>Errore Std.</i>	<i>z</i>	<i>p-value</i>	
const	-0,00326490	0,0149949	-0,2177	0,8276	
phi_1	0,338931	0,109945	3,083	0,0021	***
phi_2	0,230612	0,105966	2,176	0,0295	**
theta_1	0,998695	0,109624	9,110	<0,0001	***
theta_2	0,998695	0,0668327	14,94	<0,0001	***
theta_3	1,00000	0,105787	9,453	<0,0001	***
Media var. dipendente	-0,003285		SQM var. dipendente	0,054897	
Media innovazioni	-0,000112		SQM innovazioni	0,015877	
Log-verosimiglianza	239,0004		Criterio di Akaike	-464,0008	
Criterio di Schwarz	-446,5021		Hannan-Quinn	-456,9443	

		<i>Reale</i>	<i>Immaginario</i>	<i>Modulo</i>	<i>Frequenza</i>
AR	Radice 1	1,4734	0,0000	1,4734	0,0000
	Radice 2	-2,9431	0,0000	2,9431	0,5000
MA	Radice 1	-1,0000	0,0000	1,0000	0,5000
	Radice 2	0,0007	-1,0000	1,0000	-0,2499
	Radice 3	0,0007	1,0000	1,0000	0,2499

Test LM per l'autocorrelazione fino all'ordine 10 -
Ipotesi nulla: Non c'è autocorrelazione
Statistica test: Chi-quadro(5) = 9,3368
p-value = P(Chi-quadro(5) > 9,3368) = 0,09636

2.2 Processi a radice unitaria

Poiché il tasso di crescita dovrebbe fluttuare all'interno di una banda (o di un range di valori) si può immaginare che la serie Δy_t sia un processo stazionario, forse a media non nulla.

In questo caso y_t è un processo a *radice unitaria*.

I processi a radice unitaria non sono stazionari, quindi non lo è y_t ; però se Δy_t è stazionario allora, per definizione, y_t è *stazionario in differenza*.

Un'altra espressione che si usa frequentemente è che y_t è un processo I(1), *integrato di ordine uno*, ovvero che y_t va differenziato una sola volta affinché il risultato sia stazionario.

Si definisce come *processo I(d)* un processo la cui differenza d-esima è stazionaria.

Un processo I(1), ovvero un processo AR(1) con coefficiente $\varphi = 1$ è un *random walk*⁹:

$$\begin{aligned}y_t &= y_{t-1} + \varepsilon_t \\ &\dots \\ y_t &= y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i}\end{aligned}$$

All'interno del termine di errore confluiscono, anche, gli shock dovuti alle manovre di politica economica o gli effetti negativi/positivi di recessioni o di espansioni. Il possesso di una radice unitaria ci informa che le scelte e le riforme macroeconomiche sono in grado di alterare permanentemente l'andamento della serie storica in esame.

Qualora la serie non possedesse radice unitaria l'effetto al tempo t scemerebbe esponenzialmente nel tempo.

Un aspetto caratteristico dei random walk è quello che questi processi non sono *mean-reverting*. Se un processo è mean-reverting esso ha la tendenza a muoversi preferenzialmente verso il suo valore atteso. La mancanza di questa peculiarità provoca periodi, anche lunghi, in cui la serie presenta un andamento crescente o decrescente piuttosto marcato.

Nessuno esclude che ci possano essere effettivamente dei trend deterministici, sovrapposti a quelli puramente stocastici, questo accade nei cosiddetti *random walk con drift*.

⁹ Si veda l'appendice B per una maggiore trattazione.

Questi processi sono semplicemente processi per cui si ha:

$$\Delta y_t = \mu + \varepsilon_t$$

Quindi y_t è un random walk a cui si sovrappone una funzione lineare del tempo, con pendenza μ . Se il *drift*, ovvero la costante, è positivo si avrà un processo che tende a salire, a con fluttuazioni più marcate con il passare del tempo.

Il caso del random walk si estende, in modo piuttosto indolore, al caso in cui gli incrementi del processo non siano un white noise, ma più in generale un qualunque processo stocastico stazionario: rappresentando quest'ultimo come un processo ARMA, avremo una rappresentazione del tipo $A(L)\Delta y_t = C(L)\varepsilon_t$.

Senza sapere a priori se la serie presenta radice unitaria, ovvero se è stazionaria o meno, bisogna utilizzare regole di decisione note come *test di radice unitaria*.

A questo fine abbiamo utilizzato tre test statistici: ADF, ADF-GLS E KPSS.

I primi due sfruttano l'intuizione del *test Dickey-Fuller*.¹⁰

Si considera un processo autoregressivo di primo ordine che chiamiamo y_t :

$$y_t = \varphi y_{t-1} + \varepsilon_t$$

Per definizione

$$\Delta y_t = \rho y_{t-1} + \varepsilon_t$$

dove ε_t è un white noise e $\rho = \varphi - 1$, cosicchè il processo è stazionario solo se $\rho < 0$. Viceversa se $\rho = 0$ siamo in presenza di un processo I(1).

Questo test valuta quindi l'ipotesi nulla, H_0 , che il processo possieda radice unitaria, contro l'ipotesi alternativa H_1 , che il processo sia stazionario o trend stazionario.

- $H_0 : \rho = 0$
- $H_1 : \rho < 0$

Il test si caratterizza come unilaterale nella regione non positiva della distribuzione, che segue un andamento differente dalla classica *t-student*, quindi la sua distribuzione asintotica è diversa e per questo bisogna utilizzar delle tavole statistiche precise, i cui quantili della distribuzione sono stati calcolati *da Dickey e Fuller*. Questa distribuzione viene comunemente chiamata distribuzione DF.

La statistica test per il test di radice unitaria è, comunque, basata sul classico test di azzeramento del parametro:

¹⁰ Dickey David and Fuller Wayne, *Distribution of the estimators for autoregressive time series with a unit root*, Journal of the American statistical association, 1979.

$$t_{\rho} = \frac{\hat{\rho}}{\sqrt{\widehat{Var}(\hat{\rho})}}$$

Sono possibili tre configurazioni del test, ciascuna con differenti valori critici:

- Senza costante: $\Delta y_t = \rho y_{t-1} + \varepsilon_t$
- Con costante: $\Delta y_t = \beta_0 + \rho y_{t-1} + \varepsilon_t$
- Con costante e trend: $\Delta y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \varepsilon_t$

L'intuizione è che se il processo è stazionario, la realizzazione al tempo t dipenderà dalla precedente, in quanto gli scostamenti saranno seguiti da osservazioni del tasso di crescita di segno opposto, poiché la media di lungo periodo è costante. Il coefficiente, allora, di y_{t-1} sarà significativamente < 1 . Se invece il processo possiede radice unitaria, il termine al tempo $t-1$ non fornirà alcuna informazione sull'andamento della serie visto che dipende solo dal termine di errore. Il coefficiente, in questo caso, sarà nullo in modo significativo.

Il secondo test utilizzato parte dalla formulazione più generale:

$$y_t = \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t$$

Anche in questo caso ε_t è white noise, e la persistenza di breve periodo è catturata completamente dalla componente autoregressiva.

Sfruttando che $y_{t-k} = y_{t-1} - \sum_{i=1}^{k-1} \Delta y_{t-i}$ la formula precedente può essere riscritta come segue:

$$\Delta y_t = \rho y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_p \Delta y_{t-p} + u_t$$

Dove $\rho = (\varphi_1 + \dots + \varphi_p) - 1$ e $\gamma_i = -(\varphi_{i+1} + \dots + \varphi_p)$

In questo caso il test prende il nome di *Augmented Dickey-Fuller test (ADF)*, che utilizza la medesima logica e procedura operativa del test DF, considerando sempre il parametro ρ .

Le tre possibili tipologie di test presentate per il DF sono certamente utilizzabili anche nel ADF:

- Senza costante: $\Delta y_t = \rho y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_p \Delta y_{t-p} + u_t$
- Con costante: $\Delta y_t = \beta_0 + \rho y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_p \Delta y_{t-p} + u_t$
- Con costante e trend: $\Delta y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_p \Delta y_{t-p} + u_t$

Imponendo i vincoli $\beta_0 = 0$ e $\beta_1 = 0$ si testa un modello senza costante, ossia un random walk. Imponendo solo il vincolo $\beta_1 = 0$ si valuta un modello random walk con drift. Non imponendo vincoli si testa un modello con drift e trend. Anche in questo caso si valuta l'ipotesi nulla $H_0 : \rho = 0$, contro l'ipotesi alternativa $H_1 : \rho < 0$, considerando sempre il valore della statistica

test $t_\rho = \frac{\hat{\rho}}{\sqrt{\widehat{\text{var}}(\hat{\rho})}}$. Quando questa sia inferiore al valore critico della distribuzione di Dickey-

Fuller allora sarà rifiutata l'ipotesi nulla.

Il numero di ritardi p può essere scelto impiegando alternativamente i criteri AIC^{11} o BIC^{12} , entrambi questi criteri utilizzando il modello della massima verosimiglianza, la quale è possibile essere aumentata attraverso l'aggiunta di parametri. La differenza tra i due criteri giace proprio nel fatto che il criterio BIC introduce un elemento di penalizzazione per il maggior numero di parametri del modello. Quando $p = 0$, la formulazione è la medesima del test DF. Ulteriore strumento di analisi è il test ADF-GLS¹³, sviluppato da Elliott, Rothenberg e Stock nel 1992, come estensione del test ADF.

Il test conserva sfrutta la medesima struttura ed intuizione del test ADF: manipola la serie storica localmente per rimuovere la componente di trend, per stimare efficacemente i parametri con il metodo della massima verosimiglianza, per poi applicare sui dati trasformati il test ADF per radice unitaria.

Questi test presentati precedentemente assumono l'ipotesi nulla di esistenza di radice unitaria, ci sono invece test che partono dalla nulla di stazionarietà, il più noto è il test KPSS¹⁴.

La differenza maggiore di questo test è che l'assenza di una radice unitaria non è prova di stazionarietà, visto che il processo potrebbe essere comunque trend-stazionario. Questa è un'importante differenza poiché sia in processi a radice unitaria sia in processi trend-stazionari la media può aumentare o diminuire con il passare del tempo, ma i processi stazionari sono mean-reverting, quindi non hanno la caratteristica di permanenza degli shock economici. Anche in questo caso la distribuzione non è la classica t-student ed è necessario utilizzare valori critici specifici tabulati.

Abbiamo evidenziato fino ad ora come la presenza di radici unitarie comporti la mancata tendenza della serie a ritornare nel lungo periodo verso un percorso deterministico: gli shock producono effetti permanenti nel tempo.

Pier Perron¹⁵, invece, dimostra come la forte presenza di radici unitarie possa essere, in realtà, frutto del trattamento sbagliato dei break strutturali.

¹¹ Akaike's information criterion, Hirotugu Akaike, 1974

¹² Gideon E. Schwarz, Bayesian information criterion, *Estimating the dimension of a model*, Annals of Statistic, vol. 6, 1978

¹³ Graham Elliott, Thomas J. Rothenberg and James H. Stock, *Efficient tests for an autoregressive unit root*, Econometrica, vol. 64, no. 4, 1996

¹⁴ Kwiatkowski, Phillips, Schmidt, Shin, *Testing the null hypothesis of stationarity against the alternative of a unit root*, Journal of Econometrics, 1992

¹⁵ Pierre Perron, *The great crash, the oil price shock, and the unit root hypothesis*, Econometrica: Journal of the Econometric Society, 1989

Questi sono cambiamenti che possono comparire a causa di un netto cambiamento nei coefficienti dei parametri ad una data precisa o da una graduale evoluzione dei coefficienti nel corso di un periodo di tempo più lungo. Perron introdusse, quindi, un nuovo metodo di analisi, in cui la presenza di radice unitaria era verificata avendo precedentemente permesso la presenza nel modello di radice unitaria.

Zivot e Andrews¹⁶ evidenziano come i valori critici convenzionali dei test perdano validità qualora i punti di rottura dei dati vengano scelti a priori, basati su un'analisi preliminare: tutto ciò portava a un eccessivo rigetto della radice unitaria.

Da qui nasce la ricerca di un metodo per implementare la datazione dei break strutturali all'interno dei test per radice unitaria.

Il test di Zivot-Andrews propone una variante del test di Perron, quest'ultimo viene condotto inserendo nel modello di analisi un termine che catturi la presenza del break, la cui data è, però, sconosciuta.

Tre sono i modelli:

- $y_t = \beta_0 + \beta_1 t + \alpha y_{t-1} + \zeta_1 DU_t + \sum_{j=1}^P \delta_j \Delta y_{t-j} + \varepsilon_t$
Unico spostamento nel livello della serie (intercetta)
- $y_t = \beta_0 + \beta_1 t + \alpha y_{t-1} + \zeta_2 DT_t + \sum_{j=1}^P \delta_j \Delta y_{t-j} + \varepsilon_t$
unico spostamento nella pendenza del trend
- $y_t = \beta_0 + \beta_1 t + \alpha y_{t-1} + \zeta_1 DU_t + \zeta_2 DT_t + \sum_{j=1}^P \delta_j \Delta y_{t-j} + \varepsilon_t$
Unico spostamento nel livello della serie e nella pendenza del trend

DU_t è una *variabile dummy* per lo spostamento della media in seguito a una rottura strutturale, che vale 1 per le osservazioni successive al break strutturale; DT_t svolge la medesima funzione per il cambiamento della pendenza del trend.

In formule:

$$DU_t = \begin{cases} 1 & \text{se } t > SB \\ 0 & \text{diversamente} \end{cases}$$

$$DT_t = \begin{cases} t - SB & \text{se } t > SB \\ 0 & \text{diversamente} \end{cases}$$

L'ipotesi nulla, H_0 , è la presenza di radice unitaria con *drift*, l'ipotesi alternativa, H_1 , è un processo trend-stazionario con un break strutturale in data ignota.

¹⁶ Eric Život and Donald W.K. Andrews, *Further evidence on the Great Crash, the oil price shock, and the unit root hypothesis*, Journal of Business and Economic Statistics, 1992

Ciascuna osservazione potrebbe indicare un cambiamento strutturale e il test conduce una regressione sequenzialmente per ogni osservazione, valutando ogni volta un differente valore temporale per la variabile *SB* (*structural break*).

L'algoritmo seleziona, tra tutte le osservazioni, quella che minimizza il valore della statistica unilaterale t per testare $\alpha = 1$: l'obiettivo è determinare il punto di rottura nell'osservazione che da maggiore credito all'ipotesi di stazionarietà.

Il primo e il terzo modello sono i più indicati per l'analisi di serie storiche economiche. Le caratteristiche della distribuzione asintotica del test richiedono che gli estremi del campione non siano inclusi: è necessario ridurre, *trimming*, il campione.

2.3 Risultati

Riportiamo in tabella i risultati dei test che sono stati eseguiti: (1) ADF con costante, (2) ADF con costante e trend, (3) ADF-GLS con costante, (4) ADF-GLS con costante e trend e (5) KPSS con costante e trend.

Di ognuno vengono riportati il valore della statistica test e il valore del p-value, associato sia al criterio AIC, che al criterio BIC. Insieme al p-value viene segnalata l'accettazione o il rifiuto dell'ipotesi nulla.

Quando il livello di confidenza è \geq del 10% viene segnalato SI***, quando il livello è inferiore al 10%, ma maggiore del 5% viene indicato SI**, quando è \geq 2,5% viene indicato SI* e quando è minore del 2,5% l'ipotesi nulla viene rifiutata, indicato con NO.

Si ricordi che il test KPSS ha ipotesi contrarie rispetto agli altri test, solo in questo caso un'accettazione della nulla implica che il processo sia stazionario. Il test viene eseguito sul test di crescita annuale, utilizzando sia valori annuali, sia valori trimestrali.

L'intervallo completo annuale è l'intero dataset di 24 osservazioni; i successivi 15 campioni, comprendono cicli di 10 anni, ciascuno traslato in avanti di un anno.

L'intervallo completo trimestrale, invece, comprende in totale 94 osservazioni, i successivi 21 campioni comprendono cicli di 20 trimestri, ciascuno traslato in avanti di 4 trimestri, come visto in precedenza.

Il numero dei ritardi non è scelto a priori, ma viene calcolato dal software statistico, secondo il criterio utilizzato.

L'analisi procede, poi, con l'obiettivo di individuare l'eventualità di un processo non stazionario con il test di Zivot-Andrews, il quale ricerca la presenza di radice unitaria, valutando l'occorrenza di break strutturali.

Il test è stato condotto sulla serie storica di tasso di crescita annuale a cadenza trimestrale, applicando la cosiddetta procedura di *trimming* alle estremità, ignorando il 15% delle osservazioni estreme. Vengono, infine, riportati i valori critici delle distribuzioni, e il grafico del risultato del test. È stato impiegato il terzo modello, con spostamento nel livello della serie e nel trend.

Tabella 2.1: Risultati dei test di radice unitaria su campioni trimestrali

TEST STATISTICO					
	ADF, c	ADF, ct	ADF-GLS, c	ADF-GLS, ct	KPSS, ct
Periodo	AIC t-stat AIC p-value BIC t-stat BIC p-value	AIC t-stat AIC p-value BIC t-stat BIC p-value	AIC t-stat AIC p-value BIC t-stat BIC p-value	AIC t-stat H_0 BIC t-stat H_0	t-stat p-value
1991:1 – 2014:2	-1,569 <i>0,4984 SI***</i> -2,5257 <i>0,109 SI***</i>	-3,088 <i>0,109 SI***</i> -3,255 <i>0,073 SI**</i>	-1,578 <i>0,1078 SI***</i> -1,578 <i>0,1078 SI***</i>	-2,152 <i>SI***</i> -2,312 <i>SI***</i>	0,209 <i>0,013 NO</i>
1991:1 - 1995:4	-3,025 <i>0,0326 SI*</i> -0,754 <i>0,831 SI***</i>	-4,088 <i>0,006 NO</i> -4,088 <i>0,006 NO</i>	-0,589 <i>0,462 SI***</i> -0,589 <i>0,462 SI***</i>	-0,472 <i>SI***</i> -0,472 <i>SI***</i>	0,138 <i>0,074 SI**</i>
1992:1 – 1996:4	-2,586 <i>0,0958 SI**</i> -2,586 <i>0,0958 SI**</i>	-1,326 <i>0,881 SI***</i> -1,326 <i>0,881 SI***</i>	-0,971 <i>0,296 SI***</i> -0,971 <i>0,296 SI***</i>	-1,107 <i>SI***</i> -1,107 <i>SI***</i>	0,103 <i>> .10 SI***</i>
1993:1 – 1997:4	-2,635 <i>0,085 SI**</i> -2,635 <i>0,085 SI**</i>	-2,881 <i>0,168 SI***</i> -2,881 <i>0,168 SI***</i>	-1,616 <i>0,100 SI***</i> -1,616 <i>0,100 SI***</i>	-1,296 <i>SI***</i> -1,296 <i>SI***</i>	0,160 <i>0,042 SI*</i>
1994:1 – 1998:4	-3,732 <i>0,003 NO</i> -3,732 <i>0,003 NO</i>	-3,286 <i>0,068 SI**</i> -3,286 <i>0,068 SI**</i>	-2,084 <i>0,035 SI*</i> -2,084 <i>0,035 SI*</i>	-1,192 <i>SI***</i> -1,192 <i>SI***</i>	0,087 <i>> .10 SI***</i>
1995:1 – 1999:4	-2,967 <i>0,038 SI*</i> -2,967 <i>0,038 SI*</i>	-2,861 <i>0,175 SI***</i> -2,861 <i>0,175 SI***</i>	-1,706 <i>0,083 SI**</i> -1,706 <i>0,083 SI**</i>	-1,689 <i>SI***</i> -1,689 <i>SI***</i>	0,113 <i>> .10 SI***</i>

Continua nella pagina seguente

1996:1 – 2000:4	-2,804 <i>0,057 SI**</i> -2,804 <i>0,057 SI**</i>	-3,622 <i>0,027 SI*</i> -3,622 <i>0,027 SI*</i>	-1,801 <i>0,068 SI**</i> -1,801 <i>0,068 SI**</i>	-2,61 <i>SI***</i> -2,61 <i>SI***</i>	0,064 <i>> .10 SI***</i>
1997:1 – 2001:4	-1,477 <i>0,545 SI***</i> -1,477 <i>0,545 SI***</i>	-3,608 <i>0,029 SI*</i> -3,608 <i>0,029 SI*</i>	-1,915 <i>0,053 SI**</i> -1,915 <i>0,053 SI**</i>	-1,930 <i>SI***</i> -1,930 <i>SI***</i>	0,114 <i>> .10 SI***</i>
1998:1 – 2002:4	-1,715 <i>0,408 SI***</i> -1,715 <i>0,408 SI***</i>	-2,610 <i>0,275 SI***</i> -2,610 <i>0,275 SI***</i>	-1,704 <i>0,083 SI**</i> -1,704 <i>0,083 SI**</i>	-1,671 <i>SI***</i> -1,671 <i>SI***</i>	0,095 <i>> .10 SI***</i>
1999:1 – 2003:4	-1,032 <i>0,720 SI***</i> -1,032 <i>0,720 SI***</i>	-2,060 <i>0,534 SI***</i> -2,060 <i>0,534 SI***</i>	-1,161 <i>0,224 SI***</i> -1,161 <i>0,224 SI***</i>	-2,167 <i>SI***</i> -2,167 <i>SI***</i>	0,090 <i>> .10 SI***</i>
2000:1 – 2004:4	-1,854 <i>0,345 SI***</i> -1,854 <i>0,345 SI***</i>	-5,520 <i>0,000 NO</i> -5,520 <i>0,000 NO</i>	-1,855 <i>0,060 SI**</i> -1,855 <i>0,060 SI**</i>	-2,977 <i>SI**</i> -2,977 <i>SI**</i>	0,095 <i>> .10 SI***</i>
2001:1 – 2005:4	-3,186 <i>0,036 SI*</i> -3,186 <i>0,036 SI*</i>	-2,995 <i>0,157 SI***</i> -2,995 <i>0,157 SI***</i>	-2,786 <i>0,005 NO</i> -2,786 <i>0,005 NO</i>	-3,112 <i>SI**</i> -3,112 <i>SI**</i>	0,095 <i>> .10 SI***</i>
2002:1 – 2006:4	-2,781 <i>0,078 SI**</i> -2,781 <i>0,078 SI**</i>	-2,672 <i>0,256 SI***</i> -2,672 <i>0,256 SI***</i>	-2,661 <i>0,007 NO</i> -2,661 <i>0,007 NO</i>	-2,950 <i>SI**</i> -2,950 <i>SI**</i>	0,128 <i>0,092 SI**</i>
2003:1 – 2007:4	-3,461 <i>0,020 NO</i> -3,461 <i>0,020 NO</i>	-3,075 <i>0,112 SI***</i> -3,075 <i>0,112 SI***</i>	-2,817 <i>0,004 NO</i> -2,817 <i>0,004 NO</i>	-3,396 <i>SI*</i> -3,396 <i>SI*</i>	0,137 <i>0,074 SI**</i>

Continua nella pagina seguente

2004:1 – 2008:4	-0,236 0,918 <i>SI***</i> -0,236 0,918 <i>SI***</i>	2,859 0,999 <i>SI***</i> 2,859 0,999 <i>SI***</i>	-0,539 0,483 <i>SI***</i> -0,539 0,483 <i>SI***</i>	-1,44 <i>SI***</i> -1,44 <i>SI***</i>	0,194 <i>0,018 NO</i>
2005:1 – 2009:4	-2,806 0,057 <i>SI**</i> -2,806 0,057 <i>SI**</i>	-1,431 0,852 <i>SI***</i> -1,431 0,852 <i>SI***</i>	-0,532 0,486 <i>SI***</i> -0,532 0,486 <i>SI***</i>	-1,724 <i>SI***</i> -1,724 <i>SI***</i>	0,167 <i>0,037 SI*</i>
2006:1 – 2010:4	-1,316 0,624 <i>SI***</i> -1,316 0,624 <i>SI***</i>	-2,524 0,315 <i>SI***</i> -2,524 0,315 <i>SI***</i>	-0,462 0,515 <i>SI***</i> -0,462 0,515 <i>SI***</i>	-0,099 <i>SI***</i> -0,099 <i>SI***</i>	0,133 <i>0,082 SI**</i>
2007:1 – 2011:4	-2,616 0,089 <i>SI**</i> -2,616 0,089 <i>SI**</i>	-2,423 0,367 <i>SI***</i> -2,423 0,367 <i>SI***</i>	-1,287 0,182 <i>SI***</i> -1,287 0,182 <i>SI***</i>	-1,310 <i>SI***</i> -1,310 <i>SI***</i>	0,137 <i>0,074 SI**</i>
2008:1 – 2012:4	-3,306 0,014 <i>NO</i> -3,306 0,014 <i>NO</i>	-3,208 0,082 <i>SI**</i> -3,208 0,082 <i>SI**</i>	-1,227 0,202 <i>SI***</i> -1,227 0,202 <i>SI***</i>	-1,269 <i>SI***</i> -1,269 <i>SI***</i>	0,099 <i>> .10 SI***</i>
2009:1 – 2013:4	-3,734 0,003 <i>NO</i> -3,734 0,003 <i>NO</i>	-3,610 0,028 <i>SI*</i> -3,610 0,028 <i>SI*</i>	-1,172 0,220 <i>SI***</i> -1,172 0,220 <i>SI***</i>	-1,228 <i>SI***</i> -1,228 <i>SI***</i>	0,131 <i>0,085 SI**</i>
2009:3 – 2014:2	-4,534 0,000 <i>NO</i> -4,534 0,000 <i>NO</i>	-3,980 0,009 <i>NO</i> -3,980 0,009 <i>NO</i>	-1,571 0,109 <i>SI***</i> -1,571 0,109 <i>SI***</i>	-1,783 <i>SI***</i> -1,783 <i>SI***</i>	0,099 <i>> .10 SI***</i>

Livello di significatività	1%	2.5%	5%	10%
TEST ADF	-3.43	-3.12	-2.86	-2.57
TEST ADF-GLS	-3.77	-3.61	-3.46	-2.89
TEST KPSS	0.205	0.178	0.151	0.124

Tabella 2.2: Risultati del test Zivot-Andrews su campioni trimestrali

TEST STATISTICO ZIVOT ANDREWS			
Break	Intercetta e trend	Intercetta	Trend
AIC	-7.641	-7.693	-7.237
BIC	-7.641	-7.693	-7.237

Livello di significatività	1%	2.5%	5%	10%
TEST ZIVOT-ANDREWS	-5,57	-5,325	-5,08	-4,82

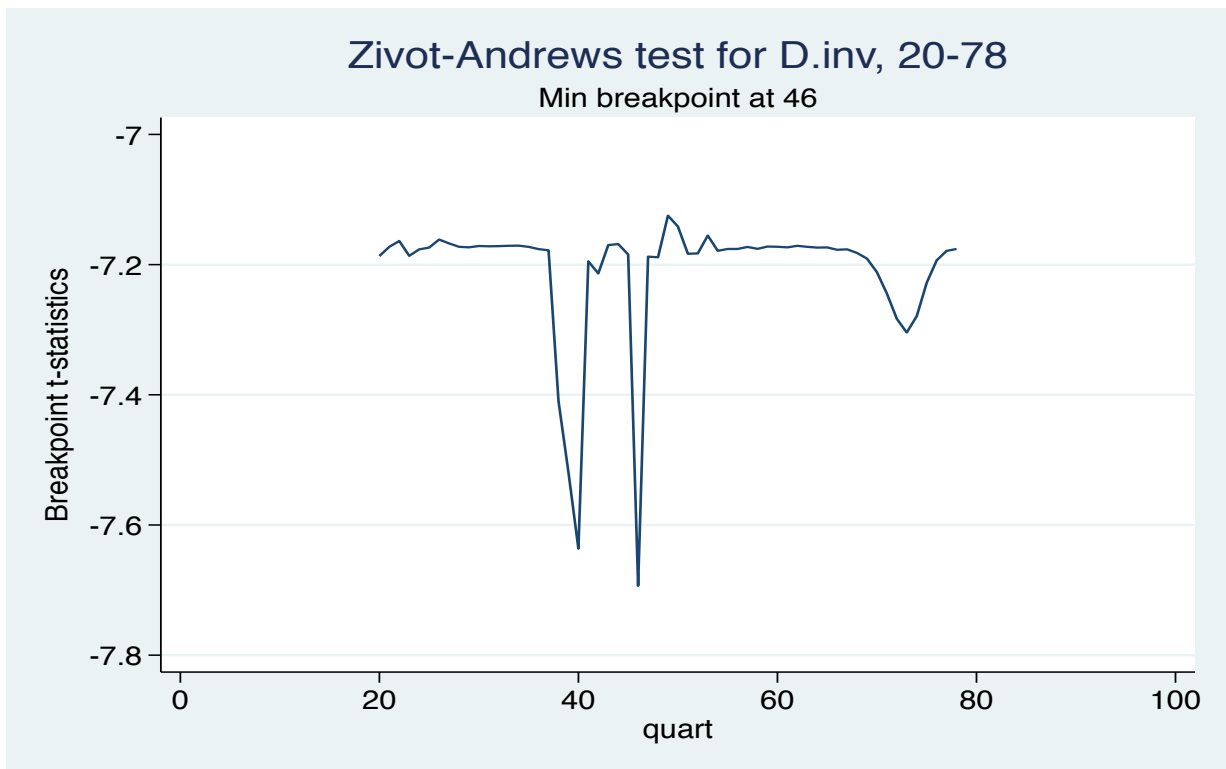


Figura 2.11: Grafico dell'andamento del valore della statistica test, riferito al test di Zivot-Andrews

Tabella 2.3: Risultati dei test di radice unitaria su campioni annuali

TEST STATISTICO					
	ADF, c	ADF, ct	ADF-GLS, c	ADF-GLS, ct	KPSS, ct
Periodo	AIC t-stat AIC p-value BIC t-stat BIC p-value	AIC t-stat AIC p-value BIC t-stat BIC p-value	AIC t-stat AIC p-value BIC t-stat BIC p-value	AIC t-stat H_0 BIC t-stat H_0	t-stat p-value
1995 – 2018	-2.195 <i>0,2131 SI***</i>	-2,153 <i>0,489 SI***</i>	-2,23 <i>0,024 SI*</i>	-2,326 <i>SI***</i>	0,129 <i>0,088 SI**</i>
	-2.195 <i>0,2131 SI***</i>	-2,153 <i>0,489 SI***</i>	-2,23 <i>0,024 SI*</i>	-2,326 <i>SI***</i>	
1995 – 2004	-2,602 <i>0,0924 SI**</i>	-1,666 <i>0,766 SI***</i>	-1,004 <i>0,283 SI***</i>	-1,316 <i>SI***</i>	0,125 <i>0,074 SI**</i>
	-2,602 <i>0,0924 SI**</i>	-1,666 <i>0,766 SI***</i>	-1,004 <i>0,283 SI***</i>	-1,316 <i>SI***</i>	
1996 – 2005	-3,745 <i>0,003 NO</i>	-3,063 <i>0,115 SI***</i>	-1,171 <i>0,221 SI***</i>	-1,661 <i>SI***</i>	0,117 <i>> .10 SI***</i>
	3,745 <i>0,003 NO</i>	-3,063 <i>0,115 SI***</i>	-1,171 <i>0,221 SI***</i>	-1,661 <i>SI***</i>	
1997 – 2006	-2,408 <i>0,139 SI***</i>	-2,989 <i>0,134 SI***</i>	-1,385 <i>0,154 SI***</i>	-1,752 <i>SI***</i>	0,096 <i>> .10 SI***</i>
	-2,408 <i>0,139 SI***</i>	-2,989 <i>0,134 SI***</i>	-1,385 <i>0,154 SI***</i>	-1,752 <i>SI***</i>	
1998 – 2007	-2,267 <i>0,182 SI***</i>	-3,527 <i>0,036 SI*</i>	-1,385 <i>0,154 SI***</i>	-1,888 <i>SI***</i>	0,094 <i>> .10 SI***</i>
	--2,267 <i>0,182 SI***</i>	-3,527 <i>0,036 SI*</i>	-1,385 <i>0,154 SI***</i>	-1,888 <i>SI***</i>	
1999 – 2008	-1,242 <i>0,610 SI***</i>	-2,834 <i>0,221 SI***</i>	-0,923 <i>0,316 SI***</i>	-3,47 <i>NO</i>	0,095 <i>> .10 SI***</i>
	-1,242 <i>0,610 SI***</i>	-2,834 <i>0,221 SI***</i>	--0,923 <i>0,316 SI***</i>	-3,47 <i>NO</i>	

Continua nella pagina seguente

2000 – 2009	-0,570 <i>0,979 SI***</i> -0,570 <i>0,979 SI***</i>	-1,590 <i>0,797 SI***</i> -1,590 <i>0,797 SI***</i>	-0,156 <i>0,629 SI***</i> -0,156 <i>0,629 SI***</i>	-1,668 <i>SI***</i> -1,668 <i>SI***</i>	0,115 <i>> .10 SI***</i>
2001 – 2010	-2,136 <i>0,236 SI***</i> 2,136 <i>0,236 SI***</i>	-3,584 <i>0,031 SI*</i> -3,584 <i>0,031 SI*</i>	-2,175 <i>0,028 SI**</i> -2,175 <i>0,028 SI**</i>	-3,308 <i>SI*</i> -1,930 <i>SI*</i>	0,109 <i>> .10 SI***</i>
2002 – 2011	-1,989 <i>0,285 SI***</i> 1,989 <i>0,285 SI***</i>	-3,475 <i>0,042 SI*</i> 3,475 <i>0,042 SI*</i>	-2,100 <i>0,0343 SI**</i> -2,100 <i>0,0343 SI**</i>	-3,00 <i>SI**</i> -3,00 <i>SI**</i>	0,100 <i>> .10 SI***</i>
2003 – 2012	-1,655 <i>0,421 SI***</i> -1,655 <i>0,421 SI***</i>	-3,147 <i>0,095 SI**</i> -3,147 <i>0,095 SI**</i>	-1,884 <i>0,056 SI**</i> -1,884 <i>0,056 SI**</i>	-3,16 <i>SI**</i> -3,16 <i>SI**</i>	0,122 <i>> .10 SI***</i>
2004 – 2013	-1,671 <i>0,414 SI***</i> -1,671 <i>0,414 SI***</i>	-3,632 <i>0,027 SI*</i> -3,632 <i>0,027 SI*</i>	-0,596 <i>0,459 SI***</i> -0,596 <i>0,459 SI***</i>	-3,482 <i>SI*</i> -3,482 <i>SI*</i>	0,130 <i>0,099 SI**</i>
2005 – 2014	-2,067 <i>0,258 SI***</i> -2,067 <i>0,258 SI***</i>	-3,661 <i>0,024 NO</i> -3,661 <i>0,024 NO</i>	-0,759 <i>0,387 SI***</i> 0,759 <i>0,387 SI***</i>	-2,990 <i>SI**</i> -2,990 <i>SI**</i>	0,158 <i>0,043 SI*</i>
2006 – 2015	-2,056 <i>0,262 SI***</i> -2,056 <i>0,262 SI***</i>	-2,130 <i>0,5281 SI***</i> -2,130 <i>0,5281 SI***</i>	-2,120 <i>0,032 SI*</i> -2,120 <i>0,032 SI*</i>	-2,122 <i>SI***</i> -2,122 <i>SI***</i>	0,150 <i>0,052 SI**</i>
2007 – 2016	-2,440 <i>0,130 SI***</i> -2,440 <i>0,130 SI***</i>	-1,770 <i>0,644 SI***</i> -1,770 <i>0,644 SI***</i>	-2,048 <i>0,038 SI*</i> -2,048 <i>0,038 SI*</i>	-2,163 <i>SI***</i> -2,163 <i>SI***</i>	0,152 <i>0,049 SI*</i>

Continua nella pagina seguente

2008 – 2017	-1,733 <i>0,387 SI***</i>	-2,465 <i>0,333 SI***</i>	-1,813 <i>0,066 SI**</i>	-2,777 <i>SI***</i>	0,129
	-1,733 <i>0,387 SI***</i>	-2,465 <i>0,333 SI***</i>	-1,813 <i>0,066 SI**</i>	-2,777 <i>SI***</i>	<i>0,100 SI***</i>
2009 – 2018	-1,473 <i>0,503 SI***</i>	-2,863 <i>0,214 SI***</i>	-1,603 <i>0,100 SI***</i>	-3,18 <i>SI**</i>	0,101
	-1,473 <i>0,503 SI***</i>	-2,863 <i>0,214 SI***</i>	-1,603 <i>0,100 SI***</i>	-3,18 <i>SI**</i>	<i>> .10 SI***</i>

Conclusioni

Arriviamo ora all'interpretazione dei risultati ottenuti, concentrandoci singolarmente sulle varie serie annuali e trimestrali per le differenti basi di dati e test condotti. Per la serie del tasso di crescita annuale, a cadenza trimestrale, degli investimenti fissi lordi i testi di radice unitaria condotti forniscono dei risultati, a primo impatto, di facile interpretazione.

Infatti, notiamo che l'ipotesi di radice unitaria viene accettata per il campione completo di dati e per la maggior parte dei sotto intervalli considerati; per gli ultimi anni, però, la serie smentisce l'ipotesi di radice unitaria, quindi ritornando alla situazione di stazionarietà.

C'è da considerare che i test condotti (ADF, ADF con trend, ADF-GLS e ADF-GLS con trend) hanno un grado di precisione crescente: se i risultati del test ADF risultano ambigui, sono più corretti e univoci i test ADF con trend e ADF-GLS con trend. Inoltre, un'altra causa per la discrepanza dei nostri risultati è la scelta del numero dei ritardi nello svolgimento del test, determinato ogni volta dal software statistico secondo il criterio applicato. Questo può generare un'incoerenza nei risultati dei dati, ma garantisce imparzialità relativamente ai ritardi considerati.

Per ultimo, l'utilizzo di cinque test differenti ci aiuta ad ottenere risultati il più realistici possibile.

Osservando più attentamente l'andamento del valore della statistica test e del p-value dei vari test, si può notare che hanno lo stesso andamento: nei primi anni considerati c'è l'accettazione dell'ipotesi di radice unitaria, per poi arrivare, durante il periodo 1994 – 2000, a un rigetto dell'ipotesi, negli anni a seguire si ritorna all'accettazione di ipotesi di non stazionarietà, fino ad una nuova conclusione di stazionarietà degli ultimi periodi considerati. Si potrebbe pensare che la serie in realtà sia stazionaria, come da accordo alle teorie statistiche ed economiche, ma che vi siano stati dei cambiamenti importanti nell'economia del paese, che siano di tipo finanziario o economico.

Infatti, come spiegato nel capitolo 2, durante il primo periodo vi è stata la crisi del 1992, con la svalutazione della lira, durante il quale l'Italia cercava di recuperare la competitività perduta nei mercati mondiali; nel secondo periodo vi è stata l'adesione dell'Italia alla moneta unica, un grande cambiamento che ha inasprito il periodo di tensione che l'Italia stava già vivendo: il

Paese era in recessione (il PIL aveva un tasso di crescita negativo del -0,1%) e secondo uno studio il 42% degli italiani registrò, durante quel periodo, un calo del reddito individuale e subì una crescita non indifferente del costo della vita.¹⁷

Vi sono poi state le due crisi più segnanti: la crisi mondiale iniziata nel tardo 2007, e la crisi del debito sovrano, avvenuta nel periodo 2010-2012.

Non vi è stata ancora una ripresa totale da questi avvenimenti, il livello antecedente la crisi non è stato recuperato, sia con riguardo alla serie presa in esame, sia con riguardo al PIL del Paese. Queste conclusioni generali possono essere considerate significative se vi è effettivamente evidenza di questi break strutturali, associati a queste situazioni appena descritte: break strutturali che modificano l'andamento della serie in modo permanente, ma non influiscono sull'ipotesi di base di stazionarietà.

Per questo abbiamo condotto il test di Zivot-Andrews, il quale testa l'ipotesi di radice unitaria, considerando la presenza di break strutturali. Vengono evidenziati, come punti di rottura, tre momenti: al terzo trimestre del 2001 vi è il primo break; al secondo trimestre del 2004, dove la statistica test raggiunge il suo valore minimo e durante il secondo trimestre del 2010. Tutte in concomitanza con cambiamenti macroeconomici importanti. Quindi il test Zivot-Andrews rifiuta l'ipotesi nulla di presenza di radice unitaria nella serie, ovvero la non-stazionarietà non dipende dalle caratteristiche intrinseche della serie storica.

Ciò ci porta a concludere che il tasso di crescita può essere considerato stazionario, e che la serie storica degli investimenti fissi lordi è una serie integrata di ordine uno. Entrambe sono molto sensibili ai cambiamenti o break strutturali dell'economia, basti osservare la loro variabilità nel corso degli anni. Nei periodi di stabilità economica, infatti, il tasso di crescita è stazionario, ma reagisce in modo brusco e considerevole a diversi eventi, soprattutto negativi. Osservando, poi, i risultati dei test per i campioni annuali, la situazione è un simile.

Visto che l'intervallo considerato è più ampio, l'ipotesi di radice unitaria viene accettata in modo più sicuro, dato che i suddetti intervalli comprendono con molta probabilità le osservazioni corrispondenti ai break strutturali. C'è da aggiungere, però, che il p-value segue lo stesso andamento, riportato sopra per i dati trimestrali, ma in modo meno deciso rispetto alla serie a cadenza trimestrale. Inoltre, negli ultimi periodi il p-value è diminuito, avvalorando la nostra ipotesi di una nuova stazionarietà degli ultimi anni considerati. Per quanto riguarda la modellazione ARMA, il modello che più si adatta è il modello ARMA(2,3).

Il modello cattura notevolmente le oscillazioni della variabile dipendente: infatti da una

¹⁷ <https://www.repubblica.it/2004/a/sezioni/economia/crisi2003/crisi2003/crisi2003.html>

variabilità del 0,054897 si arriva ad una variabilità del 0,015877; inoltre tutti i coefficienti vengono considerati significativi al livello di significatività, almeno, del 5%. Il modello non presenta autocorrelazione dei residui, come dimostrato dal test di Test di Ljung-Box, il quale ha un p-value del 0,09, che ci porta ad accettare l'ipotesi nulla al 5%.

Per ultimo, osservando il grafico dei residui rispetto al tempo (figura 2.12) e il correlogramma dei residui (figura 2.13) si può notare come questi siano incorrelati e presentino un comportamento aleatorio, come si voleva dimostrare. Questo permette di ritenere valido il modello e, essendo la serie presa in esame stazionaria, diviene possibile derivarne i momenti e fare previsioni ed inferenza sui dati stessi, e calcolare i possibili valori futuri della serie e i rispettivi intervalli di confidenza.

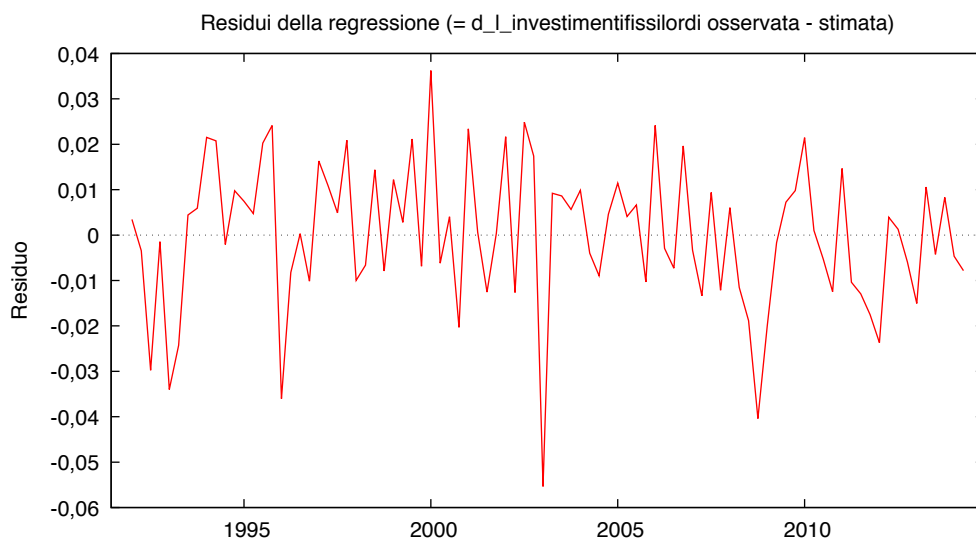


Figura 2.12: grafico dei residui della regressione rispetto al tempo

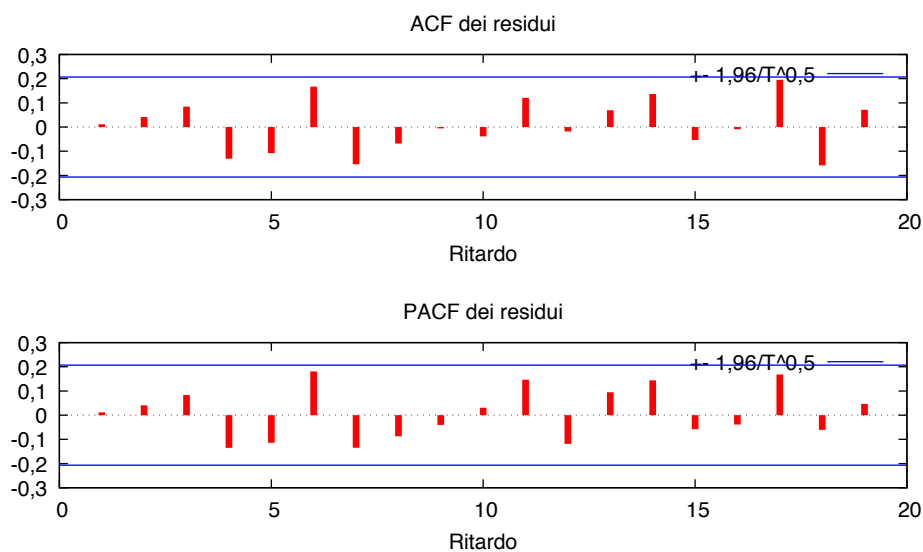


Figure 2.13: correlogramma dei residui rispetto al tempo

Appendice A

Correlogramma e la funzione di autocorrelazione

L'autocorrelazione di un processo stocastico corrisponde al coefficiente di correlazione di Pearson, ρ , tra realizzazione di uno stesso processo in istanti differenti. Dato un processo y_t e considerate due realizzazioni al tempo t e s , si definisce l'autocorrelazione come:

$$\rho_{t,s} = \frac{E[(y_t - \mu_t)(y_s - \mu_s)]}{\sigma_t \sigma_s}$$

Qualora i parametri fossero definiti, questa assume valori nell'intervallo $[-1,1]$.

Ove il processo stocastico sia stazionario, quindi media e varianza non dipendono dal tempo, l'autocorrelazione dipende solo dall'intervallo temporale $\tau = t - s$:

$$\rho_\tau = \frac{E[(y_t - \mu)(y_s - \mu)]}{\sigma^2} = \frac{\gamma_\tau}{\gamma_0}$$

La funzione di autocorrelazione presenta le seguenti proprietà:

- $\rho_0 = 1$
- $\rho_\tau = \rho_{-\tau}$
- $|\rho_\tau| \leq \rho_0$

Lo stimatore della funzione di autocorrelazione risulta essere:

$$\widehat{\rho}_k = \frac{1}{(n-k)\widehat{\sigma}^2} \sum_{t=1}^{n-k} (y_t - \widehat{\mu})(y_{t+k} - \widehat{\mu})$$

Un modo abbastanza semplice per vedere se una serie presenta autocorrelazione è quella di tracciarne il correlogramma con la funzione ACF, si tratta di un grafico cartesiano in cui per ciascun ritardo viene indicato il valore dell'autocorrelazione.

La procedura è un classico test statistico del coefficiente con $H_0 : \rho = 0$, e distribuzione approssimata alla normale.

Nello specifico l'intervallo di confidenza risulta essere

$$-z_{1-\frac{\alpha}{2}}SE(\rho_\tau) \leq \rho_\tau \leq z_{1-\frac{\alpha}{2}}SE(\rho_\tau)$$

Dove $z_{1-\frac{\alpha}{2}}$ indica il valore del percentile della distribuzione, scelto α , e $SE(\rho_\tau)$ errore standard della distribuzione.

Valori esterni a tale intervallo indicano la presenza di autocorrelazione significativa.

La funzione di autocorrelazione parziale fornisce, invece, il grado di autocorrelazione tra due osservazioni della serie storica, al netto delle osservazioni intermedie.

In termini più formali, l'autocorrelazione parziale del ritardo τ è la correlazione tra l'osservazione y_t e $y_{t-\tau}$, a cui viene rimossa la dipendenza lineare di y_t , sui ritardi $y_{t-1}, \dots, y_{t-\tau+1}$.

Per calcolare la PACF, bisogna partire dalla nozione di autocovarianza di un processo AR(p).

Sappiamo che $\gamma_\tau = E[y_{t-\tau}y_t]$ per processi con media nulla, allora

$$\gamma_\tau = E [y_{t-\tau}(\alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p})]$$

Per $\tau = 0, 1, \dots, p$

$$\gamma_0 = \alpha_1 \gamma_1 + \dots + \alpha_p \gamma_p = \sigma^2$$

$$\gamma_1 = \alpha_1 \gamma_0 + \alpha_2 \gamma_1 + \dots + \alpha_p \gamma_{p-1}$$

...

$$\gamma_p = \alpha_1 \gamma_{p-1} + \alpha_2 \gamma_{p-2} + \dots + \alpha_p \gamma_0$$

Dividendo le ultime equazioni per γ_0 otteniamo:

$$\rho_1 = \alpha_1 + \alpha_2 \rho_1 + \dots + \alpha_p \rho_{p-1}$$

$$\rho_2 = \alpha_1 \rho_1 + \alpha_2 + \dots + \alpha_p \rho_{p-2}$$

...

$$\rho_p = \alpha_1 \rho_{p-1} + \alpha_2 \rho_{p-2} + \dots + \alpha_p$$

Il calcolo della PACF si basa sull'individuazione dei coefficienti dei parametri di un AR(p), attraverso i valori $\rho_1 \dots \rho_p$.

Test di Ljung-Box e di Box-Pierce

Una statistica che può essere utilizzata per verificare l'assenza di autocorrelazione è una opportuna combinazione lineare dei coefficienti di autocorrelazione dei residui ε_t :

$$LB = n(n+2) \sum_{t=1}^k \frac{\varepsilon_t^2}{n-t}$$

dove k è un intero prescelto. Se è vera l'ipotesi nulla, cioè assenza di autocorrelazione, la statistica LB si distribuisce asintoticamente secondo una variabile casuale *chi-quadro* con k gradi di libertà.

Un test analogo a quello di Ljung-Box è quello che si basa sulla statistica proposta da Box-Pierce:

$$BP = n \sum_{t=1}^k \varepsilon_t^2$$

Le due statistiche differiscono semplicemente per il diverso sistema di ponderazione adoperato, ma asintoticamente hanno la medesima distribuzione. Si dimostra, tuttavia, che LB ha una convergenza più rapida alla distribuzione asintotica e, per tale motivo, risulta preferibile al test di Box-Pierce.

Appendice B

White noise

Un *white noise*, o rumore bianco, è un processo aleatorio che descrive un segnale caratterizzato da una traiettoria costante.

Il nome *bianco* deriva dalle proprietà dello spettro della luce bianca, che è costante in quanto questa è formata dalla sovrapposizione di onde elettromagnetiche di tutte le frequenze visibili e di intensità simile a ogni frequenza.¹⁸

Questo presenta determinate caratteristiche:

$$\begin{aligned}E(\varepsilon_t) &= 0 \\E(\varepsilon_t^2) &= V(\varepsilon_t) = \sigma^2 \\ \gamma_k &= 0 \quad \text{per } |k| > 0\end{aligned}$$

Il processo è quindi stazionario. Dal punto di vista probabilistico, si distingue in genere tra rumore bianco in senso forte e in senso debole. Un processo aleatorio ε_t è un *white noise in senso forte* se la serie è composta da variabili aleatorie indipendenti e identicamente distribuite, caratterizzate da media nulla e varianza costante. Il white noise è il processo stocastico più semplice che si possa immaginare, infatti, è un processo che possiede momenti fino al secondo ordine, costanti nel tempo e non danno al processo alcuna memoria di sé. Il processo ε_t è, invece, un *white noise in senso debole* se viene meno l'assunzione di indipendenza.

Si ha un white noise gaussiano se, inoltre, ciascun ε_t è distribuito secondo una distribuzione normale; in tal caso non c'è distinzione tra rumore bianco forte o debole, poiché indipendenza e incorrelazione coincidono.

In un contesto di serie storiche il white noise è generalmente usato per la costruzione di modelli più complessi. Infatti, nel modello a medie mobili $MA(1)$, la variabile y_t dipende soltanto dal termine di errore ai tempi t e $t-1$. Più in generale, un processo $MA(q)$ è costituito dalla combinazione di q variabili di un white noise in tempi successivi. Il caso di un processo autoregressivo, o di un $ARMA(p,q)$ si riconduce, sotto opportune condizioni che garantiscono la stazionarietà del processo, a un processo a media mobile $MA(\infty)$. Tali processi possono cioè essere espressi come combinazione lineare infinita di elementi di un rumore bianco.

¹⁸ http://www.treccani.it/enciclopedia/rumore-bianco_%28Dizionario-di-Economia-e-Finanza%29/

Random walk

Si dice *random walk*, o passeggiata aleatoria, un processo stocastico definito come

$$y_t = y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN$$

Supponendo $y_0 = 0$, si ha

$$y_1 = 0 + \varepsilon_1 = \varepsilon_1$$

$$y_2 = 0 + \varepsilon_2 = \varepsilon_1 + \varepsilon_2$$

...

$$y_T = \sum_{t=1}^T \varepsilon_t$$

Quindi:

$$E(y_t) = 0 \quad V(y_t) = V \sum_{t=1}^T \varepsilon_t = T\sigma^2$$

Essendo la varianza non costante, ma aumenta nel tempo, questo processo non è stazionario.

Appendice C

Script

Includiamo lo script utilizzato per calcolare i test di radice unitaria.

```
include gig.gfn
/*ADF TEST WITH CONSTANT*/
smpl 1991:1 2014:2
adf -1 d_l_investmentifissilordi --c --verbose --test-down=AIC
adf -1 d_l_investmentifissilordi --c --verbose --test-down=BIC
loop i=1991:1..2005:2
smpl i i+39
adf -1 d_l_investmentifissilordi --c --verbose --test-down=AIC
adf -1 d_l_investmentifissilordi --c --verbose --test-down=BIC
endloop

/*ADF TEST WITH CONSTANT AND TREND*/
smpl 1991:1 2014:2
adf -8 d_l_investmentifissilordi --ct --test-down=AIC
adf -1 d_l_investmentifissilordi --ct --test-down=BIC
loop i=1991:1..2005:2
smpl i i+39
adf -8 d_l_investmentifissilordi --ct --test-down=AIC
adf -1 d_l_investmentifissilordi --ct --test-down=BIC
endloop

/*ADF-GLS WITH CONSTANT*/
smpl 1991:1 2014:2
adf -1 d_l_investmentifissilordi --c --test-down=AIC --gls --
    perron-qu
adf -1 d_l_investmentifissilordi --c --test-down=BIC --gls --
    perron-qu
loop i=1991:1..2005:2
smpl i i+39
adf -8 d_l_investmentifissilordi --c --test-down=AIC --gls --
    perron-qu
```

```
adf -1 d_l_investmentifissilordi --c --test-down=BIC --glS --  
perron-qu  
endloop
```

/*ADF-GLS WITH CONSTANT AND TREND*/

```
smpl 1991:1 2014:2  
adf -1 d_l_investmentifissilordi --ct --test-down=AIC --glS --  
perron-qu  
adf -1 d_l_investmentifissilordi --ct --test-down=BIC --glS --  
perron-qu  
loop i=1991:1..2005:2  
smpl i i+39  
adf -1 d_l_investmentifissilordi --ct --test-down=AIC --glS --  
perron-qu  
adf -1 d_l_investmentifissilordi --ct --test-down=BIC --glS --  
perron-qu  
endloop
```

/*KPSS TEST*/

```
smpl 1991:1 2014:2  
kpss -5 d_l_investmentifissilordi --trend  
  
loop i=1991:1..2005:2  
smpl i i+39  
kpss -5 d_l_investmentifissilordi --trend  
endloop
```

Bibliografia

1. Warren Milton Persons, *Indices of General Business Conditions*. Harvard Univ. Committee on Economics research, 1919.
2. Gottman, J.M., *Time series analysis. A comprehensive introduction for social scientists*, Cambridge 1981
3. Brockwell Peter and Davis Richard. *Time Series: Theory and Methods, Second Edition*, Springer-Verlag 1991
4. Peter J. Brockwell Richard A. Davis, *Introduction to Time Series and Forecasting*, Springer, 1996
5. *Sistema europeo nazionale dei conti SEC*, Istat, Istituto nazionale di statistica.
6. Gómez Victor e Maravall Agustin, *Estimation, Prediction and Interpolation for Nonstationary Series With the Kalman Filter*, 1996
7. Dickey David and Fuller Wayne, *Distribution of the estimators for autoregressive time series with a unit root*, Journal of the American statistical association, 1979.
8. Gideon E. Schwarz, Bayesian information criterion, *Estimating the dimension of a model*, Annals of Statistic, vol. 6, 1978
9. Graham Elliott, Thomas J. Rothenberg and James H. Stock, *Efficient tests for an autoregressive unit root*, Econometrica, vol. 64, no. 4, 1996
10. Kwiatkowski, Phillips, Schmidt, Shin, *Testing the null hypothesis of stationarity against the alternative of a unit root*, Journal of Econometrics, 1992
11. Pierre Perron, *The great crash, the oil price shock, and the unit root hypothesis*, *Econometrical*, Journal of the Econometric Society, 1989
12. Eric Život and Donald W.K. Andrews, *Further evidence on the Great Crash, the oil price shock, and the unit root hypothesis*, Journal of Business and Economic Statistics, 1992

9722 parole.