



Università
Ca'Foscari
Venezia

Corso di Laurea magistrale
in Economia e Finanza

Tesi di Laurea

Un approccio econometrico sul NASDAQ Composite

Modello di regressione per la previsione dell'indice
statunitense

Relatore

Ch.mo Prof. Domenico Sartore

Correlatore

Ch.ma Prof.ssa Monica Billio

Laureando

Riccardo Livio Toffoli

Matricola

887586

Anno Accademico 2021 / 2022

INDICE

Introduzione.....	6
CAPITOLO I - Analisi descrittiva delle variabili.....	9
1.1 L'indice NASDAQ Composite	10
1.1.1 Andamento ed evoluzione del NASDAQ Composite	11
1.2 Le variabili esplicative.....	14
1.2.1 TB: 3-Month Treasury Bill Secondary Market Rate	14
1.2.2 CC: Consumer sentiment index	14
1.2.3 BOND: Long-Term Government Bond Yields: 10-year	15
1.2.4 PCE: Personal consumption expenditures.....	15
1.2.5 PPIT: Producer price index by industry: data processing, hosting and related services	16
1.2.6 EM: Employment level	16
1.2.7 IPT: Industrial production, construction supplies.....	16
1.2.8 RET: Total retail trade.....	17
1.3 Analisi grafica delle variabili esplicative	18
1.4 Analisi statistica della serie endogena.....	20
1.5 Introduzione alla teoria dei modelli per l'eteroschedasticità condizionale	31
1.5.1 Modello per l'eteroschedasticità della serie endogena	33
1.6 Analisi di integrazione delle variabili esplicative	38
1.6.1 La variabile "TB"	38
1.6.2 La variabile "CC"	44
1.6.3 La variabile "BOND"	47
1.6.4 La variabile "PCE"	51
1.6.5 La variabile "PPIT"	54
1.6.6 La variabile "EM"	57

1.6.7 La variabile "IPT"	61
1.6.8 La variabile "RET"	64
1.7 Ulteriori variabili utili ai fini dell'equazione di breve periodo	68
1.7.1 La variabile "LF"	68
1.7.2 La variabile "USD"	71
1.7.3 La variabile "CHUS"	74
CAPITOLO II - Procedura a due passi di Engle-Granger	77
2.1 Analisi di cointegrazione e stima dell'equazione statica	77
2.2 L'equazione di breve periodo	94
CAPITOLO III - Previsioni dal modello econometrico e confronto con il benchmark	109
3.1 Previsione statica dall'equazione statica	113
3.2 Previsione statica dall'equazione dinamica	115
3.3 Previsione dinamica dal modello dinamico	118
3.4 Confronto della previsione dal modello econometrico con il modello ARIMA del benchmark	122
3.4.1 Previsione statica del benchmark	122
3.4.2 Previsione dinamica del benchmark	125
Conclusioni	129
Appendice	134
Bibliografia	151
Sitografia	153
Ringraziamenti	155

Introduzione

Il NASDAQ Composite è uno tra gli indici più rilevanti a livello mondiale, nonché il paniere di riferimento delle società quotate del settore tech statunitense.

Il presente elaborato intende proporre un approccio econometrico per la descrizione dell'indice NASDAQ Composite, finalizzato a sviluppare l'aspetto della cointegrazione attraverso la procedura a due passi di Engle-Granger e a fornire delle previsioni efficaci per il primo anno della pandemia da Covid-19.

Nel primo capitolo vengono analizzate le variabili utilizzate per la caratterizzazione dell'indice attraverso un approccio descrittivo delle stesse. In prima istanza vengono selezionate quelle serie storiche utili ai fini della spiegazione dell'indice. Queste variabili verranno definite esplicative (o esogene), mentre il NASDAQ Composite verrà considerato come serie dipendente (o endogena).

Entrambe le categorie vengono studiate sia secondo un criterio qualitativo che quantitativo. Nel primo caso viene approfondito il significato della serie, ne viene inquadrato l'andamento e viene individuato il possibile effetto sulla variabile endogena. A seguire, viene effettuata un'analisi statistica delle serie storiche, le quali vengono sottoposte a molteplici test, tra i quali quello inerente alla normalità, alla linearità e alla radice unitaria. La disamina de quo sarà più approfondita nel caso della variabile endogena, in corrispondenza della quale vengono individuati dei modelli per l'interpretazione della serie storica e dell'eteroschedasticità condizionale, che verranno ripresi nell'ultimo paragrafo.

Il secondo capitolo rappresenta il corpo centrale dell'elaborato. Accertata infatti la non stazionarietà della serie endogena e della maggior parte delle esplicative, le variabili vengono impiegate nella procedura a due passi di Engle-Granger. Tale procedura è volta alla stima dell'equazione di lungo periodo (statica) e di breve periodo (dinamica) in forma Error Correction Mechanism (ECM). La prima delinea le relazioni di lungo periodo tra la

variabile dipendente e le variabili esogene, andando a descrivere come queste ultime impattano sull'indice. Dall'equazione in questione viene generata la componente dei residui, sulla quale vengono effettuati test volti a verificarne la stazionarietà.

Nel paragrafo successivo la componente dei residui viene inserita nell'equazione dinamica in forma ECM, rappresentativa di una relazione di breve periodo tra le variabili esplicative e quella dipendente.

Verificata la cointegrazione dell'equazione statica, nonché l'impatto delle variabili esogene sull'endogena e determinata l'equazione dinamica in forma ECM, si verifica l'assenza di eteroschedasticità e di autocorrelazione seriale nei residui dell'equazione di breve periodo.

Le fasi sopraelencate sono una prerogativa fondamentale ai fini della previsione della serie endogena in un "futuro virtuale", che interessa il periodo relativo al primo anno della crisi pandemica. A livello del terzo e ultimo capitolo si stimano, dunque, le previsioni derivanti dall'equazione dinamica e dalla curva di lungo periodo per l'intervallo compreso tra il gennaio 2020 e il gennaio 2021. La bontà previsiva è valutata in riferimento a misure come il coefficiente di disuguaglianza di Theil, gli errori percentuali assoluti in media e la Covariance Proportion. Le previsioni generate vengono opportunamente confrontate con l'andamento effettivo della variabile dipendente, in modo tale da osservare graficamente quanto le stime nel "futuro virtuale" generate dai modelli econometrici divergano dal trend registrato dall'indice. Sarà di estremo interesse osservare in quale misura le curve stimate siano in grado di intercettare l'andamento della serie endogena in uno dei periodi più nefasti della storia più recente.

L'ultimo paragrafo riprende quanto era stato proposto con l'analisi statistica della serie endogena nel primo capitolo, con il cosiddetto modello "benchmark".

Si propongono infatti due previsioni, una statica e una dinamica, basate sulla sola variabile endogena definita da un processo di tipo ARIMA (p,d,q) e depurata dagli effetti di dipendenza rispetto alle variabili esplicative della regressione econometrica. Le curve stimate vengono valutate con gli stessi criteri relativi alle previsioni effettuate in precedenza. La bontà previsiva risultante dalle previsioni dal modello "benchmark" viene confrontata con quanto ottenuto dalle curve generate dal modello econometrico e da quanto effettivamente registrato dalla serie endogena oggetto di studio. Quest'ultimo

punto sarà essenziale per valutare l'efficacia e la robustezza del modello econometrico e delle relative previsioni.

CAPITOLO I - Analisi descrittiva delle variabili

L'analisi descrittiva delle variabili inserite nel modello, si sostanzia in un inquadramento generale delle variabili ritenute utili ai fini della regressione uniequazionale. Nello specifico si tratta di esplicitare le caratteristiche delle variabili supposte esogene, come la fonte di provenienza nonché il database al quale si accede per scaricare la serie storica, una breve descrizione della stessa e i possibili effetti sull'endogena. Tutte le variabili vengono scaricate su base mensile, per un periodo campionario compreso tra l'ottobre del 2002 e il gennaio del 2021. Per l'analisi delle variabili verranno utilizzati i software EViews9 e R.

Le analisi preliminari, che condurranno alla stima dell'equazione di lungo e di breve periodo, sono basate sull'intervallo compreso tra l'ottobre del 2002 e il gennaio del 2020 per un totale di 208 osservazioni mensili. Le previsioni effettuate si concentrano nei 12 mesi che seguono il suddetto intervallo di stima, in un futuro virtuale che si protrae dal gennaio 2020 al gennaio 2021, per un totale di 12 osservazioni.

Le suddivisioni temporali di cui sopra non sono affatto casuali, ma giustificate dal verificarsi di alcuni eventi storici che hanno impattato significativamente sulle variabili. Il punto di partenza della disamina si colloca alla fine del 2002, per ragioni che risiedono nella variabile endogena del NASDAQ Composite. L'indice, come riportato nel paragrafo dedicato, negli anni che precedono il sample, ha subito un impatto notevole causato dalla bolla delle "Dot-com". Lo shock ha generato un allontanamento rispetto ai fondamentali non indifferente delle società tech (o legate al settore dell'internet in generale), portando conseguenti distorsioni di panieri come il NASDAQ. Quest'ultimo ha risentito pesantemente degli effetti speculativi della bolla. L'inserimento di un intervallo periodale inclusivo di questi shock avrebbe potuto indurre a delle distorsioni nell'interpretazione delle equazioni stimate e, più in generale, nella relazione tra variabili esplicative e l'endogena. Il periodo del futuro virtuale è stato selezionato in coerenza alle finalità dello studio proposto, ovvero a determinare delle previsioni efficaci per il periodo successivo

allo scoppio della crisi pandemica, esaminando e osservando eventuali gap rispetto all'effettiva realizzazione dell'indice.

1.1 L'indice NASDAQ Composite

Il NASDAQ (“National Association of Securities Dealers Automated Quotation”) è un mercato di borsa statunitense istituito nel 1971. Nel complesso conta più di tremila società quotate, attestandosi come la borsa valori con maggiore numero di titoli dopo il NYSE. L'indice rappresentativo di questo listino è il NASDAQ Composite, che include la maggioranza dei titoli quotati, circa 3000. Aspetto caratterizzante dell'indice risiede nel fatto che i titoli sono afferenti al settore tecnologico. Ciò suggerisce quanto siano fortemente rilevanti gli andamenti a livello macroeconomico di questo settore e quanto l'indice risenta di notizie relative a ciò in maniera importante.

Il NASDAQ è stato il primo sistema borsistico esclusivamente elettronico. Il passaggio graduale ad un sistema di questo tipo, portato a termine alla fine degli anni 80, ebbe il merito principale di ridurre le asimmetrie informative, di garantire maggiore accessibilità ai prodotti finanziari e di ridurre le inefficienze dei mercati.

Si tratta inoltre di un paniere calcolato sulla base della ponderazione per la capitalizzazione dei titoli. In altre parole, il suo valore è basato sulla sommatoria dei prezzi di ogni quotazione ponderata per ogni singolo peso. In quanto ai criteri di inclusione, un titolo per essere quotato nel NASDAQ, deve essere presente esclusivamente su questo listino, eccezion fatta per quei titoli che sono già stati quotati in modo permanente e continuativo in un'altra borsa da prima del 2004. È solo questa la discriminante per la quale può essere concessa la doppia quotazione.

Nel paniere in questione non è compresa soltanto la classe di attività mobiliari che si esaurisce con le azioni ordinarie, differentemente da quanto avviene in altri casi. Vengono inclusi nell'indice, infatti, anche partecipazioni in limited partnership, fondi comuni di investimento immobiliari, azioni di interesse benefico e tracking stocks. In linea con le altre borse, le società quotate sono soggette a revisione contabile certificata e al deposito annuale dei relativi bilanci.

1.1.1 Andamento ed evoluzione del NASDAQ Composite

Il grafico sottostante mostra l'andamento della serie endogena oggetto di studio, dagli anni 90' ad oggi:

Figura 1.1- Andamento dell'indice NASDAQ Composite dal 1990/10 al 2022/10



Si può constatare un trend nel complesso crescente, che ha raggiunto un punto di massimo nei primi mesi del 2020, per poi affrontare una fase decrescente a partire da marzo dello stesso anno.

Guardando nel dettaglio alle variazioni registrate dal NASDAQ, fino ai primi mesi del 2000 l'indice ha manifestato un andamento lineare e crescente, arrivando ad un massimo relativo il 10 marzo del 2000. Qui, infatti, la serie ha subito gli effetti della bolla speculativa delle Dot-com, che ha impattato nelle società legate al tech e al digitale. Il fatto che il NASDAQ abbia risentito più severamente di questa bolla rispetto ad altri panieri è coerente con la composizione delle società che lo caratterizzano. A partire dall'autunno dello stesso anno i prezzi delle società del settore digitale crollano

drammaticamente poiché sopravvalutate rispetto al loro valore intrinseco. Il NASDAQ subisce questa contrazione, per poi assestarsi a partire dal 2002. Da qui realizza un trend in leggera crescita sino alla grande recessione. L’impatto sull’indice non sarà tuttavia significativo come quello precedente, mostrando una fase di decrescita a partire dalla fine del 2007 fino a marzo 2009. Il default di Lehman Brothers e delle altre istituzioni finanziarie del 2008 porta l’economia US ai minimi storici, nonostante la serie in oggetto ne risenta in misura decisamente minore rispetto ad altri indici USA, così come in relazione ad altre misure macroeconomiche. Il NASDAQ ha riportato delle variazioni piuttosto singolari, sia in termini percentuali che assoluti, mostrando uno shock maggiore in corrispondenza della bolla del 2000 rispetto a quello generato dalla crisi finanziaria del 2007-2008. A titolo esemplificativo si riporta il grafico del NYSE che, così come il Wilshire 5000 o il Dow Jones, presenta un comportamento differente rispetto a quanto emerge dalla figura 1.1.

Figura 1.2- Andamento dell’indice NYSE dal 1999/01 al 10/2022



Come riportato poc'anzi, tuttavia, l'evoluzione oggetto di discussione è in linea con quanto atteso da un qualsivoglia paniere i cui pesi maggiori derivano da società del settore tecnologico e digitale.

A seguire, il NASDAQ affronta quindi una fase di crescita pressochè costante dalla prima metà del 2009 sino al 2020, anno della crisi pandemica. Anche in questa circostanza l'indice è rappresentativo di una contrazione minore rispetto ad altri panieri statunitensi (come lo stesso NYSE) e mondiali. Dopo il minimo registrato nel marzo del 2020, ancor più degna di nota è la tendenza successiva dell'indice che passa da un prezzo di 6879\$ (2020/03/15) a 13215\$ (2021/03/14) in un anno. La variazione percentuale nello stesso intervallo è maggiore rispetto a quella degli altri indici statunitensi menzionati, in coerenza con la forte espansione delle società tech degli ultimi anni. Se da un lato si è assistito ad una tendenza rialzista maggiore dal minimo del 2020 per tutto il 2021, con il 2022 si è verificata una contrazione più significativa, dovuta ad una somma di molteplici fattori. Tra questi l'aumento dell'inflazione e il conseguente aumento dei tassi della Fed (in una prima fase le sole aspettative del suddetto rialzo) ne sono gli elementi fondanti. Senza entrare nelle specifiche macroeconomiche, altre motivazioni risiedono nell'incertezza geopolitica dovuta al conflitto Russia-Ucraina e nel rallentamento più cospicuo rispetto ad altri, del settore tecnologico.

Concentrando l'attenzione nella serie endogena analizzata, il periodo d'analisi del NASDAQ Composite ai fini dell'analisi econometrica proposta, non corrisponde con l'anno in cui è stato introdotto l'indice stesso nel 1971 e nemmeno con l'analisi grafica di cui sopra. Questo per diversi ordini di ragioni.

In prima istanza per la maggiore disponibilità di dati relativi alle variabili esplicative negli anni più recenti.

Un altro fattore che ha condotto alla scelta di un periodo campionario differente è relativo all'evoluzione storica dell'indice.

Da ultimo il fatto che i pesi delle società nell'indice sono cambiati considerevolmente a seguito della crisi delle Dot-com.

Ancora una volta si è voluto isolare dalla regressione l'effetto di variabili che hanno inciso sulla bolla delle Dot-com, dando priorità a quelle che di norma hanno effetti sull'indice.

1.2 Le variabili esplicative

Di seguito si riporta un elenco e una breve descrizione delle variabili esplicative utilizzate nel presente progetto¹.

1.2.1 TB: 3-Month Treasury Bill Secondary Market Rate

La variabile in oggetto riguarda il rendimento a breve dei titoli di stato statunitensi. Misura quindi il rendimento nel mercato secondario di obbligazioni a breve termine emesse dagli Stati Uniti. L'andamento nel corso del tempo della variabile "TB" è indicativo del grado di rischio del paese emittente. Un aumento del tasso segnala un'esposizione ad un rischio più elevato per l'investitore e, di conseguenza, una remunerazione maggiore per lo stesso. Proprio per questo ci si attende un impatto negativo sull'endogena quando si andrà a stimare l'equazione statica.

La variabile "TB" è stata scaricata dal database "FRED", misurata in percentuale e con frequenza mensile. Non è destagionalizzata, quindi può essere necessaria una destagionalizzazione nel proseguo dell'analisi.

1.2.2 CC: Consumer sentiment index

È un indice di fiducia riferito ai consumatori. Un "sentiment" maggiore si riscontra in corrispondenza di fasi economiche in espansione, una variazione negativa nel caso di fasi di rallentamento ed incertezza, nelle quali vi è maggiore pessimismo tra i consumatori. Tale indicatore viene utilizzato per stimare la spesa o il risparmio futuro dei cittadini, a seconda che ci si trovi in una fase espansiva o di contrazione. Per le ragioni di cui sopra ci si attende un impatto negativo sull'endogena nel caso di variazioni negative dell'esplicativa, positivo laddove la fiducia e dunque, la spesa prevista dai consumatori, si prospetta in aumento.

¹ Le serie sono scaricate dal sito ufficiale della Federal Reserve Bank of St. Louis, nell'ambiente FRED Economic Data <http://research.stlouisfed.org>.

La variabile è denominata con l'acronimo "CC" ed è espressa come un indice su base mensile attingendo dal database di Fred. Anche in questo caso la variabile non è stata stagionalizzata, processo che verrà valutato successivamente qualora ne sussistano i presupposti.

1.2.3 BOND: Long-Term Government Bond Yields: 10-year

La variabile "BOND" si riferisce al rendimento a dieci anni dei titoli di stato statunitensi. Analogamente a quanto descritto per la variabile "TB", l'andamento del rendimento dei titoli di stato US è rappresentativo del rischio sopportato dagli investitori originato dalla sottoscrizione dei titoli di debito in discussione. L'emissione in questo caso riguarda titoli di debito a lungo termine e, di conseguenza, è comprensivo di un premio per il rischio derivante da una sottoscrizione per un arco temporale più ampio. Anche in questa fattispecie si prevede un impatto nel complesso negativo sull'endogena, sebbene gli effetti nel lungo periodo possano rivelarsi ambivalenti su un indice di borsa come il NASDAQ Composite.

La variabile verrà denominata con l'appellativo "BOND", misurata in percentuale su osservazioni mensili non stagionalizzate.

1.2.4 PCE: Personal consumption expenditures

Il "PCE" rappresenta una tra le variabili più rilevanti da un punto di vista macroeconomico per un Paese. Tale acronimo viene utilizzato per descrivere la spesa individuale per il consumo corrente di tutti i residenti degli Stati Uniti. In altri termini la variabile misura la quantità del reddito individuale destinato al consumo corrente per tutti i residenti del Paese in analisi. Per sottrazione, fornisce altresì una proxy dell'ammontare del risparmio complessivo individuale. La variabile è comprensiva della spesa nei beni durevoli e non durevoli del Paese di riferimento.

La serie "personal consumption expenditures" verrà indicata dall'acronimo "PCE", è scaricata dal database di FRED su base mensile, in miliardi di dollari. Vengono utilizzati

i prezzi reali e la serie è stata stagionalizzata con procedura S.A.A.R. (“seasonally adjusted annual rate”).

1.2.5 PPIT: Producer price index by industry: data processing, hosting and related services

La variabile supposta esogena “PPIT” è un indice dei prezzi alla produzione che misura le variazioni in media dei prezzi riferiti ai produttori nazionali per i loro output. In questo caso la variabile si concentra in un cluster specifico di prodotti, con particolare riferimento all'area data processing e dei servizi o dei prodotti informatico-tecnologici.

La presente variabile proviene dal database Fred ed è espressa come un indice con dati mensili. Non è stata stagionalizzata.

1.2.6 EM: Employment level

La variabile in questione misura in valore assoluto il numero di popolazione occupata. Viene stimata per la fascia di popolazione statunitense tra i 16 e i 64 anni. Il dato è sensibile al ciclo economico di riferimento e, nel lungo termine, alla politica economica dei governi sul sistema educativo e alle politiche di inserimento nel mondo del lavoro. Comprende i soli residenti negli Stati Uniti per la fascia di età indicata, che non sono impegnati nelle forze armate e ad esclusione della popolazione in carcere.

La variabile “EM” è misurata in migliaia, con osservazioni su base mensile aggiornate sul sito Fred e stagionalizzate.

1.2.7 IPT: Industrial production, construction supplies

La produzione industriale è un indice rappresentativo di tutti gli output reali del Paese di riferimento e copre settori che vanno da quello manifatturiero, energetico, fino alla catena di costruzioni e infrastrutture. Nel caso in discussione, ci si è concentrati su quest'ultime

serie, che oltre a fornire una proxy adeguata dell'indice generale si concentra sull'impianto delle costruzioni degli Stati Uniti.

La variabile denominata “IPT” è scaricata dalla banca dati Fred si presenta come un indice su base mensile ed è già stata destagionalizzata.

1.2.8 RET: Total retail trade

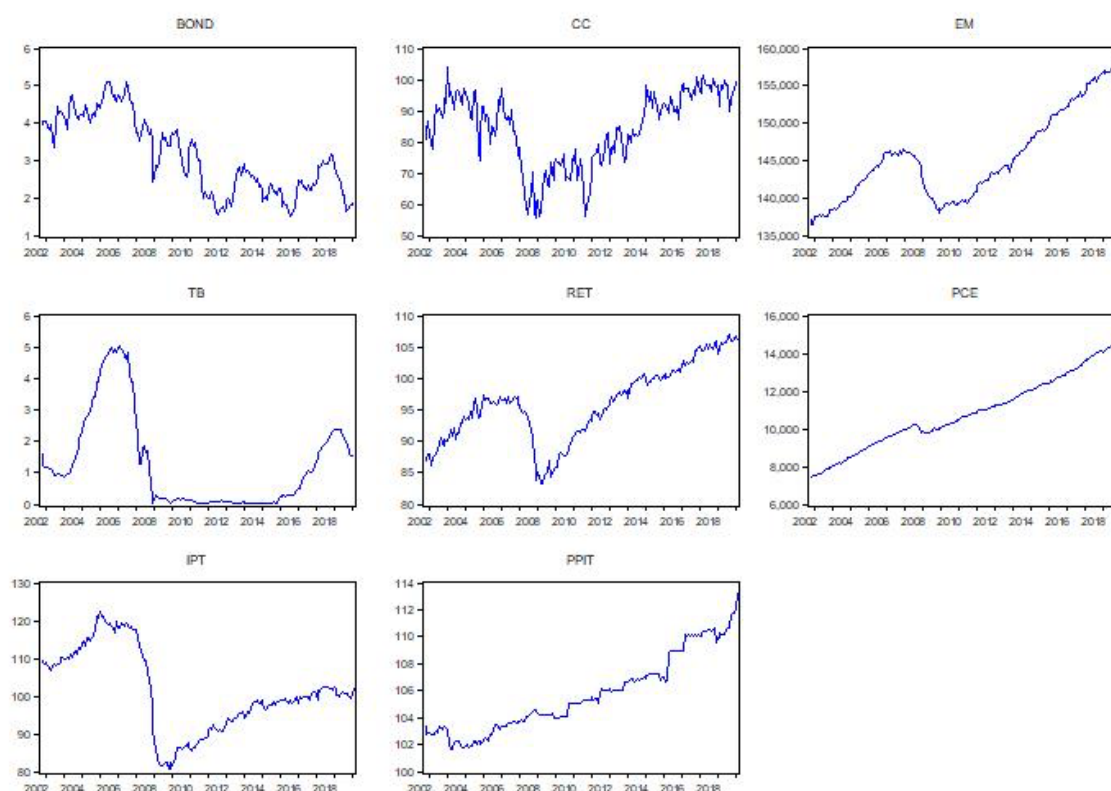
Anche la variabile “RET” costituisce una fonte di interesse per la definizione dell'indice. Nello specifico questa grandezza rappresenta il totale del commercio al dettaglio negli Stati Uniti. Negli ultimi anni l'e-commerce ha avuto un ruolo di primo ordine nel panorama economico mondiale, che ha impattato fortemente nel computo della variabile in oggetto.

Tale misura è indicata con l'appellativo “RET”. È scaricata dal database Fred, come un indice su dati mensili, con procedura di destagionalizzazione già realizzata.

1.3 Analisi grafica delle variabili esplicative

In questo paragrafo vengono riportati i grafici delle nove variabili esplicative che entreranno potenzialmente a far parte della definizione dell'endogena NASDAQ Composite. Verrà presentato un breve confronto dell'andamento complessivo delle serie storiche. Differentemente da quanto proposto di consueto nel prosieguo, le nove variabili vengono presentate attraverso un'unica rappresentazione grafica finalizzata ad un confronto diretto ed immediato.

Figura 1.3- Confronto grafico variabili esplicative dal 2002/10 al 2021/01



Entrando nel merito della disamina, si può notare che la variabile “CC” mostra un trend tipico degli indici di confidenza o degli altri indicatori di incertezza economica. Nello specifico si può apprezzare una fase ribassista del “sentiment” conseguente alla grande recessione statunitense e alla crisi del 2011. Negli anni che seguono questi eventi vi è una ripresa graduale dell’indice, senza variazioni particolarmente significative.

Sono rappresentativi di un trend nel complesso simile le variabili “IPT”, “EM” e “RET”, sebbene con variazioni di entità differente. La variabile relativa alla produzione industriale denota la variabilità maggiore in seguito alla grande recessione, con una decrescita significativa dal 2008 al 2010, che non verrà compensata dal successivo rialzo negli anni a venire. Le misure relative al commercio al dettaglio e all'occupazione sono caratterizzate da un andamento affine, con una crescita più sostanziale negli anni che seguono la grande recessione. Negli anni più recenti la prima delle due mostra un leggero rallentamento.

Guardando alla struttura dei grafici inerenti ai rendimenti obbligazionari “TB” e “BOND”, si segnala un andamento nel complesso decrescente per quelli a lungo termine, con una ripresa contenuta dal 2002 al 2007 e in corrispondenza del 2018. I rendimenti a breve sono caratterizzati da un grafico crescente dal 2003 con un massimo tra il 2006 e il 2007 e una contrazione che mantiene i valori a livelli minimi fino al 2016. Da qui al 2018 si può constatare un'ulteriore fase di crescita.

Infine, il “PCE” e il “PPIT” mostrano un trend che dal 2002 al 2020 si mantiene nel complesso crescente. La spesa individuale nella fattispecie dimostra una crescita lineare e costante, con una lieve contrazione in prossimità della grande recessione. In seguito a questo evento anche la seconda delle variabili segnala un leggero rallentamento per poi aumentare con maggiore intensità dal 2018.

1.4 Analisi statistica della serie endogena

In sede di analisi della serie endogena, come anticipato, si considera il periodo tra il 2002/10/01 e il 2020/01/01, su dati mensili. Si prenderanno in considerazione le chiusure aggiustate dell'indice per ogni mese della variabile NASDAQ Composite, denominata per semplicità con "N".

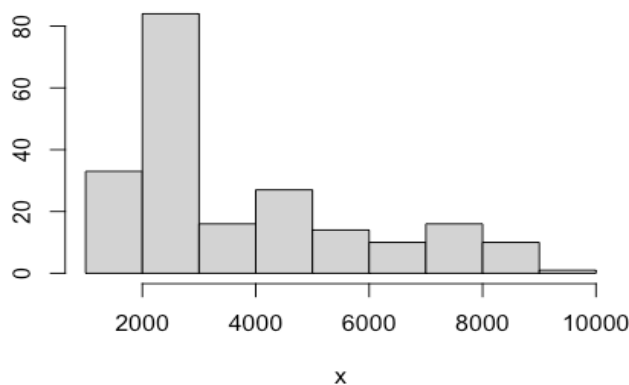
Dalla computazione delle statistiche basilari si deducono alcune peculiarità della serie originaria rispetto innanzitutto alla normale: nella fattispecie la serie presenta una leggera asimmetria e valori di curtosi che si allontanano rispetto a quelli assunti dalla gaussiana, seppur non significativamente.

Figura 1.4- Statistiche normalità della variabile N

Variance	4.048139e+06
Stdev	2.011999e+03
Skewness	1.000207e+00
Kurtosis	-1.801210e-01

Si ricorda che la normale, infatti, assume valori di asimmetria e curtosi rispettivamente di 0 e 3. Sull'originaria abbiamo un allontanamento non trascurabile per il momento terzo. Dall'istogramma della serie si ha la conferma di quanto premesso:

Figura 1.5- Istogramma della variabile N



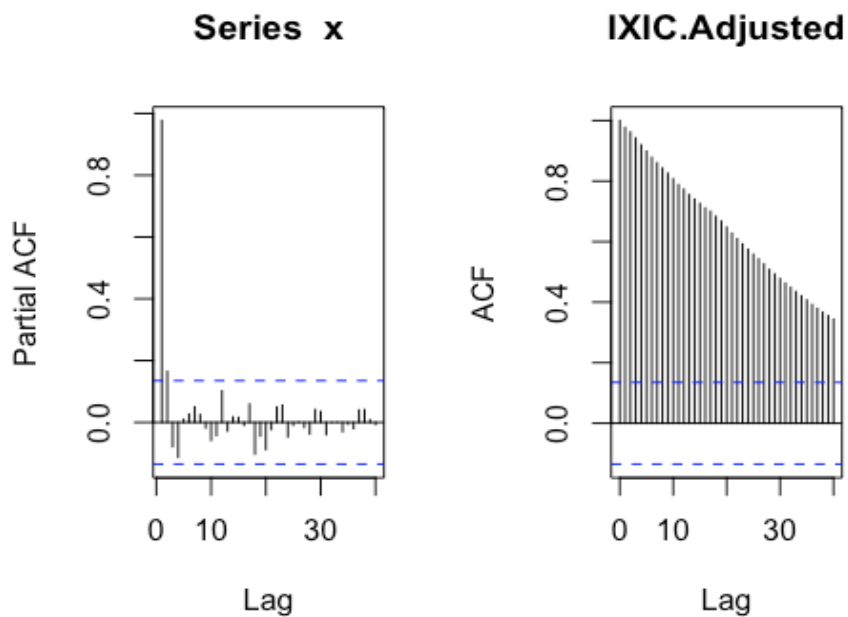
Notiamo infatti un'asimmetria pronunciata dell'istogramma rispetto alla conformazione della normale. Infine, il test di Jarque Bera conferma la non normalità della serie, portando al rifiuto dell'ipotesi nulla H_0 .

Figura 1.6- Jarque Bera test della variabile N

```
Jarque Bera Test  
  
data: x  
X-squared = 35.388, df = 2, p-value = 2.068e-08
```

Dai correlogrammi ai livelli emerge una rappresentazione con una marcata autocorrelazione per i vari lag e possibile non stazionarietà:

Figura 1.7- Correlogrammi ai livelli della variabile N



Vista la non normalità si procede facendo la logaritmica sulla variabile ripetendo il procedimento visto poc'anzi e generando i valori per i vari momenti della serie:

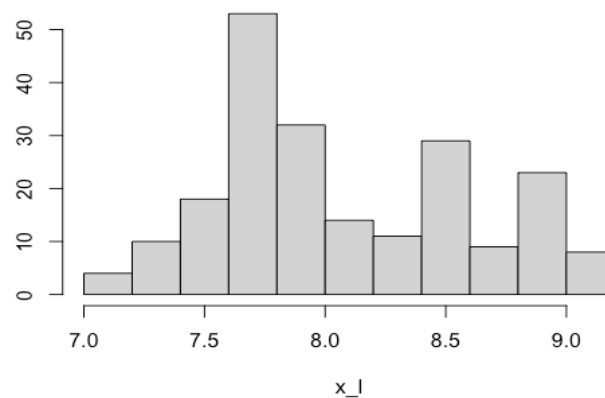
Figura 1.8- Statistiche normalità sulla trasformata logaritmica di N

Variance	0.261198
Stdev	0.511076
Skewness	0.400778
Kurtosis	-1.001840

Si nota anzitutto che la serie presenta dei valori di asimmetria vicini a quelli di una gaussiana, nonostante il momento quarto non proceda nella stessa direzione.

L'istogramma è indicativo di una variazione rilevante, soprattutto relativamente all'asimmetria.

Figura 1.9- Istogramma trasformata logaritmica della variabile N



Migliora anche la probabilità assegnata al test per la normalità di Jarque Bera, sebbene non segnali un p-value tale da portare all'accettazione dell'ipotesi nulla H0.

Figura 1.10- Jarque Bera test sulla trasformata logaritmica N_L

Jarque Bera Test
data: x_1
X-squared = 14.015, df = 2, p-value = 0.000905

Per le ragioni di cui sopra si procede applicando il logaritmo all'endogene, come di consueto per le serie finanziarie. A seguito della logaritmica sull'originaria i correlogrammi relativi l'autocorrelazione parziale e l'autocorrelazione semplice non riportano variazioni ai livelli.

Assume centralità in una disamina di una qualsivoglia serie storica l'analisi di integrazione, che mira a studiare la stazionarietà o, al contrario, la non stazionarietà della stessa. I test che verranno condotti a questo fine sono l'ADF e il Phillips Perron. In questo caso entrambi portano allo stesso risultato. Il primo dei due, nello specifico riporta i seguenti valori:

Figura 1.11- Test ADF sulla variabile N_L con trend

```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.196279 -0.025023  0.003289  0.029932  0.100385
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.2562724  0.1358313   1.887  0.0607 .
## z.lag.1      -0.0352303  0.0187359  -1.880  0.0616 .
## tt           0.0003374  0.0001592   2.119  0.0353 *
## z.diff.lag   0.0816435  0.0717692   1.138  0.2567
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

##
## Residual standard error: 0.04706 on 193 degrees of freedom
## Multiple R-squared:  0.02749,    Adjusted R-squared:  0.01237
## F-statistic: 1.818 on 3 and 193 DF,  p-value: 0.1452
##
##
## Value of test-statistic is: -1.8804 3.2039 2.2735
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47

```

Effettuando il test sulla variabile logaritmica ai livelli il trend appare significativo. Osservando il valore critico del test si andrà ad accettare l'ipotesi nulla di radice unitaria, a confermare dunque la non stazionarietà già anticipata dai correlogrammi. La serie si presenta come un'integrata di ordine uno ($I(1)$) e andrà differenziata per essere ricondotta alla stazionarietà.

I test successivi verranno proposti sulla differenziata logaritmica e saranno finalizzati ad inquadrarne i caratteri essenziali per arrivare ad una stima di un modello per la media corretto ed efficiente. A tale scopo i primi test effettuati sono quelli inerenti alla linearità della serie, nello specifico il Teraesvirta, così come il White test, che portano al rifiuto dell'ipotesi nulla di linearità in media.

Figura 1.12- Teraesvirta Test sulla variabile N_L

Teraesvirta Neural Network Test

```

data: diff(x_1)
X-squared = 12.995, df = 2, p-value = 0.001507

```

Anche Jarque Bera conduce a risultati pressoché analoghi a quanto visto per l'originaria, confermando la non normalità della serie differenziata.

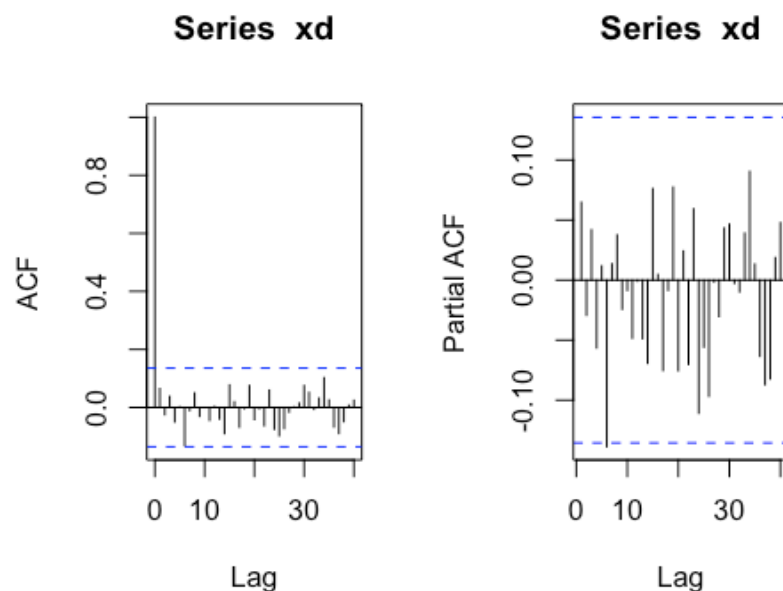
Figura 1.13- Jarque Bera test sulla serie differenziata di N_L

```
Jarque Bera Test

data:  xd
X-squared = 34.326, df = 2, p-value = 3.518e-08
```

Ad assumere particolare rilevanza sono i correlogrammi dell'autocorrelazione semplice e dell'autocorrelazione parziale, sempre sulla differenziata, i quali da un lato possono confermare se la serie sia stata inquadrata correttamente come non stazionaria e eventualmente se vi sia presenza di stagionalità, dall'altro suggeriscono la formulazione del processo ARIMA (p,d,q) per interpretare la serie.

Figura 1.14- Correlogrammi in differenza prima sulla trasformata logaritmica di N



Si può notare che la serie differenziata non assume più i caratteri di un integrata di ordine uno, non necessita dunque di ulteriori differenziazioni. Non è inoltre necessaria una destagionalizzazione della stessa.

Appurata la non linearità la non normalità e la non stazionarietà della serie si procede con la stima del modello ARIMA (p,d,q) più funzionale al NASDAQ Composite, modello che in seguito verrà denominato “benchmark”. In genere si ricorre al correlogramma dell'autocorrelazione semplice e dell'autocorrelazione parziale per disporre delle informazioni necessarie a tal fine. Sulla base della struttura di questi ultimi in differenza prima, e sulla base della significatività nei diversi lag, è possibile risalire alla tipologia all'ordine del processo. La significatività dell'autocorrelazione semplice sarà indicatrice dei modelli a media mobile quella dell'autocorrelazione parziale degli autoregressivi. Nel caso di un comportamento sovrapponibile o quantomeno simile tra i lag significativi dei correlogrammi, spesso accompagnato dall'uscita regolare dalle bande, si otterranno modelli misti come gli ARMA(p,q). Le rappresentazioni dei correlogrammi in differenza precedentemente mostrati a livello della figura 1.14, non conducono ad una netta interpretazione in quanto a uscita e significatività rispetto ai lag, sia per quanto concerne l'autocorrelazione semplice che per l'auto correlazione parziale, assumendo i tratti tipici di un random walk. Si andrà quindi a stimare un modello ARIMA (0,0,0) sulla differenziata logaritmica, del tipo:

$$X_T = X_{t-1} + u$$

È possibile studiare anche alcuni modelli alternativi purché vengano rispettate le condizioni sul rapporto tra coefficienti del processo e standard error.

Figura 1.15- Stima modello ARIMA sulla variabile N_L differenziata

```

arma(x = xd, order = c(0, 0, 0))

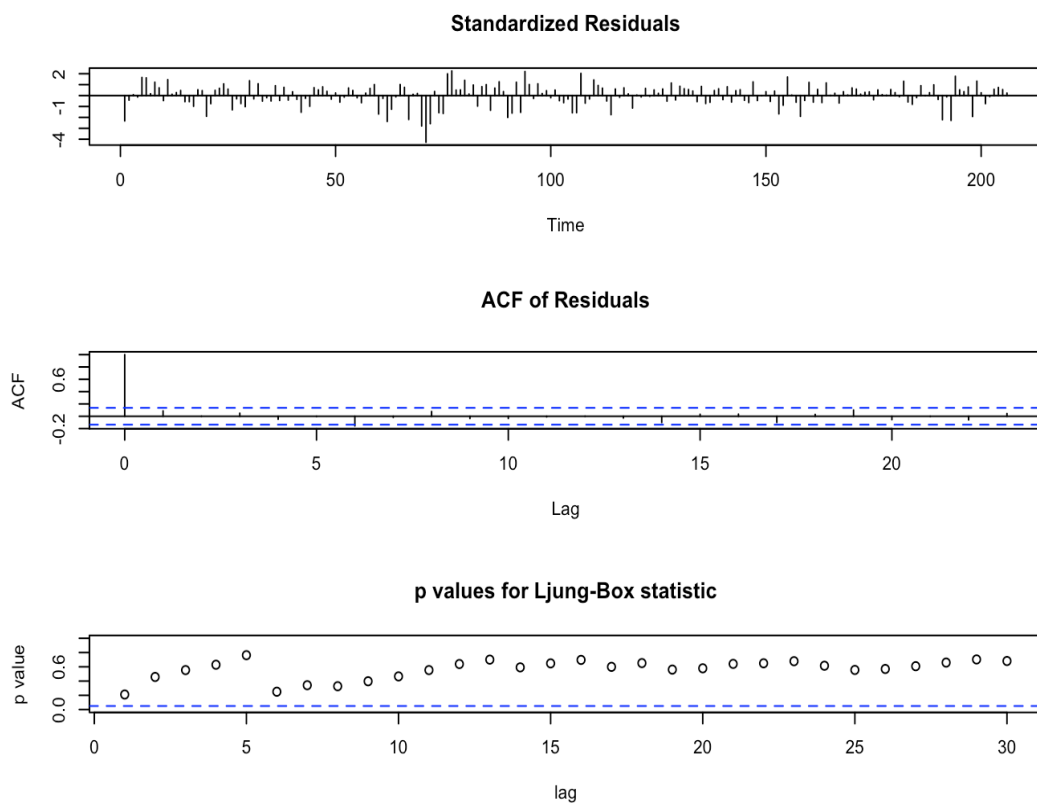
Coefficients:
  intercept
      0.0088
s.e.      0.0033

sigma^2 estimated as 0.002277:  log likelihood = 334.46,  aic = -664.
92

```


La valutazione del modello stimato in relazione a quelli alternativi viene vagliata dai criteri informativi come le statistiche AIC (Akaike Information Criterion) e BIC (Bayesian Information Criterion). La lettura di questi ultimi va nella stessa direzione, nonostante il primo tenda ad avere alcuni limiti sulla sovra parametrizzazione. Di norma si tenderà a scegliere il modello che riporta valori più contenuti in seno ai criteri appena menzionati. Appena stimato il modello ARIMA (p,d,q), si prende in esame l'indipendenza dei residui. L'assenza di correlazione seriale viene confermata innanzitutto dalla struttura del Ljung Box test statistic:

Figura 1.16- Rappresentazioni sui residui del modello stimato sulla variabile N_L



Quest'ultimo evidenzia dei livelli di p-value significativi, grazie ai quali si accetta l'ipotesi nulla di indipendenza per la serie. Una volta generata la serie dei residui del modello stimato è possibile ripetere il test per ogni lag specifico. A titolo esemplificativo, la figura sottostante non fa altro che confermare per il decimo lag l'indipendenza.

Figura 1.17- Box-Ljung test sui residui della variabile N_L

```
Box-Ljung test

data:  res
X-squared = 9.7167, df = 10, p-value = 0.4657
```

La componente dei residui generata dal processo random walk stimato sulle serie endogene è utile ai fini della stima di un modello per la volatilità.

Prima di verificare la presenza di eventuali effetti ARCH sulla componente generata, va indagata, ed eventualmente accertata, l'ipotesi nulla di residui non i.i.d. (identicamente e indipendentemente distribuiti). Il rifiuto dell'ipotesi nulla in questione viene suffragato dal BDS test:

Figura 1.18- BDS Test sui residui sulla variabile N_L

```
BDS Test

data:  res

Embedding dimension = 2 3

Epsilon for close points = 0.0239 0.0478 0.0717 0.0957

Standard Normal =
  [ 0.0239 ] [ 0.0478 ] [ 0.0717 ] [ 0.0957 ]
[ 2 ]      3.4276      3.7620      4.1556      4.9366
[ 3 ]      3.4828      4.1801      4.7911      5.3413

p-value =
  [ 0.0239 ] [ 0.0478 ] [ 0.0717 ] [ 0.0957 ]
[ 2 ]      6e-04      2e-04          0          0
[ 3 ]      5e-04      0e+00          0          0
```

A seguire viene verificata la presenza di effetti ARCH e la non linearità della serie rispettivamente ricorrendo al test LM ARCH e al Teraesvirta test:

Figura 1.19- ARCH Test sui residui sulla variabile N_L

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: res
```

```
Chi-squared = 28.94, df = 2, p-value = 5.196e-07
```

Figura 1.20- Teraesvirta Test sui residui sulla variabile N_L

```
Teraesvirta Neural Network Test
```

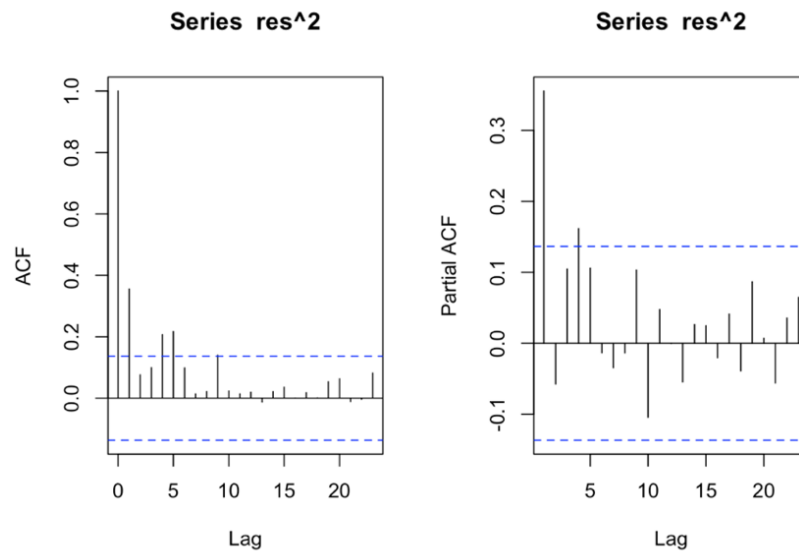
```
data: res
```

```
X-squared = 14.361, df = 2, p-value = 0.0007613
```

Dal primo si rifiuta l'ipotesi nulla di assenza di effetti ARCH sulla serie, dal secondo si rifiuta H_0 sulla linearità in media dei residui. Si procede quindi alla stima di un modello per l'eteroschedasticità condizionale sulla componente dell'innovazione.

Uno sviluppo in tal senso viene suggerito anche dai correlogrammi dei residui al quadrato:

Figura 1.21- Correlogramma residui al quadrato della variabile N_L



Questi, oltre a confermare la presenza di effetti ARCH, possono essere interpretati per intercettare l'ordine e la tipologia del modello per l'eteroschedasticità più conforme alla serie. Anche in quanto ad efficacia e funzionalità del modello, la lettura è analoga a quanto riportato precedentemente per la ricerca del modello ARIMA: verranno preferiti quei modelli che registrano livelli più bassi con riferimento alle statistiche “AIC” e “BIC”. Accanto a ciò, i parametri che scaturiscono dalla stima devono rispettare alcune condizioni. Nello specifico questi ultimi dovranno essere significativi, con $\omega > 0$, con $\alpha > 0$, $\beta > 0$ e $\alpha + \beta < 1$.

1.5 Introduzione alla teoria dei modelli per l'eteroschedasticità condizionale

Nel presente capitolo verrà approfondito il modello per la volatilità che descrive i residui di una serie storica.

Questo passaggio è essenziale, come riporta uno dei massimi esponenti della materia econometrica Robert Engle². Infatti, oltre a comprendere come possano influire delle variabili esplicative su una variabile endogena nei modelli univariati, è d'uopo porsi delle domande sulla volatilità che insiste sui medesimi.

La versione standard dei minimi quadrati poggia le proprie fondamenta su un'ipotesi: il valore atteso al quadrato dell'errore è eguale in ogni punto, in altre parole è omoschedastico. Le serie storiche in cui, tuttavia, la varianza dell'errore non è costante ed uguale in determinati punti, soffrono di eteroschedasticità, come avviene di norma nell'equazione di lungo periodo che emerge dalla statica.

I modelli ARCH(p) e GARCH (p,q) sono utili proprio ai fini di intercettare e ricavare dei modelli per la varianza.

Il primo fra i due ipotizza che la varianza di domani (in $t + 1$) sarà una media ponderata dei residui al quadrato per il tempo t considerato, secondo la seguente formula:

$$h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

Ovviamente l'assunto dei pesi uguali in un intervallo "t" è una semplificazione e non trova riscontri frequenti nella realtà.

Una generalizzazione che diede maggiore rilievo empirico fu quella proposta da T. P. Bollerslev³, con il modello GARCH(p,q):

² Engle R., *GARCH 101: An Introduction to the Use of ARCH/GARCH models in Applied Econometrics*, NYU Working Paper, 2001/10.

³ T.P. Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity*, Journal of Econometrics 31, 1986/02.

$$h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$$

Quest'ultimo viene comunemente declinato secondo un ordine di p e q che corrisponde a 1, assumendo la forma di GARCH(1,1). Tale specificazione suggerisce che la previsione più robusta per la varianza in un periodo futuro t+1 non è altro che una media ponderata della varianza stessa e l'errore al quadrato più recente.

In letteratura sono stati proposti anche molteplici modelli alternativi e più complessi che, in generale, tendono ad interpretare con maggiore robustezza la volatilità di serie finanziarie che reagiscono asimmetricamente agli shock. In altri termini, nella pratica econometrica, si è riscontrata la tendenza insita in alcune variabili a reagire diversamente in risposta ad impulsi negativi o positivi. Vi sono serie finanziarie che reagiscono amplificando l'effetto delle "buone notizie" a discapito delle notizie negative (un caso frequente è quello di una società quotata che realizza degli utili maggiori rispetto a quelli attesi) e viceversa.

Ulteriori specificazioni di modelli per l'eteroschedasticità condizionale sono quelli tipici di condizioni particolari sui parametri.

Da un punto di vista computazionale, i modelli in oggetto tendono a riprendere la struttura tipica dei GARCH (p,q), introducendo alcuni vincoli o compensazioni. Tra i più diffusi si ricordano il TGARCH, IGARCH, EGARCH ed NGARCH. Tralasciando la formalizzazione dei diversi modelli per l'eteroschedasticità condizionale (non necessaria ai fini di questo case study), nel proseguo tornerà utile la struttura dell'IGARCH (Integrated Generalized Autoregressive conditional heteroskedasticity), introdotto da Engle e Bollerslev (1986). Quest'ultimo è un caso singolare di GARCH(p,q), in cui la sommatoria dei parametri α e β deve essere pari all'unità. Nei modelli standard GARCH(p,q) la condizione è del tipo:

$$\sum_{i=1}^p \alpha_i + \sum_{i=1}^q \beta_i < 1$$

Nel modello GARCH(p,q) i parametri sono tutti positivi ed il polinomio associato al processo del tipo $1 - A(L) - B(L)$ ha radici esterne al cerchio unitario e si dimostra, dunque, stazionario.

Il modello IGARCH presenta invece la seguente condizione:

$$\sum_{i=1}^p \alpha_i + \sum_{i=1}^q \beta_i = 1$$

Mentre nel modello GARCH si riscontrava una condizione del tipo:

- $A(L) + B(L) < 1$,

nell' IGARCH avremo che:

- $A(L) + B(L) = 1$.

Nel primo caso gli impulsi futuri sulla varianza del processo tendono dunque ad esaurirsi al crescere della variabile temporale: più è contenuto il risultato che scaturisce dalla somma tra $A(L)$ e $B(L)$ più rapidamente si esauriranno gli shock. Nel caso dell'IGARCH avremo invece la persistenza degli shock futuri sulla varianza: l'effetto degli impulsi sulla dispersione tenderà a permanere nel corso del tempo. La base teorica del modello richiama fortemente la formulazione dei processi random walk, in cui la variabile X_t è dipendente da X_{t-1} , ma differentemente da quest'ultimo il processo è stazionario in senso stretto, coerentemente con quanto riportato da D. B. Nelson (1990). Analogamente al caso dei random walk, nei modelli IGARCH la varianza di un periodo andrà ad influire sulla varianza dei periodi futuri.

1.5.1 Modello per l'eteroschedasticità della serie endogena

Nel prosieguo verranno proposti alcuni modelli alternativi per la volatilità che rispettano queste condizioni e che sono in grado di interpretare efficacemente la volatilità della serie. Il primo è quello di un ARCH(1). Quest'ultimo pur essendo il più semplice tra i modelli proposti in letteratura, ottiene risultati notevoli alla luce delle statistiche e dei test relativi all'indipendenza e all'eteroschedasticità. La finalità di questi modelli è quella di intercettare gli effetti di volatility clustering mantenendo l'indipendenza della serie e, possibilmente, migliorando le statistiche sulla normalità. Come emerge dai Ljung Box

test effettuati per l'assenza di autocorrelazione seriale e soprattutto dall'LM ARCH test, il modello soddisfa pienamente questi requisiti e conduce all'accettazione di H0 per quanto concerne indipendenza ed eteroschedasticità.

Figura 1.22- Stima e statistiche ARCH(1) per la variabile N_L

Error Analysis:					
Estimate	Std. Error	t value	Pr(> t)		
omega	0.0016777	0.0002201	7.622	2.51e-14	***
alpha1	0.2526683	0.1010143	2.501	0.0124	*
skew	0.6738634	0.0607790	11.087	< 2e-16	***
shape	1.7128062	0.2813417	6.088	1.14e-09	***
Standardised Residuals Tests:					
Statistic	p-Value				
Jarque-Bera Test	R	Chi ²	17.32166	0.0001732405	
Shapiro-Wilk Test	R	W	0.9662057	7.622345e-05	
Ljung-Box Test	R	Q(10)	6.523791	0.7695064	
Ljung-Box Test	R	Q(15)	8.300669	0.9111635	
Ljung-Box Test	R	Q(20)	14.55582	0.8012346	
Ljung-Box Test	R ²	Q(10)	16.66244	0.08217512	
Ljung-Box Test	R ²	Q(15)	17.78791	0.2739801	
Ljung-Box Test	R ²	Q(20)	20.00086	0.4578757	
LM Arch Test	R	TR ²	13.25031	0.3511126	

Information Criterion Statistics:			
AIC	BIC	SIC	HQIC
-3.359895	-3.295276	-3.360630	-3.333761

Il QQ plot generato per i residui standardizzati è indicativo di una struttura lineare e riporta valori allineati. Tutte le condizioni vengono interamente soddisfatte eccezion fatta per il parametro “mu” che verrà rimosso.

Figura 1.23- QQ Plot residui standardizzati modello ARCH(1) per la variabile N_L



Il modello più diffuso in letteratura, nonché quello che ha trovato un maggiore riscontro empirico nella pratica econometrica, è quello di un GARCH(1,1). La struttura stessa dei correlogrammi dei residui sembrerebbe suggerire un modello misto. Si procede dunque in questa direzione, andando a stimare un GARCH(1,1):

Figura 1.24- Stima e statistiche GARCH(1,1) per la variabile N_L

```
Error Analysis:
Estimate  Std. Error  t value Pr(>|t|)
omega    0.0002854  0.0001200   2.379  0.01738 *
alpha1   0.2080658  0.0728971   2.854  0.00431 **
beta1    0.6624555  0.0925653   7.157  8.27e-13 ***
skew     0.6301802  0.0700073   9.002  < 2e-16 ***
shape    1.9149779  0.3174965   6.031  1.62e-09 ***
```

Standardised Residuals Tests:

Statistic p-Value

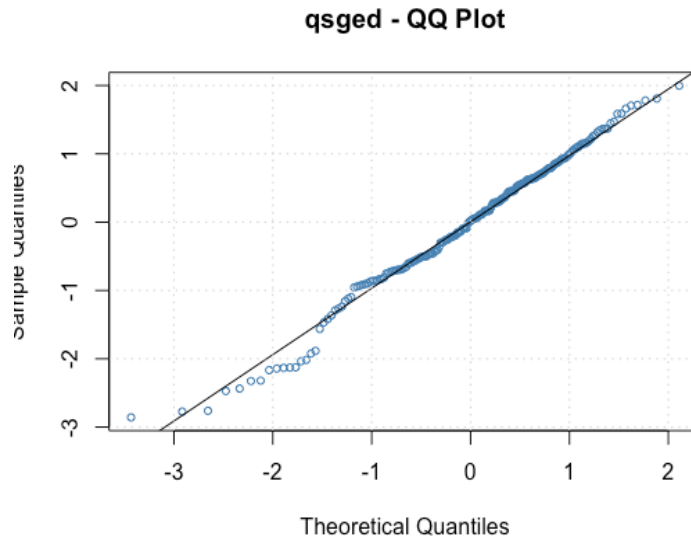
Jarque-Bera Test	R	Chi ²	14.87748	0.0005880251
Shapiro-Wilk Test	R	W	0.9644836	4.778356e-05
Ljung-Box Test	R	Q(10)	4.63759	0.9140369
Ljung-Box Test	R	Q(15)	6.602875	0.9677481
Ljung-Box Test	R	Q(20)	11.91042	0.9191246
Ljung-Box Test	R ²	Q(10)	5.174983	0.8791865
Ljung-Box Test	R ²	Q(15)	8.174523	0.9165953
Ljung-Box Test	R ²	Q(20)	9.072851	0.982048
LM Arch Test	R	TR ²	6.120554	0.9098834

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-3.418318	-3.337544	-3.419459	-3.385650

Anche in questo caso α e β rispettano le condizioni e sono significativi. Il parametro “mu” verrà nuovamente rimosso in quanto non significativo. Analogamente alla rappresentazione dell’ARCH(1), i test portano all'accettazione dell'ipotesi nulla sia per il Ljung Box che per LM ARCH. Dunque, anche in questo contesto il modello per l'eteroschedasticità soddisfa tutte le prerogative elencate precedentemente. Rispetto all’ARCH(1) registra delle statistiche più performanti a livello di AIC e BIC, che sono più contenute rispetto alle precedenti. Il QQ plot è pressochè analogo a quanto visto prima:

Figura 1.25- QQ Plot residui standardizzati GARCH(1,1) per variabile N_L



Verrà preferito quest'ultimo per i motivi di cui sopra, che attribuiscono al modello di Bollerslev un'interpretazione più efficace. Sarebbe possibile proporre ulteriori estensioni del GARCH, riprendendo quanto premesso in merito alla sensibilità differente di alcune variabili in presenza di shock negativi e/o positivi, volti alla stima di modelli asimmetrici per intercettare l'effetto provocato dalle “buone” e/o “cattive” notizie. Tuttavia il “bias test” applicato a questa variabile nonché la stima di alcuni modelli alternativi in tal senso, non sembrerebbero indicare questa tendenza, confermando l'adozione del GARCH(1,1) per la serie dei residui dell'endogena.

Figura 1.26- Bias test sulla variabile endogena

ESTIMATE		P Value
Sign Bias Test	0.5237002	0.01021596
Negative Size Bias Test	-3.8067769	0.23170132
Positive Size Bias Test	-9.9579958	0.01835493
Test Congiunto	2.5808431	0.05463843
-sign	0.4694742	0.15304133
-neg.size	2.4955990	0.54533571
-pos-size	-4.7681709	0.40886132

1.6 Analisi di integrazione delle variabili esplicative

Nel seguente paragrafo verranno riportati alcuni test effettuati sulle variabili esplicative, con la finalità di descrivere le medesime e di inquadrarle correttamente per l'inserimento nelle equazioni di lungo e di breve periodo.

Di norma si procederà analizzando le statistiche di base che le caratterizzano (come la simmetria e la curtosi), effettuando i test sulla normalità della serie ed -eventualmente- applicando il logaritmo alla serie studiata. A seguire verrà approfondita la possibile stagionalità e la stazionarietà, attraverso la struttura dei correlogrammi e, nel secondo caso, sfruttando test come l'ADF, il Phillips Perron o il KPSS. Da ultimo si studierà la struttura dei correlogrammi differenziati (nel caso di serie integrate) per accertare che il processo di differenziazione sia avvenuto correttamente.

Le analisi in questione vengono effettuate sull'ambiente software R e considerando il periodo campionario precedentemente specificato. I risultati ottenuti per l'ADF test vengono appositamente riportati in *Appendice* per rendere questo capitolo più scorrevole.

1.6.1 La variabile "TB"

La variabile "TB", riferita ai rendimenti a breve termine dei titoli di stato statunitensi, presenta delle statistiche di base che non sono in linea con quelle tipiche della normale, caratterizzate da una curtosi in prossimità del valore 3 e una asimmetria pari a 0.

Figura 1.27- Statistiche normalità sulla variabile TB

Variance	2.252756
Stdev	1.500918
Skewness	1.193790
Kurtosis	0.304936

In questo caso si nota un discostamento da entrambi i valori, che viene confermato dal test di Jarque Bera sulla normalità. Il suddetto test porta infatti al rifiuto dell'ipotesi nulla H_0 :

Figura 1.28- Jarque Bera test sulla variabile TB

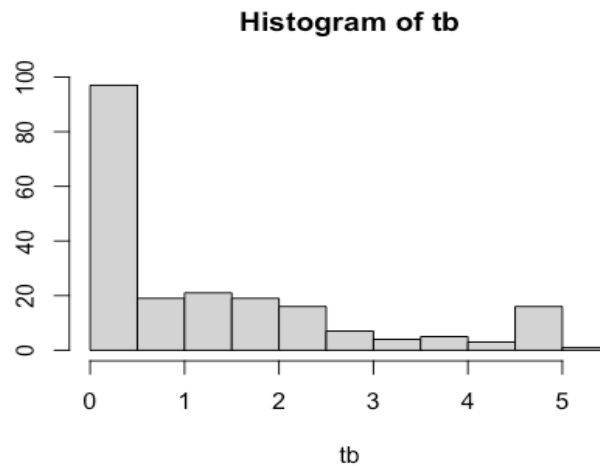
Jarque Bera Test

data: tb

X-squared = 51.108, df = 2, p-value = 7.98e-12

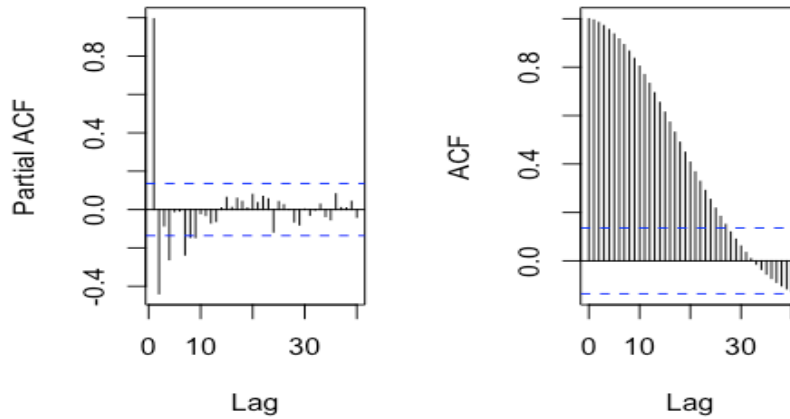
Da un punto di vista grafico la non normalità viene illustrata anche dalla rappresentazione dell'istogramma della serie in analisi, che mostra una struttura decisamente lontana dalla gaussiana:

Figura 1.29- Istogramma della variabile TB



Accertata la non normalità, si potrebbe procedere rendendo logaritmica la serie, con la finalità di riportare la stessa ad una struttura più vicina alla normale. Tuttavia, seguendo tale processo, la serie "TB" non porta ad un miglioramento sostanziale delle statistiche di cui sopra e di test come Jarque Bera. Si proseguirà mantenendo dunque la serie originaria. La disamina procede studiando i correlogrammi dell'originaria, con lo scopo di individuare la possibile non stazionarietà della serie.

Figura 1.30- Correlogrammi della variabile TB



Dall'autocorrelazione semplice emerge una struttura indice, per i vari lag, di un'elevata autocorrelazione, raffigurazione tipica di una serie non stazionaria.

Il Teraesvirta test conduce invece all'accettazione dell'ipotesi nulla H_0 : la serie è lineare in media:

Figura 1.31- Teraesvirta Test sulla variabile TB

```
Teraesvirta Neural Network Test

data:  tb
X-squared = 0.039906, df = 2, p-value = 0.9802
```

Al fine di verificare formalmente la presunta non stazionarietà della serie (anticipata dalla struttura dell'autocorrelazione parziale), si effettua il test ADF, avente come ipotesi nulla H_0 la presenza di radice unitaria:

Figura 1.32- Test ADF sulla variabile TB con presenza di trend

```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression trend
```

```

##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.77556 -0.08319  0.03773  0.10632  1.41988
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0379419  0.0505569  -0.750  0.453862
## z.lag.1      -0.0131404  0.0141039  -0.932  0.352638
## tt           0.0002593  0.0004376   0.592  0.554205
## z.diff.lag1  0.1640612  0.0689598   2.379  0.018311 *
## z.diff.lag2 -0.2552733  0.0692085  -3.688  0.000292 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3442 on 197 degrees of freedom
## Multiple R-squared:  0.09103,    Adjusted R-squared:  0.07258
## F-statistic: 4.932 on 4 and 197 DF,  p-value: 0.0008233
##
##
## Value of test-statistic is: -0.9317 0.5741 0.859
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47

```

Osservando la probabilità del trend “tt”, si segnala che il valore associato non è significativo: sarà necessario ripetere il test rimuovendo il trend e verificando la significatività relativa all’intercetta (“drift”).

Figura 1.33- Test ADF sulla variabile TB con presenza di drift

```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.86050 -0.05367  0.02616  0.09079  1.73423
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.01693    0.02725  -0.621   0.535
## z.lag.1     -0.01903    0.01353  -1.406   0.161
##
## Residual standard error: 0.3527 on 205 degrees of freedom
## Multiple R-squared:  0.009555,    Adjusted R-squared:  0.004723
## F-statistic: 1.978 on 1 and 205 DF,  p-value: 0.1612
##
##
## Value of test-statistic is: -1.4063 0.9888
##
## Critical values for test statistics:
##      1pct  5pct 10pct
```



```
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

Nel qual caso il trend o l'intercetta non si rivelassero significativi, si procederà nuovamente alla loro rimozione riportando il test definitivo per verificare l'ipotesi nulla di radice unitaria.

Figura 1.34– Test ADF sulla variabile TB senza trend e drift

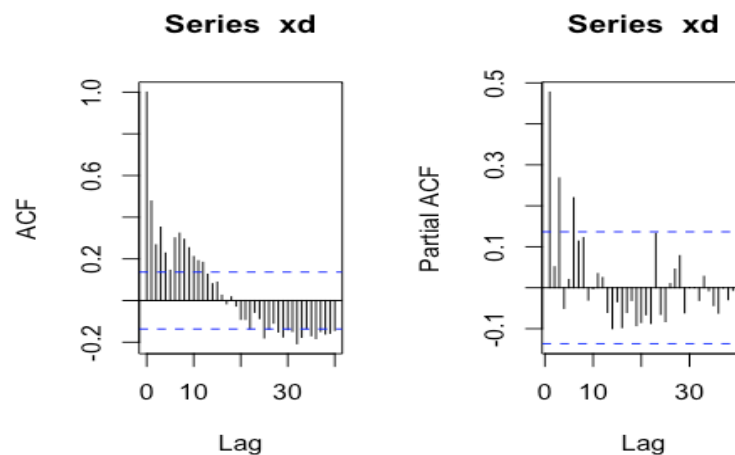
```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.85337 -0.01928  0.00353  0.06314  0.44403
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.003126   0.005402  -0.579   0.563
##
## Residual standard error: 0.1534 on 206 degrees of freedom
## Multiple R-squared:  0.001623, Adjusted R-squared:  -0.003224
## F-statistic: 0.3349 on 1 and 206 DF, p-value: 0.5634
##
##
## Value of test-statistic is: -0.5787
##
```

```
## Critical values for test statistics:  
##      1pct  5pct 10pct  
## tau1 -2.58 -1.95 -1.62
```

Il valore assunto dalla statistica test conduce formalmente all'accettazione dell'ipotesi nulla di radice unitaria: la serie "TB" è non stazionaria e $I(1)$. Per essere ricondotta alla stazionarietà dovrà essere differenziata.

I correlogrammi in differenza prima sottostanti confermano che la differenziazione è stata realizzata correttamente e che la serie non necessita di una differenza seconda. Quest'ultimo caso porterebbe infatti ad una sovra-differenziazione e ad una conseguente stima errata delle equazioni nella procedura a due passi di Engle-Granger. La serie non presenta un'uscita regolare dalle bande, indice del fatto che non necessita di alcuna destagionalizzazione.

Figura 1.35- Correlogrammi in differenza prima della serie TB



1.6.2 La variabile "CC"

La variabile "CC", riguardante il "sentiment" dei consumatori statunitensi, riporta delle statistiche di base più vicine a quelle di una gaussiana, rispetto al caso della variabile "TB". L'asimmetria in particolare, segnala dei valori prossimi a 0, seppur negativi. Il momento quarto delinea invece valori più vicini a quelli di una distribuzione platicurtica:

Figura 1.36- Statistiche normalità della variabile CC

Variance	139.937561
Stdev	11.829521
Skewness	-0.580115
Kurtosis	-0.652558

La non normalità viene formalizzata dal rifiuto dell'ipotesi nulla del test di Jarque Bera. Nella serie in oggetto non si è applicato il logaritmo, nonostante i valori assunti in corrispondenza del momento terzo e quarto, poiché la presenza di una asimmetria negativa non avrebbe fatto altro che allontanare maggiormente lo stesso valore rispetto a 0.

Figura 1.37- Jarque Bera test variabile sulla CC

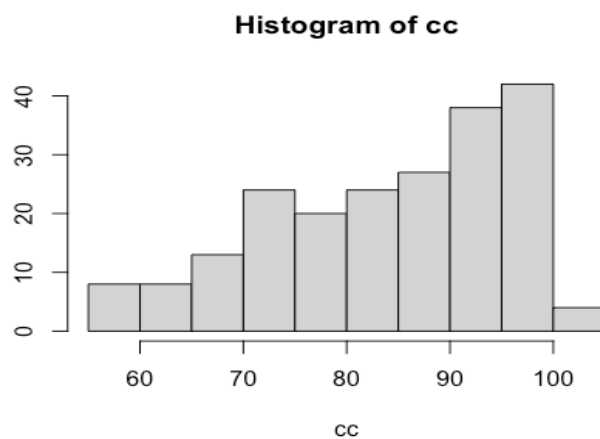
Jarque Bera Test

data: cc

X-squared = 15.274, df = 2, p-value = 0.0004822

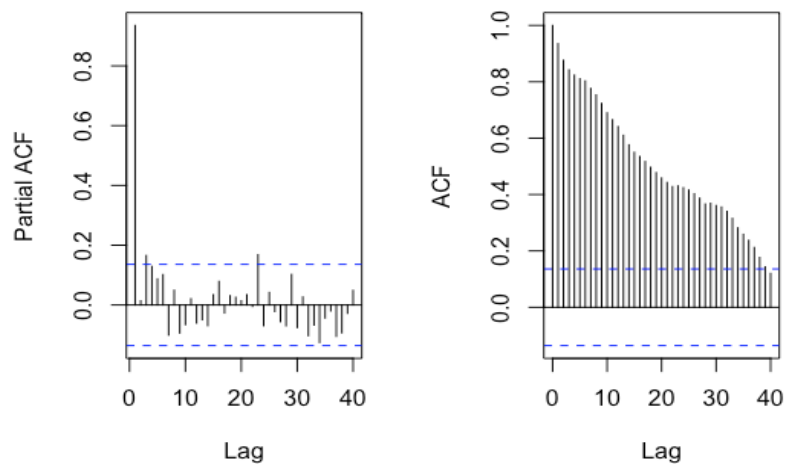
L'istogramma riportato di seguito conferma la struttura di una distribuzione con asimmetria negativa e non normale:

Figura 1.38- Istogramma della variabile CC



Posto che anche in questo caso verrà mantenuta la serie originaria, si procede con l'analisi dell'autocorrelazione semplice e parziale, volta a verificare la presenza di radici unitarie. Anche in questo caso l'autocorrelazione semplice decresce lentamente e segnala una forte autocorrelazione in seno alla variabile originaria "CC". Ciò suggerisce la presenza di una radice unitaria:

Figura 1.39 Correlogrammi della variabile CC



Concentrando l'attenzione sulla linearità, si osserva come in questo caso l'ipotesi nulla di linearità in media non venga accettata, dopo aver eseguito il Teraesvirta test:

Figura 1.40- Teraesvirta test sulla variabile CC

```
Teraesvirta Neural Network Test

data: cc
X-squared = 6.3762, df = 2, p-value = 0.04125
```

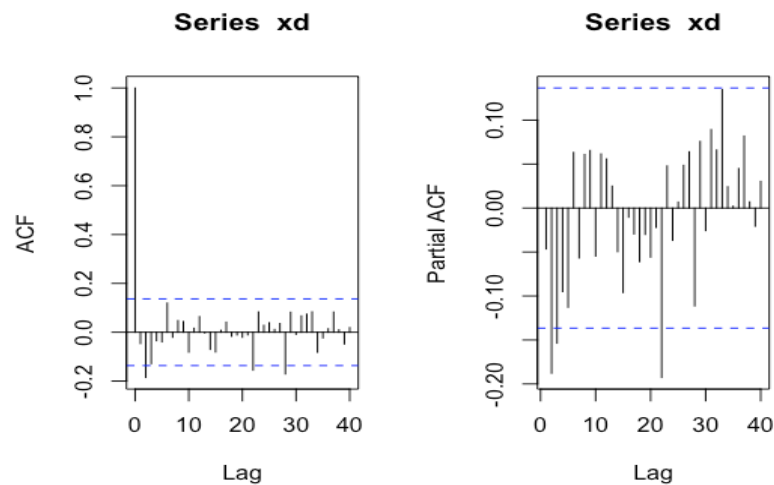
La presenza o l'assenza di una radice unitaria viene nuovamente accertata attraverso il test ADF. Dopo aver ripetuto il test senza la componente del trend, il test viene effettuato con l'intercetta che, come si osserva dalla rappresentazione in *Appendice*, risulta significativa.

Il valore della statistica test risulta pari a -2.2549 e porta ad accettare l'ipotesi nulla H_0 di radice unitaria: la serie "CC", per il periodo campionario specificato risulta non stazionaria e $I(1)$.

A seguito della differenza prima sulla serie originaria si può osservare che i correlogrammi assumono nuovamente una struttura tipica delle serie stazionarie, con un'uscita sporadica dalle bande dell'autocorrelazione semplice.

Anche in questo caso la serie non presenta stagionalità e non sarà necessario alcun ulteriore aggiustamento.

Figura 1.41- Correlogramma in differenza prima della variabile CC



1.6.3 La variabile "BOND"

La variabile esplicativa "BOND" evidenzia dei valori di asimmetria prossimi a quelli di una gaussiana, con valori inerenti alla curtosi che sono tuttavia tipici di una distribuzione appiattita o platicurtica.

Figura 1.42- Statistiche normalità della variabile BOND

Variance	1.041940
Stdev	1.020755

Skewness	0.221612
Kurtosis	-1.254991

Dalla realizzazione di questi valori sulla curtosi deriva il rifiuto dell'ipotesi nulla H_0 , contestualmente al test di Jarque Bera.

Figura 1.43- Jarque Bera test sulla variabile BOND

```

Jarque Bera Test

data: bond
X-squared = 15.012, df = 2, p-value = 0.0005497

```

Ciononostante, i valori centrali dell'asimmetria non rendono strettamente necessaria l'applicazione e l'utilizzo della trasformata logaritmica sulla serie originaria. Ciò si evince anche dal test di D'Agostino sulla normalità, che conferma la bontà dei valori riguardanti l'asimmetria, sebbene il momento quarto non tenda ai valori tipici della normale.

Figura 1.44- D'Agostino test sulla variabile BOND

```

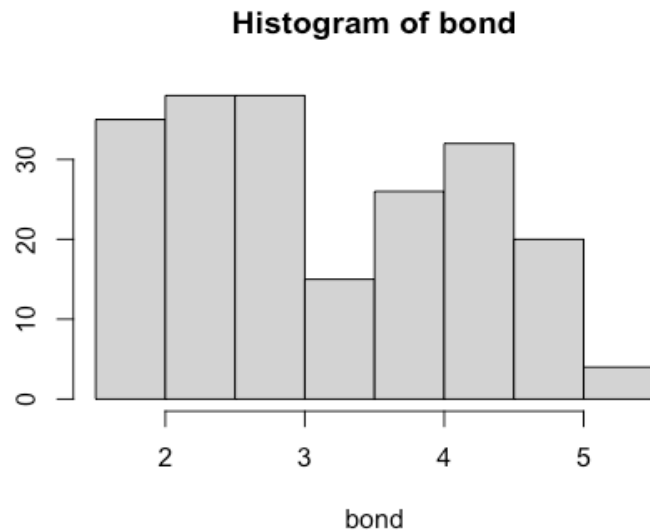
D'Agostino Normality Test

Test Results:
STATISTIC:
  Chi2 | Omnibus: 124.2076
  Z3   | Skewness: 1.3432
  Z4   | Kurtosis: -11.0636
P VALUE:
  Omnibus Test: < 2.2e-16
  Skewness Test: 0.1792
  Kurtosis Test: < 2.2e-16

```

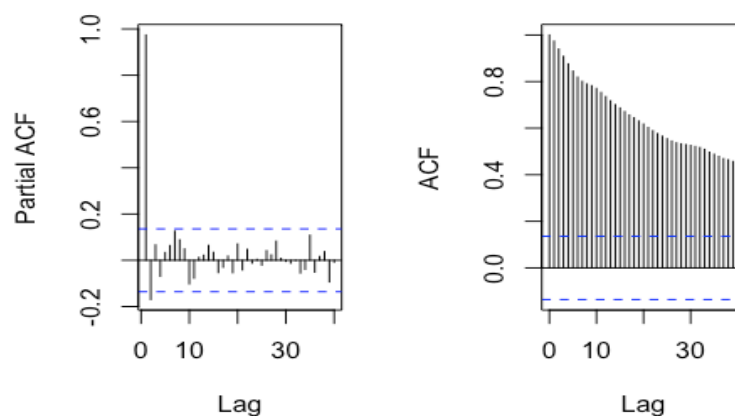
L'istogramma fornisce una sintesi grafica di quanto riportato, delineando una distribuzione appiattita e con una lieve asimmetria positiva:

Figura 1.45- Istogramma della variabile BOND



L'analisi descrittiva prosegue con lo studio dei correlogrammi ai livelli sulla serie "BOND" originaria. La struttura del grafico relativa all'autocorrelazione semplice denota una forte autocorrelazione e suggerisce la non stazionarietà:

Figura 1.46- Correlogrammi della variabile BOND



Con riferimento alla linearità, il Teraesvirta test conduce all'accettazione di H_0 , ovverosia dell'ipotesi di linearità in media, come riportato di seguito:

Figura 1.47- Teraesvirta test sulla variabile BOND

```
Teraesvirta Neural Network Test

data: bond
X-squared = 1.0147, df = 2, p-value = 0.6021
```

Approfondendo l'aspetto della stazionarietà, viene eseguito il test ADF, inserendo in un primo momento il trend.

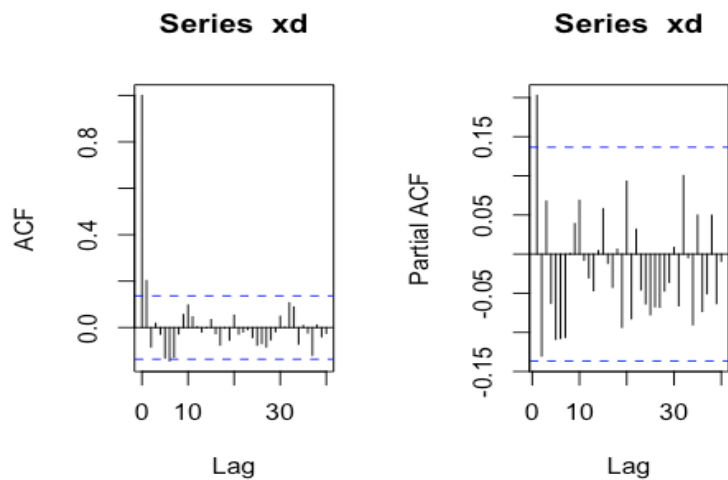
Il test segnala una probabilità significativa per la componente del trend e un valore della statistica test che, letto contestualmente ai valori critici, comporta l'accettazione dell'ipotesi nulla H_0 .

Ciò certifica la non stazionarietà della variabile "BOND", che sarà un'integrata di ordine uno ($I(1)$). La serie originaria dovrà essere differenziata per essere ricondotta alla stazionarietà in media e varianza.

Il correlogramma riportato successivamente mostra che la differenza prima sulla serie raggiunge questa finalità e che la stessa viene interpretata correttamente come una $I(1)$.

Si precisa che, anche sulla variabile in oggetto, non insiste alcuna stagionalità e di conseguenza non verrà applicato alcun aggiustamento finalizzato alla destagionalizzazione.

Figura 1.48- Correlogramma in differenza prima della variabile BOND



1.6.4 La variabile "PCE"

Per quanto concerne la variabile "PCE", le statistiche preliminari su asimmetria e curtosi sono indicative di una distribuzione che approssima la normale, seppur con lievi discostamenti rispetto ai valori esatti propri della gaussiana.

Figura 1.49- Statistiche normalità della variabile PCE

Variance	3.777021e+06
Stdev	1.943456e+03
Skewness	1.847750e-01
Kurtosis	-8.984070e-01

Nonostante la prossimità dei valori soprariportati, la serie porta al rifiuto dell'ipotesi di normalità per Jarque Bera considerando un livello di significatività del 5%.

Figura 1.50- Jarque Bera test sulla variabile PCE

```
Jarque Bera Test  
  
data: pce  
X-squared = 7.8826, df = 2, p-value = 0.01942
```

Il test di D'Agostino denota come il rifiuto dell'ipotesi di normalità precedente sia dovuto ai valori assunti dal momento quarto, che effettivamente sono tipici di una distribuzione leggermente appiattita.

Figura 1.51- D'Agostino test sulla variabile PCE

```
D'Agostino Normality Test  
  
Test Results:
```

STATISTIC:

Chi2 | Omnibus: 21.8066

Z3 | Skewness: 1.1237

Z4 | Kurtosis: -4.5325

P VALUE:

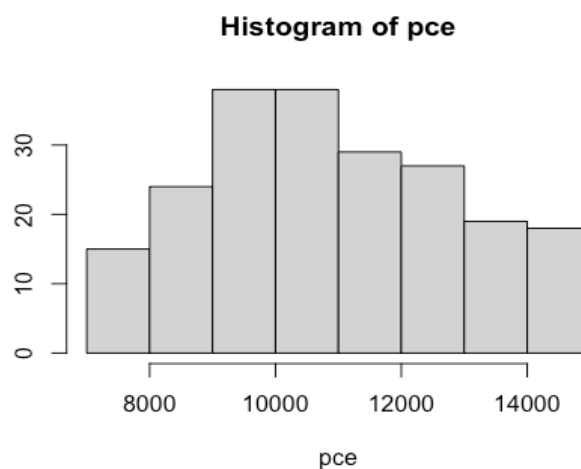
Omnibus Test: 1.84e-05

Skewness Test: 0.2612

Kurtosis Test: 5.828e-06

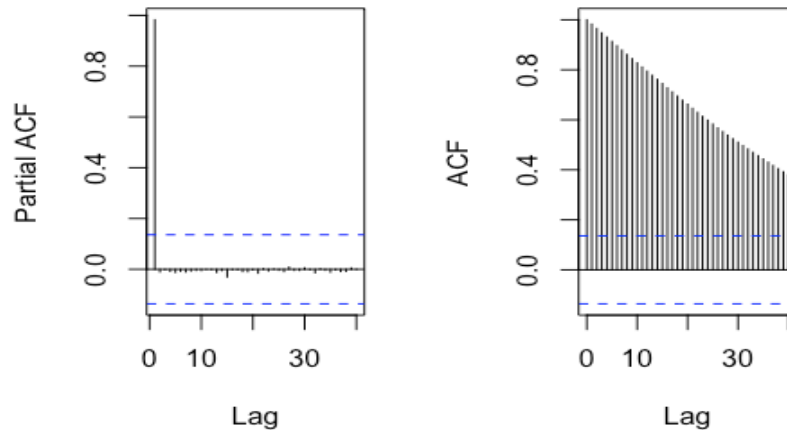
L'istogramma riassume quanto detto in precedenza e, come i test effettuati, non è indicativo di una serie che esige una trasformazione logaritmica.

Figura 1.52- Istogramma della variabile PCE



Si generano dunque i correlogrammi della serie "PCE", volti a sondare la possibile non stazionarietà della serie originaria. Tale ipotesi sembra essere confermata dall'autocorrelazione semplice della serie, che esce con continuità dalle bande e decresce lentamente per i vari lag:

Figura 1.53- Correlogrammi della variabile PCE



La serie oggetto di studio, considerando un livello di significatività al 5%, non sembrerebbe soddisfare la linearità in media. Il Teraesvirta test porta infatti al rifiuto dell'ipotesi nulla H_0 .

Figura 1.54- Teraesvirta test sulla variabile PCE

Teraesvirta Neural Network Test

```
data: pce
X-squared = 7.1877, df = 2, p-value = 0.02749
```

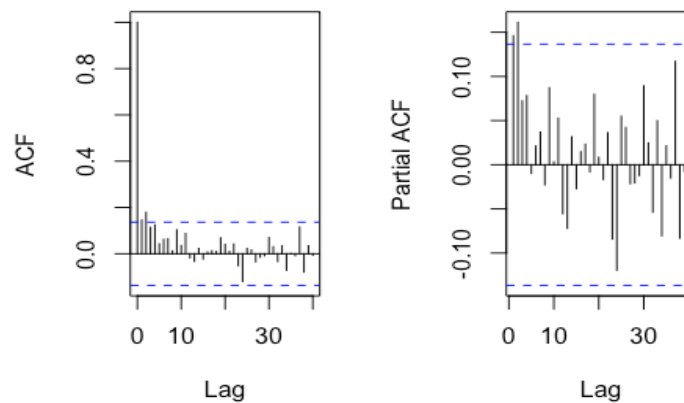
Tornando a quanto premesso con i correlogrammi ai livelli della variabile, il test ADF non fa altro che confermare la non stazionarietà della serie. Quest'ultimo riporta dei valori della statistica test che conducono all'accettazione dell'ipotesi nulla H_0 : anche la serie "PCE" è un'integrata di ordine uno $I(1)$.

Il valore della statistica test, così come la struttura fortemente autocorrelata dell'autocorrelazione semplice ai livelli e, da ultima, la rappresentazione grafica della serie (per la quale si rimanda al paragrafo 1.3) potrebbero suggerire l'esigenza di una ulteriore differenziazione nonché la sussistenza di una variabile $I(2)$.

In realtà questa ipotesi viene scartata dopo aver generato i correlogrammi della serie in differenza prima, che confermano l'interpretazione data alla variabile "PCE", per la quale

è sufficiente una sola differenziazione sull'originaria per la stazionarietà. Non si richiede alcuna destagionalizzazione della serie in fase preliminare.

Figura 1.55- Correlogrammi in differenza prima della variabile PCE



1.6.5 La variabile "PPIT"

La serie storica "PPIT" denota delle statistiche di base leggermente sfasate rispetto a quelle che caratterizzano una gaussiana. Nella fattispecie, la variabile si discosta da una curva normocurtica segnando dei valori a livello del momento terzo e quarto che non si avvicinano rispettivamente a 0 e 3, come illustrato dalle statistiche generate:

Figura 1.56- Statistiche normalità della variabile PPIT

Variance	8.040111
Stdev	2.835509
Skewness	0.564020
Kurtosis	-0.771498

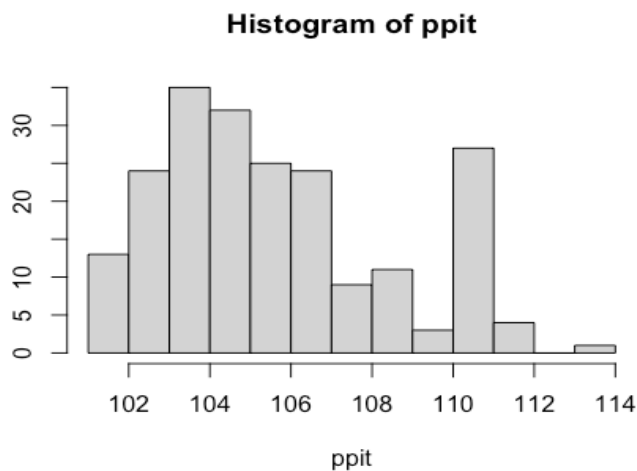
A conferma di quanto premesso il test di Jarque Bera porta a rigettare l'ipotesi nulla di normalità, come si può osservare dalla rappresentazione sottoindicata:

Figura 1.57- Jarque Bera test della variabile PPIT

```
Jarque Bera Test  
  
data: ppit  
X-squared = 16.063, df = 2, p-value = 0.0003251
```

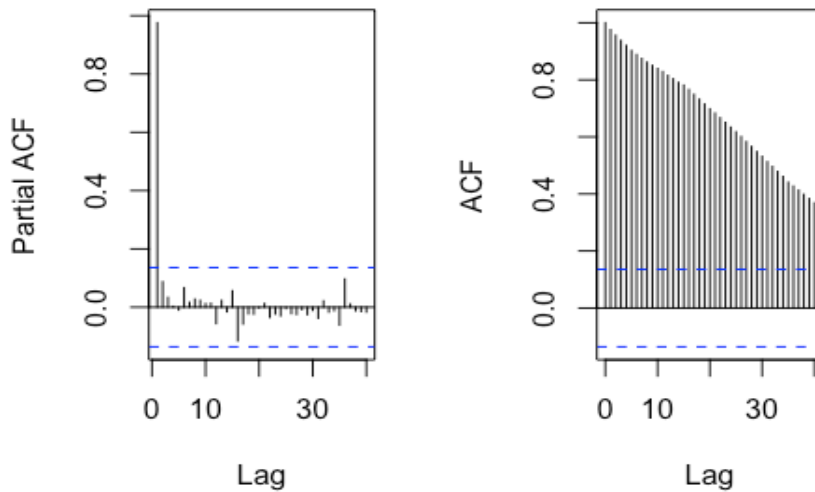
L'istogramma generato dalla variabile originaria raffigura una distribuzione con asimmetria positiva e platicurtica. L'applicazione del logaritmo, anche in questo caso, non porta ad un miglioramento sostanziale di asimmetria e curtosi tale da rendere necessaria l'operazione in questione. Si è optato per il mantenimento della serie originaria a discapito della trasformata logaritmica.

Figura 1.58- Istogramma della variabile PPIT



Ancora una volta assume centralità l'analisi di integrazione della serie, per determinare se quest'ultima presenti delle radici unitarie. La struttura dell'autocorrelazione semplice riprende quanto discusso e riprodotto in seno alle altre variabili esplicative: il correlogramma ai livelli assume i connotati tipici di una serie che presenta radici unitarie, come mostrato nella figura a seguire:

Figura 1.59- Correlogrammi della variabile PPIT



La linearità in media della serie viene indagata dal test di Teraesvirta, che convalida tale ipotesi accettando H_0 . La realizzazione è lineare:

Figura 1.60- Teraesvirta test sulla variabile PPIT

Teraesvirta Neural Network Test

```
data: ppit
```

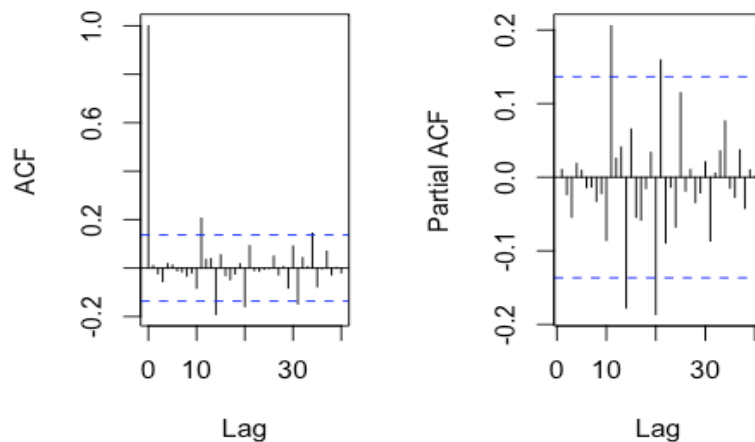
```
X-squared = 1.5466, df = 2, p-value = 0.4615
```

Riprendendo lo studio relativo alla presenza di radici unitarie sulla serie in discussione, la stazionarietà viene vagliata dal test ADF, come di consueto riportato in *Appendice*.

L'esito è in linea con quanto atteso dopo la disamina dei correlogrammi ai livelli. Nello specifico il trend appare significativo e il test va ad accettare l'ipotesi nulla H_0 . La serie "PPIT" verrà considerata come una $I(1)$ e verrà differenziata per essere riportata alla stazionarietà.

Come di consueto vengono proposti i correlogrammi in differenza prima per verificare la correttezza della differenziazione e sondare la presenza di una seconda radice unitaria. La rappresentazione sottostante rigetta questa possibilità definitivamente:

Figura 1.61- Correlogramma in differenza prima della variabile PPIT



1.6.6 La variabile “EM”

La variabile “EM” per il periodo campionario considerato, riporta valori sulle statistiche di base che si discostano dalla distribuzione normale in modo significativo. Dalle statistiche generate emerge infatti una distribuzione asimmetrica e platicurtica:

Figura 1.62- Statistiche normalità variabile EM

Variance	3.659367e+07
Stdev	6.049270e+03
Skewness	6.021310e-01
Kurtosis	-7.313130e-01

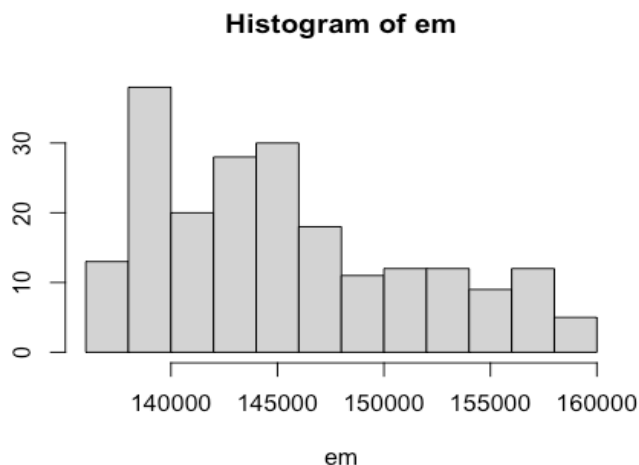
Il test di Jarque Bera rifiuta formalmente l’ipotesi H0 di normalità della serie originaria:

Figura 1.63- Jarque Bera test sulla variabile EM

Jarque Bera Test	
data:	em
X-squared =	17.113, df = 2, p-value = 0.0001923

Anche l'istogramma riporta un comportamento lontano da quello che definisce una curva normocurtica, oltre ad essere indicativo di una asimmetria piuttosto significativa:

Figura 1.64- Iistogramma della variabile EM



Nel caso in analisi, l'asimmetria positiva consente una trasformazione logaritmica della serie originaria, atta ad avvicinare la variabile al momento terzo e quarto della gaussiana. Applicando il logaritmo si registra un lieve aggiustamento della statistica inerente all'asimmetria della variabile "EM", che si avvicina maggiormente alle realizzazioni tipiche di una normale. La curtosi risente di variazioni minime rispetto alla serie storica originaria:

Figura 1.65- Statistiche normalità trasformata logaritmica della variabile EM

Variance	0.001692
Stdev	0.041139
Skewness	0.545762
Kurtosis	-0.796858

Il test di Jarque Bera sulla trasformata logaritmica, pur non portando ad accettare l'ipotesi nulla H_0 , denota un p-value maggiore rispetto a quello generato precedentemente:

Figura 1.66- Jarque Bera test sulla variabile EM_L

Jarque Bera Test

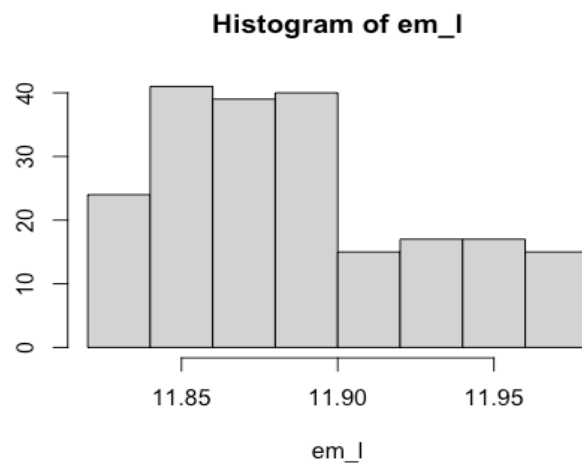
```
data: em_l
```

```
X-squared = 15.688, df = 2, p-value = 0.000392
```

Infine, anche l'istogramma della trasformata logaritmica mostra una distribuzione con una asimmetria positiva, ma meno accentuata rispetto al grafico dell'originaria.

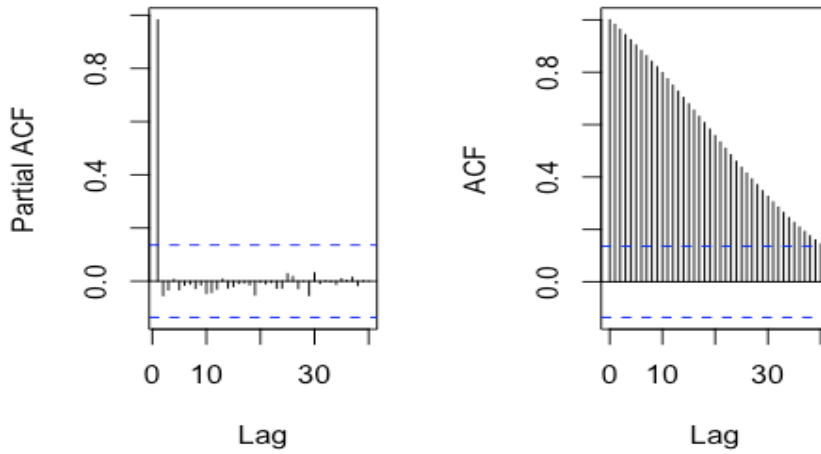
Per le ragioni sopraelencate, con particolare riguardo all'asimmetria e al test di Jarque Bera, si è deciso di proseguire nello studio proposto ricorrendo alla trasformata logaritmica, definita come "EM_L".

Figura 1.67- Iistogramma della variabile EM_L



L'autocorrelazione semplice ai livelli che scaturisce dalla trasformata logaritmica rispetto alla variabile originaria, è indicativa di una serie non stazionaria che presenta, dunque, una o più radici unitarie. La decrescita a livello dei vari lag, pur essendo lenta, mostra l'andamento tipico di un'integrata di ordine uno.

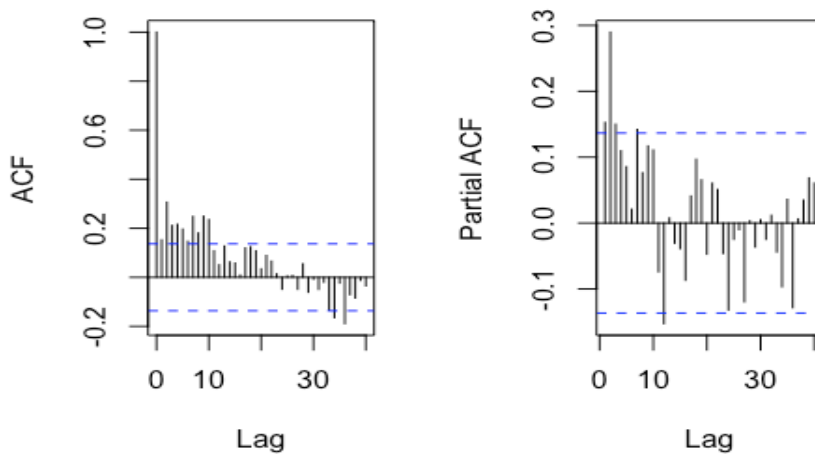
Figura 1.68- Correlogramma della variabile EM_L



Il test ADF, a seguito della rimozione del trend e dell'intercetta (poiché non significativi), conferma l'impressione data dai correlogrammi. La variabile, dopo la trasformazione logaritmica, è caratterizzata da valori della statistica test che, posti in relazione con i valori critici, conducono all'accettazione dell'ipotesi nulla H_0 : la serie "EM_L" ha una radice unitaria e non è stazionaria in media e varianza.

Al fine di verificare l'adeguatezza dell'analisi realizzata vengono riportati i grafici dell'autocorrelazione in differenza prima, che confermano la presenza di una radice unitaria e rigettano la presenza di stagionalità sulla serie:

Figura 1.69- Correlogrammi in differenza prima della variabile EM_L



Concludendo l'analisi descrittiva della serie "EM_L", quest'ultima soddisfa la linearità in media, come dimostrato dal Teraesvirta test che ottiene un p-value che porta all'accettazione dell'ipotesi nulla.

Figura 1.70- Teraesvirta test sulla variabile EM_L

```
Teraesvirta Neural Network Test  
  
data: em_l  
X-squared = 1.7308, df = 2, p-value = 0.4209
```

1.6.7 La variabile "IPT"

La suddetta variabile, relativa alla produzione industriale degli Stati Uniti, è indicativa di una asimmetria che approssima la normale, con valori prossimi a 0. Il momento quarto si discosta significativamente dal caso di una mesocurtica, segnalando valori decisamente più contenuti:

Figura 1.71- Statistiche normalità della variabile IPT

Variance	121.712821
Stdev	11.032353
Skewness	0.130697
Kurtosis	-0.922565

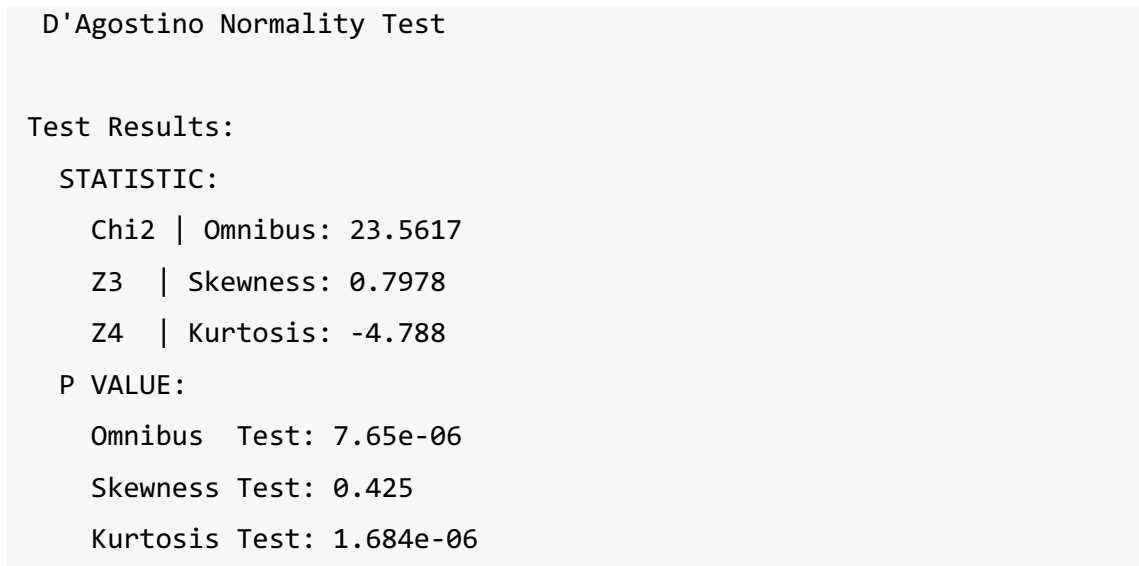
Il test di Jarque Bera, nonostante i valori assunti dalla curtosi, registra un p-value per il quale si rifiuta l'ipotesi nulla solo per un livello di significatività al 5%:

Figura 1.72- Jarque Bera test sulla variabile IPT

```
Jarque Bera Test  
  
data: ipt  
X-squared = 7.659, df = 2, p-value = 0.02172
```

Il test di D'Agostino conferma la prossimità della skewness rispetto alla gaussiana e, per contro, la distanza del momento quarto dalla stessa.

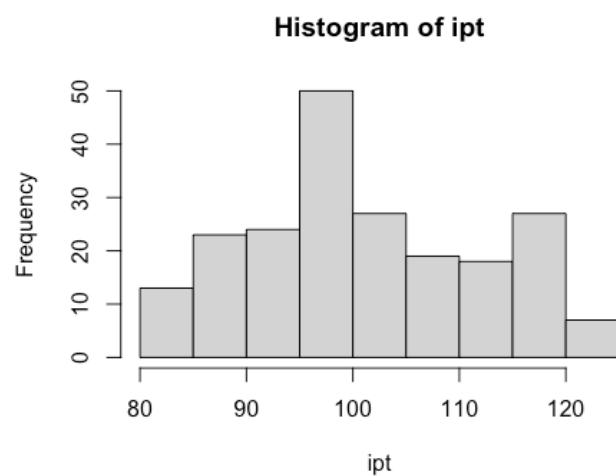
Figura 1.73- D'Agostino test sulla variabile IPT



L'istogramma fornisce un'interpretazione grafica di quanto premesso ed evidenzia una distribuzione platicurtica con una leggera asimmetria positiva.

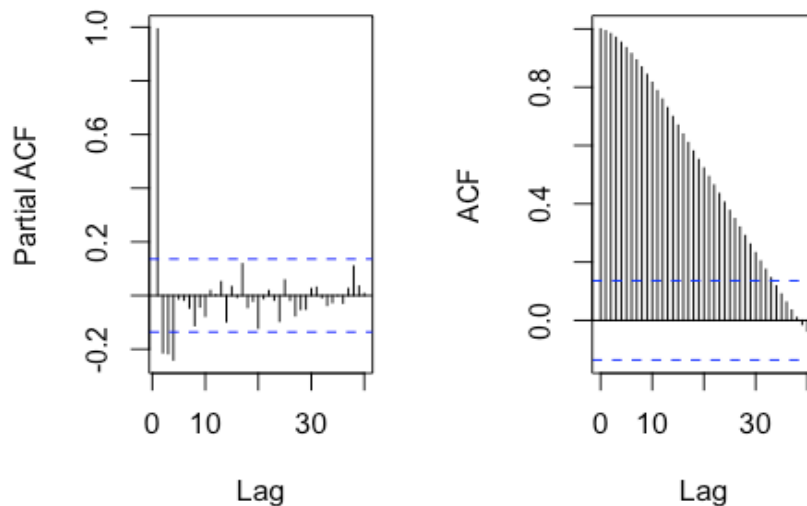
Tali valori consentono la prosecuzione dell'analisi mantenendo l'originaria e senza la necessità di applicare il logaritmo.

Figura 1.73- Istogramma della variabile IPT



Guardando alla stazionarietà della serie originaria “IPT”, i correlogrammi generati sono rappresentativi di una variabile presumibilmente non stazionaria in media e varianza. Si noti infatti come la decrescita dell’autocorrelazione semplice ai livelli sia tutt’altro che rapida:

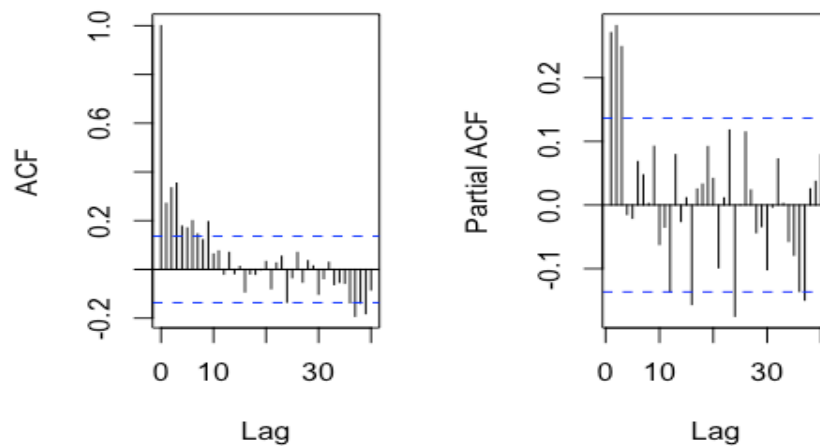
Figura 1.74- Correlogramma della variabile IPT



Per rendere formale la non stazionarietà viene effettuato il test ADF per la presenza di radice unitaria. Dai test emerge un valore di statistica test tale da accettare l’ipotesi nulla H_0 , comprovando dunque la presenza di una radice unitaria. La variabile “IPT” è non stazionaria, nello specifico integrata di ordine uno (I(1)) e ne andrà differenziata l’originaria per essere interpretata correttamente.

I correlogrammi alle differenze prime sono tipici di una serie che è stata riportata alla stazionarietà a seguito della differenziazione, così come delineato dalla figura sottostante:

Figura 1.75- Correlogramma in differenza prima della variabile IPT



Non si segnala inoltre l'esigenza di effettuare alcuna procedura di destagionalizzazione della serie.

In quanto alla linearità la variabile in discussione risulta essere lineare in media. Viene infatti accettata l'ipotesi nulla del test di Teraesvirta relativo alla linearità in media della realizzazione:

Figura 1.76- Teraesvirta test sulla variabile IPT

```
Teraesvirta Neural Network Test

data: ipt
X-squared = 0.30163, df = 2, p-value = 0.86
```

1.6.8 La variabile "RET"

La variabile "RET" riporta delle statistiche preliminari sommariamente in linea con le peculiarità della gaussiana. Il momento terzo tende a zero pur registrando una lieve asimmetria negativa; il momento quarto è inferiore rispetto al valore 3 della normale, segnale di una curva leggermente appiattita:

Figura 1.77- Statistiche normalità della variabile RET

Variance	36.643805
Stdev	6.053413
Skewness	-0.101144
Kurtosis	-0.812150

Come di consueto si procede eseguendo il test di Jarque Bera, con il quale verrà accettata l'ipotesi nulla di normalità per la serie in questione:

Figura 1.78- Jarque Bera test sulla variabile RET

Jarque Bera Test
data: ret
X-squared = 5.7818, df = 2, p-value = 0.05553

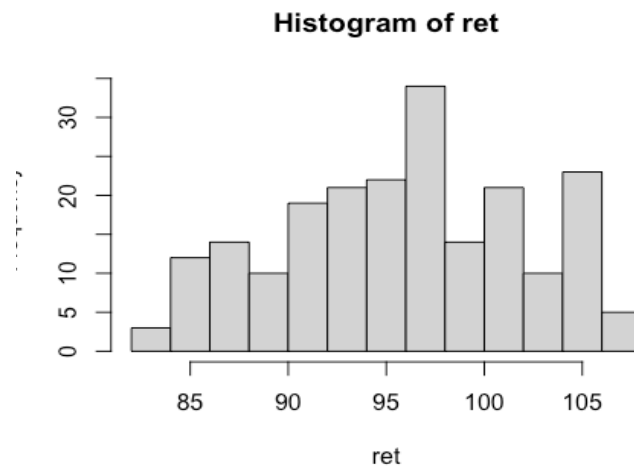
Il test di D'Agostino dà una panoramica più approfondita sulle ragioni che portano ad accettare H0. Nella fattispecie le motivazioni per cui la variabile "RET" approssima la distribuzione normale si riscontrano proprio a livello del momento terzo della serie.

Figura 1.79- D'Agostino test sulla variabile RET

D'Agostino Normality Test
Test Results:
STATISTIC:
Chi2 Omnibus: 14.2361
Z3 Skewness: -0.6184
Z4 Kurtosis: -3.7221
P VALUE:
Omnibus Test: 0.0008104
Skewness Test: 0.5363
Kurtosis Test: 0.0001976

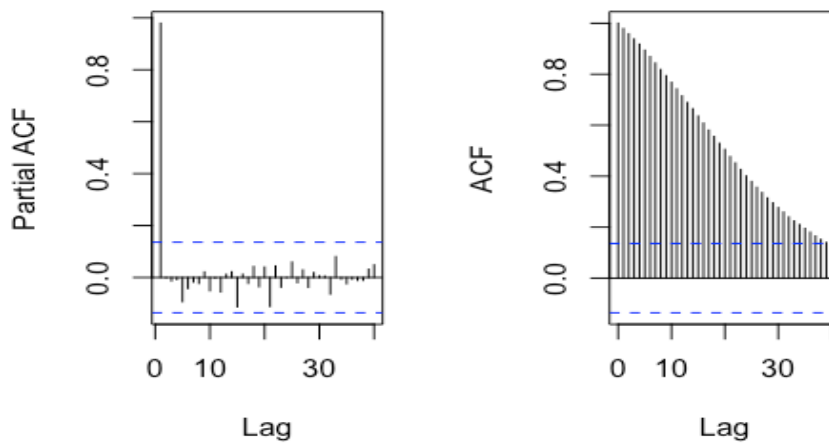
La distribuzione che si manifesta nel seguente istogramma riflette quanto detto precedentemente:

Figura 1.80- Istogramma della variabile RET



L'interpretazione dei correlogrammi risulta fondamentale per inquadrare correttamente la serie alla luce della possibile non stazionarietà. Analogamente alle variabili precedentemente discusse, "RET" mostra una struttura ai livelli dell'autocorrelazione semplice propria di una serie non stazionaria.

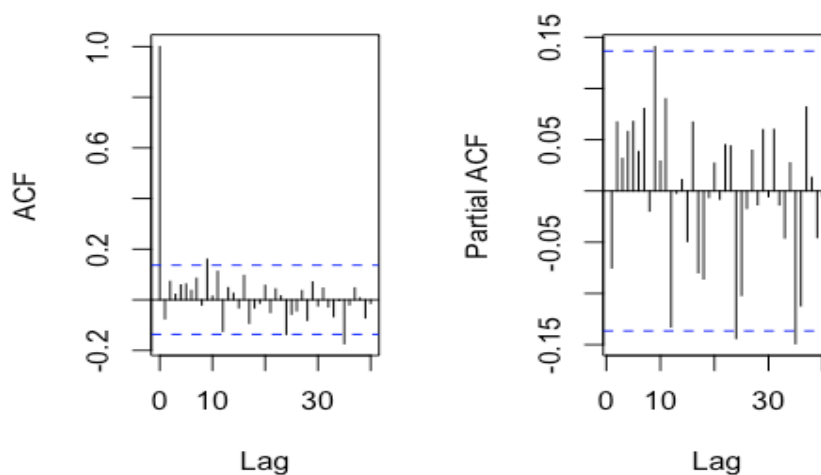
Figura 1.81- Correlogramma della variabile RET



La struttura dei correlogrammi anticipa quanto viene avvalorato dagli esiti dell'ADF test. Quest'ultimo, a seguito della rimozione di trend e drift (poiché non significativi), realizza una statistica test che, confrontata con i valori critici, porta ad accettare l'ipotesi nulla di radice unitaria. La serie "RET" sarà dunque non stazionaria in media e varianza e presenta una radice unitaria. Ancora una volta si procederà alla differenziazione per riportare la stessa alla stazionarietà.

Un accertamento dell'avvenuta differenziazione viene rilevato generando i correlogrammi in differenza prima. Questi ultimi sono coerenti con i risultati ottenuti a livello dei test.

Figura 1.82- Correlogramma in differenza prima della variabile RET



Così come per il test ADF e il Jarque Bera, anche il Teraesvirta assume un p-value da cui consegue l'accettazione dell'ipotesi nulla di linearità in media:

Figura 1.83- Teraesvirta test sulla variabile RET

Teraesvirta Neural Network Test

data: ret

X-squared = 1.0933, df = 2, p-value = 0.5789

1.7 Ulteriori variabili utili ai fini dell'equazione di breve periodo

Le variabili descritte nel paragrafo appena concluso, se non stazionarie rientrano (come si vedrà nel secondo capitolo) sia nel primo che nel secondo step della procedura a due passi di Engle-Granger. In fase di stima dell'equazione di breve periodo, altresì definita “dinamica”, influiranno ulteriori serie storiche che si sono rivelate fondamentali per la realizzazione di previsioni attendibili ed efficaci, pur non essendo presenti all'interno dell'equazione statica. Ciò precisato, si proporrà una descrizione delle variabili in oggetto, seguendo lo stesso procedimento effettuato sulle serie che caratterizzano la totalità del progetto discusso.

1.7.1 La variabile “LF”

La variabile “LF” rappresenta la forza lavoro negli Stati Uniti (“Labor Force Participation Rate”). È espressa come un tasso ed è stata scaricata dal database FRED con osservazioni mensili, già destagionalizzata. Nonostante la destagionalizzazione effettuata, come si vedrà, la serie presenta un residuo di stagionalità (sebbene non particolarmente significativa) nell'autocorrelazione in differenza prima. Ciò può essere indice di una destagionalizzazione che può non essere stata effettuata correttamente. Applicando la destagionalizzazione secondo la modalità “TRAMO SEATS” su R e Eviews, la procedura non rimuove completamente la stagionalità residua. Si è deciso dunque di adottare la serie originaria valutando come non significativa la stagionalità residua che emerge dai correlogrammi in differenza.

Entrando nel merito dell'analisi statistica della variabile, emerge un'asimmetria perfettamente allineata con quella di una gaussiana, mentre i valori decisamente più contenuti della curtosi sono indicativi di una distribuzione platicurtica e/o di osservazioni simmetriche ma poste agli estremi:

Figura 1.84- Statistiche normalità variabile LF

Variance	2.080242
Stdev	1.442304

Skewness	0.099618
Kurtosis	-1.733364

In ragione di ciò viene rifiutata l'ipotesi nulla del test sulla normalità di Jarque Bera:

Figura 1.85- Jarque Bera test sulla variabile LF

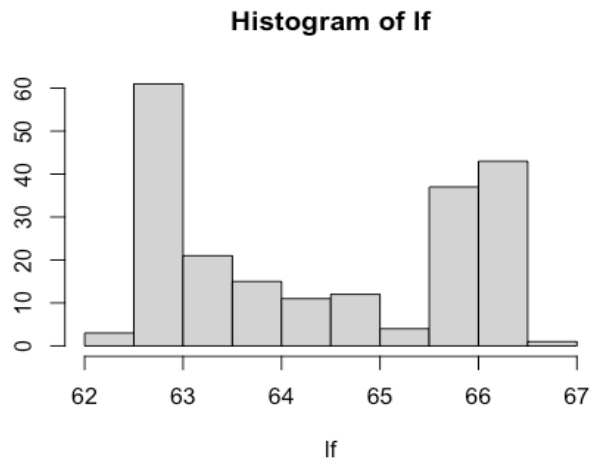
Jarque Bera Test

data: lf

X-squared = 26.021, df = 2, p-value = 2.237e-06

L'istogramma mostra una situazione atipica rispetto ad una curva normocurtica ed emerge la presenza di una forte componente distributiva in corrispondenza delle code:

Figura 1.86- Istogramma della variabile LF



La linearità in media viene rigettata dal test di Teraesvirta secondo un livello di significatività al 5%, come rappresentato di seguito:

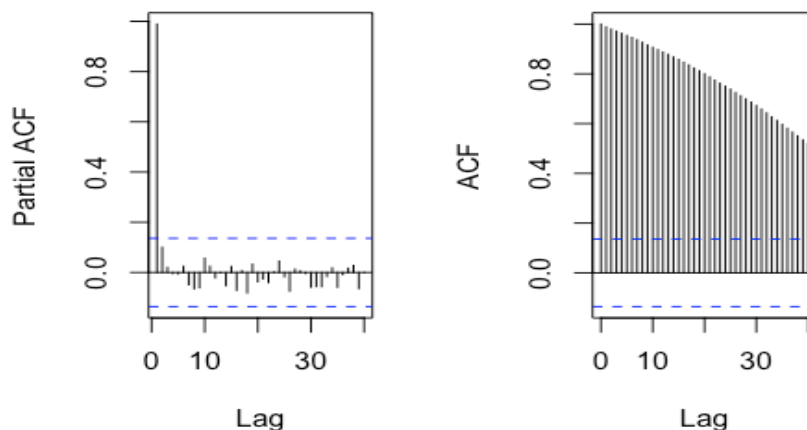
Figura 1.87- Teraesvirta test sulla variabile LF

```
Teraesvirta Neural Network Test

data: lf
X-squared = 9.1515, df = 2, p-value = 0.0103
```

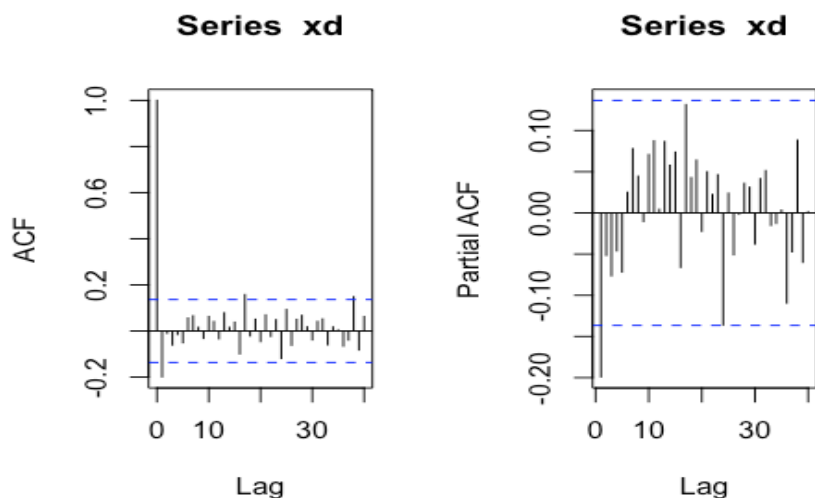
In quanto allo studio di integrazione della serie, l'autocorrelazione semplice ai livelli ha una decrescita piuttosto lenta nei vari lag, con una struttura analoga ai casi già analizzati indicativi di non stazionarietà.

Figura 1.88- Correlogramma della variabile LF



Quanto anticipato dalla figura di cui sopra viene comprovato dal test ADF sulla serie originaria. A seguito della rimozione del trend e del drift dal test poiché non significativi, la variabile in discussione risulta non stazionaria e, nello specifico, la serie “LF”, dopo aver appurato la non linearità e la non normalità sarà considerata I(1). Si propongono dunque i correlogrammi della variabile cui è stata applicata la differenza prima, processo necessario per riportare la serie alla stazionarietà. Tale interpretazione della serie “LF” trova riscontro nelle rappresentazioni sottostanti, nelle quali si risolve la non stazionarietà:

Figura 1.89- Correlogramma in differenza prima della variabile LF



1.7.2 La variabile “USD”

La serie storica “USD” rappresenta il tasso di cambio dollaro/euro (U.S. Dollars to Euro Spot Exchange Rate). Mostra dunque il rapporto tra le due valute espresso come il quantitativo di dollari per un euro. Il tasso di cambio in questione denota un interesse notevole da un punto di vista macroeconomico, essendo espressione del rapporto che intercorre tra le valute di due continenti e partner commerciali. La serie è stata scaricata dalla banca dati FRED su osservazioni mensili. Sulla destagionalizzazione della serie vale lo stesso ragionamento della variabile “LF”.

Approfondendo l’analisi preliminare di “USD” emergono le seguenti indicazioni di base:

Figura 1.90- Statistiche normalità variabile USD

Variance	0.015663
Stdev	0.125153
Skewness	0.240928
Kurtosis	-0.574776

I dati sembrano comprovare la normalità della serie, sebbene siano indice di una curtosi leggermente inferiore al valore pari a 3 tipico della gaussiana. La bontà del momento

terzo e quarto rispetto alla normale viene convalidata dal p-value che emerge dal test di Jarque Bera:

Figura 1.91- Jarque Bera test sulla variabile USD

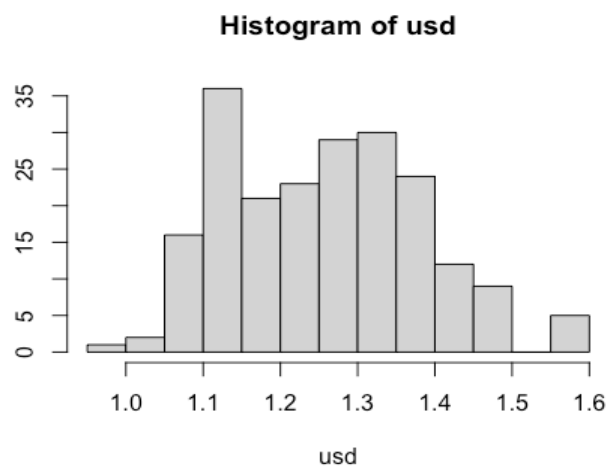
Jarque Bera Test

data: usd

X-squared = 4.6755, df = 2, p-value = 0.09654

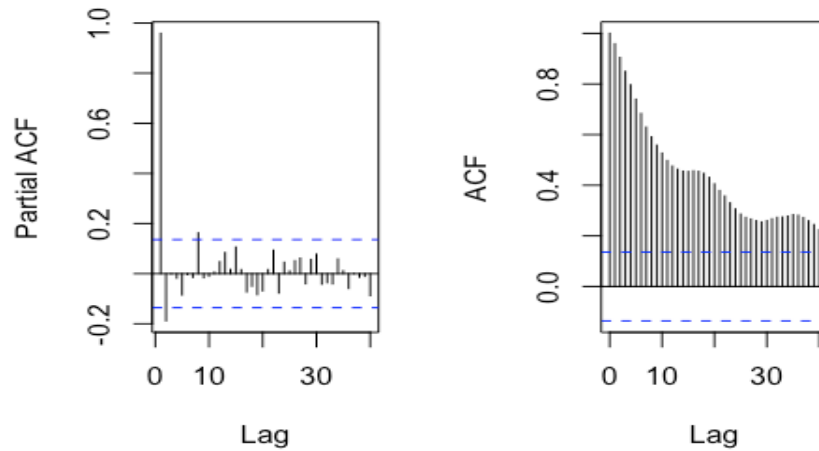
Anche l'istogramma nel complesso è indicativo di una fattispecie approssimabile alla distribuzione normale:

Figura 1.92- Istogramma della variabile USD



Volgendo l'attenzione all'analisi di integrazione della variabile, si nota come i correlogrammi ai livelli di "USD" siano indicativi di una situazione che presenta non stazionarietà:

Figura 1.93- Correlogramma della variabile USD



Emerge infatti una struttura dell'autocorrelazione semplice che decresce lentamente, tipica di questa situazione. Verrà confermata formalmente la presenza di radice unitaria effettuando il test ADF. L'intercetta appare significativa considerando un livello di significatività al 5%. Il valore che scaturisce dalla statistica test, confrontato con i valori critici, indica la presenza di una radice unitaria: la serie "USD" è non stazionaria e verrà differenziata come di consueto.

Prima di ciò volgendo uno sguardo alla linearità, si riscontra che la variabile "USD" soddisfa in media questa proprietà, come si può desumere dall'accettazione di H0 dal Teraesvirta test:

Figura 1.94- Teraesvirta test sulla variabile USD

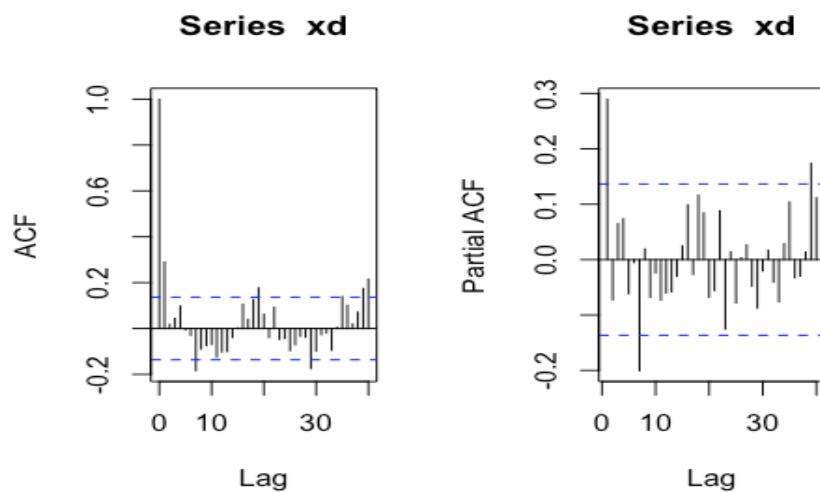
Teraesvirta Neural Network Test

```
data:  usd
```

```
X-squared = 1.5028, df = 2, p-value = 0.4717
```

Tornando alla non stazionarietà di questa serie, a seguito della differenza prima sulla variabile esplicativa per ricondurre la stessa alla stazionarietà, avremo i seguenti correlogrammi:

Figura 1.95- Correlogrammi in differenza prima della variabile USD



1.7.3 La variabile “CHUS”

Come la variabile di cui si è appena conclusa la disamina, anche la serie denominata “CHUS” rappresenta un tasso di cambio, nella fattispecie il tasso di cambio a spot tra lo Yuan cinese e il dollaro, espresso come ammontare di valuta cinese per un dollaro.

Le statistiche di base denotano una leggera asimmetria positiva, con una curtosi più bassa rispetto a quella tipica di una normale:

Figura 1.96- Statistiche normalità della variabile CHUS

Variance	0.556597
Stdev	0.746055
Skewness	0.621145
Kurtosis	-1.071805

Il test di Jarque Bera conferma il fatto che la distribuzione realizzata dalla serie non si avvicina a quella di una normale, rifiutando l'ipotesi nulla di normalità H_0 :

Figura 1.97- Jarque Bera test sulla variabile CHUS

Jarque Bera Test

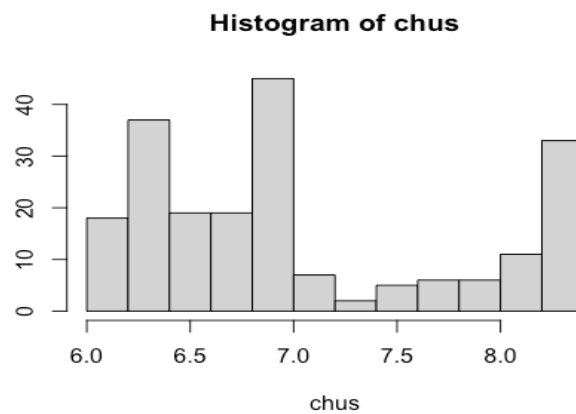
data: chus

X-squared = 23.182, df = 2, p-value = 9.249e-06

Applicando il logaritmo sulla serie originaria la trasformata non riporta alcun avvicinamento alla normale, registrando valori pressochè analoghi ai precedenti.

Si è optato quindi per la continuazione dello studio utilizzando la variabile originaria, il cui istogramma sottostante mostra un comportamento lontano da una curva mesocurtica:

Figura 1.98- Istogramma della variabile CHUS



Oltre alla non normalità della serie, il test di Teraesvirta comprova la non linearità in media, attraverso il rifiuto di H_0 :

Figura 1.99- Teraesvirta test sulla variabile CHUS

Teraesvirta Neural Network Test

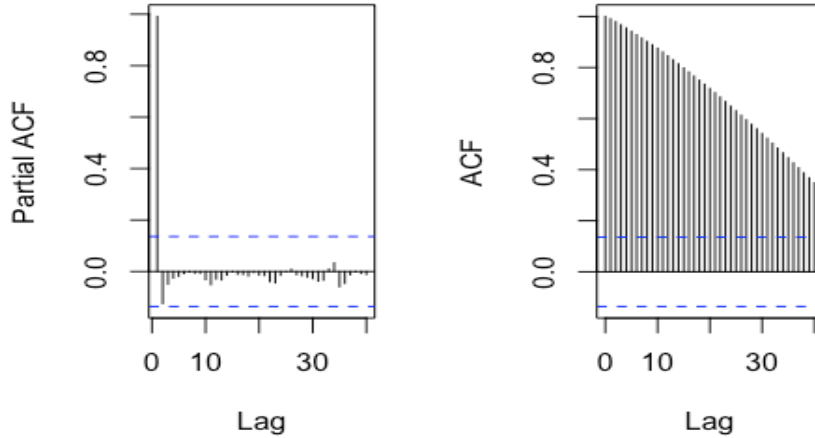
data: chus

X-squared = 12.187, df = 2, p-value = 0.002258

Come di consueto si pone lo sguardo alla stazionarietà della serie, affinché quest'ultima venga inquadrata correttamente nella procedura a due passi di Engle-Granger. I

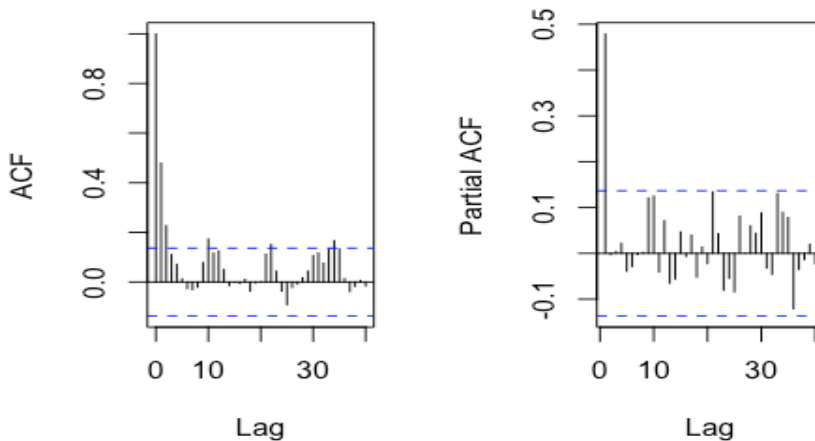
correlogrammi suggeriscono la non stazionarietà della serie ai livelli, vista la decrescita piuttosto modesta dell'autocorrelazione semplice:

Figura 1.100- Correlogrammi della variabile CHUS



La conferma di ciò risiede nel test ADF, volto a sondare la presenza di radici unitarie. Qui si andrà ad accettare l'ipotesi nulla H_0 di radice unitaria, dopo aver rimosso trend e intercetta poiché non significativi. La variabile "CHUS" è dunque non normale, non lineare e non stazionaria (nello specifico $I(1)$). Ancora una volta per essere ricondotta alla stazionarietà la serie originaria deve essere differenziata. La possibilità di una seconda radice unitaria viene rigettata dalla struttura dei correlogrammi in differenza prima, la cui autocorrelazione semplice esce dalle bande di Bartlett occasionalmente.

Figura 1.101- Correlogramma in differenza prima della variabile CHUS



CAPITOLO II - Procedura a due passi di Engle-Granger

2.1 Analisi di cointegrazione e stima dell'equazione statica

Una volta ultimata l'analisi di integrazione delle singole variabili si può procedere alla stima delle equazioni attraverso la procedura a due passi di Engle-Granger, finalizzata a descrivere la relazione che intercorre tra le variabili esogene e quella endogena. Tale processo, si articola in due fasi principali, a loro volta costituite da ulteriori passaggi e test che verranno giustificati di volta in volta.

La prima fase prevede la stima dell'equazione statica, ovverosia di un modello di regressione lineare, ottenuto ricorrendo al metodo OLS (minimi quadrati ordinari) con correzione HAC.

Generata la serie dei residui dall'equazione di cui sopra, si passerà al secondo step, volto a stimare l'equazione dinamica in forma ECM.

La stima dell'equazione statica e dinamica fa affidamento ad alcuni elementi teorici che risultano essenziali per la specificazione del modello.

Il primo è il concetto di cointegrazione⁴. Nella teoria econometrica quest'ultimo prelude ad un legame tra due o più variabili integrate di ordine uno ($I(1)$), che si diranno cointegrate quando la combinazione lineare delle stesse darà origine ad un grado di integrazione $I(0)$. In altri termini, si ammette l'esistenza di una combinazione lineare avente un grado di integrazione ridotto rispetto a quello delle singole variabili che la caratterizzano. Tale concetto viene utilizzato in seguito alla stima dell'equazione statica e dopo aver generato la serie dei residui. A questo livello la presenza di cointegrazione consentirà di legare l'equazione statica a quella dinamica, proprio attraverso la serie ECM originata, che verrà inserita nel secondo passo in sede di stima dell'equazione di breve periodo.

⁴ Engle, Robert F., Granger, Clive W. J. (1987), *Co-integration and error correction: Representation, estimation and testing*, *Econometrica*; 55.

Una proprietà che riveste un ruolo fondamentale è quella di superconsistenza, sviluppata da J. H. Stock⁵. Riprendendo l'autore, date due variabili X e Y integrate di ordine 1, se X è cointegrata con Y lo stimatore convergerà verso il valore vero più velocemente rispetto a quanto accade per due variabili stazionarie e non cointegrate. Si deve a questa proprietà il motivo per il quale nella stima dell'equazione statica verranno inserite le sole variabili non stazionarie, individuate in sede di analisi preliminare.

Sempre in sede di stima dell'equazione statica verrà applicata la correzione HAC di Newey–West sulla matrice di covarianza degli stimatori. Questa procedura è giustificata dal fatto che nella pratica econometrica, il più delle volte i residui generati presentano eteroschedasticità e autocorrelazione. La stima effettuata sarà di tipo “HAC standard error and covariance estimator” e andrà a correggere i soli errori standard stimati, senza influire sui singoli parametri.

Entrando nel merito del primo dei due passi di Engle-Granger, l'equazione statica viene stimata secondo un approccio deduttivo, che va dal generale al particolare.

Ciò avviene per diversi ordini di ragioni, la più rilevante riguarda il fatto che in una qualsivoglia regressione, l'esclusione di una variabile rilevante dal modello di specificazione ha una gravità maggiore rispetto all'inserimento di una variabile irrilevante. Nel primo caso, infatti, il modello di specificazione porta ad uno stimatore non solo non efficiente, ma anche distorto. Nel caso di un modello di specificazione che include una variabile non rilevante rispetto al DGP, avremo invece delle stime non distorte, seppur non efficienti. Conseguenza ne è che il rischio di incorrere in una regressione spuria è propria della prima fattispecie, che prevede l'esclusione di variabili rilevanti. La procedura seguita nel presente studio prevede l'inserimento di tutte le variabili che si presume siano esogene, per poi eliminare singolarmente quelle che si rivelano irrilevanti.

Per comprendere quali serie storiche siano non rilevanti nella spiegazione del NASDAQ Composite e per arrivare all'equazione finale di lungo periodo si ricorre ai criteri di selezione elencati di seguito:

⁵ Stock J., (1987), *Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors*. *Econometric*, 55 :1035-1056.

- Significatività al 5% del regressore
- Adjusted R-Squared
- Akaike info criterion, denominato “AIC”
- Schwarz info criterion, denominato “BIC” o “SIC”

I criteri seguono un ordine cronologico in quanto ad applicazione: nella fattispecie sono state inizialmente rimosse quelle variabili che superavano il 5% di significatività, per poi rivolgere l’attenzione ai valori assunti dall’Adjusted R-Squared.

I criteri informativi di Akaike (“AIC”) e Schwarz (“BIC”) sono soggetti ad un’interpretazione analoga, con il secondo tra i due che penalizza maggiormente il numero di parametri. In ambo i casi la lettura delle statistiche BIC e AIC è favorevole in corrispondenza a valori più bassi rispetto alla rimozione di variabili dalla regressione o a modelli alternativi rispetto a quello definito.

L’equazione statica generale proposta in prima battuta, per il periodo compreso tra 2002/10/01 e 2020/01/01, è la seguente:

Figura 2.1- Stima dell'equazione statica generale

Dependent Variable: N_L
Method: Least Squares
Date: 12/11/22 Time: 19:16
Sample: 2002M10 2020M01
Included observations: 208
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	56.83913	15.50001	3.667039	0.0003
BOND	0.053249	0.017784	2.994176	0.0031
CC	0.005882	0.001046	5.622093	0.0000
TB	-0.027431	0.007604	-3.607220	0.0004
RET	0.021552	0.007343	2.935190	0.0037
PCE	0.000309	0.000102	3.020179	0.0029
EM_L	-4.842258	1.331540	-3.636584	0.0004
IPT	0.008954	0.002687	3.332210	0.0010
PPIT	0.023384	0.006637	3.523167	0.0005
CHUS	0.008882	0.030163	0.294477	0.7687
CPI	-0.017701	0.012945	-1.367343	0.1731
DEBG	-2.06E-06	1.83E-05	-0.112568	0.9105
UN_L	-0.078343	0.114069	-0.686800	0.4930
FF_L	-0.027011	0.014895	-1.813455	0.0713
W	0.043100	0.079025	0.545403	0.5861
WTI_L	-0.023531	0.042310	-0.556163	0.5788
USD	0.273323	0.096673	2.827291	0.0052
PSV_L	0.017694	0.028853	0.613256	0.5404

R-squared	0.991955	Mean dependent var	8.069092
Adjusted R-squared	0.991235	S.D. dependent var	0.511076
S.E. of regression	0.047847	Akaike info criterion	-3.159037
Sum squared resid	0.434981	Schwarz criterion	-2.870212
Log likelihood	346.5399	Hannan-Quinn criter.	-3.042251
F-statistic	1378.057	Durbin-Watson stat	1.198120
Prob(F-statistic)	0.000000	Wald F-statistic	1456.410
Prob(Wald F-statistic)	0.000000		

Le serie storiche di cui non si è proposta l'analisi descrittiva e di integrazione⁶ sono considerate a tutti gli effetti irrilevanti nel contesto della procedura a due passi di Engle-Granger e avrebbero portato a stime non efficaci. Per completezza si riporta di seguito un breve elenco delle variabili in questione:

⁶ Tutte le serie storiche che sono state incluse in un primo momento nell'equazione statica generale hanno subito gli stessi processi delle variabili esplicative che sono state effettivamente selezionate per la statica finale. Si è evitato di riportare in seno ai paragrafi relativi all'analisi descrittiva e di integrazione lo stesso grado di approfondimento riservato ai regressori che si sono rivelati effettivamente significativi e che, conseguentemente a ciò, sono stati inseriti nell'equazione finale di breve o di lungo periodo.

- CPI: “Consumer Price Index: Total All Items for the United States”; indice su dati mensili; destagionalizzata. Fonte: banca dati FRED.
- DEBG: “Market Value of Gross Federal Debt”; miliardi di dollari su dati mensili; non destagionalizzata. Fonte: banca dati FRED.
- UN_L: Trasformata logaritmica di “Unemployment Rate”; percentuale su dati mensili; destagionalizzata. Fonte: banca dati FRED.
- FF_L: Trasformata logaritmica di “Federal Funds Effective Rate”; percentuale su dati mensili; non destagionalizzata. Fonte: banca dati FRED.
- W: “Average Hourly Earnings Total Private”; dollari all’ora su dati mensili; destagionalizzata. Fonte: banca dati FRED.
- WTI_L: Trasformata logaritmica di “Spot Crude Oil Price: West Texas Intermediate”; dollari al barile su dati mensili; non destagionalizzata. Fonte: banca dati FRED.
- PSV_L: Trasformata logaritmica di “Personal Saving Rate”; percentuale su dati mensili; destagionalizzata. Fonte: banca dati FRED.

Ovviamente le serie elencate, dopo il test per la radice unitaria ADF o Phillips Perron sono risultate integrate di ordine 1 (I(1)). I regressori “CHUS” e “USD”, di cui si è invece proposta l’analisi di integrazione, sono stati rimossi dall’equazione di lungo periodo, ma risulteranno significativi nell’equazione di breve.

Si può osservare che le variabili inserite in un primo momento come ad esempio “DEBG”, “WTI_L” o “PSV_L” presentano un livello di significatività superiore al 5%, non sono dunque significative e verranno rimosse dalla stima dell’equazione statica. Si procede alla rimozione dei regressori iniziando da quelli meno significativi. Tali variabili non appaiono infatti utili nella relazione di lungo periodo per la variabile endogena “N_L”, ma potranno fornire un apporto in fase di stima dell’equazione dinamica, dopo aver verificato la cointegrazione.

Come anticipato nell’introduzione del capitolo in oggetto, nella pratica econometrica si includono, nell’equazione di lungo periodo, le sole variabili che presentano radici

unitarie, nel caso di I(1) senza la necessità di alcuna differenziazione⁷. Eventuali variabili stazionarie verranno riprese sempre a livello dell'equazione dinamica, dove le serie che presentano radici unitarie verranno invece differenziate coerentemente con il loro grado di integrazione.

Rimuovendo i regressori secondo quanto indicato dalla significatività e dai criteri di selezione, si otterrà la seguente equazione statica finale:

Figura 2.2- Stima dell'equazione statica finale

Dependent Variable: N_L
Method: Least Squares
Date: 11/26/22 Time: 18:08
Sample: 2002M10 2020M01
Included observations: 208
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

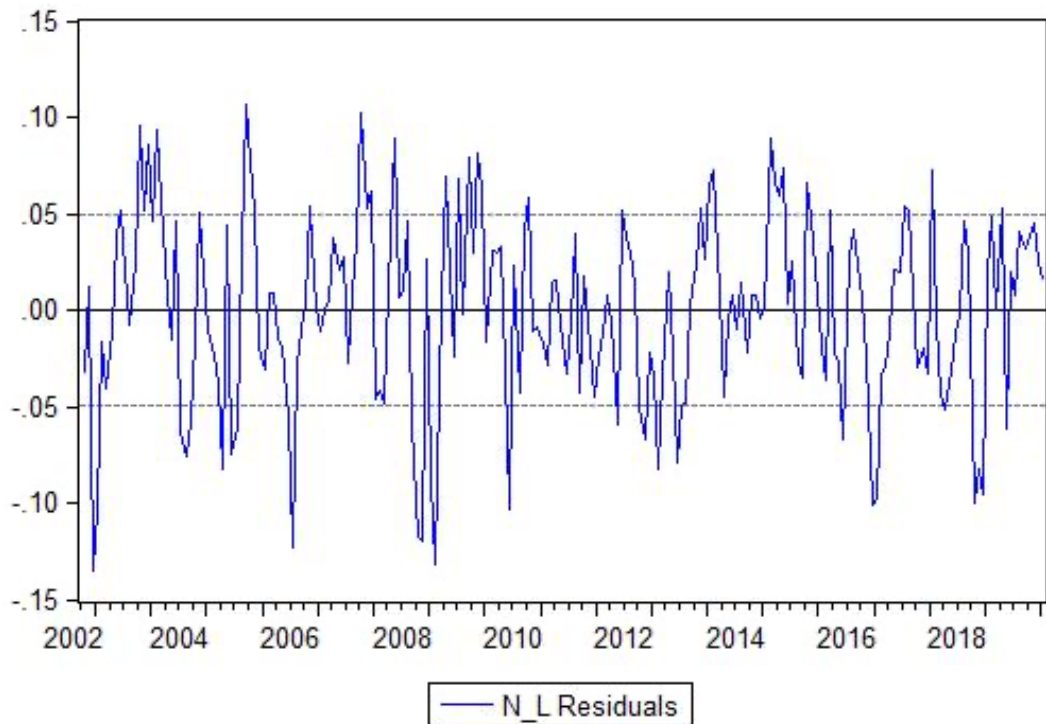
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	45.56460	10.06726	4.526017	0.0000
BOND	0.070400	0.009411	7.480527	0.0000
CC	0.006041	0.000777	7.779124	0.0000
TB	-0.041635	0.007025	-5.927050	0.0000
RET	0.023354	0.005415	4.312869	0.0000
PCE	0.000260	3.12E-05	8.331926	0.0000
EM_L	-3.859475	0.876468	-4.403442	0.0000
IPT	0.006443	0.002820	2.285031	0.0234
PPIT	0.018819	0.007414	2.538309	0.0119

R-squared	0.990816	Mean dependent var	8.069092
Adjusted R-squared	0.990447	S.D. dependent var	0.511076
S.E. of regression	0.049953	Akaike info criterion	-3.113158
Sum squared resid	0.496568	Schwarz criterion	-2.968746
Log likelihood	332.7685	Hannan-Quinn criter.	-3.054765
F-statistic	2683.605	Durbin-Watson stat	1.165669
Prob(F-statistic)	0.000000	Wald F-statistic	1666.733
Prob(Wald F-statistic)	0.000000		

Dall'equazione statica ottenuta si andrà a generare sin da subito la serie dei residui, il cui grafico è riportato nella figura sottostante:

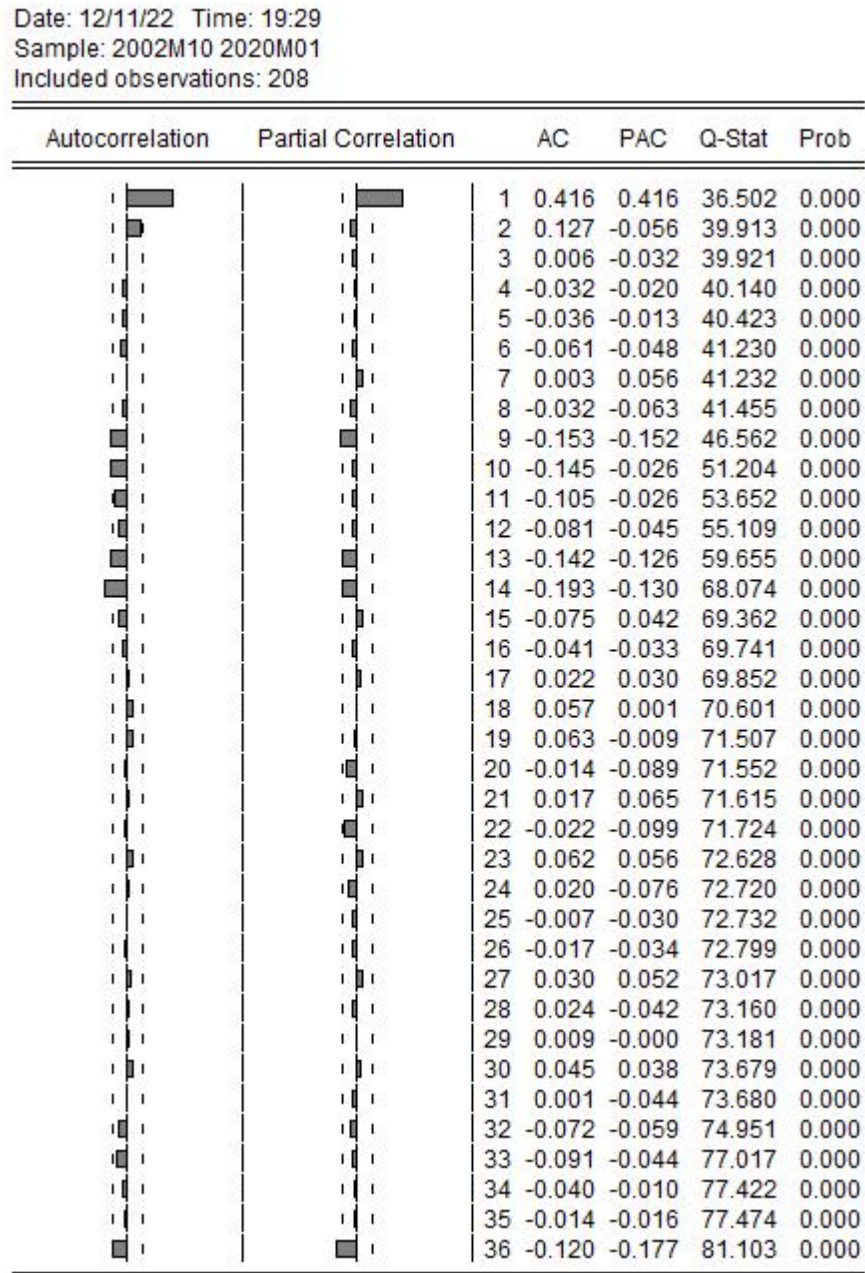
⁷ Qualora la variabile fosse I(2) il regressore andava inserito in differenza prima.

Figura 2.3- Residui generati dall'equazione statica finale



In primis andrà verificata, attraverso il correlogramma della stessa serie la stazionarietà, conditio sine qua non per passare al secondo passo della procedura di Engle-Granger, per evitare il rischio di regressioni spurie e, in definitiva, per accertare la cointegrazione:

Figura 2.4- Correlogramma della serie dei residui dell'equazione statica



L'autocorrelazione semplice dei residui (denominati "ECM") decresce rapidamente, anticipando che con ogni probabilità la serie risulterà stazionaria. Per formalizzare la stazionarietà dei residui viene effettuato il test ADF sulla componente ai livelli:

Figura 2.5- ADF test sulla serie dei residui dell'equazione statica

Null Hypothesis: ECM_ALT has a unit root
 Exogenous: None
 Lag Length: 0 (Automatic - based on SIC, maxlag=14)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-9.224679	0.0000
Test critical values:		
1% level	-2.576181	
5% level	-1.942368	
10% level	-1.615679	

*Mackinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(ECM_ALT)
 Method: Least Squares
 Date: 12/11/22 Time: 19:27
 Sample (adjusted): 2002M11 2020M01
 Included observations: 207 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ECM_ALT(-1)	-0.583884	0.063296	-9.224679	0.0000

R-squared	0.292313	Mean dependent var	0.000228
Adjusted R-squared	0.292313	S.D. dependent var	0.053008
S.E. of regression	0.044592	Akaike info criterion	-3.377693
Sum squared resid	0.409625	Schwarz criterion	-3.361593
Log likelihood	350.5912	Hannan-Quinn criter.	-3.371182
Durbin-Watson stat	1.948189		

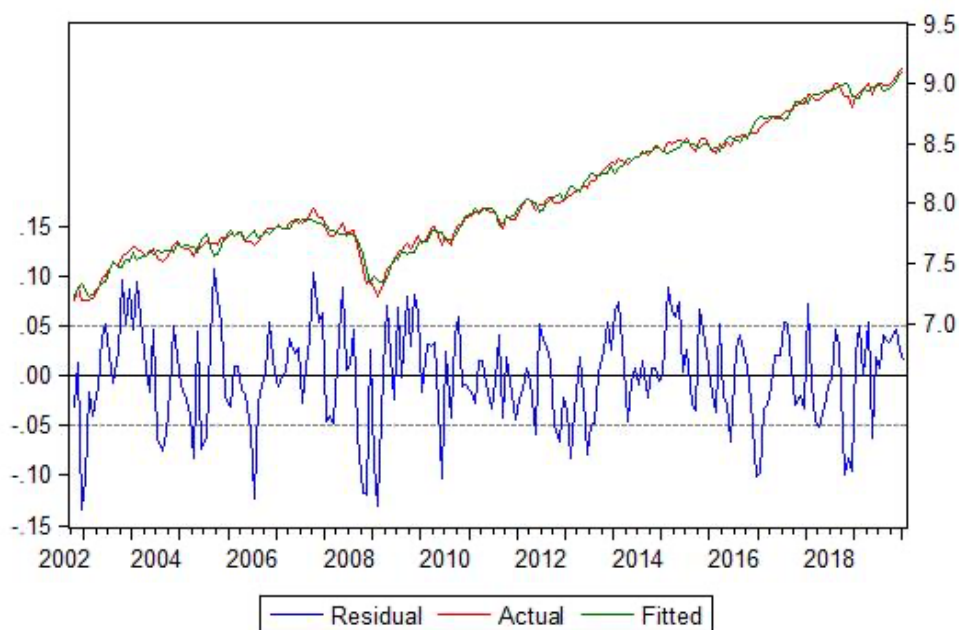
Il test conduce al rifiuto dell'ipotesi nulla H_0 di radice unitaria, confermando la stazionarietà della serie ECM. Oltre alla probabilità associata al test, è stato analizzato il valore in corrispondenza alla t-Statistic, in questo caso pari a -9.224679. Soffermarsi esclusivamente sul valore probabilistico per un livello di significatività del 5% (o anche inferiore), potrebbe risultare alquanto riduttivo essendo che il test non include il numero dei regressori che caratterizzano l'equazione statica, ma la singola serie sul quale viene applicato. Nel caso in oggetto abbiamo otto variabili esplicative, oltre alla costante e il valore della statistica porta abbondantemente al rifiuto dell'ipotesi nulla di radice unitaria, secondo quanto emerge dal manuale "Econometria delle serie storiche"⁸. La serie ECM generata sarà dunque $I(0)$: ciò sancisce la cointegrazione tra le variabili incluse nella

⁸ Hamilton J. D., (1995), *Econometria delle serie storiche*, Monduzzi Editore.

statica, rigettando il rischio di incorrere in una regressione spuria e suffragando la validità dell'equazione di lungo periodo proposta.

Ponendo l'attenzione su quest'ultima, ed in particolare su quanto emerge dalla figura 2.2, l'equazione statica finale registra dei valori di R- squared e Adjusted R- squared vicini all'unità, rispettivamente di 0.99081 e 0.99044. Ciò sta a significare che la regressione è in grado di intercettare oltre il 99% della variabilità della serie endogena "N_L". Anche i criteri informativi di Akaike e Schwarz sembrano migliorare a seguito della rimozione di ciascun regressore non significativo. La validità dell'equazione di lungo periodo nell'intercettare la relazione tra variabili esplicative e dipendente può essere apprezzata nel grafico che segue:

Figura 2.6- Confronto tra curva effettiva e stimata dall'equazione statica finale



La figura mostra l'andamento per l'intervallo tra 2002/10/01 e 2020/01/01 della curva reale realizzata dalla variabile dipendente (in rosso "Actual"), di quella stimata attraverso la regressione che ha origine dall'equazione statica (in verde "Fitted") e dei residui (in blu "Residual"). Nel complesso la curva stimata coglie le fasi di maggiore crescita e decrescita della curva effettiva del NASDAQ Composite, pur mostrando alcuni limiti nell'intercettare le variazioni meno significative.

Si può notare come sin dai primi mesi del periodo campionario la regressione sia in grado di cogliere la fase di espansione realizzata dalla variabile dipendente a seguito della bolla delle Dot-com. Sino alla grande recessione la serie “N_L” è soggetta a variazioni minime, così come anche la curva stimata che, tuttavia, non riesce ad intercettare la totalità delle variazioni della variabile dipendente. L’equazione di lungo periodo riesce a spiegare la contrazione dovuta alla crisi del 2007/2008, pur non sovrapponendosi perfettamente al punto di minimo della variabile dipendente. Anche per il periodo di crescita successivo alla recessione la curva stimata ottiene un andamento più che soddisfacente, così come in corrispondenza alla contrazione registrata nel 2011/2012 e nelle variazioni successive. I residui rappresentano il differenziale tra le due curve e mostrano un livello di dispersione maggiore nei primi mesi del 2003 e, altresì, in prossimità alla crisi della grande recessione.

Con riferimento alle variabili esogene incluse nell’equazione di lungo periodo della Figura 2.2, si possono commentare i coefficienti delle variabili esogene. La lettura di questi ultimi non va interpretata guardando alla dimensione del valore associato, ma bensì al segno che li caratterizza. Il valore in modulo è infatti sensibile all’unità di misura tra le variabili ed indica una propensione marginale.

Il segno negativo sarà proprio di una variabile esogena che ha un impatto negativo sulla serie dipendente, viceversa nel caso di un segno positivo. Si noti come i segni delle serie esplicative sono nel complesso coerenti con quanto ci si può attendere in ottica macroeconomica e quanto anticipato in sede di analisi descrittiva delle variabili.

Entrando nel merito delle più rappresentative, si osserva che l’indice di sentiment dei consumatori (“CC”) ha un impatto positivo sul NASDAQ: un incremento della variabile esogena porta ad un aumento della variabile dipendente. Tale indicatore in corrispondenza di un aumento prospetta una riduzione del risparmio privato e, più in generale, maggiore inclinazione alla spesa da parte dei cittadini, un quadro in linea con l’impatto positivo sull’indice che emerge dalla regressione.

Un ragionamento pressochè analogo può essere esteso alla variabile “PCE”. L’aumento della spesa individuale per il consumo è indicativo di una fase espansiva sulla suddetta componente del PIL e si traduce in un minor risparmio da parte dei cittadini. Anche questa variabile è in linea con quanto atteso, segnalando una relazione positiva rispetto alla serie del NASDAQ Composite.

Un impatto negativo viene apportato dai rendimenti a breve termine sui titoli di debito statunitensi nel mercato secondario (“TB”). Un rialzo dei rendimenti è indicativo di un maggiore rischio per gli investitori, da cui deriva maggiore incertezza nei mercati finanziari. In coerenza con quanto atteso, un incremento dei rendimenti a breve termine ha un impatto negativo sulla variabile dipendente.

Anche la variabile “IPT” relativa alla produzione industriale non può che avere un impatto positivo sulla variabile endogena. Da un aumento della produzione industriale di un settore collegato a quello tecnologico (che, come anticipato, è il settore trainante del NASDAQ Composite) non può che derivare un conseguente effetto positivo dell’indice oggetto di studio. Ragionamento pressochè analogo per le variabili “PPIT” e “RET”.

In quanto ai rendimenti di stato delle obbligazioni a lungo termine, denominati con “BOND”, il coefficiente generato dall’equazione statica può essere soggetto ad alcune ambiguità. Se da una parte è vero che ad un aumento dei rendimenti obbligazionari corrisponde un aumento del rischio, dall’altra va sottolineato che vi sono alcuni fattori, primo fra tutti il premio per il rischio, che insistono maggiormente nei titoli a lungo termine. Inoltre, la relazione che scaturisce dalla statica va interpretata in un contesto di lungo periodo.

Una visione più completa relativamente all’impatto delle variabili esplicative, viene fornita dai coefficienti standardizzati generati a livello dell’equazione statica:

Figura 2.7- Coefficienti standardizzati dell’equazione statica finale

Scaled Coefficients
Date: 12/11/22 Time: 19:18
Sample: 2002M10 2020M01
Included observations: 208

Variable	Coefficient	Standardized Coefficient	Elasticity at Means
C	45.56460	NA	5.646807
BOND	0.070400	0.140607	0.027245
CC	0.006041	0.139836	0.063000
TB	-0.041635	-0.122273	-0.006624
RET	0.023354	0.276611	0.277444
PCE	0.000260	0.987952	0.350163
EM_L	-3.859475	-0.310671	-5.685330
IPT	0.006443	0.139089	0.080817
PPIT	0.018819	0.104413	0.246478

La lettura di questi ultimi risulta rilevante alla luce degli effetti sulla variabile dipendente: più sono vicini a 0, meno la variabile esogena impatta sull'endogena, più sono prossimi a 1 maggiore sarà l'impatto sulla medesima⁹. Va evidenziato che i valori in corrispondenza a questi coefficienti non devono essere letti percentualmente, ma in ottica unitaria¹⁰. Si noti come la variabile che esercita un'influenza maggiore sull'indice NASDAQ risulta essere la spesa individuale per il consumo "PCE". Quest'ultima riporta valori positivi prossimi all'unità, pari a 0,987952. Valori invece prossimi a 0 e, di conseguenza, meno impattanti, sono propri dei regressori "PPIT" e "TB".

L'elasticità riporta la variazione percentuale in media dell'endogena alla variazione dell'1% dell'esogena. A titolo esemplificativo, ipotizzando un aumento dell'1% della variabile "BOND", la variabile endogena aumenta del 2,7%, un incremento dell'1% di "IPT" fa aumentare dell'8% il NASDAQ Composite e via dicendo. Il regressore più significativo sotto questo punto di vista risulta essere "EM_L", ovvero il tasso di occupazione.

Di seguito si propone una stima dell'equazione statica attraverso un metodo differente, che mira ad intercettare un modello per l'eteroschedasticità condizionale della regressione. Nella fattispecie sulla stessa equazione di lungo periodo si è effettuata la stima della regressione attraverso il metodo ML ARCH.

In una prima fase si sono verificati i modelli più comuni, tra i quali il GARCH e l'ARCH. In prima battuta questi ultimi hanno ottenuto un riscontro positivo, soddisfacendo le condizioni sui coefficienti e sulla significatività associata¹¹. La stima più adeguata sembra essere quella di un ARCH(1):

⁹ Tali coefficienti non sono necessariamente compresi tra -1 e 1, possono anche superare questi valori e avere impatti ancor più estremi

¹⁰ Il ragionamento sottostante a questa interpretazione è legato al concetto di impulso unitario applicato ad uno "steady state" (stato di quiete).

¹¹ La sommatoria dei coefficienti deve essere inferiore all'unità e, al contempo, questi ultimi devono risultare significativi.

Figura 2.8- Stima modello ARCH(1) per l'eteroschedasticità condizionale

Dependent Variable: N_L
 Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)
 Date: 02/11/23 Time: 22:02
 Sample: 2002M10 2020M01
 Included observations: 208
 Convergence achieved after 104 iterations
 Coefficient covariance computed using outer product of gradients
 Presample variance: backcast (parameter = 0.7)
 GARCH = C(10) + C(11)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	31.41004	9.090949	3.455089	0.0006
BOND	0.069910	0.009051	7.723629	0.0000
CC	0.005370	0.000565	9.499924	0.0000
TB	-0.043653	0.006270	-6.962517	0.0000
RET	0.021169	0.003611	5.862573	0.0000
PCE	0.000247	2.63E-05	9.401429	0.0000
EM_L	-2.590019	0.784557	-3.301248	0.0010
IPT	0.005443	0.002127	2.558808	0.0105
PPIT	0.014810	0.004784	3.095522	0.0020

Variance Equation				
C	0.001475	0.000240	6.155830	0.0000
RESID(-1)^2	0.441179	0.162285	2.718550	0.0066

R-squared	0.990008	Mean dependent var	8.069092
Adjusted R-squared	0.989606	S.D. dependent var	0.511076
S.E. of regression	0.052105	Akaike info criterion	-3.147774
Sum squared resid	0.540274	Schwarz criterion	-2.971270
Log likelihood	338.3685	Hannan-Quinn criter.	-3.076405
Durbin-Watson stat	1.004505		

Ovviamente si possono stimare diversi modelli alternativi e più complessi, purché vengano soddisfatte le condizioni accennate in precedenza sui coefficienti.

A titolo esemplificativo si riporta un modello per l'eteroschedasticità condizionale di tipo IGARCH(2,1), già inquadrato da un punto di vista teorico nel primo capitolo dell'elaborato:

Figura 2.9- Stima modello IGARCH(2,1) per l'eteroschedasticità condizionale

Dependent Variable: N_L
 Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)
 Date: 12/15/22 Time: 08:16
 Sample: 2002M10 2020M01
 Included observations: 208
 Convergence not achieved after 500 iterations
 Coefficient covariance computed using outer product of gradients
 Presample variance: backcast (parameter = 0.7)
 GARCH = C(10)*RESID(-1)^2 + C(11)*RESID(-2)^2 + (1 - C(10) - C(11))
 *GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	45.62377	8.015068	5.692249	0.0000
BOND	0.068593	0.009695	7.074690	0.0000
CC	0.005858	0.000610	9.610461	0.0000
TB	-0.042111	0.006693	-6.291538	0.0000
RET	0.019285	0.004112	4.689643	0.0000
PCE	0.000275	2.47E-05	11.12856	0.0000
EM_L	-3.843818	0.691705	-5.557016	0.0000
IPT	0.007823	0.002186	3.578707	0.0003
PPIT	0.017505	0.005296	3.305601	0.0009

Variance Equation				
	Coefficient	Std. Error	z-Statistic	Prob.
RESID(-1)^2	0.428366	0.120899	3.543177	0.0004
RESID(-2)^2	-0.411465	0.130217	-3.159837	0.0016
GARCH(-1)	0.983098	0.013997	70.23645	0.0000

R-squared	0.990580	Mean dependent var	8.069092
Adjusted R-squared	0.990201	S.D. dependent var	0.511076
S.E. of regression	0.050591	Akaike info criterion	-3.139351
Sum squared resid	0.509330	Schwarz criterion	-2.962847
Log likelihood	337.4925	Hannan-Quinn criter.	-3.067982
Durbin-Watson stat	1.095789		

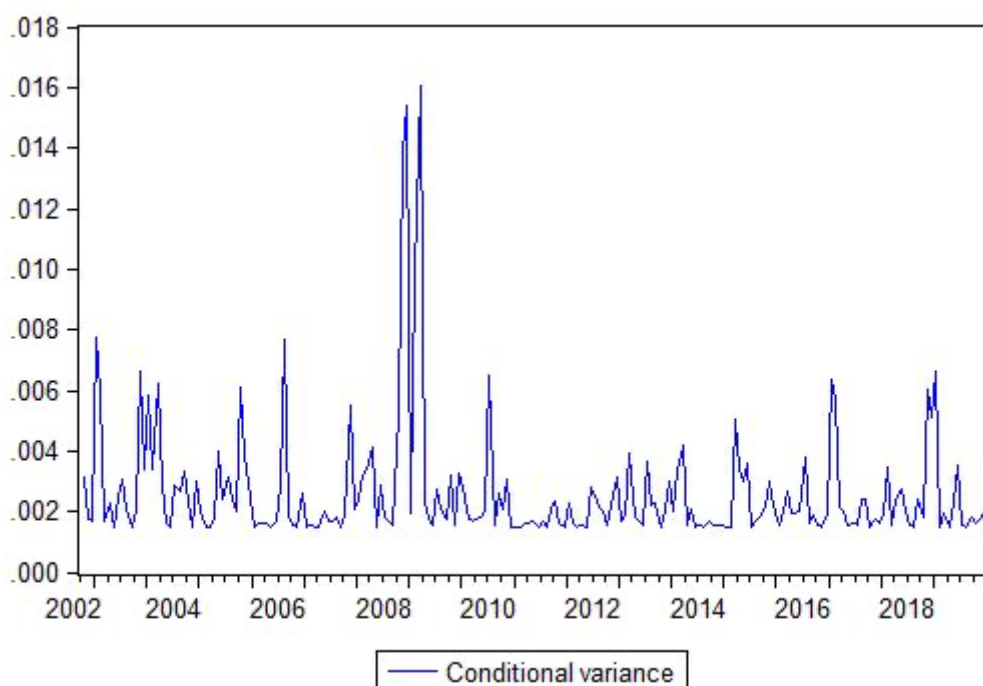
Commentando brevemente il risultato, si può apprezzare anzitutto che la somma dei coefficienti del modello IGARCH(2,1) stimato soddisfa pienamente la condizione introdotta precedentemente, risultando pari all'unità. Un altro fattore a conferma della validità del modello stimato per la varianza risiede nella probabilità associata ai coefficienti, che risulta in tutti i casi significativa. Concentrando l'attenzione sui criteri informativi del modello, si può osservare che Akaike e Schwarz non riportano variazioni sostanziali rispetto alla stima della regressione attraverso il metodo OLS, così come l'Adjusted R-squared. I coefficienti, la significatività e il segno dei regressori rimangono coerenti con quanto visto nella prima equazione statica. Nello specifico tutte le variabili esplicative risultano significative e la natura degli impatti sul NASDAQ Composite non

segnala variazioni. I regressori “EM_L” e “TB” mantengono un segno negativo, viceversa le altre variabili.

Le medesime osservazioni possono essere estese a quanto individuato dal modello ARCH(1), in corrispondenza del quale i criteri informativi AIC e SIC riportano valori leggermente migliori.

Guardando a quanto ottenuto dal modello ARCH in termini di varianza:

Figura 2.10- Varianza modello ARCH(1) stimato



Dal grafico si può apprezzare che quanto ottenuto è coerente con l’evoluzione storica della variabile endogena. In primo luogo, si segnala un picco iniziale nella varianza della serie negli ultimi mesi del 2002, dovuto ad un effetto residuo della bolla delle Dot-com. Il picco più consistente si verifica tra la fine del 2008 e i primi mesi del 2009, in linea con l’impatto sul presente indice a seguito della grande recessione e alla forte volatilità dei mercati. Come riportato nel primo capitolo, infatti, il NASDAQ ha avuto una variazione ed una volatilità maggiore l’anno successivo allo scoppio della recessione statunitense, per poi avviare un trend crescente e pressochè lineare sino al 2020. La varianza non

riporta ulteriori segnali evidenti fino ad oggi, mostrando un leggero rialzo soltanto in corrispondenza dei primi mesi del 2016 e del 2018.

2.2 L'equazione di breve periodo

Appurata la stazionarietà della serie dei residui generati dall'equazione di lungo periodo si può procedere con la stima dell'equazione di breve periodo in forma ECM. Come si può desumere dalla denominazione attribuita all'equazione in questione, in questo caso si andrà ad analizzare la relazione di breve periodo tra le variabili esplicative e l'endogena. Ciò viene realizzato tenendo in considerazione alcuni punti cardine:

- In primis l'equazione dinamica verrà stimata sullo stesso periodo della statica, ovverosia nell'intervallo tra il 2002/10/01 e il 2020/01/01.
- Si terranno 12 ritardi temporali per ogni variabile, essendo che la frequenza delle serie storiche è su base mensile.
- Differentemente dalla regressione statica, le variabili inserite devono essere stazionarie. Ciò comporta l'applicazione di una differenza prima per i regressori con una radice unitaria, mentre le variabili stazionarie verranno inserite senza la differenza.
- Non verrà applicata la correzione HAC poiché per considerare valida l'equazione dinamica è essenziale che i residui di quest'ultima siano white noise. Il fatto che un modello per essere considerato valido debba avere dei disturbi con assenza di eteroschedasticità e autocorrelazione rende superflua l'applicazione di metodi di correzione come la procedura HAC.
- Verranno inserite non soltanto le variabili stazionarie, ma anche le variabili rimosse dall'equazione statica generale in quanto non significative o non particolarmente migliorative dei criteri informativi di Akaike (AIC) e di Schwarz (BIC) o dell'Adjusted R-squared.
- Ai fini di un'equazione dinamica corretta, sarà ideale, anche se non strettamente necessario, ottenere una distribuzione della componente dei residui normale. Questi ultimi in corrispondenza di valori di asimmetria prossimi a 0 e di una curtosi vicina a 3 saranno la conferma di una stima efficiente.
- La componente ECM viene inserita ritardata di un periodo (quindi verrà inserita la serie ECM(-1)).

- I residui nelle varie fasi di stima dovranno essere conformi ad una realizzazione di un processo white noise.
- Tale componente dovrà risultare significativa e compresa nell'intervallo (-2;0) in sede di equazione dinamica generale e, soprattutto, nella dinamica finale.

Si riporta di seguito la rappresentazione derivante dalla stima dell'equazione dinamica generale:

Figura 2.11- Stima dell'equazione dinamica generale

Dependent Variable: D(N_L)
Method: Least Squares
Date: 02/12/23 Time: 19:24
Sample (adjusted): 2003M11 2020M01
Included observations: 195 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.004569	0.053203	0.085873	0.9323
D(CPI(-1))	-0.018819	0.035029	-0.537250	0.5960
D(CPI(-2))	-0.062035	0.035659	-1.739645	0.0947
D(CPI(-3))	0.020064	0.036309	0.552582	0.5857
D(CPI(-4))	-0.030859	0.036785	-0.838912	0.4098
D(CPI(-5))	-0.047593	0.031907	-1.491602	0.1488
D(CPI(-6))	0.037355	0.031321	1.192666	0.2447
D(CPI(-7))	0.000981	0.036337	0.027004	0.9787
D(CPI(-8))	-0.003033	0.038059	-0.079682	0.9372
D(CPI(-9))	0.052378	0.041675	1.256829	0.2209
D(CPI(-10))	-0.000602	0.040262	-0.014951	0.9882
D(CPI(-11))	-0.075191	0.041891	-1.794923	0.0853
D(CPI(-12))	0.028068	0.040648	0.690506	0.4965
D(CPI)	0.081573	0.037691	2.164244	0.0406
D(PCE)	1.49E-05	0.000373	0.040032	0.9684
D(PCE(-1))	0.000165	0.000334	0.493263	0.6263
D(PCE(-2))	0.000614	0.000314	1.957798	0.0620
D(PCE(-3))	6.87E-05	0.000334	0.205576	0.8389
D(PCE(-4))	-0.000122	0.000316	-0.388216	0.7013
D(PCE(-5))	0.000223	0.000307	0.727315	0.4741
D(PCE(-6))	-0.000103	0.000343	-0.301299	0.7658
D(PCE(-7))	0.000423	0.000326	1.296438	0.2072
D(PCE(-8))	5.12E-05	0.000392	0.130804	0.8970
D(PCE(-9))	-0.000450	0.000341	-1.320504	0.1991
D(PCE(-10))	-0.000260	0.000354	-0.734432	0.4698
D(PCE(-11))	0.000182	0.000361	0.504686	0.6184
D(PCE(-12))	0.000190	0.000392	0.484653	0.6323
D(EM_L)	2.459225	4.537570	0.541970	0.5928
D(EM_L(-1))	11.33365	4.973447	2.278832	0.0319
D(EM_L(-2))	4.406466	4.471210	0.985520	0.3342
D(EM_L(-3))	-1.938685	4.586412	-0.422702	0.6763
D(EM_L(-4))	-5.475362	5.039016	-1.086594	0.2880
D(EM_L(-5))	-3.844718	4.962672	-0.774727	0.4461
D(EM_L(-6))	-7.538451	4.981845	-1.513184	0.1433

D(EM_L(-7))	-11.09614	5.351284	-2.073548	0.0490
D(EM_L(-8))	-7.389758	6.358530	-1.162180	0.2566
D(EM_L(-9))	-0.108401	4.516815	-0.023999	0.9811
D(EM_L(-10))	2.765084	4.465009	0.619279	0.5416
D(EM_L(-11))	-3.102500	5.002868	-0.620144	0.5410
D(EM_L(-12))	0.885807	4.383015	0.202100	0.8415
D(TB)	0.047889	0.053844	0.889408	0.3826
D(TB(-1))	-0.135481	0.071107	-1.905306	0.0688
D(TB(-2))	-0.074221	0.070753	-1.049014	0.3046
D(TB(-3))	0.053308	0.070862	0.752280	0.4592
D(TB(-4))	0.137309	0.081101	1.693060	0.1034
D(TB(-5))	-0.070761	0.085943	-0.823341	0.4184
D(TB(-6))	-0.033080	0.082065	-0.403100	0.6904
D(TB(-7))	-0.086910	0.085639	-1.014837	0.3203
D(TB(-8))	-0.041025	0.075192	-0.545608	0.5904
D(TB(-9))	-0.068276	0.063323	-1.078227	0.2917
D(TB(-10))	0.086880	0.070104	1.239305	0.2272
D(TB(-11))	-0.082715	0.067375	-1.227687	0.2315
D(TB(-12))	0.107649	0.059839	1.798987	0.0846
D(CC)	-0.000234	0.001880	-0.124621	0.9019
D(CC(-1))	-0.002969	0.001792	-1.657229	0.1105
D(CC(-2))	0.001401	0.001799	0.778881	0.4437
D(CC(-3))	0.000110	0.001829	0.060000	0.9527
D(CC(-4))	-0.000322	0.001599	-0.201239	0.8422
D(CC(-5))	1.35E-05	0.001805	0.007499	0.9941
D(CC(-6))	0.004614	0.001634	2.824273	0.0094
D(CC(-7))	0.000962	0.001839	0.523093	0.6057
D(CC(-8))	-0.000242	0.001778	-0.136031	0.8929
D(CC(-9))	-0.001814	0.001852	-0.979119	0.3373
D(CC(-10))	-0.001080	0.001814	-0.595714	0.5569
D(CC(-11))	-0.002460	0.001579	-1.557727	0.1324
D(CC(-12))	0.001005	0.001639	0.613156	0.5455
D(RET)	0.009819	0.013655	0.719090	0.4790
D(RET(-1))	-0.003599	0.016229	-0.221752	0.8264
D(RET(-2))	-0.040987	0.015449	-2.653116	0.0139
D(RET(-3))	-0.025619	0.019910	-1.286704	0.2105
D(RET(-4))	-0.011780	0.019933	-0.590966	0.5601
D(RET(-5))	-0.012961	0.019191	-0.675396	0.5059
D(RET(-6))	-0.016277	0.020876	-0.779718	0.4432
D(RET(-7))	-0.024732	0.019340	-1.278789	0.2132
D(RET(-8))	0.000647	0.020485	0.031567	0.9751
D(RET(-9))	0.026655	0.019836	1.343728	0.1916
D(RET(-10))	0.027505	0.017691	1.554791	0.1331
D(RET(-11))	0.011846	0.016640	0.711868	0.4834
D(RET(-12))	-0.006222	0.015755	-0.394896	0.6964
D(PPIT)	0.009700	0.021297	0.455472	0.6529
D(PPIT(-1))	0.038175	0.024206	1.577078	0.1279
D(PPIT(-2))	-0.019193	0.021456	-0.894503	0.3799
D(PPIT(-3))	0.002885	0.022266	0.129572	0.8980
D(PPIT(-4))	0.012339	0.020755	0.594528	0.5577
D(PPIT(-5))	-0.008883	0.019872	-0.447005	0.6589
D(PPIT(-6))	0.001512	0.021446	0.070521	0.9444
D(PPIT(-7))	0.030363	0.020401	1.488302	0.1497
D(PPIT(-8))	0.006367	0.022254	0.286126	0.7772
D(PPIT(-9))	-0.020914	0.026238	-0.797101	0.4332
D(PPIT(-10))	0.023101	0.021277	1.085747	0.2884
D(PPIT(-11))	-0.007285	0.025119	-0.290022	0.7743
D(PPIT(-12))	-0.008766	0.026851	-0.326465	0.7469

UNC_L	-0.027871	0.045711	-0.609724	0.5478
UNC_L(-1)	-0.030127	0.047467	-0.634697	0.5316
UNC_L(-2)	0.031888	0.047982	0.664570	0.5127
UNC_L(-3)	-0.041574	0.047459	-0.876004	0.3897
UNC_L(-4)	-0.035653	0.052340	-0.681182	0.5023
UNC_L(-5)	-0.014709	0.064417	-0.228339	0.8213
UNC_L(-6)	0.086988	0.057491	1.513079	0.1433
UNC_L(-7)	-0.016498	0.063803	-0.258576	0.7982
UNC_L(-8)	-0.084041	0.066232	-1.268875	0.2167
UNC_L(-9)	0.042620	0.048158	0.884997	0.3849
UNC_L(-10)	0.040627	0.055327	0.734307	0.4699
UNC_L(-11)	0.008247	0.054277	0.151936	0.8805
UNC_L(-12)	0.040107	0.044028	0.910952	0.3714
D(BOND)	0.029604	0.034549	0.856871	0.4000
D(BOND(-1))	-0.037348	0.040818	-0.914986	0.3693
D(BOND(-2))	0.000487	0.036859	0.013209	0.9896
D(BOND(-3))	-0.044815	0.039216	-1.142757	0.2644
D(BOND(-4))	-0.055659	0.036802	-1.512382	0.1435
D(BOND(-5))	-0.015173	0.042548	-0.356607	0.7245
D(BOND(-6))	0.053601	0.034160	1.569117	0.1297
D(BOND(-7))	-0.035873	0.037402	-0.959123	0.3471
D(BOND(-8))	-0.017434	0.036789	-0.473902	0.6399
D(BOND(-9))	-0.036640	0.035297	-1.038050	0.3096
D(BOND(-10))	-0.011413	0.037536	-0.304061	0.7637
D(BOND(-11))	0.005927	0.028658	0.206810	0.8379
D(BOND(-12))	0.025652	0.031557	0.812868	0.4243
D(USD)	0.191652	0.275176	0.696469	0.4928
D(USD(-1))	-0.081519	0.292711	-0.278497	0.7830
D(USD(-2))	0.050207	0.267785	0.187491	0.8529
D(USD(-3))	-0.094284	0.270490	-0.348567	0.7305
D(USD(-4))	-0.170982	0.240766	-0.710160	0.4844
D(USD(-5))	0.491406	0.236407	2.078646	0.0485
D(USD(-6))	-0.111365	0.277365	-0.401511	0.6916
D(USD(-7))	-0.036408	0.287920	-0.126453	0.9004
D(USD(-8))	-0.190645	0.269059	-0.708562	0.4854
D(USD(-9))	-0.296281	0.235443	-1.258401	0.2204
D(USD(-10))	-0.190033	0.231428	-0.821133	0.4197
D(USD(-11))	-0.212088	0.233014	-0.910194	0.3718
D(USD(-12))	0.176184	0.256563	0.686707	0.4988
D(LF)	-0.070673	0.072713	-0.971952	0.3408
D(LF(-1))	-0.123933	0.071961	-1.722228	0.0979
D(LF(-2))	-0.171116	0.070918	-2.412865	0.0238
D(LF(-3))	-0.011286	0.076833	-0.146891	0.8844
D(LF(-4))	-0.036231	0.072776	-0.497839	0.6231
D(LF(-5))	0.048697	0.078970	0.616658	0.5433
D(LF(-6))	0.129164	0.074829	1.726131	0.0972
D(LF(-7))	0.180278	0.082484	2.185620	0.0388
D(LF(-8))	0.191492	0.093159	2.055535	0.0509
D(LF(-9))	0.147812	0.081067	1.823338	0.0807
D(LF(-10))	-0.000689	0.077168	-0.008935	0.9929
D(LF(-11))	-0.002924	0.062907	-0.046473	0.9633
D(LF(-12))	-0.015224	0.067576	-0.225284	0.8237
D(CHUS)	-0.146825	0.140607	-1.044220	0.3068
D(CHUS(-1))	0.167972	0.188944	0.889003	0.3828
D(CHUS(-2))	0.127234	0.187831	0.677383	0.5046
D(CHUS(-3))	0.166403	0.176686	0.941800	0.3557
D(CHUS(-4))	-0.125496	0.158491	-0.791819	0.4362
D(CHUS(-5))	-0.084112	0.156096	-0.538852	0.5950
D(CHUS(-6))	-0.031508	0.177271	-0.177738	0.8604

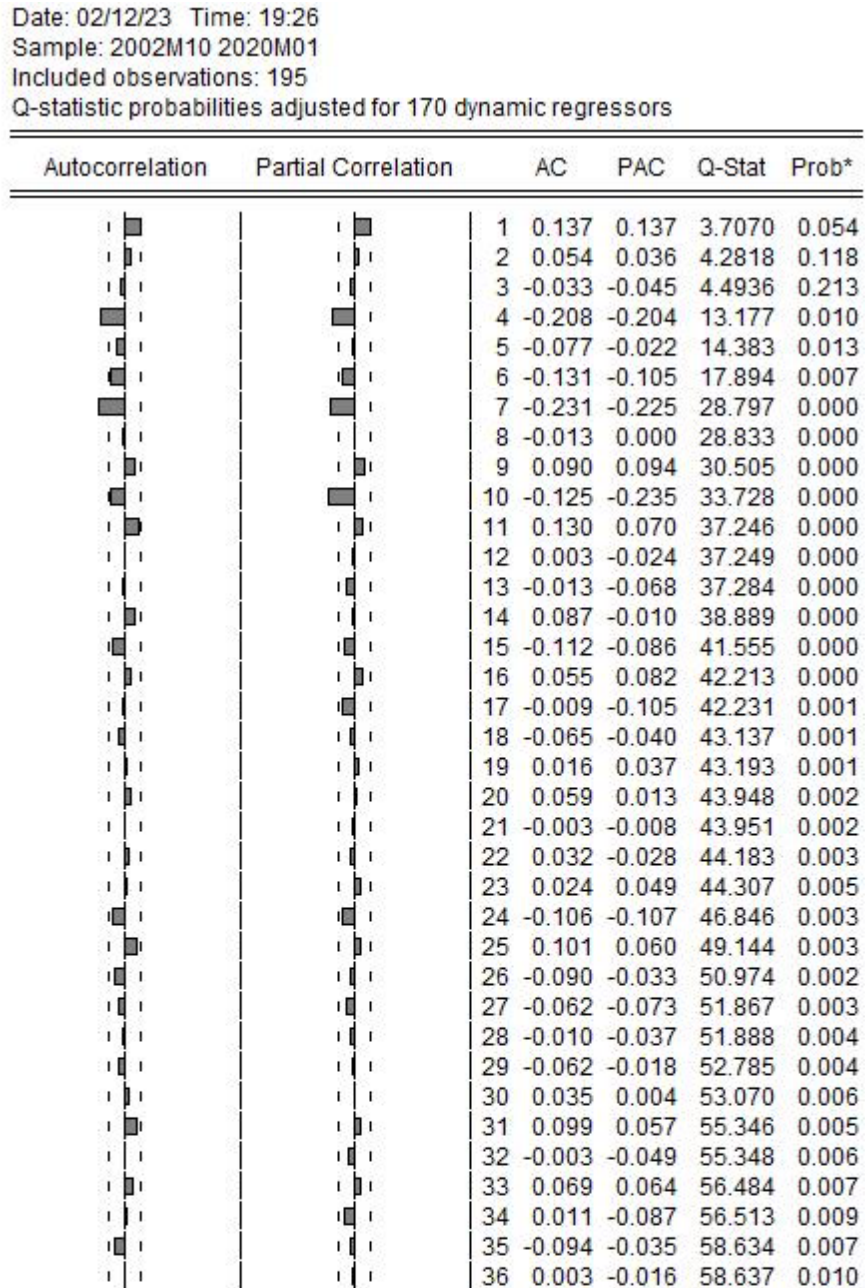
D(CHUS(-7))	0.194177	0.191427	1.014364	0.3205
D(CHUS(-8))	0.104572	0.209884	0.498239	0.6229
D(CHUS(-9))	0.384337	0.215391	1.784365	0.0870
D(CHUS(-10))	-0.189421	0.202414	-0.935808	0.3587
D(CHUS(-11))	0.058961	0.227772	0.258858	0.7980
D(CHUS(-12))	-0.427830	0.215849	-1.982076	0.0590
D(IPT)	0.009744	0.008806	1.106474	0.2795
D(IPT(-1))	0.017800	0.008625	2.063754	0.0500
D(IPT(-2))	0.014586	0.008405	1.735379	0.0955
D(IPT(-3))	-0.007188	0.007994	-0.899059	0.3776
D(IPT(-4))	-0.005140	0.007909	-0.649971	0.5219
D(IPT(-5))	0.004449	0.008695	0.511630	0.6136
D(IPT(-6))	0.013586	0.008416	1.614250	0.1195
D(IPT(-7))	0.002683	0.008735	0.307179	0.7614
D(IPT(-8))	0.003128	0.008995	0.347783	0.7310
D(IPT(-9))	-0.001653	0.010045	-0.164575	0.8707
D(IPT(-10))	0.006610	0.008352	0.791403	0.4365
D(IPT(-11))	0.000978	0.009198	0.106358	0.9162
D(IPT(-12))	-0.000958	0.008584	-0.111659	0.9120
ECM_ALT(-1)	-0.640009	0.186261	-3.436094	0.0022
<hr/>				
R-squared	0.939327	Mean dependent var	0.007975	
Adjusted R-squared	0.509561	S.D. dependent var	0.047302	
S.E. of regression	0.033126	Akaike info criterion	-4.318091	
Sum squared resid	0.026336	Schwarz criterion	-1.447923	
Log likelihood	592.0139	Hannan-Quinn criter.	-3.155994	
F-statistic	2.185668	Durbin-Watson stat	1.722346	
Prob(F-statistic)	0.013436			

Nella rappresentazione di cui sopra si è deciso di non includere i ritardi dell'endogena stessa. L'immissione di questi ritardi non sarebbe errata da un punto di vista computazionale, ma nella pratica econometrica si tendono a privilegiare quei modelli che spiegano la variabile dipendente senza ricorrere ai ritardi della stessa. L'approccio inverso rispetto a quello utilizzato in questo studio, specie nel qual caso tali ritardi fossero stati significativi, avrebbe potuto ottenere delle previsioni migliori, ma nel contempo avrebbe portato ad una stima ad hoc meno credibile da un punto di vista teorico. Si sarebbe infatti attribuita un'interpretazione auto esplicativa alla variabile endogena

Analogamente a quanto visto nel primo passo della procedura di Engle – Granger, alcune variabili verranno rimosse dalla presente equazione. La scrematura prevede dei passaggi in parte differenti rispetto a quelli effettuati per arrivare all'equazione di lungo periodo finale. Ancor prima di passare alla rimozione delle variabili non significative, si effettuano i test volti ad indagare l'eventuale presenza di eteroschedasticità e autocorrelazione nella componente dei residui, oltre a verificare

che i coefficienti associati al disturbo siano compresi tra -2 e 0. Di seguito si riportano i correlogrammi dei residui:

Figura 2.12- Correlogramma dei residui dell'equazione dinamica generale



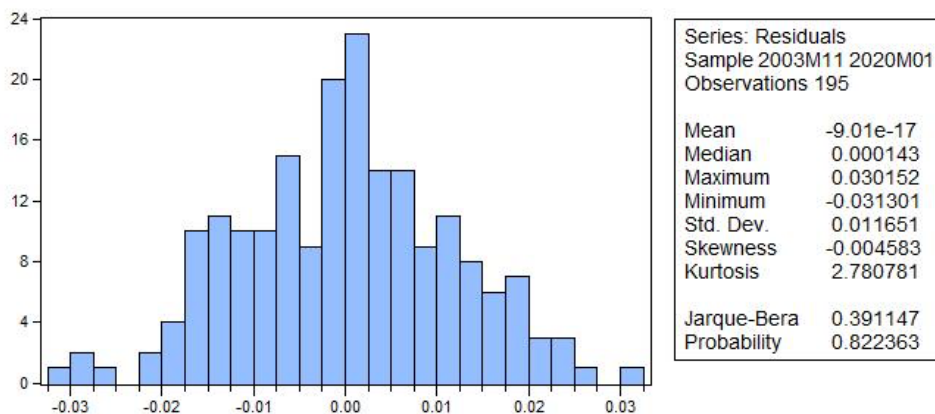
Dalle rappresentazioni sembra emergere l'assenza di autocorrelazione. Per accertare l'assenza di eteroschedasticità anche nell'equazione dinamica generale, viene effettuato il test ARCH.

Figura 2.13- Test ARCH per l'eteroschedasticità dei residui della dinamica generale

Heteroskedasticity Test: ARCH			
F-statistic	0.956477	Prob. F(5,184)	0.4459
Obs*R-squared	4.813230	Prob. Chi-Square(5)	0.4391

Quest'ultimo prevede come ipotesi nulla H_0 l'assenza di eteroschedasticità. In questo caso la probabilità associata al test è di 0,4459 e conduce all'accettazione dell'ipotesi nulla a favore di residui omoschedastici. I risultati ottenuti dall'analisi effettuata sulla componente di disturbo sono indicativi di stime non distorte anche nell'equazione generale stimata. In questa fase si può analizzare anche la distribuzione dei residui per valutare l'efficienza della stima effettuata:

Figura 2.14- Distribuzione e valori sulla normalità dei residui della dinamica generale



Si precisa che, anche qualora alcuni dei risultati ottenuti venissero disattesi, la situazione non sarebbe particolarmente allarmante: nella dinamica generale viene inserito un elevato numero di variabili e di ritardi del tutto irrilevanti. Ciò comporta stime che di norma non sono affatto efficienti.

Sarà dunque essenziale ripetere i medesimi test nella fase finale della stima, a seguito della rimozione dei regressori non significativi.

Come anticipato poc'anzi, la procedura di esclusione delle variabili dall'equazione di breve periodo non è basata sulla stessa logica che aveva portato alla statica finale. Anzitutto i test relativi alla componente dei residui verranno ripetuti man mano che si andranno a rimuovere le variabili. Queste ultime verranno rimosse partendo dai ritardi più lontani e dalle variabili contemporanee che non risultano significative. Nel primo caso, infatti, includere regressori più recenti sarebbe in linea con la finalità di un'equazione di breve periodo, ovvero tipizzare la variabile dipendente attraverso i ritardi meno lontani nel tempo. Le variabili esogene contemporanee invece porterebbero a delle stime in fase di previsione meno credibili (fermo restando che ciò non andrebbe ad inficiare la validità del modello).

Si andranno quindi a rimuovere quei ritardi delle variabili esogene partendo da quelle la cui probabilità associata è maggiore e seguendo i criteri di cui sopra. Nelle fasi finali, in caso di modelli alternativi o di ambiguità nella significatività dei regressori, si andranno a privilegiare le specificazioni ricorrendo ai criteri informativi introdotti nel primo passo della procedura di Engle-Granger, con riguardo al criterio di Akaike, a quello di Schwarz e all'Adjusted R Squared.

L'equazione di breve periodo ottenuta sarà:

Figura 2.15- Stima dell'equazione dinamica finale

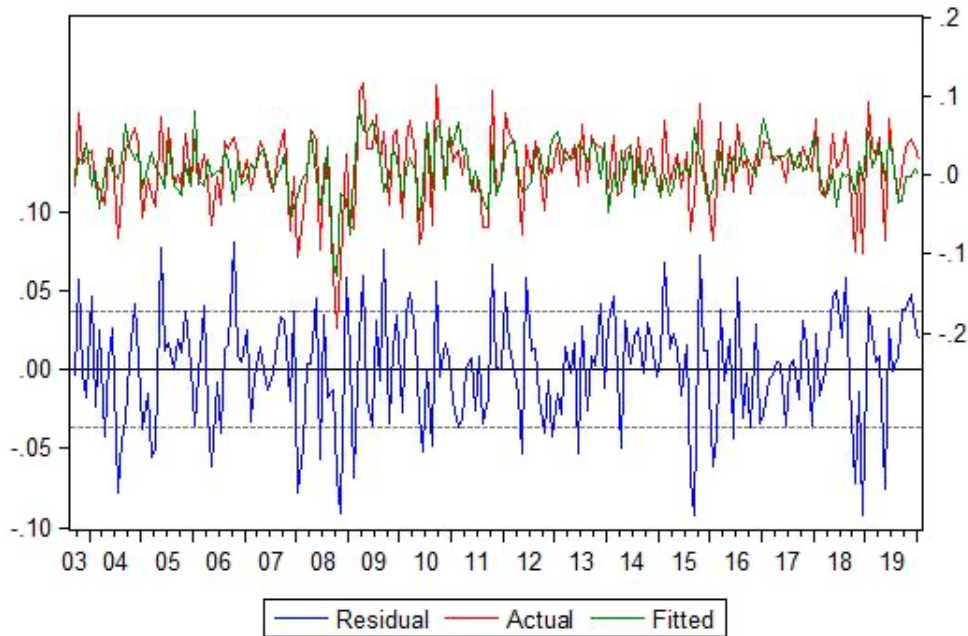
Dependent Variable: D(N_L)
 Method: Least Squares
 Date: 02/19/23 Time: 14:27
 Sample (adjusted): 2003M09 2020M01
 Included observations: 197 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007569	0.002859	2.647110	0.0088
D(TB(-3))	0.038654	0.019313	2.001424	0.0468
D(TB(-5))	-0.054434	0.018543	-2.935567	0.0038
D(TB(-9))	-0.057793	0.020940	-2.759995	0.0064
D(TB(-10))	0.067781	0.020717	3.271739	0.0013
D(BOND)	0.046346	0.014615	3.171197	0.0018
D(BOND(-7))	-0.025916	0.013296	-1.949134	0.0528
D(RET)	0.015207	0.003646	4.170546	0.0000
D(USD)	0.445756	0.097538	4.570098	0.0000
D(CHUS)	-0.140669	0.065423	-2.150133	0.0329
D(CHUS(-9))	0.176049	0.062774	2.804504	0.0056
D(CHUS(-1))	0.161258	0.064443	2.502353	0.0132
D(CHUS(-7))	0.102673	0.061074	1.681129	0.0945
D(LF)	-0.053205	0.020380	-2.610618	0.0098
D(LF(-2))	-0.064962	0.020302	-3.199825	0.0016
D(LF(-4))	-0.039577	0.019838	-1.995005	0.0476
ECM_ALT(-1)	-0.259009	0.061311	-4.224525	0.0000
R-squared	0.437762	Mean dependent var		0.008225
Adjusted R-squared	0.387786	S.D. dependent var		0.047349
S.E. of regression	0.037048	Akaike info criterion		-3.670861
Sum squared resid	0.247060	Schwarz criterion		-3.387539
Log likelihood	378.5798	Hannan-Quinn criter.		-3.556170
F-statistic	8.759337	Durbin-Watson stat		2.045395
Prob(F-statistic)	0.000000			

Con riferimento all'equazione dinamica finale si può apprezzare la negatività della componente ECM, il cui coefficiente risulta interno all'intervallo (-2;0). L'eccessiva prossimità al valore 0 potrebbe essere indicatrice di un modello instabile, fattispecie che non viene registrata nel caso oggetto di studio, in cui tale valore è di -0.259009. La componente ECM risulta inoltre significativa.

La rappresentazione sottostante guarda alla relazione tra la curva stimata (in verde) e quella effettiva (in rosso):

Figura 2.16- Curva stimata dall'equazione dinamica, curva effettiva e relativi residui



Una lettura della distanza tra le due curve, ovvero della deviazione della curva stimata rispetto a quella reale, è fornita dalla serie “residual” (in blu). Quest’ultima segnala valori non particolarmente difforni tra le due curve e suggerisce stazionarietà in media e varianza della componente residuale. Tra le deviazioni più significative vi sono quelle che si verificano in prossimità del 2004, in cui l’equazione stimata non sembra cogliere perfettamente due fasi di minimo dell’indice. Ulteriori variazioni degne di nota e ancor più evidenti, sono quelle in prossimità della recessione statunitense del 2008, giustificate dalla volatilità più elevata nei mercati finanziari. Infine, si segnala un ulteriore scostamento tra curva reale e stimata in prossimità dell’anno 2015 e nel 2018. Nel complesso la figura ottenuta rappresenta risultati più che soddisfacenti.

Ancor prima di commentare i regressori vengono approfondite l’autocorrelazione, l’eteroschedasticità e la normalità dei residui, in modo tale da rigettare il rischio di una stima distorta e/o non efficiente. Andando per ordine l’immagine sottostante illustra la struttura dei residui; la successiva dei residui al quadrato:

Figura 2.17- Correlogramma dei residui dell'equazione dinamica finale

Date: 02/19/23 Time: 14:34

Sample: 2002M10 2020M01

Included observations: 197

Q-statistic probabilities adjusted for 16 dynamic regressors

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*	
		1	-0.023	-0.023	0.1100	0.740
		2	-0.014	-0.015	0.1521	0.927
		3	-0.028	-0.029	0.3120	0.958
		4	-0.099	-0.101	2.3155	0.678
		5	0.038	0.033	2.6165	0.759
		6	-0.107	-0.111	4.9689	0.548
		7	0.044	0.036	5.3713	0.615
		8	0.094	0.085	7.2098	0.514
		9	0.023	0.030	7.3231	0.604
		10	-0.106	-0.125	9.6726	0.470
		11	-0.135	-0.124	13.532	0.260
		12	0.002	-0.003	13.533	0.331
		13	-0.050	-0.055	14.074	0.369
		14	-0.060	-0.082	14.858	0.388
		15	-0.031	-0.060	15.066	0.447
		16	0.068	0.037	16.072	0.448
		17	-0.010	-0.051	16.096	0.517
		18	-0.049	-0.040	16.625	0.549
		19	-0.017	-0.007	16.688	0.611
		20	-0.064	-0.080	17.602	0.614
		21	0.120	0.080	20.786	0.472
		22	-0.029	-0.025	20.979	0.522
		23	0.049	0.038	21.529	0.549
		24	0.018	-0.038	21.601	0.603
		25	-0.072	-0.081	22.797	0.589
		26	-0.047	-0.078	23.301	0.616
		27	-0.048	-0.021	23.822	0.640
		28	0.071	0.034	24.987	0.629
		29	-0.010	-0.052	25.011	0.678
		30	0.059	0.032	25.827	0.684
		31	0.005	-0.018	25.833	0.729
		32	-0.092	-0.092	27.860	0.676
		33	0.003	-0.014	27.862	0.721
		34	0.062	0.117	28.777	0.721
		35	-0.060	-0.090	29.650	0.724
		36	-0.027	-0.077	29.822	0.756

Figura 2.18- Correlogramma dei residui al quadrato dell'equazione dinamica finale

Date: 02/19/23 Time: 14:35
 Sample: 2002M10 2020M01
 Included observations: 197

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.086	0.086	1.4814	0.224
		2	-0.007	-0.015	1.4917	0.474
		3	-0.089	-0.088	3.0928	0.378
		4	0.064	0.081	3.9318	0.415
		5	0.150	0.139	8.5102	0.130
		6	0.046	0.014	8.9417	0.177
		7	-0.042	-0.036	9.2974	0.232
		8	-0.079	-0.054	10.608	0.225
		9	0.039	0.039	10.932	0.280
		10	0.098	0.066	12.966	0.226
		11	0.045	0.017	13.401	0.268
		12	-0.039	-0.023	13.722	0.319
		13	-0.089	-0.059	15.418	0.282
		14	-0.073	-0.077	16.550	0.281
		15	-0.052	-0.080	17.130	0.311
		16	0.005	-0.007	17.135	0.377
		17	-0.036	-0.021	17.413	0.427
		18	-0.108	-0.080	19.991	0.333
		19	-0.057	-0.020	20.703	0.353
		20	0.021	0.030	20.801	0.409
		21	-0.015	-0.043	20.851	0.468
		22	-0.032	-0.022	21.081	0.516
		23	-0.095	-0.051	23.127	0.453
		24	-0.084	-0.053	24.718	0.421
		25	-0.002	0.006	24.719	0.478
		26	-0.050	-0.071	25.286	0.503
		27	0.005	0.014	25.291	0.558
		28	0.022	0.060	25.404	0.606
		29	-0.112	-0.126	28.308	0.501
		30	-0.041	-0.041	28.702	0.533
		31	-0.042	-0.041	29.110	0.563
		32	0.041	0.008	29.503	0.594
		33	0.076	0.075	30.886	0.573
		34	-0.057	-0.053	31.678	0.582
		35	0.065	0.104	32.705	0.579
		36	-0.038	-0.055	33.063	0.609

La struttura non sembra tipica di serie autocorrelate, ma si procede al test LM *Breusch-Godfrey Serial Correlation* per avvalorare quanto osservato:

Figura 2.19- Test di Breusch-Godfrey per la correlazione dei residui della dinamica finale

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.132348	Prob. F(2,178)	0.8761
Obs*R-squared	0.292514	Prob. Chi-Square(2)	0.8639

Il test in questione ha per ipotesi nulla H_0 l'assenza di autocorrelazione. In questo caso la probabilità associata è indicativa di residui abbondantemente non autocorrelati e conferma quanto suggerivano i correlogrammi.

Per quanto concerne l'eteroschedasticità si effettua il test ARCH, avente come ipotesi nulla H_0 l'assenza di eteroschedasticità e, dunque, l'omoschedasticità della serie sottoposta al test:

Figura 2.20- Test ARCH per l'eteroschedasticità dei residui della dinamica finale

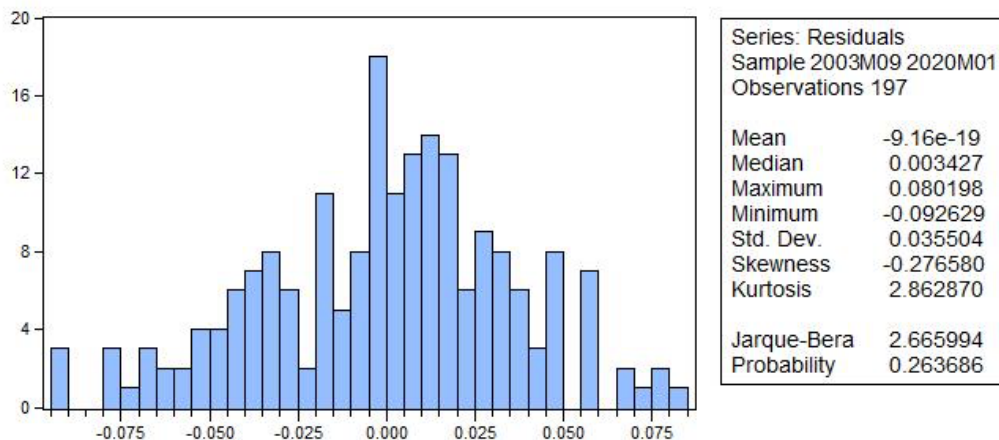
Heteroskedasticity Test: ARCH			
F-statistic	1.108762	Prob. F(3,190)	0.3468
Obs*R-squared	3.337877	Prob. Chi-Square(3)	0.3424

Anche questo test ottiene i risultati attesi e afferma l'assenza di eteroschedasticità sulla componente dei residui.

Il risultato che emerge dai test effettuati è fondamentale e suffraga la presenza di stime non distorte nella totalità dei processi che hanno riguardato la procedura a due passi di Engle-Granger.

Un ulteriore passaggio è quello volto ad indagare la normalità della serie attraverso l'istogramma, i valori in corrispondenza al momento terzo e quarto e, soprattutto, il test di Jarque Bera:

Figura 2.21- Distribuzione e valori sulla normalità dei residui della dinamica finale



Dalla figura 2.21 si può notare una struttura leggermente asimmetrica, ma con una curtosi prossima a 3 ed una probabilità associata al test di Jarque Bera che segnala l'accettazione dell'ipotesi nulla di normalità.

Le analisi effettuate sui residui sono rappresentative di una situazione ideale: nello specifico il correlogramma e i test conducono ad un modello per l'interpretazione dei residui di tipo white noise.

Riprendendo ora l'equazione dinamica finale, si possono analizzare i regressori restanti, le statistiche AIC (criterio di Akaike) e BIC (criterio di Schwarz), e l'Adjusted R-Squared rispetto al modello generale. Anzitutto si osserva che l'equazione in oggetto è costituita da sei variabili ("BOND", "RET", "TB", "LF", "CHUS", "USD") ed i relativi ritardi, oltre alla componente ECM(-1).

Partendo dalle variabili presenti anche a livello dell'equazione di lungo periodo si può apprezzare che i regressori "BOND", "RET", "TB" vengono mantenuti anche per la spiegazione dell'equazione di breve periodo. Questi ultimi si rivelano fondamentali per definire gli effetti nella relazione di breve periodo sull'endogena. Guardando alla stima effettuata, tutte le variabili citate sono significative e migliorative dei criteri di Akaike e Schwarz ed una rimozione delle medesime non porta ad un esito positivo.

Lo stesso ragionamento può essere esteso alle variabili che non erano state precedentemente incluse nell'equazione statica, ma che risultano presenti in

corrispondenza dell'equazione dinamica. Tra queste si osservano “USD”, “CHUS” ed “LF”. Le variabili “USD” e “CHUS” indicano un tasso di cambio relativo al dollaro statunitense rapportato rispettivamente all'Euro e allo Yuan cinese (quest'ultimo sulla valuta cinese rispetto al dollaro). Coerentemente con quanto atteso in sede di analisi introduttiva, anche i tassi di cambio hanno un ruolo esplicativo sul NASDAQ Composite, sebbene il loro apporto sia limitato alla sola relazione di breve periodo.

È da sottolineare che, accanto alle variabili incluse sia nella statica che nell'equazione dinamica, ve ne siano alcune che pur essendo significative nella prima tra le relazioni, non siano incluse a livello del passo successivo. Tra queste si ricordano i regressori “PCE”, “EM_L”, “IPT” e “PPIT”. Questi ultimi sono stati rimossi nei passaggi intermedi che hanno condotto all'equazione dinamica finale poiché non significativi.

Nel complesso l'Adjusted R-Squared unitamente ai criteri informativi migliorano nel corso delle varie fasi di rimozione dei ritardi delle variabili.

In *Appendice* sono state riportate le funzioni di risposta impulsiva per tutte le variabili coinvolte nell'equazione statica e/o nella dinamica.

Tali funzioni mostrano la dinamica di aggiustamento della variabile dipendente a seguito di un impulso dato dalle variabili esogene. In altri termini, attraverso queste funzioni, si può apprezzare il comportamento di una variabile il cui stato di quiete (“steady state”) viene perturbato. Coerentemente con quanto previsto dalla teoria econometrica, ogni funzione di risposta impulsiva tende a 0 con l'aumentare del tempo e la cumulata tende al valore del coefficiente di lungo periodo. Ovviamente per quelle variabili inserite esclusivamente in sede di equazione di breve periodo (come per “CHUS”, “USD” e “LF”), anche la funzione cumulata tenderà a 0, essendo nullo anche il coefficiente di lungo periodo.

In generale, nelle funzioni di risposta impulsiva più rapidamente la variabile converge a 0 e più velocemente tenderà ad una condizione di stabilità.

CAPITOLO III - Previsioni dal modello econometrico e confronto con il benchmark

In seno al presente capitolo si andranno a proporre diverse previsioni sull'indice NASDAQ Composite. Come specificato a livello introduttivo, la previsione della variabile endogena in questione rappresenta la finalità principale del progetto, unitamente a tutte le considerazioni che derivano dalla bontà delle analisi effettuate. L'orizzonte temporale che verrà sottoposto a questi processi sarà quello di un "futuro virtuale" di 12 mesi, compreso tra il gennaio 2020 e il gennaio 2021. La decisione di questo intervallo è motivata dall'interesse che possono suscitare gli output delle previsioni rispetto al periodo della crisi pandemica SARS CoV-2.

Si ricorda che l'OMS ha dichiarato ufficialmente lo stato di crisi pandemica l'11 marzo 2020, sebbene il primo caso negli Stati Uniti venne confermato il 20 gennaio 2020. In ragione di ciò la fase ribassista dei mercati fu precedente all'annuncio ufficiale da parte dell'OMS. La Borsa statunitense assorbì prima la notizia, mostrando un trend decrescente già nel mese precedente (nello specifico a partire dal 21-02-2020).

L'immagine sottostante rappresenta quanto anticipato, segnalando uno shock dell'indice NASDAQ Composite nella seconda metà di febbraio, che raggiunge un punto di minimo nella seconda metà del mese successivo:

Figura 3.1- Andamento dell'indice NASDAQ Composite dal 2019/10 al 2021/01



La bontà delle previsioni effettuate andrà interpretata seguendo proprio quanto premesso e considerando l'aumento generalizzato della volatilità nei mercati finanziari per l'intervallo considerato.

La domanda implicita che ci si pone è se il modello sia in grado di cogliere ed intercettare adeguatamente il declino dell'indice conseguente all'evento pandemico, attraverso le variabili esplicative definite nel contesto della procedura a due passi di Engle-Granger. Attingendo proprio dalle stime dell'equazione statica e dinamica si effettueranno tre previsioni dal modello econometrico:

- Previsione statica dal modello di lungo periodo
- Previsione statica dal modello di breve periodo
- Previsione dinamica dal modello di breve periodo

Una previsione statica differisce da una dinamica poiché prevede le realizzazioni future della variabile dipendente con i ritardi delle esplicative unitamente a quelli della variabile dipendente stessa. L'apporto della variabile endogena in questa procedura consente da un lato di ottenere, di norma, delle previsioni migliori, dall'altro di avere dei risultati che sono meno robusti da un punto di vista teorico, essendo che si tenderà a sfruttare la tendenza auto-esplicativa della variabile stessa.

La previsione dinamica, invece, viene effettuata ricorrendo alle serie storiche delle variabili esogene della regressione unitamente a quelli previsti ai ritardi precedenti dalla dipendente (ovvero dal NASDAQ Composite stesso). In questo modo la previsione risulterà normalmente meno efficace, conseguentemente al fatto che si rimuove la relazione con i dati storici dell'endogena stessa.

Nell'introdurre questo capitolo si è più volte fatto riferimento al concetto di previsione "migliore". Di seguito si andrà a specificare in quali termini alcune realizzazioni future risultano più efficaci rispetto ad altre. Tali considerazioni trovano riscontro con quanto implementato a livello pratico nel progetto. Nello specifico, la bontà previsiva è stata valutata con riferimento alle seguenti misure:

- *Root Mean Squared Error* (Errore Quadratico Medio), calcolato come:

$$RMSE = \sqrt{\sum_{t=T+1}^{T+h} (\hat{x}_t - x_t)^2 / h}$$

- *Mean Absolute Error* (Errore Assoluto Medio), calcolato come:

$$MAE = \sqrt{\sum_{t=T+1}^{T+h} |\hat{x}_t - x_t| / h}$$

- *Mean Percentage Absolute Error* (Errore Percentuale Medio Assoluto), calcolato come:

$$MAPE = 100 \sum_{t=T+1}^{T+h} \left| \frac{\hat{x}_t - x_t}{x_t} \right| / h$$

- *Theil Inequality Coefficient* (Coefficiente di Disuguaglianza di Theil), calcolato come:

$$TIC = \frac{\sqrt{\sum_{t=T+1}^{T+h} (\hat{x}_t - x_t)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} (\hat{x}_t)^2 / h + \sum_{t=T+1}^{T+h} (x_t)^2 / h}}$$

Soffermandosi sui primi due indicatori, il RMSE e il MAE non vanno osservati in valore assoluto, ma devono essere confrontati a seconda di eventuali modelli alternativi stimati sulla stessa variabile dipendente. Questi ultimi dipendono, infatti, dalla scala di misura della variabile endogena stessa e perdono di utilità nel momento in cui vengono osservati singolarmente o con modelli riguardanti variabili differenti. Se osservati correttamente, il

MAE e il RMSE attribuiscono una bontà previsiva migliore in prossimità di valori più contenuti rispetto a eventuali modelli alternativi stimati sulla stessa variabile.

Il MAPE e il TIC, al contrario, possono essere interpretati anche in assenza di modelli alternativi, poiché portano ad un valore invariante rispetto alla scala di misura dell'endogenea. Il coefficiente di disuguaglianza di Theil è compreso in un intervallo tra 0 e 1 e risulta fondamentale nell'analisi della singola previsione. Quest'ultima sarà migliore in corrispondenza di valori prossimi allo 0, meno efficace quando i valori tendono all'unità.

È possibile valutare la bontà previsiva considerando alcune declinazioni delle formule precedenti. Nella fattispecie è spesso riportato anche il MSE (Mean Squared Error), ottenuto semplicemente elevando al quadrato il RMSE. Lo stesso Theil ha proposto un'ulteriore scomposizione del MSE in tre componenti per misurare la performance delle stime effettuate:

- *Bias Proportion*: indica la misura in cui la previsione si allontana dai valori empirici futuri. In altri termini indica il gap tra la previsione effettuata e la curva effettiva futura e, di conseguenza, si otterranno risultati migliori per un indicatore che assume valori prossimi a 0. Per contro, valori più elevati saranno indicatori di una curva prevista significativamente sopra (o sotto) la curva effettiva futura.
- *Variance Proportion*: indica il gap in varianza tra curva prevista ed effettiva. È dunque una misura dell'errore in varianza in seno alla previsione effettuata e, come per il Bias Proportion sarà indice di una stima migliore in corrispondenza di valori che tendono a zero (ovvero quando si minimizza l'errore in varianza).
- *Covariance Proportion*: indica la correlazione tra curva prevista e curva effettiva futura. Essendo che la somma delle tre componenti deve risultare pari a 1, la Covariance Proportion sarà indicatrice di una previsione ottimale in corrispondenza di valori che tendono all'unità.

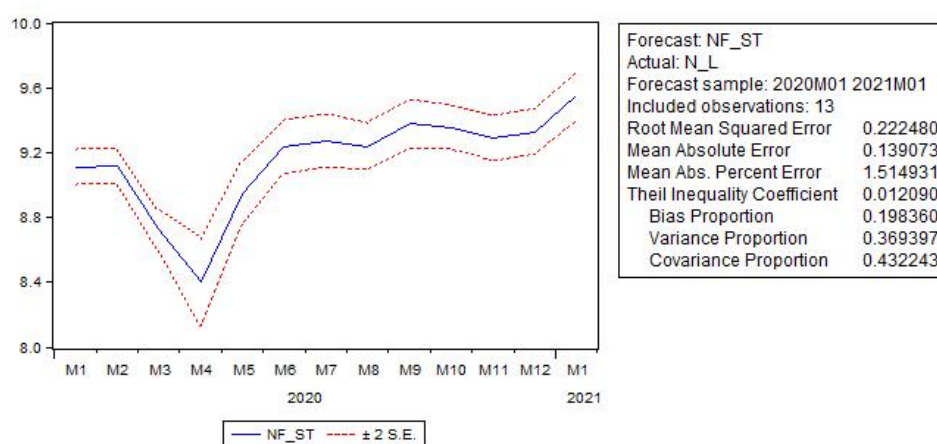
Di norma ad un indice di Theil piuttosto contenuto seguiranno delle misure di performance, come quelle di cui sopra, altrettanto soddisfacenti. Ciò andrà ad indicare un'elevata efficacia previsiva, oltre a certificare che l'insieme dei processi che ha preceduto questa fase ha portato all'esito desiderato.

3.1 Previsione statica dall'equazione statica

Andando per ordine, la prima tra le previsioni realizzate è quella che trae origine dall'equazione statica finale del primo passo della procedura di Engle-Granger. La previsione viene effettuata sulla variabile dipendente "N_L" (la trasformata logaritmica della serie del NASDAQ Composite) per il periodo compreso tra 01-01-2020 e 01-01-2021¹². Si sono stimati anche gli standard error per la previsione, utili a guardare se la curva effettiva futura sia compresa nell'intervallo generato dagli errori. Gli intervalli generati si basano su una confidenza al 95%, fissando quindi un α al 5%.

Dalla figura di seguito si possono apprezzare i risultati ottenuti in sede di stima:

Figura 3.2- Previsione statica dal modello statico



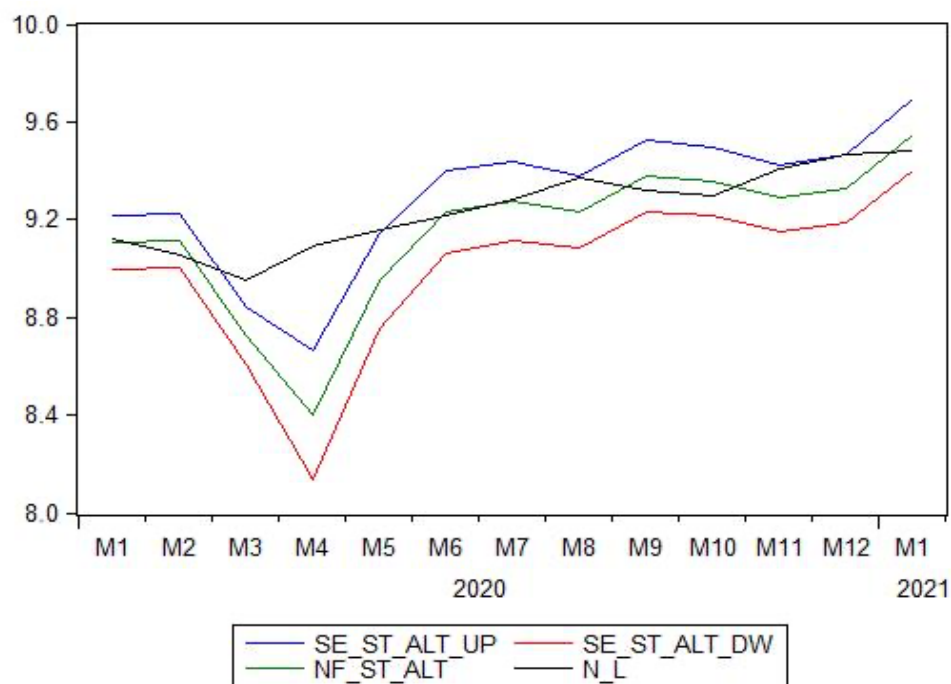
Dalla rappresentazione grafica si può notare un trend inizialmente decrescente, che sembrerebbe cogliere la contrazione della variabile endogena in corrispondenza dello scoppio della crisi pandemica nel mese di marzo e aprile. Anche la fase rialzista dei mesi appena successivi sembrerebbe intercettare l'andamento effettivo del NASDAQ Composite. Concentrando l'attenzione nelle misure di bontà previsiva si può osservare che:

¹² Si precisa che in questo caso l'unica previsione possibile è quella statica, essendo che il modello stesso d'origine non ha una componente dinamica.

- Coefficiente di disuguaglianza di Theil (TIC): segnala valori soddisfacenti che tendono a 0, nello specifico pari a 0.012090. Indica che nel complesso la previsione ha ottenuto dei risultati in linea con l'andamento della curva effettiva.
- Bias Proportion: non ottiene risultati ottimali, pur non attestandosi nemmeno su valori lontani da 0 (che è il valore ottimale). Sta ad indicare che probabilmente la curva effettiva non risulterà sempre interna alla banda di dispersione generata dagli standard error.
- Variance Proportion: vale lo stesso ragionamento fatto sul Bias Proportion, essendo che il valore assunto è pari a 0.369397. Indica un errore in varianza non trascurabile.
- Covariance Proportion: essendo che la somma delle tre componenti analizzate deve essere pari a 1, anche questo indicatore non ottiene risultati ottimali e tendenti all'unità. Il valore si attesta a 0.432243.

Una sintesi di quanto mostrano le misure di cui sopra è fornita dal confronto tra la previsione effettuata dal modello di lungo periodo (in verde) e l'andamento realizzato della curva effettiva della variabile endogena (in nero):

Figura 3.3- Confronto tra curva effettiva e statica prevista dall'equazione statica



Le curve denominate con “SE_ST_ALT_UP” e “SE_ST_ALT_DW”, indicano rispettivamente l’intervallo superiore e quello inferiore generati dagli standard error rispetto alla curva prevista. Questi intervalli sono utili per valutare gli scostamenti e la distanza di quanto previsto rispetto all’effettivo andamento della variabile endogena.

Il grafico conferma quanto ci si poteva attendere dalle misure di bontà previsiva precedentemente analizzate. La previsione non coglie correttamente la fase ribassista dell’indice in corrispondenza dei mesi di marzo e aprile, andando ad sovrastimare la contrazione del NASDAQ Composite. Ciò è dovuto realisticamente ad alcuni regressori che hanno risentito maggiormente della depressione economica. Va considerato che la stima viene effettuata per il periodo che, unitamente alla bolla delle Dot.com, ha avuto un impatto maggiore in termini di variazione e di volatilità dell’indice oggetto di studio. Nei mesi successivi la curva prevista rientra nelle bande di dispersione e sembra dare un’immagine in linea con quanto registrato dalla curva effettiva. La previsione statica rimane nel complesso soddisfacente seppur non ottimale.

3.2 Previsione statica dall’equazione dinamica

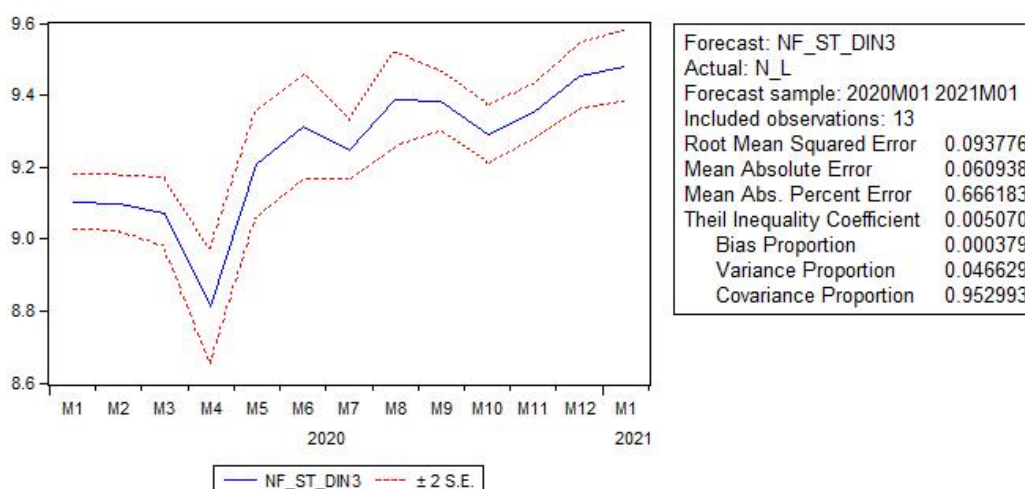
In questo caso si riprende la stima del modello di breve periodo, ovvero dell’equazione dinamica finale in forma ECM. Analogamente a quanto visto nel paragrafo precedente dall’equazione si andrà a generare una previsione statica, con i relativi standard error considerando sempre un intervallo di confidenza del 95%. Viene utilizzato lo stesso periodo campionario e lo stesso intervallo per il periodo di futuro virtuale, ovverosia quello compreso tra 2020/01/01 e 2021/01/01.

Rispetto alla previsione dalla regressione statica in questo caso bisogna considerare l’inserimento della variabile ECM nella stima dell’equazione anche per il periodo di “futuro virtuale”. La componente ECM era stata generata esclusivamente sul periodo campionario, mentre per la previsione sarà essenziale risalire alla componente in questione anche per il periodo di “futuro virtuale”. È dunque necessario generare la serie ECM per l’intero periodo, andando a calcolare tale componente come differenza tra la variabile dipendente ed equazione statica finale. Una volta concluso questo processo, si

andrà a sostituire la componente ECM precedente con quella appena calcolata all'interno della regressione dinamica.

Si procederà dunque alla stima della previsione statica dal modello dinamico e dei relativi standard error, dalla quale si otterrà il seguente output:

Figura 3.4- Previsione statica dal modello dinamico



Il grafico sembra rappresentare una situazione in linea con la curva effettiva del NASDAQ Composite nello stesso periodo. In particolare si può osservare un trend ribassista in corrispondenza dei mesi di marzo e aprile 2020, cui segue una fase rialzista e di stabilità della serie. Commentando le misure di bontà previsiva si può apprezzare che:

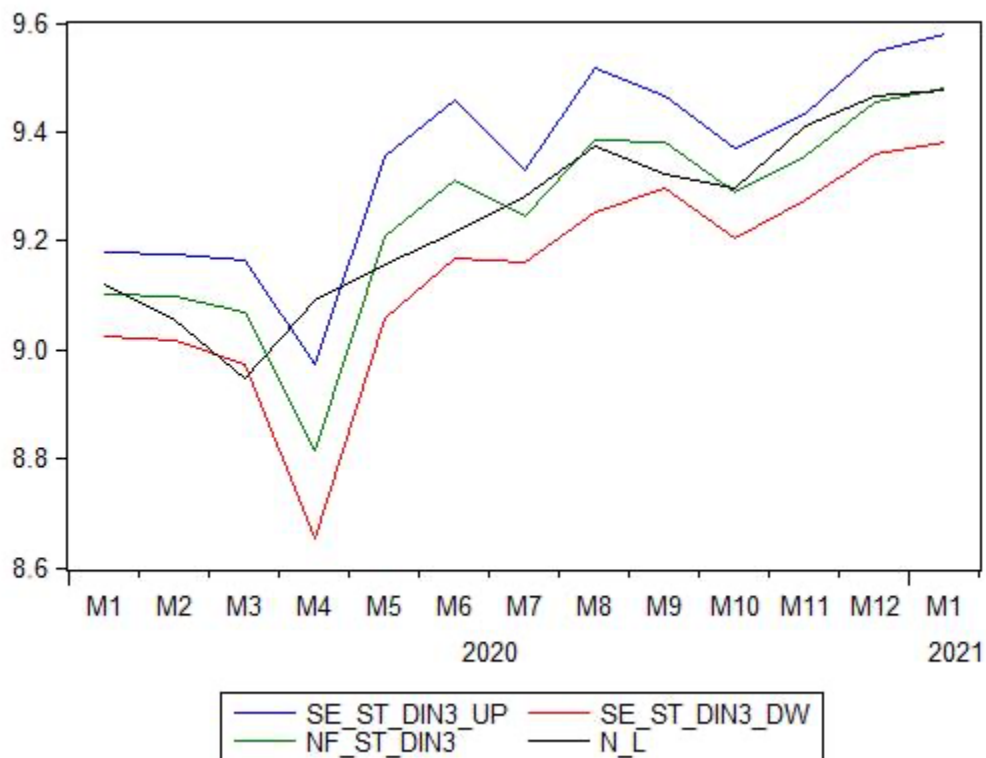
- Coefficiente di disuguaglianza di Theil: si attesta su valori decisamente bassi, pari a 0.00507, ottenendo peraltro risultati migliori rispetto alla previsione dalla regressione statica. Ciò sta ad indicare un'elevata efficacia previsiva della curva stimata.
- Bias Proportion: assume un valore sostanzialmente nullo. Anche in questo caso il risultato ottenuto è più che positivo e non dà segnali di una curva stimata distante da quella effettiva.

- Variance Proportion: segnala un valore di 0.046629. In coerenza con le misure di cui sopra, mostra un valore tipico di una previsione che non si discosta in varianza dalla curva effettiva.
- Covariance Proportion: conseguentemente ai valori numerici assunti dagli altri indicatori porta ad un valore approssimabile all'unità, nello specifico pari a 0.952993.

Differentemente da quanto osservato nella previsione sull'equazione di lungo periodo, in questo caso si ottengono dei valori associati ad un'ottima bontà previsiva per la variabile dipendente.

Tutto ciò si traduce in un esito decisamente migliore rispetto a prima nel confronto tra curva prevista (in verde) ed effettiva (in nero) della variabile endogena:

Figura 3.5- Confronto tra curva effettiva e statica prevista dall'equazione dinamica



Dalla figura si può apprezzare come la curva risultante dalla previsione segua nel complesso quella effettiva, intercettandone l'andamento e, in parte, le fasi di massimo e minimo. La stima effettuata non riesce ancora a cogliere con esattezza la fase ribassista cominciata nel marzo 2020: anche in questo caso si può osservare come la curva stimata abbia una contrazione più significativa in prossimità dello scoppio della crisi pandemica. Nei mesi successivi la curva effettiva è costantemente compresa nelle bande di dispersione generate e il suo andamento viene intercettato con maggiore efficacia nella seconda metà del 2020.

3.3 Previsione dinamica dal modello dinamico

Nel presente paragrafo verrà trattata la previsione dinamica generata dalla regressione di breve periodo. Oltre alla dinamicità propria di questa previsione, diversamente da quanto mostrato nella previsione statica in questo caso si possono utilizzare due modalità alternative per l'inserimento della componente ECM.

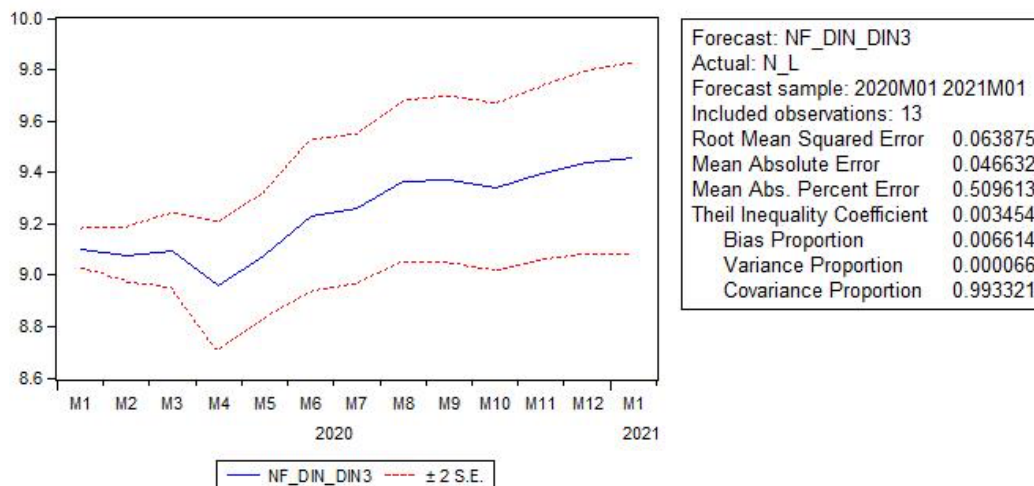
Un primo approccio è quello di immettere la serie ECM che contiene, per il periodo di "futuro virtuale", i valori stimati generati dalla regressione statica finale. Questo permette di avere tutte le osservazioni necessarie per la serie dei residui sia per il periodo campionario che per quello futuro.

Un percorso alternativo e più comune è quello di inserire nell'equazione dinamica finale la differenza tra i valori effettivi della variabile dipendente e la regressione statica, andando a ritardare tutti i regressori di un periodo (compresa la variabile endogena).

Nel progetto in discussione si è deciso di optare per quest'ultima modalità, che ha portato a reinserire la componente $ECM(-1)$ in forma estesa nell'equazione di breve periodo, così come spiegato poc'anzi. Il risultato sarà quello di una regressione dinamica estesa, sulla quale verrà realizzata una previsione per l'intervallo compreso tra 2020/01/01 e il 2021/01/01.

L'output ottenuto sarà una previsione dinamica di questo tipo:

Figura 3.6- Previsione dinamica dal modello dinamico



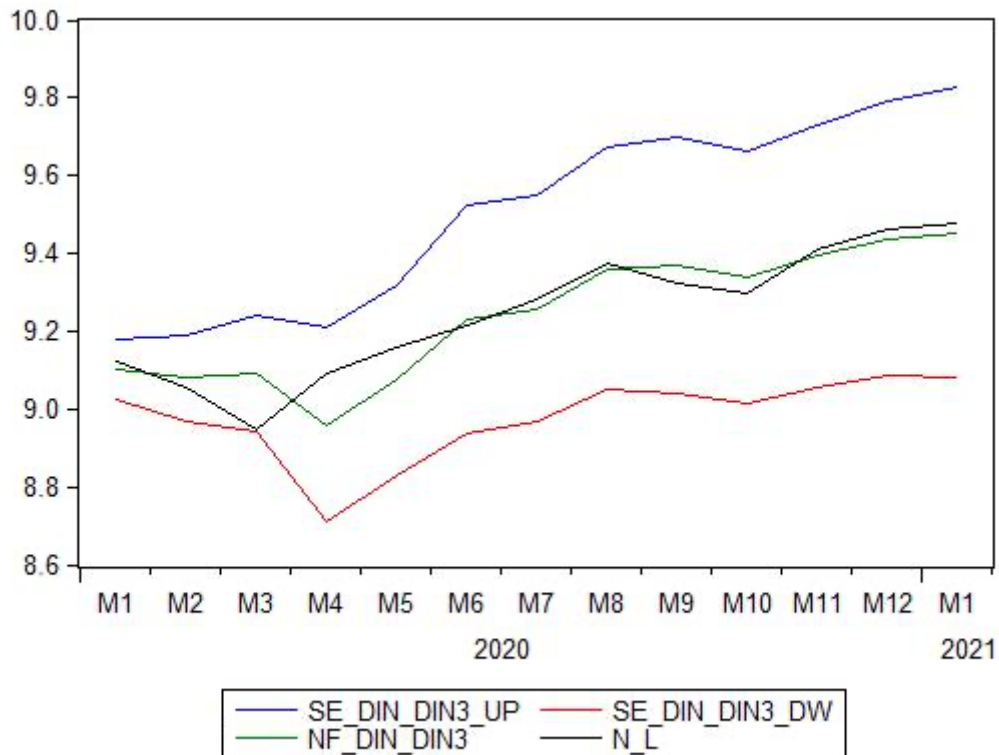
Dal grafico si può apprezzare un andamento che sembra essere più in linea con la curva effettivamente realizzata dall'indice NASDAQ Composite. Secondo le stesse modalità viste nei paragrafi precedenti sono stati creati gli intervalli superiori ed inferiori generando lo standard error.

In merito alle misure di bontà previsiva si possono fare le seguenti osservazioni:

- Coefficiente di disuguaglianza di Theil: registra un valore decisamente basso, pari a 0.003454. Ciò è segnale di una buona previsione e anticipa la bontà previsiva delle misure di seguito.
- Bias Proportion: assume un valore decisamente contenuto, di 0.006614. Questo valore è la conferma di una previsione più che soddisfacente ed indica che la curva stimata non si collocherà al di sopra (o al di sotto) della curva effettiva, ma che seguirà l'andamento medio di quest'ultima.
- Variance Proportion: si attesta ad un valore di 0,000066. L'output indica un valore ottimale, con un errore in varianza sostanzialmente nullo.
- Covariance Proportion: riporta un valore che tende all'unità, pari a 0.993321. Indica un valore ottimale in quanto a correlazione della curva stimata con quella effettiva.

Nella previsione dinamica dall'equazione dinamica si sono ottenuti i risultati migliori in termini di bontà previsiva. Tali valori trovano pieno riscontro nel confronto tra la curva stimata (in verde) e quello che ha effettivamente realizzato l'indice della variabile endogena (in nero) per lo stesso periodo:

Figura 3.7- Confronto tra curva effettiva e dinamica prevista dall'equazione dinamica



Si può osservare come la previsione effettuata dall'equazione di breve periodo sia in grado di intercettare l'andamento realizzato dalla serie effettiva "N_L". Quest'ultima è, infatti, costantemente compresa nell'intervallo generato dagli standard error, sebbene non tenda a sovrapporsi in tutti i periodi. Nello specifico continua a persistere un gap tra quanto stimato in previsione e quanto empiricamente realizzato dalla variabile dipendente nei mesi di marzo e aprile, dove la prima tra le due curve coglie in ritardo la fase di contrazione dell'indice. Nei mesi che seguono, la previsione riprende l'andamento effettivo del Nasdaq Composite, cogliendo in parte la fase ribassista che caratterizza la seconda metà di settembre e il mese di ottobre. Infine, a novembre e dicembre 2020 la curva stimata dalla previsione si sovrappone perfettamente alla curva effettiva. Ciò

conferma empiricamente la bontà previsiva a distanza di 10 mesi rispetto al periodo campionario.

Essendo che la curva stimata è costantemente compresa nelle bande generate dall'intervallo di confidenza, è possibile accettare per una significatività al 95%, l'ipotesi di una previsione statisticamente uguale alla curva effettiva della variabile dipendente.

3.4 Confronto della previsione dal modello econometrico con il modello ARIMA del benchmark

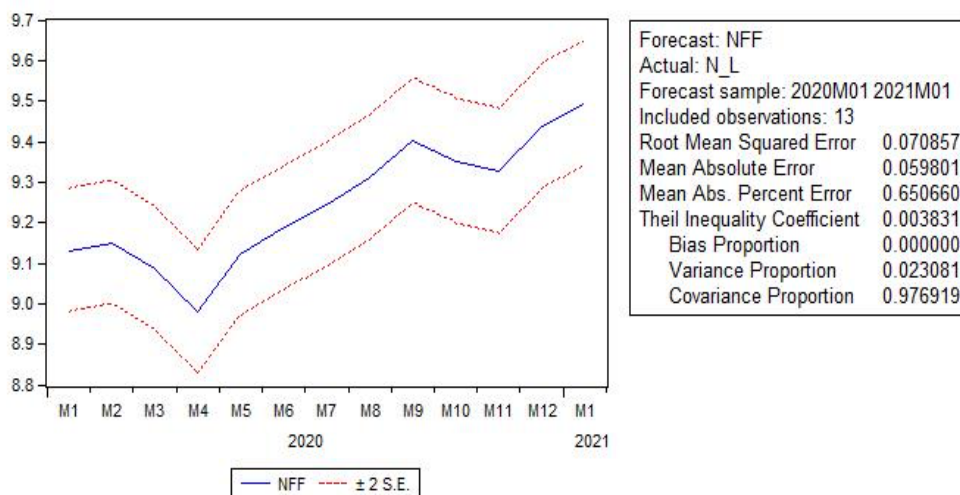
In questo paragrafo viene ripreso quanto visto nel primo capitolo con l'analisi statistica sulla variabile endogena, che verrà denominata "benchmark". Il capitolo richiamato individuava come modello per la variabile dipendente NASDAQ Composite un ARIMA(0,1,0), ovvero un processo random walk.

Si procederà alla stima di una previsione statica e ad una dinamica sul modello per la media della variabile dipendente, al fine di confrontare i risultati ottenuti con quanto effettivamente realizzato dalla serie NASDAQ Composite.

3.4.1 Previsione statica del benchmark

Di seguito viene effettuata la previsione statica del benchmark, generando i relativi standard error, per il periodo compreso tra 2020/01/01 e 2021/01/01:

Figura 3.8- Previsione statica del benchmark



Rispetto alla previsione statica dal modello econometrico dinamico notiamo che le misure di bontà previsiva sono lievemente migliori. La tabella a seguire, riporta un confronto diretto tra i due modelli:

Figura 3.9- Confronto della bontà previsiva tra la previsione statica dall'equazione dinamica econometrica e la previsione statica del benchmark

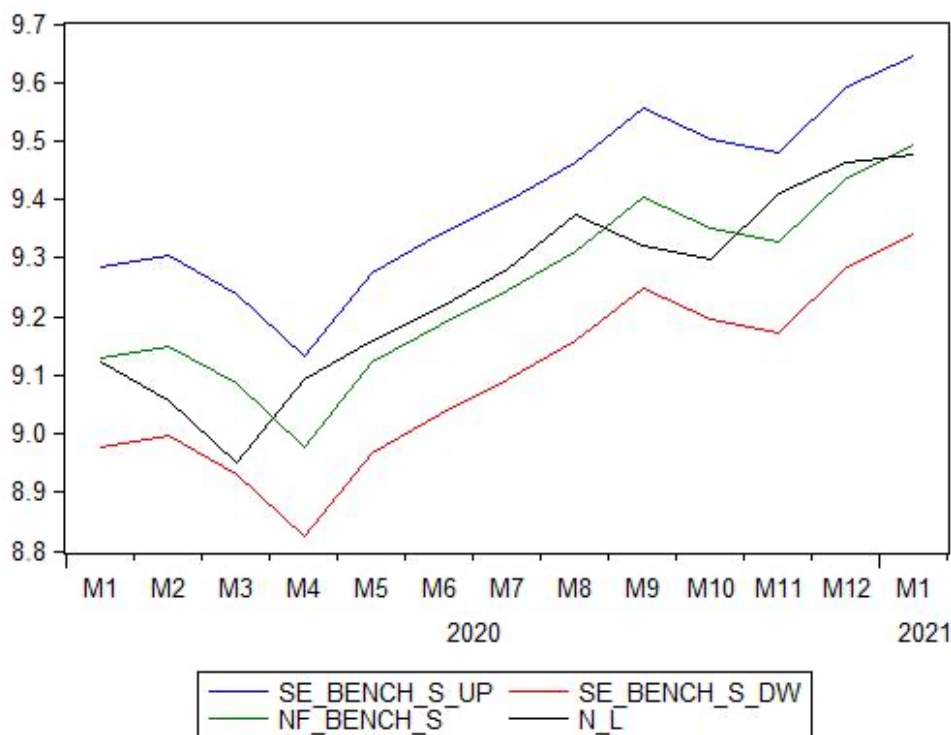
	RMSE	MAE	MAPE	TIC	BIAS	VAR	COV
Modello econometrico	0.09377	0.0609	0.6662	0.00507	0.00038	0.0466	0.953
Modello ARIMA	0.07086	0.0598	0.6507	0.0038	0	0.0231	0.9769

Commentando brevemente quanto rappresentato:

- RMSE, MAE e MAPE: risultano in linea rispetto a quelli ottenuti dal modello econometrico, andando a segnalare una previsione dal modello ARIMA sostanzialmente uguale in termini di bontà previsiva.
- Coefficiente di disuguaglianza di Theil (indicato con "TIC"): nel modello benchmark è leggermente inferiore rispetto a quello del modello econometrico, nello specifico di 0.003831, indice del fatto che anche la previsione ottenuta dal modello ARIMA ottiene ottimi risultati.
- Bias, Variance, Covariance Proportion (indicati rispettivamente da "BIAS", "VAR" e "COV"): confermano la bontà del modello ARIMA in previsione. I valori ottenuti sono decisamente bassi in prossimità delle prime due misure, si collocano vicino all'unità nel caso della Covariance Proportion, con un valore pari a 0.976919. Quanto ottenuto dal modello econometrico non si discosta rispetto a quanto generato dalla stima del benchmark.

Nel complesso la previsione statica dal modello ARIMA sembrerebbe in linea con il modello econometrico, con dei risultati lievemente migliori a favore del primo tra i due. Come di consueto, il grafico sottostante mette a confronto quanto ottenuto in previsione dal benchmark rispetto all'effettivo andamento della variabile endogena "N_L":

Figura 3.10- Confronto tra curva effettiva e statica del benchmark



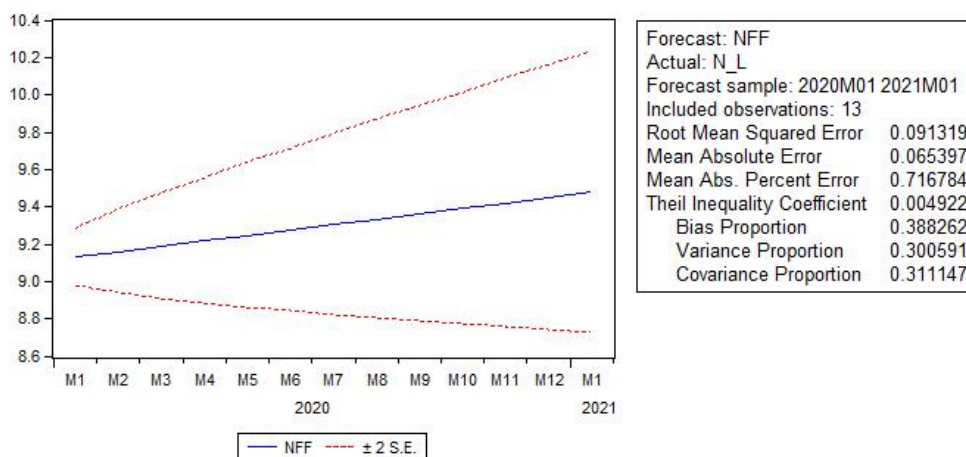
La curva prevista evidenzia un comportamento coerente con la natura di una serie random walk rispetto alla variabile dipendente. Quest'ultima (in nero) mostra un andamento che anticipa quanto realizzato dalla stima della previsione (in verde). In realtà, sebbene il risultato sia indicativo di una curva piuttosto artificiosa, l'output rivela una stima che è costantemente compresa all'interno delle bande degli standard error. Tuttavia, quanto stimato coglie in ritardo tutte le variazioni realizzate dalla variabile endogena.

Rispetto alla previsione statica dal modello di breve periodo econometrico, i risultati in corrispondenza alle misure di bontà previsiva sono leggermente migliori, ma la curva ottenuta denota una scarsa utilità essendo costantemente in ritardo nell'intercettare le variazioni dell'indice oggetto di discussione. Conseguentemente a ciò, nel complesso sarà ritenuta migliore la previsione statica dal modello econometrico.

3.4.2 Previsione dinamica del benchmark

Per la previsione dinamica del benchmark si prende sempre in analisi il modello ARIMA specificato per la variabile endogena, con eguale periodo campionario e generando una previsione per l'intervallo compreso tra 2020/01/01 e 2021/01/01. L'output sarà il seguente:

Figura 3.11- Previsione dinamica del benchmark



Dal grafico si può apprezzare un andamento lineare, con delle bande generate dagli standard error che tendono ad allargarsi con l'aumentare del tempo. Ciò è indicativo di una previsione piuttosto incerta e sommaria. Confrontando le misure di bontà previsiva rispetto all'analogo modello econometrico si riscontra:

Figura 3.12- Confronto della bontà previsiva tra la previsione dinamica dall'equazione dinamica econometrica e la previsione statica del benchmark

	RMSE	MAE	MAPE	TIC	BIAS	VAR	COV
Modello econometrico	0.0639	0.0466	0.5096	0.00345	0.0066	0	0.9933
Modello ARIMA	0.0913	0.0654	0.7168	0.00492	0.3883	0.3006	0.3111

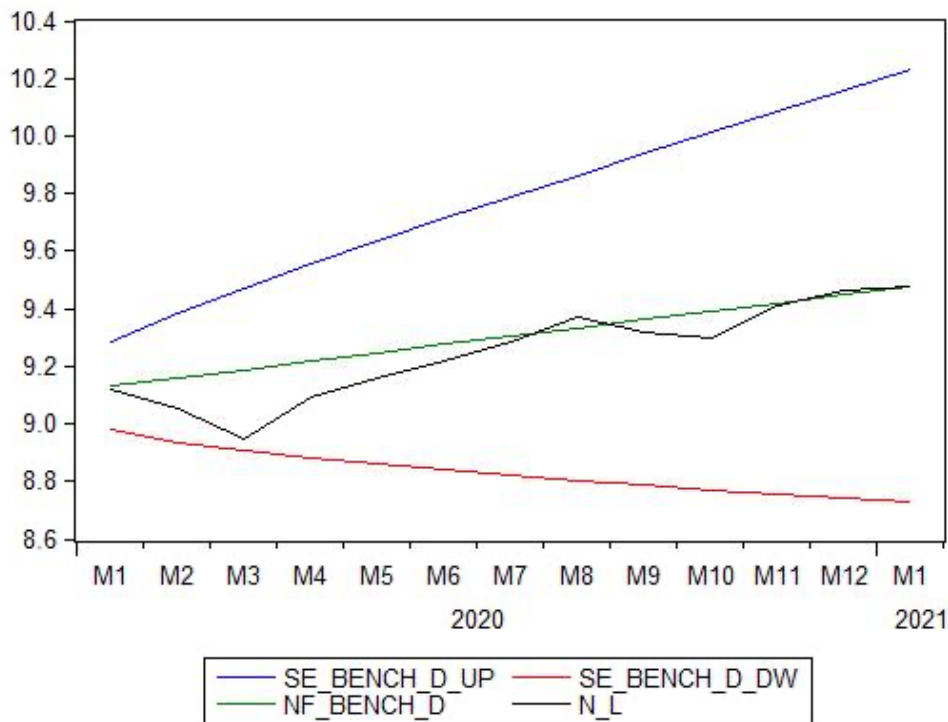
Dalla rappresentazione di cui sopra emerge:

- RMSE MAE e MAPE sono più contenuti per il modello econometrico sulla previsione dinamica dall'equazione di breve periodo. Ciò sta ad indicare una bontà previsiva a favore della stima econometrica.
- Coefficiente di disuguaglianza di Theil: si attesta ad un valore di 0.004922 per le stime sul modello ARIMA. La prossimità a 0 suggerirebbe un modello valido ed efficace per la previsione dell'indice. Tuttavia, questo valore è più elevato rispetto a quanto scaturisce dal modello econometrico, che realizza uno 0.00345, indice di un risultato migliore.
- Bias, Variance, Covariance Proportion: riportano valori decisamente migliori nel caso della previsione dinamica dal modello econometrico. Il modello ARIMA mostra valori tutt'altro che soddisfacenti, in coerenza con quanto emerge graficamente. Le tre misure assumono valori intermedi e suffragano la possibilità che la curva empirica della variabile dipendente non segua la previsione dinamica generata dal benchmark

Nel complesso i risultati, specie quelli ottenuti in corrispondenza delle misure di Bias, Variance e Covariance Proportion, sono indicativi di una previsione derivante dal modello ARIMA(0,1,0) piuttosto sommaria e approssimativa, nonché decisamente meno efficace rispetto alla analoga previsione dal modello econometrico. In questo caso il modello generato dalla regressione di breve periodo “batte” il benchmark, ottenendo risultati di bontà previsiva migliori.

Ciò si può riscontrare altresì attraverso un confronto diretto tra la curva effettiva e quanto scaturisce dalla previsione del modello ARIMA (0,1,0):

Figura 3.13- Confronto tra curva effettiva e dinamica del benchmark



La previsione (in verde) ha un andamento che soltanto in media riesce ad intercettare il trend della variabile dipendente (in nero). La curva stimata non riesce a cogliere alcuna variazione della variabile endogena, confermando quanto suggeriva la figura 3.11. La curva generata è comunque in grado di seguire lo stesso trend del NASDAQ Composite, caratterizzato da un andamento nel complesso crescente, arrivando a sovrapporsi alla curva effettiva negli ultimi mesi del 2020.

Il benchmark è rappresentativo di una previsione statica soddisfacente, ma al contempo artificiosa, e di una previsione dinamica piuttosto approssimativa.

Nel complesso quanto ottenuto in seno alla previsione econometrica ottiene dei risultati preferibili rispetto a quanto stimato per il “futuro virtuale” dal modello random walk sulla sola variabile endogena. Il modello derivante dalla procedura a due passi di Engle-Granger sarà dunque più efficace nell’interpretare l’indice oggetto di studio, anche in previsione.

Conclusioni

Nel presente elaborato si è proposta una regressione uniequazionale del NASDAQ Composite, finalizzata alla stima di modelli di previsione validi per descrivere il comportamento dell'indice nel primo anno della crisi pandemica.

In un primo momento, in coerenza con la natura della variabile endogena, si sono opportunamente selezionate le variabili esplicative atte a caratterizzare l'indice per il periodo compreso tra ottobre 2002 e gennaio 2020. Si sono scelte a tal fine serie storiche macroeconomiche e finanziarie, da inserire nella procedura a due passi di Engle-Granger. Per il primo passo, a livello dell'equazione statica, si sono analizzate, oltre alla dipendente, otto variabili esogene. L'analisi descrittiva ha riguardato le serie storiche: "3-Month Treasury Bill Secondary Market Rate", "Consumer Sentiment index", "Long-Term Government Bond Yields 10-year", "Personal Consumption Expenditures", "Producer price index by industry data processing hosting and related services", "Employment Level", "Industrial production Construction supplies" e "Total retail trade". Le variabili di cui sopra sono state sottoposte a test preliminari volti ad identificarne la natura e a comprenderne le caratteristiche. Dai test di integrazione come l'ADF, tutte le serie elencate sono risultate non stazionarie e con una radice unitaria.

Accertato ciò, i regressori menzionati, unitamente ad altre variabili potenzialmente esplicative, sono stati inseriti nell'equazione statica. In seguito, sono state gradualmente rimosse quelle variabili che non risultavano significative o che non erano migliorative dei criteri informativi, arrivando alla struttura finale dell'equazione di lungo periodo.

Questa prima fase è utile per studiare la relazione di lungo periodo tra le serie storiche esplicative e l'indice NASDAQ Composite. I risultati emersi dall'equazione statica dimostrano un impatto significativo delle variabili macroeconomiche sulla serie endogena. Nello specifico, la variabile della spesa individuale per il consumo ha registrato l'impatto più significativo sull'indice statunitense, nel quale hanno influito fortemente anche i rendimenti dei titoli di stato a lungo termine e il tasso di occupazione.

Dall'equazione statica si sono generati i residui che, coerentemente con quanto previsto dalla teoria econometrica, sono risultati stazionari: il modello si è rivelato stabile e conferma la cointegrazione delle variabili. Il test ADF sui residui (denominati "ECM"),

non solo conferma che la serie è stazionaria, ma riporta un valore di t-Student corrispondente a -9.224679 , ampiamente in zona di rifiuto dell'ipotesi nulla sulla presenza di radice unitaria.

A questo livello si è altresì approfondita l'eteroschedasticità condizionale, fornendo una possibile interpretazione alla componente attraverso i modelli ARCH e GARCH. Il modello ARCH(1) è risultato il più utile a questo fine, fornendo una spiegazione efficace per la varianza.

Ultimata l'equazione statica si è effettuata la stima dell'equazione di breve periodo (o dinamica). In un primo momento si sono richiamate nella stima tutte le variabili relative al primo passo, comprese quelle che erano state escluse poichè non significative o stazionarie, oltre alla componente dei residui ECM ritardata di un periodo. Accanto a queste ultime, si sono altresì inserite ulteriori variabili, potenzialmente utili a definire una relazione di breve periodo. Tra queste si ricordano le serie "Labor Force Participation Rate", "U.S. Dollars to Euro Spot Exchange Rate" e lo "Yuan to U.S. Dollar Spot Exchange Rate". Al fine di arrivare all'equazione dinamica finale, le variabili sono state rimosse a seconda della significatività, dando precedenza ai ritardi temporali più lontani e alle serie contemporanee. Nel corso di questa procedura si è prestata estrema attenzione alla componente ECM che, oltre ad essere significativa e ad avere un coefficiente compreso tra $(-2;0)$, deve risultare conforme ad un processo white noise. L'assenza di queste condizioni porterebbe infatti ad un modello instabile e con eteroschedasticità residua. Queste condizioni sono state costantemente soddisfatte, in particolare:

- il test LM ARCH ha condotto all'accettazione dell'ipotesi nulla di assenza di eteroschedasticità.
- il test LM Breusch-Godfrey Serial Correlation ha portato ad accettare l'ipotesi nulla di assenza di autocorrelazione seriale.
- la componente ECM(-1) è risultata significativa e pari a -0.259009 .

Nell'equazione dinamica finale sono state mantenute le variabili denominate "TB", "BOND", "RET", "USD", "CHUS" ed "LF". Da quanto ottenuto si può constatare come a livello dell'equazione di breve periodo vengano meno alcune relazioni macroeconomiche (presenti nell'equazione di lungo periodo) a favore di variabili

finanziare, come quelle relative ai tassi o ai rendimenti. In altri termini, nel breve periodo si perde la dipendenza dalle variabili macroeconomiche, tra le quali “PCE”, “IPT”, “EM_L” e “PPIT”, che impattano sull’indice nel lungo periodo.

Il passo successivo è stato quello di generare le previsioni derivanti dal modello econometrico de quo. Dall’equazione di lungo periodo si è generata una previsione statica, dall’equazione di breve periodo una dinamica e una statica, per l’intervallo compreso tra il 2020/01/01 e il 2021/01/01, definito “futuro virtuale”.

Gli output delle previsioni econometriche sono stati confrontati in primis con la curva effettiva realizzata dal NASDAQ Composite per lo stesso periodo e, in seconda battuta, con una previsione basata su un modello di tipo ARIMA (0,1,0) della sola variabile endogena (definito “modello benchmark”).

Le stime hanno confermato la robustezza del modello econometrico e hanno ottenuto delle misure di bontà previsiva efficienti. In particolare, dalla previsione statica dal modello dinamico il coefficiente di disuguaglianza di Theil è risultato pari a 0.00507 ed ha intercettato adeguatamente l’andamento della variabile dipendente in previsione, pur non cogliendo con precisione il crollo di marzo 2020.

Con riguardo alla previsione dinamica dal modello dinamico si ottengono stime ancor più efficaci. Nello specifico il coefficiente di disuguaglianza di Theil assume valori prossimi a zero, pari a 0.003454, e la curva effettiva della variabile endogena viene intercettata con ottimi risultati. Rispetto al “modello benchmark”, in quest’ultimo caso la previsione econometrica si dimostra migliore.

In conclusione, con il presente elaborato, accanto al modello di regressione sul NASDAQ Composite, si sono proposte delle curve di previsione efficaci per il primo anno della crisi pandemica. Questi risultati sono stati ottenuti a seguito di un’analisi econometrica della variabile dipendente attraverso la procedura a due passi di Engle-Granger, che ha fornito una descrizione dell’indice in uno dei periodi più nefasti della storia.

Appendice

Relativamente all'analisi descrittiva dei regressori, vengono riportati di seguito i risultati emersi dai test di integrazione, nello specifico dal test ADF (con trend o intercetta o none):

ADF test per la variabile CC con intercetta

```
## #####  
## # Augmented Dickey-Fuller Test Unit Root Test #  
## #####  
##  
## Test regression drift  
##  
##  
## Call:  
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -13.4331  -1.9309  -0.0691   2.4614  11.5776   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)   
## (Intercept)  4.72509     2.07966   2.272  0.0241 *   
## z.lag.1      -0.05528     0.02451  -2.255  0.0252 *   
## z.diff.lag   -0.01452     0.07035  -0.206  0.8367   
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 4.082 on 203 degrees of freedom  
## Multiple R-squared:  0.02632,    Adjusted R-squared:  0.01673   
## F-statistic: 2.744 on 2 and 203 DF,  p-value: 0.0667  
##  
##  
## Value of test-statistic is: -2.2549 2.5816
```

```
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

ADF test per la variabile BOND con trend

```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.10217 -0.12399  0.00913  0.11150  0.58955
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.2628625  0.1087186   2.418  0.0165 *
## z.lag.1     -0.0584322  0.0231064  -2.529  0.0122 *
## tt          -0.0008706  0.0003930  -2.215  0.0279 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2033 on 204 degrees of freedom
## Multiple R-squared:  0.03088,    Adjusted R-squared:  0.02138
## F-statistic:  3.25 on 2 and 204 DF,  p-value: 0.04078
##
##
```

```
## Value of test-statistic is: -2.5288 2.352 3.2502
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

ADF test per la variabile PCE senza trend e intercetta

```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -181.02  -15.59    2.44   22.17   86.15
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1  0.0032010  0.0002338   13.69  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 37.09 on 206 degrees of freedom
## Multiple R-squared:  0.4764, Adjusted R-squared:  0.4739
## F-statistic: 187.5 on 1 and 206 DF, p-value: < 2.2e-16
##
##
```

```
## Value of test-statistic is: 13.6918
```

```
##
```

```
## Critical values for test statistics:
```

```
##      1pct  5pct 10pct
```

```
## tau1 -2.58 -1.95 -1.62
```

ADF test per la variabile PPIT con trend

```
## #####
```

```
## # Augmented Dickey-Fuller Test Unit Root Test #
```

```
## #####
```

```
##
```

```
## Test regression trend
```

```
##
```

```
##
```

```
## Call:
```

```
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -1.06085 -0.10433 -0.04717  0.07328  1.90510
```

```
##
```

```
## Coefficients:
```

```
##           Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)  6.465839   2.902991   2.227  0.02708 *
```

```
## z.lag.1      -0.064565   0.028846  -2.238  0.02635 *
```

```
## tt           0.003798   0.001405   2.703  0.00748 **
```

```
## z.diff.lag   0.028196   0.075496   0.373  0.70920
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Residual standard error: 0.28 on 193 degrees of freedom
```

```
## Multiple R-squared:  0.04758,    Adjusted R-squared:  0.03278
```

```
## F-statistic: 3.214 on 3 and 193 DF,  p-value: 0.02405
```

```
##
```

```
##
## Value of test-statistic is: -2.2382 5.257 4.7827
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

ADF test per la variabile EM_L senza trend e intercetta

```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min        1Q      Median        3Q       Max
## -0.0068457 -0.0013180 -0.0000916  0.0013994  0.0066545
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      3.452e-05  1.353e-05   2.550  0.0115 *
## z.diff.lag1  1.174e-01  6.754e-02   1.739  0.0836 .
## z.diff.lag2  3.069e-01  6.757e-02   4.543 9.62e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002111 on 199 degrees of freedom
## Multiple R-squared:  0.2012, Adjusted R-squared:  0.1891
```



```

## F-statistic: 16.7 on 3 and 199 DF, p-value: 1.017e-09
##
##
## Value of test-statistic is: 2.5501
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62

```

ADF test per la variabile IPT senza trend e intercetta

```

## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.7015 -0.5909  0.0068  0.6489  3.4717
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -0.0001694  0.0007144  -0.237  0.812851
## z.diff.lag1   0.1258419  0.0691945   1.819  0.070458 .
## z.diff.lag2   0.2393949  0.0674393   3.550  0.000481 ***
## z.diff.lag3   0.2532603  0.0692581   3.657  0.000327 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.037 on 200 degrees of freedom

```

```

## Multiple R-squared:  0.2055, Adjusted R-squared:  0.1897
## F-statistic: 12.94 on 4 and 200 DF,  p-value: 2.19e-09
##
##
## Value of test-statistic is: -0.2371
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62

```

ADF test per la variabile RET senza trend e intercetta

```

## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.00492 -0.41883  0.07551  0.46372  2.22655
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      0.0009351  0.0005906   1.583   0.115
## z.diff.lag -0.0606420  0.0712313  -0.851   0.396
##
## Residual standard error: 0.7963 on 197 degrees of freedom
## Multiple R-squared:  0.01471,    Adjusted R-squared:  0.004708
## F-statistic: 1.471 on 2 and 197 DF,  p-value: 0.2323
##

```

```
##
## Value of test-statistic is: 1.5833
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

ADF test per la variabile LF senza trend e intercetta

```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.40254 -0.08243  0.01734  0.09816  0.31778
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -0.0002740  0.0001467  -1.867  0.06334 .
## z.diff.lag -0.1985495  0.0693446  -2.863  0.00464 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1334 on 200 degrees of freedom
## Multiple R-squared:  0.05041,    Adjusted R-squared:  0.04091
## F-statistic: 5.308 on 2 and 200 DF,  p-value: 0.005671
```

```
##
##
## Value of test-statistic is: -1.8672
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

ADF test per la variabile USD con intercetta

```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.08319 -0.01512 -0.00231  0.01849  0.09292
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.04295     0.02043   2.102  0.0368 *
## z.lag.1     -0.03419     0.01610  -2.124  0.0349 *
## z.diff.lag   0.29876     0.06694   4.463 1.36e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02715 on 197 degrees of freedom
## Multiple R-squared:  0.1027, Adjusted R-squared:  0.09364
## F-statistic: 11.28 on 2 and 197 DF,  p-value: 2.303e-05
```

```
##
##
## Value of test-statistic is: -2.1244 2.266
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

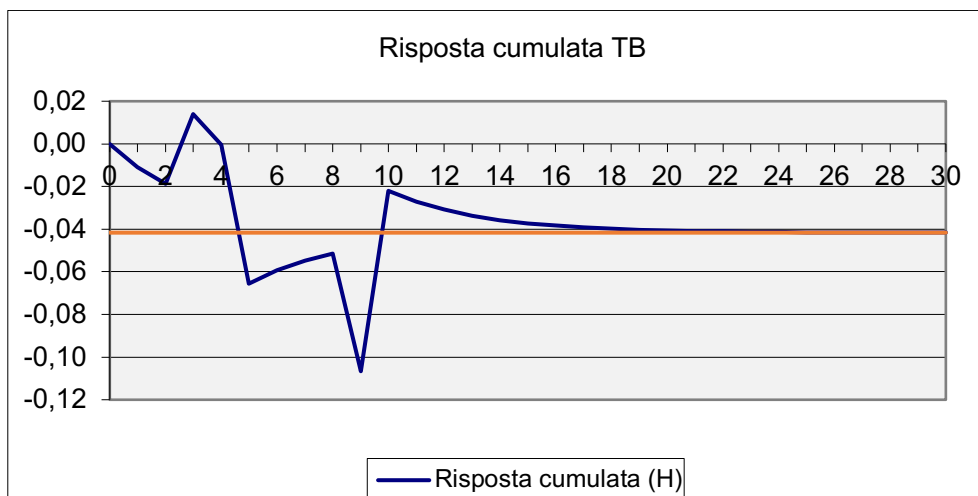
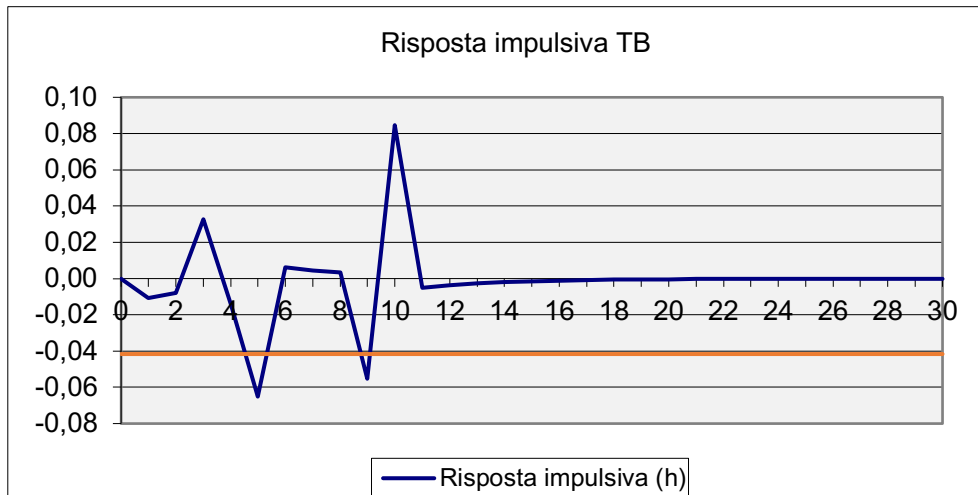
ADF test per la variabile CHUS senza trend e intercetta

```
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.153459 -0.018339  0.000950  0.009163  0.208775
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -0.0005280  0.0004482  -1.178    0.24
## z.diff.lag  0.4835617  0.0628822   7.690 6.86e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04383 on 197 degrees of freedom
## Multiple R-squared:  0.2466, Adjusted R-squared:  0.239
## F-statistic: 32.25 on 2 and 197 DF,  p-value: 7.668e-13
```

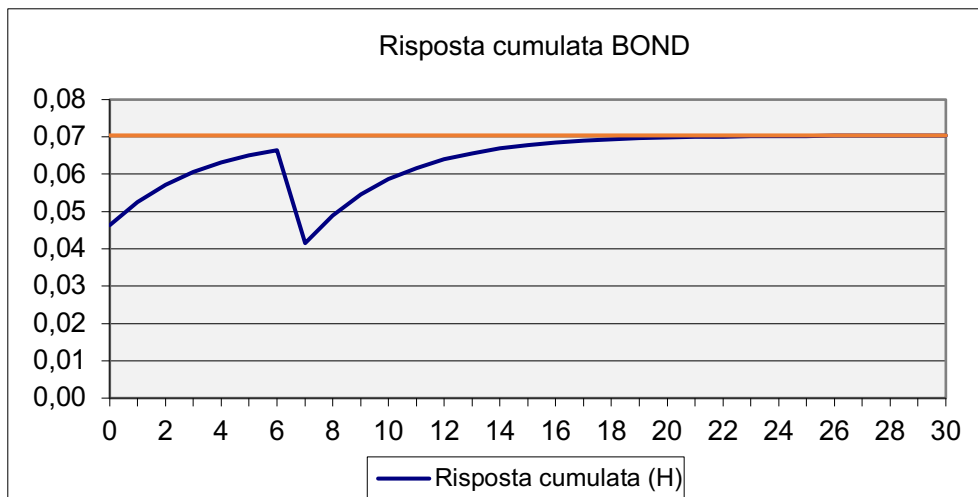
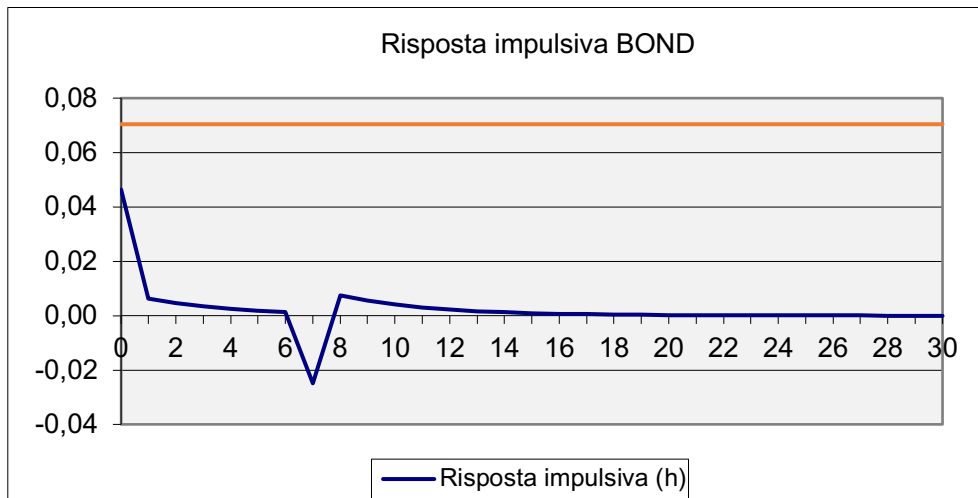
```
##  
##  
## Value of test-statistic is: -1.1782  
##  
## Critical values for test statistics:  
##      1pct  5pct 10pct  
## tau1 -2.58 -1.95 -1.62
```

In quanto all'equazione statica e dinamica si riportano per completezza tutti i grafici di risposta impulsiva delle variabili:

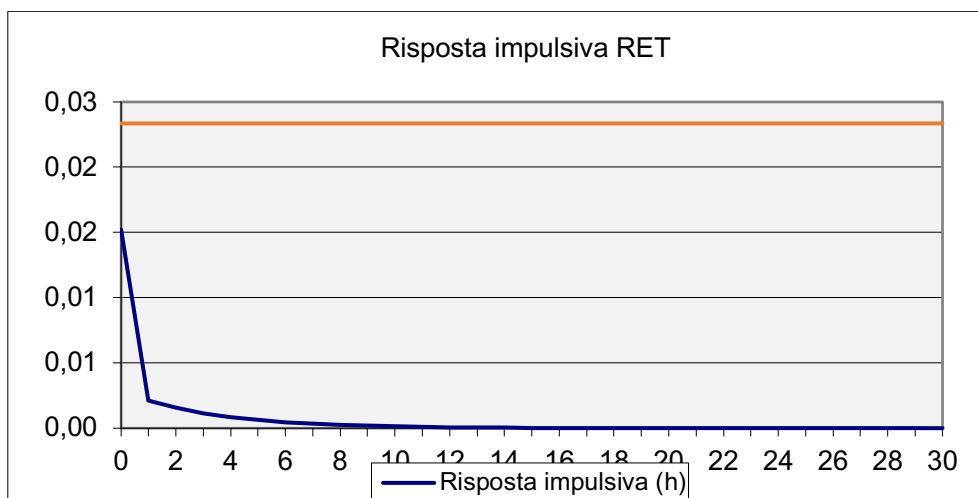
Funzione di risposta impulsiva e cumulata per la variabile "TB"

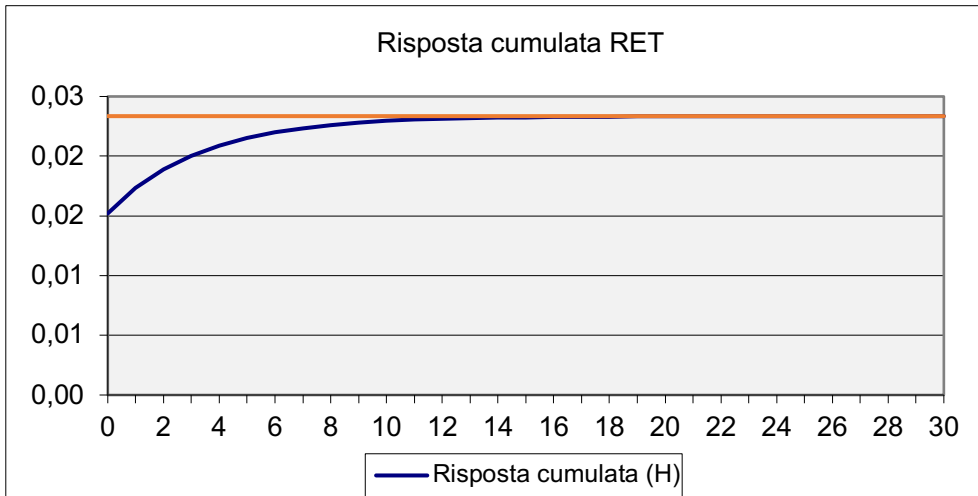


Funzione di risposta impulsiva e cumulata per la variabile "BOND"

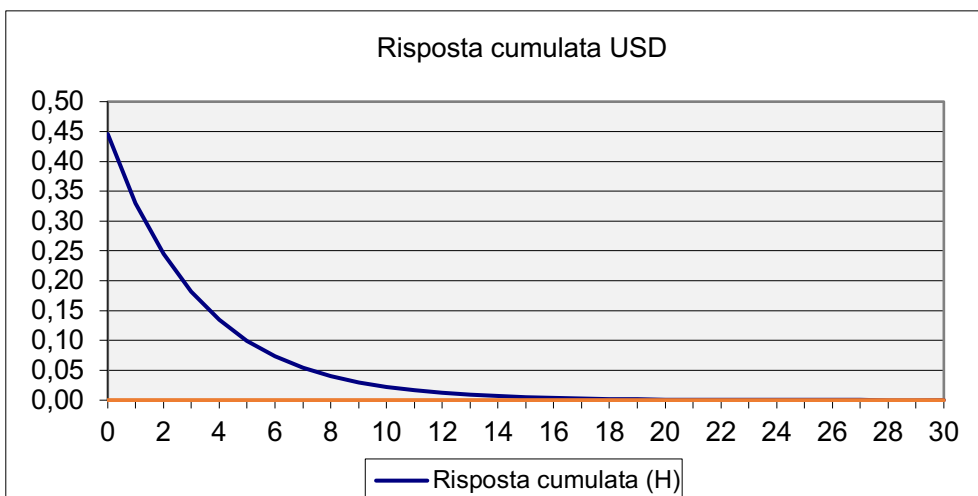
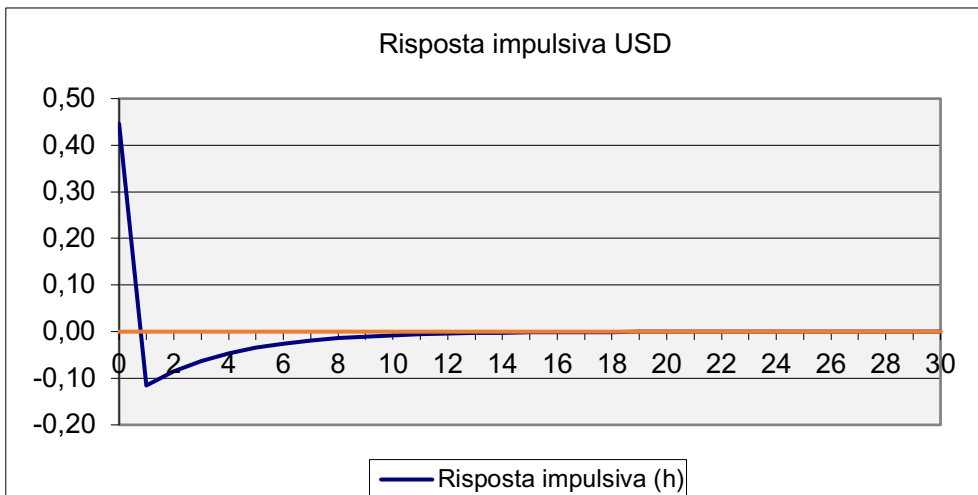


Funzione di risposta impulsiva e cumulata per la variabile "RET"

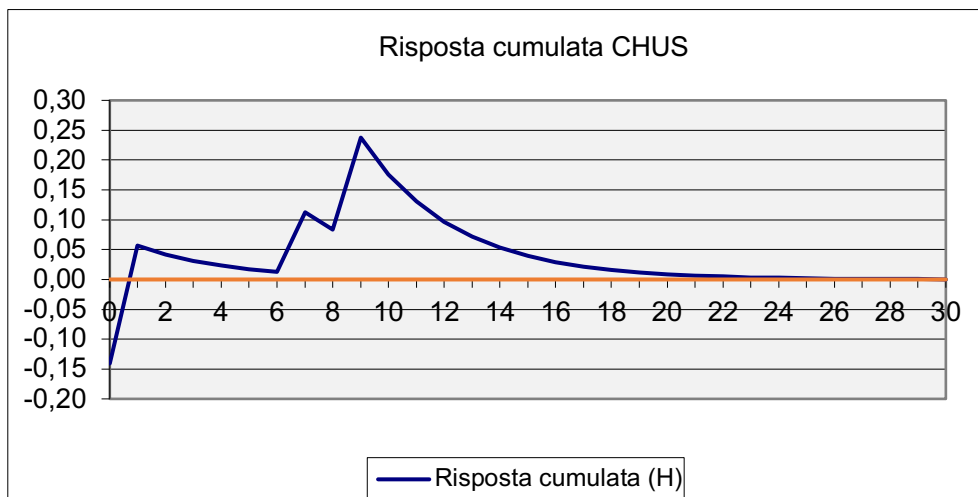
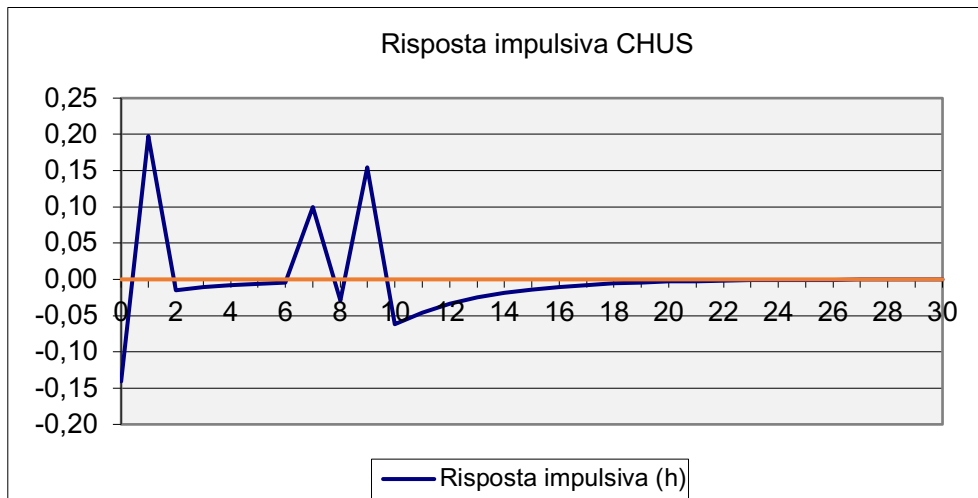




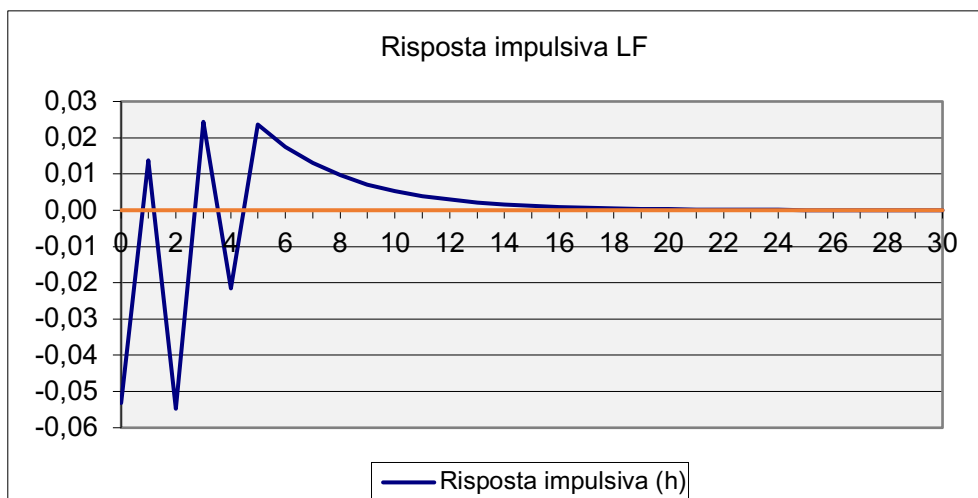
Funzione di risposta impulsiva e cumulata per la variabile "USD"

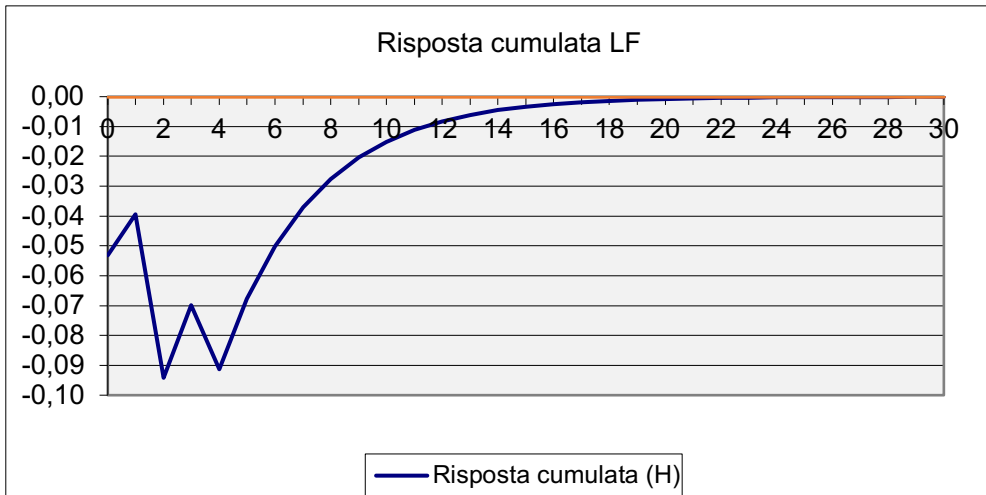


Funzione di risposta impulsiva e cumulata per la variabile "CHUS"



Funzione di risposta impulsiva e cumulata per la variabile "LF"





Bibliografia

Bollerslev T. P., (1986), *Generalized Autoregressive Conditional Heteroskedasticity*, Journal of Econometrics 31.

Bollerslev T., Chou R.Y., Kroner K.F., (1992), *ARCH Modeling in Finance. A Review of the Theory and Empirical Evidence*, Journal of Econometrics 52, 5–59.

Cavaliere G., Taylor A. M. R., (2005) *Stationarity tests under time-varying second moments Econometric Theory* 21, 1112 – 1129.

Engle R., (2001), *GARCH 101: An Introduction to the Use of ARCH/GARCH models in Applied Econometrics*, NYU Working Paper.

Engle R., Granger C. W. J., (1987), *Co-integration and error correction: Representation, estimation and testing*, Econometrica; 55.

Fama E. F., (1995) *Random Walks in Stock Market Prices*, Financial Analysts Journal, 51:1, 75-80.

Hamilton J. D., (1995), *Econometria delle serie storiche*, Monduzzi Editore.

Jarque C. M., Bera A. K., (1987), *A Test for Normality of Observations and Regression Residuals*. *International Statistical Review*, *Revue Internationale de Statistique*, 55(2), 163-172.

Johnston J., (2001), *Econometria*, FrancoAngeli, Milano.

Stock J., (1987), *Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors*. *Econometric*, 55, 1035-1056.

Terasvirta T., (2009), *An Introduction to univariate GARCH models*, Springer.

Verbeek M., (2008), *A Guide to Modern Econometrics*, John Wiley & Sons.

Sitografia

<https://www.borsaitaliana.it>

<https://www.bloomberg.com>

<https://www.cnbc.com>

<https://fred.stlouisfed.org>

<https://it.finance.yahoo.com/>

<https://scholar.google.com/>

<https://www.ilsole24ore.com/>

Ringraziamenti

A margine di questa tesi ho pensato diverse volte a quante analogie ci potrebbero essere tra l'esperienza formativa che ho affrontato e questo progetto. Se da una parte è vero che questi anni sono frutto di un percorso che ho realizzato in prima persona, dall'altra ho avuto alcune variabili "esogene" che vi hanno contribuito.

A tal proposito desidero ringraziare per il contributo più recente, il mio relatore, il professor Domenico Sartore, che oltre ad avermi sostenuto nelle incertezze, mi ha dato fiducia e mi ha insegnato una materia che continuerò ad approfondire.

Ringrazio Emma, che è sempre stata al mio fianco, facendomi ritrovare ogni qualvolta mi sentivo perso.

Grazie a mio papà, Valter, che ha sempre riposto in me quella stima e quella forza per proseguire.

Grazie a Martina, che ha fatto più di quanto doveva, sostenendomi e comprendendomi in tutti questi anni.

Grazie a mio nonno, Livio, un esempio in quanto a caparbietà e determinazione.

Grazie a tutti i miei amici e in particolare a Marco Moro, Alessandra, Federico, Fabio, Giuseppe e Jack per avermi accompagnato ed arricchito.