



UNIVERSITÀ DEGLI STUDI DI PADOVA

**DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI
“M.FANNO”**

CORSO DI LAUREA IN ECONOMIA

PROVA FINALE

**“ANALISI DELLA SERIE STORICA DEL PREZZO DEL TITOLO
UNICREDIT”**

RELATORE:

CH.MO PROF. NUNZIO CAPPUCCIO

LAUREANDO: MATTEO ALTOE’

MATRICOLA N. 1160855

ANNO ACCADEMICO 2019 – 2020

Indice

Introduzione.....	3
Capitolo 1: Il gruppo Unicredit.....	4
Capitolo 2: Concetti teorici	7
2.1 Introduzione alle serie storiche	7
2.2 Funzione di autocorrelazione e correlogramma.....	8
2.3 Test di Radice Unitaria	9
2.4 Modello ARMA	12
2.5 Modello GARCH	14
Capitolo 3: Analisi delle variabili	16
3.1 Analisi della variabile prezzi	16
3.2 Analisi della variabile rendimenti	22
Conclusioni.....	33

Numero parole: 7542

Introduzione

In questo elaborato ci troveremo ad affrontare l'analisi del prezzo del titolo di Unicredit, con l'obiettivo di generare un modello statistico in grado di spiegare l'andamento azionario del gruppo.

Nel primo capitolo verrà presentato il Gruppo Unicredit, con particolare riferimento al modello di business, alla composizione dell'azionariato e alla struttura organizzativa.

Il secondo capitolo tratterà i concetti teorici relativi all'ambito delle serie temporali. In particolare, il paragrafo 2.1 darà un'introduzione alla teoria econometrica che ha portato a sviluppare test e modelli impiegati nell'elaborato. Il paragrafo 2.2 descriverà un concetto fondamentale per il campo illustrato, ossia quello di stazionarietà del processo stocastico e presenterà le funzioni di autocorrelazione (ACF) e di autocorrelazione parziale (PACF). Nel paragrafo successivo verrà introdotto un gruppo di test econometrici (i test di radice unitaria) che studia la stazionarietà della serie storica; in particolare sono stati approfonditi il test DF, una sua estensione, ossia il test DF-GLS e il test Zivot-Andrews per ricercare eventuali break strutturali nell'andamento dei valori. I paragrafi 2.4 e 2.5 esporranno rispettivamente il modello ARMA, che spiega il processo stocastico di una serie storica, e il modello GARCH, che descrive la volatilità dei valori.

Nel terzo e ultimo capitolo verrà esposta l'elaborazione dei dati relativi alla serie storica presa in analisi. Essa verrà effettuata applicando ai dati i test e i modelli presentati nel capitolo precedente. Il capitolo 3 è formato da due paragrafi, divisi in funzione della variabile approfondita. Il primo studierà la variabile Prezzi e il secondo, invece, si occuperà di analizzare la variabile Rendimenti. Tutte le elaborazioni e i grafici pertinenti effettuati nel corso dell'analisi saranno trattati attraverso l'impiego del software R.

L'elaborato terminerà con una breve conclusione, che riassumerà i risultati dello studio effettuato e le relative implicazioni.

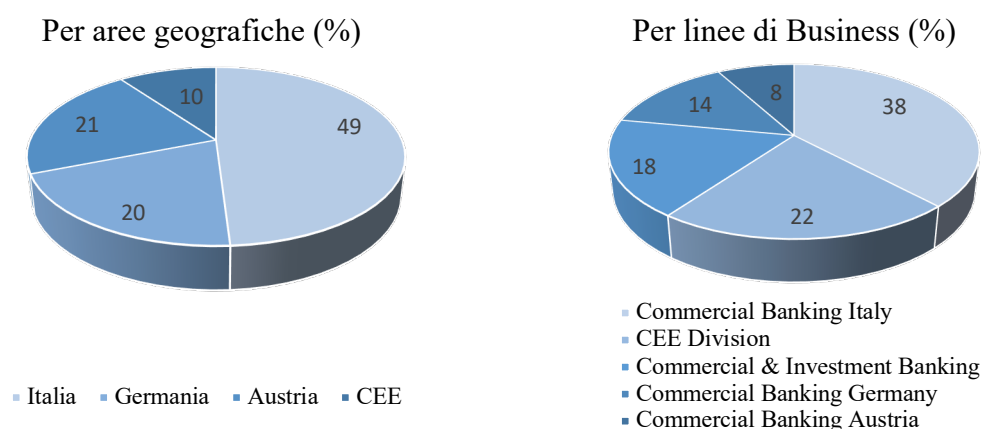
Capitolo 1: Il gruppo Unicredit

Unicredit S.p.a. è uno dei principali gruppi finanziari italiani ed europei (secondo in Italia e settimo in Europa per capitalizzazione di mercato), la sua sede principale è situata a Torre Unicredit a Milano, vanta 16 milioni di clienti in Europa e circa 84.000 dipendenti.

Essa è quotata alla Borsa di Milano ed è contenuta nel FTSE MIB, l'indice azionario che comprende le 40 società italiane con la maggiore capitalizzazione di mercato.

Unicredit Group nasce nel 1998, inizialmente con il nome di Unicredito Italiano, dalla fusione tra i gruppi Credito Italiano (formato a sua volta da Credito Italiano e Rolo Banca 1473) e Unicredito (composto da Cassa di Risparmio di Verona, Vicenza, Belluno e Ancona, Cassa di Risparmio di Torino e Cassamarca).

Immagine 1.1¹ Ricavi del gruppo Unicredit



Il principale mercato di Unicredit è quello italiano, che da solo genera il 49% del fatturato; in Italia è elevata anche la presenza nel territorio, con oltre 4000 filiali che consentono di essere vicini ai clienti e alle comunità locali. Per rafforzare questi legami sono inoltre state costituite sette Aree Territoriali, i cui responsabili hanno la funzione di dialogo con le istituzioni locali. È significativa anche la presenza internazionale: Unicredit è infatti presente con banche del proprio gruppo e uffici di rappresentanza in 38 nazioni in tutto il mondo.

La banca è dotata di un capitale sociale di 21.059.536.950,48 € diviso in 2.237.261.803 azioni ordinarie. La maggiore azionista è la società d'investimento statunitense BlackRock, che, con 113.550.196 azioni, detiene il 5,075% del capitale, seguita da Capital Research and Management Company, proprietaria del 5,022% del capitale di Unicredit.

¹ Dati aggiornati all'8 giugno 2020 reperibili nel sito <https://www.unicreditgroup.eu/it>

Tabella 1.1²: Principali Azionisti del Gruppo Unicredit

Maggiori azionisti	Azioni ordinarie	% di possesso
1. Gruppo BlackRock	113.550.196	5,075
2. Capital Research and Management Company	112.363.870	5,022 ⁽²⁾
- di cui per conto di EuroPacific Growth Fund	78.373.584	3,503
3. Norges Bank	67.366.057	3,011
4. Delfin S.a.r.l.	43.056.324	1,925
5. Fondazione Cassa di Risparmio di Ve-Vi-Bi e An	40.097.626	1,792
6. Fondazione Cassa di Risparmio di Torino	36.757.449	1,643
7. Gruppo Allianz SE	25.273.986	1,130

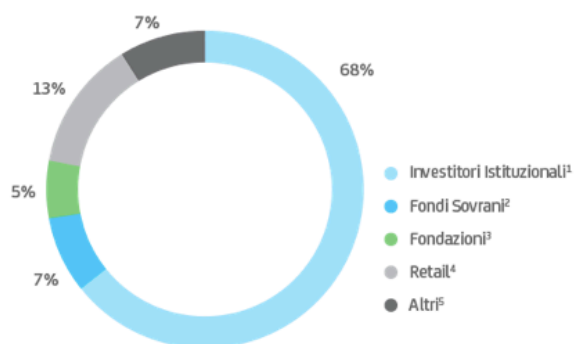
Il 68% delle azioni sono in mano a investitori istituzionali, provenienti per il 27% dagli Stati Uniti, per 12% dal Regno Unito, per il 9% da stati europei, per il 2% dall'Italia e per il restante 17% da altri stati.

Le fondazioni bancarie, che in passato figuravano tra i principali azionisti, anche in seguito ad aumenti di capitale non sottoscritti integralmente, ora controllano il 5% del capitale; tra queste le principali sono la Fondazione Cassa di Risparmio di Venezia, Vicenza, Belluno e Ancona (1,792%) e la Fondazione Cassa di Risparmio di Torino (1,643%).

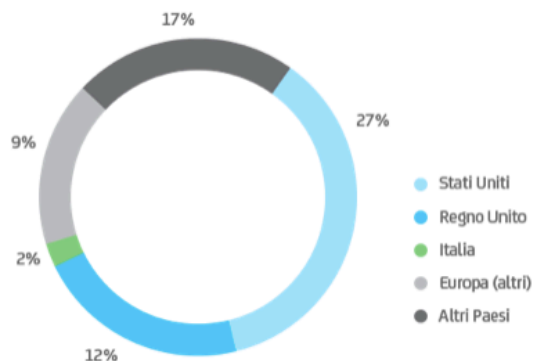
Il 7% del capitale è di proprietà di fondi sovrani mentre il 13% appartiene ai clienti retail; l'ultimo 7% appartiene ad altre tipologie di proprietari.

Immagine 1.2³: Composizione dell'Azionariato

Composizione per tipologia di azionisti



Composizione per geografia degli investitori istituzionali



² Dati aggiornati all'8 giugno 2020; fonte: <https://www.unicreditgroup.eu/it>

³ Dati aggiornati all'8 giugno 2020 reperibili nel sito <https://www.unicreditgroup.eu/it>

La struttura del gruppo Unicredit è formata da tre divisioni:

- Western Europe Commercial Banking
- Central & Eastern Europe Commercial Banking
- Corporate & Investment Banking

La prima branca è costituita dalle banche commerciali presenti in Austria, Germania e Italia e dal Wealth Management Group. Le prime hanno la funzione di fornire a clienti privati prodotti bancari (servizi di pagamento, mutui, credito al consumo e risparmio gestito) e di supportare la clientela corporate locale. Il secondo si occupa di gestione della ricchezza, fornendo a imprenditori e privati la possibilità di proteggere e accrescere il proprio patrimonio tramite servizi di finanziamento, di investimento e di pianificazione patrimoniale.

Attraverso la divisione Central & Eastern Europe Commercial Banking, il gruppo opera in dieci Paesi dell'Europa centrale e orientale: Bosnia ed Erzegovina, Bulgaria, Croazia, Repubblica Ceca, Romania, Russia, Serbia, Slovacchia, Slovenia e Ungheria; opera inoltre anche nel campo del Leasing nelle Repubbliche Baltiche.

Il ramo Corporate & Investment Banking ha un ruolo di relazione con i clienti multinazionali ("Multinational"), con selezionati clienti "Large Corporate", caratterizzati da una forte domanda potenziale di prodotti di investment banking e con le istituzioni finanziarie. Questo ramo offre i propri servizi a 1500 società multinazionali e istituzioni finanziarie e a oltre 600.000 clienti corporate e pubblici.

Capitolo 2: Concetti teorici

2.1 Introduzione alle serie storiche

L'andamento del prezzo di un'azione può essere studiato attraverso la sua serie storica.

Secondo Kirchgässner e Wolters (2007) una serie storica è definita come un insieme di osservazioni quantitative disposte in ordine cronologico, solitamente misurate ad intervalli equidistanti. La variabile tempo in genere viene considerata come una variabile discreta.

In altre parole, si parla di serie storiche quando si considera un fenomeno in relazione alla sua evoluzione nel corso del tempo.

L'approccio classico all'analisi delle serie storiche è basato sul concetto che quest'ultime siano formate da un numero finito di componenti tra loro indipendenti e non direttamente osservabili. In particolare, questo metodo è stato sviluppato da Warren M. Pearson (1919) che individuava quattro componenti differenti. Esse sono:

1. Il Trend, che misura la componente tendenziale di lungo periodo, dovuto a elementi come lo sviluppo tecnologico.
2. Il Business Cycle, che è determinato dagli scostamenti rispetto al trend di durata superiore ad un anno, causati da fattori collegati al ciclo economico quali la crescita e la recessione.
3. Il Seasonal Cycle, che contiene le oscillazioni di durata inferiore all'anno, causate da fattori che hanno a che fare con la stagionalità, ad esempio il clima.
4. Il Residuo, il componente che contiene tutti i movimenti che non sono causati né dal trend, né dal ciclo economico, né dal ciclo stagionale.

L'approccio classico all'analisi di serie storiche si basa sull'assunto che i primi tre componenti siano funzioni deterministiche del tempo; solamente i residui vengono considerati una variabile stocastica, in particolare questi ultimi formano un processo stocastico white noise, ovvero una serie di variabili casuali indipendenti con media zero e varianza costante.

A partire dagli anni '70, il metodo classico è stato abbandonato e contemporaneamente è stato sviluppato un nuovo approccio che abbandona l'idea di differenti componenti, sia deterministici che stocastici, e assume che l'intero processo di generazione della serie storica sia basato su un processo di natura stocastica. Particolarmente rilevanti in questo ambito sono stati gli studi di George E. P. Box e Gwilym M. Jenkins (1970).

Recentemente è stata nuovamente presa in considerazione la possibilità di scomporre la serie storica in più componenti, tuttavia, a differenza dell'approccio classico, essi sono rappresentati

da processi stocastici, descritti da Parzen E. (1962) come fenomeni casuali che derivano da un processo che si sviluppa nel tempo attraverso una legge probabilistica.

2.2 Funzione di autocorrelazione e correlogramma

Per poter effettuare un'analisi su una serie storica essa deve essere basata su un processo stocastico stazionario, cioè un processo i cui i momenti sono costanti nel tempo, questo perché un processo non stazionario porterebbe a previsioni con risultati distorti. Ci sono due differenti definizioni di stazionarietà.

Facendo riferimento al concetto di stazionarietà in senso stretto o forte, la distribuzione di un determinato numero di variabili casuali del processo stocastico è la stessa in caso della loro traslazione lungo l'asse temporale. Questa situazione è difficile da raggiungere nella pratica, perciò, usualmente si impiega il concetto di stazionarietà in senso largo o debole, che richiede che i momenti primi e secondi non varino al variare del tempo.

È quindi richiesto che:

- $E[x_t] = \mu$ sia costante per ogni valore di t
- $V[x_t] = E[(x_t - \mu_t)^2] = \sigma^2$ sia finito e costante per ogni t
- $Cov[x_t, x_s] = E[(x_t - \mu_t)(x_s - \mu_s)] = \gamma(|s - t|)$ dipenda solamente dalla distanza $|s - t|$ e non dai due istanti t ed s .

Per studiare la stazionarietà della serie storica è possibile realizzare un'analisi dell'autocorrelazione, che indica in che misura i valori della serie storica sono stati determinati dai valori ritardati. La funzione di autocorrelazione (ACF) può essere definita come:

$$\rho_\tau = \frac{E(x_\tau - \mu)(x_{t+\tau} - \mu)}{E[(x_\tau - \mu)^2]}$$

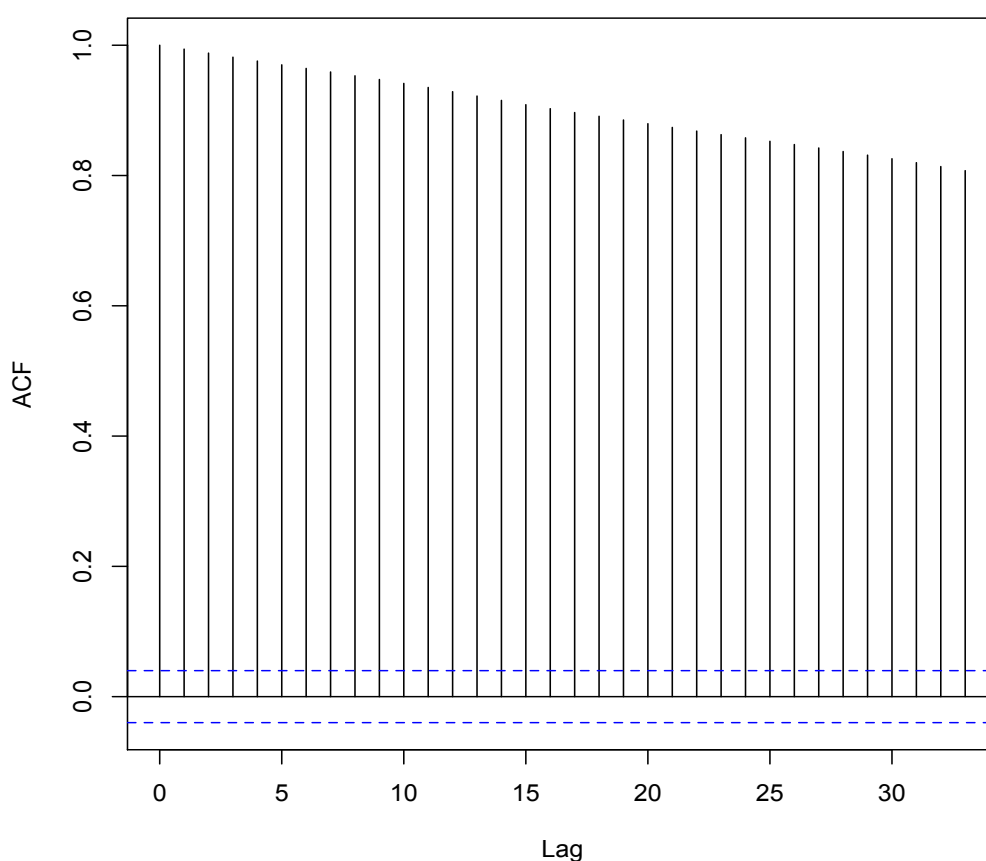
Dove $E(x_\tau - \mu)(x_{t+\tau} - \mu)$ è la covarianza tra i termini x_τ e $x_{t+\tau}$ e $E[(x_\tau - \mu)^2]$ è la varianza.

Un ulteriore strumento di analisi è dato dalla funzione di autocorrelazione parziale (PACF). Essa è definita da Kirchgässner e Wolters come “la correlazione che rimane se il possibile impatto di tutte le altre variabili casuali è stato eliminato”. Questa funzione viene utilizzata per ovviare al fatto che le funzioni di autocorrelazione di processi autoregressivi di ordine finito

convergono sempre a 0 senza però interrompersi improvvisamente, ciò che rende difficile discriminare processi di ordine simile.

Per studiare queste funzioni si fa uso del correlogramma, il grafico che rappresenta il valore, compreso tra -1 e 1, di ρ , il coefficiente di correlazione, in relazione alla lunghezza del ritardo. Si può dire che l'autocorrelazione tra un termine al tempo t e un suo termine ritardato è significativamente diversa da 0 quando il valore di ρ è superiore in valore assoluto agli estremi dell'intervallo di confidenza.

Immagine 2.2.1: Esempio di Correlogramma



Nella rappresentazione presa come esempio si può notare una forte presenza di autocorrelazione positiva e di conseguenza il processo stocastico non sarà stazionario. Graficamente i coefficienti di correlazione, tutti positivi e lentamente decrescenti all'aumentare del ritardo, superano le linee blu tratteggiate, le quali rappresentano gli estremi dell'intervallo di confidenza.

2.3 Test di Radice Unitaria

I test di Radice Unitaria sono un gruppo di test econometrici che hanno in comune l'obiettivo di analizzare la stazionarietà della serie storica. Essi sono così chiamati perché sono basati sulla ricerca di una radice unitaria, che renderebbe non stazionario il processo stocastico.

Per trasformare un processo non stazionario in uno stazionario è necessario rimuovere il trend. Questo è possibile attraverso il processo di differenziazione, utilizzando cioè la differenza tra le osservazioni.

Il primo test di questo gruppo è stato sviluppato da David A. Dickey e Wayne A. Fuller (1979), dai quali prende il nome di test DF. Il test verifica la presenza di un trend stocastico nella serie storica e si basa su un modello autoregressivo di ordine 1:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$$

Il test si basa sul seguente sistema di ipotesi unilaterale:

$$\begin{cases} H_0 : \beta_1 = 1 \\ H_1 : \beta_1 < 1 \end{cases}$$

È altresì possibile modificare il modello AR(1) sottraendo da entrambi i lati Y_{t-1} , ottenendo così:

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + \varepsilon_t$$

Con $\delta = \beta_1 - 1$ e $\Delta Y_t = Y_t - Y_{t-1}$

In questo caso il test d'ipotesi sarà:

$$\begin{cases} H_0 : \delta = 0 \\ H_1 : \delta < 0 \end{cases}$$

Se si accetta l'ipotesi nulla H_0 nel modello autoregressivo è presente radice unitaria, mentre se l'ipotesi nulla viene rifiutata il processo è stazionario.

Il test DF si applica solamente nel caso di autoregressione del primo ordine, quindi non è in grado di inglobare la correlazione di ordine superiore del modello. (Stock e Watson, 2016)

Per ovviare a questo problema si fa uso del test di Dickey-Fuller Aumentato (ADF), sviluppato grazie agli studi di Said e Dickey del 1984. Il test si basa sul seguente processo autoregressivo di ordine p (>1):

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + \sum_{i=1}^p \gamma_i \Delta Y_{t-i} + \varepsilon_t$$

L'ipotesi nulla H_0 è, come nel test precedente, $\delta = 0$ ed indica la presenza di radice unitaria, mentre l'ipotesi alternativa $H_1 : \delta < 0$ evidenzia la stazionarietà del processo.

I modelli finora presentati considerano che l'ipotesi alternativa sia una serie stazionaria, tuttavia è anche possibile che la serie storica segua un trend deterministico, perciò si può implementare ai modelli sopracitati la presenza di una funzione deterministica del tempo. Nei modelli con trend deterministico l'ipotesi alternativa è che la serie sia stazionaria lungo un trend deterministico. La statistica per la verifica dell'ipotesi nulla dei due test sopracitati è una statistica t ed è chiamata statistica di Dickey-Fuller. Essa ha una distribuzione asimmetrica verso sinistra e non tende alla distribuzione normale nemmeno per grandi campioni. Per questo motivo è necessario utilizzare i valori critici individuati da MacKinnon nei suoi studi del 1990.

Tabella 2.3.1: Valori critici della statistica DF

Tipo di regressore deterministico	Senza costante			Con costante			Con costante e trend		
	90%	95%	99%	90%	95%	99%	90%	95%	99%
Valore critico	-1.62	-1.94	-2.57	-2.57	-2.86	-3.43	-3.13	-3.41	-3.96

Il test ADF-GLS, sviluppato da Elliot, Graham e Rothenberg nel 1996, è la versione efficiente del test ADF. Si tratta di un test più potente del test ADF normale, quindi rifiuta più spesso l'ipotesi nulla di presenza di radice unitaria quando è vera l'ipotesi alternativa; in altre parole è in grado di discriminare con più precisione una radice unitaria e una radice vicina ad 1.

I parametri deterministici (intercetta e trend) sono stimati tramite una regressione Generalized Least Squares (GLS) per poi essere sottratti alla regressione iniziale e alla regressione detrendizzata viene poi effettuato il test ADF. Se tra i regressori deterministici è presente solo l'intercetta, la distribuzione della statistica del test ADF-GLS converge asintoticamente alla distribuzione della statistica DF, di conseguenza si possono utilizzare gli stessi valori critici. Se invece è presente anche un trend deterministico i valori critici sono rispettivamente: -2,57 per $\alpha = 10\%$, -2,89 per $\alpha = 5\%$ e -3,48 per $\alpha = 1\%$.

L'ultimo test di radice unitaria preso in considerazione è il test Zivot-Andrews, che ricerca la presenza di break strutturali, cioè variazioni inattese dei parametri della regressione nel corso del tempo, nella serie storica (Zivot e Andrews, 1992). Il test, per ricercare i break, fa uso delle variabili dummy, ovvero una particolare variabile che assume valore 1 se una determinata condizione è realizzata e 0 altrimenti. L'ipotesi nulla del test è la presenza di un processo non

stazionario, mentre l'ipotesi alternativa è la presenza di un processo stazionario lungo un trend deterministico con un break strutturale. La data in cui avviene il break è stimata dal test ed è quella in cui il test ha valore minimo, cioè nel punto in cui esso è più distante dall'accettare H_0 . Il break può riguardare una variazione nella pendenza del trend, una traslazione della serie o un cambiamento combinato nella pendenza e nel livello della serie storica, a seconda della variante del modello che viene impiegata.

2.4 Modello ARMA

Il principale obiettivo dell'analisi delle serie storiche è identificare i processi stocastici che le hanno realizzate. Per ottenere ciò è necessario fare uso di opportuni modelli statistici in grado di rappresentare adeguatamente la realtà.

Nel campo dell'analisi delle serie storiche, un modello a cui si fa riferimento è quello autoregressivo a media mobile, chiamato anche ARMA. Esso è basato su un processo misto, formato sia da termini autoregressivi (AR) che da termini a media mobile (MA).

Il modello autoregressivo si basa su un processo in cui la variabile dipende dai suoi valori precedenti. Esso analizza quindi la variabile dipendente effettuando una regressione sui suoi valori ritardati.

Un processo autoregressivo del primo ordine, chiamato anche processo AR(1), può essere definito come:

$$x_t = \delta + \alpha x_{t-1} + u_t$$

dove u_t è un processo White Noise con media 0 e varianza σ^2 .

La formula generale per descrivere un modello autoregressivo di ordine p è

$$x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + u_t$$

dove $\alpha_p \neq 0$ e u_t è un processo WN $(0, \sigma^2)$.

Questa definizione descrive un processo in cui la variabile dipendente x_t viene regredita su p ritardi.

Il modello a media mobile (MA) genera una serie storica tramite una regressione sui termini di errore ritardati u , in particolare, il processo stocastico viene prodotto dalla somma di q termini di errore u .

Il modello MA(1) è dato dalla seguente equazione:

$$x_t = \mu + u_t - \beta u_{t-1}$$

con u_t uguale a un processo WN $(0, \sigma^2)$. Dall'equazione si può vedere che il valore di x_t è calcolato utilizzando solamente i valori dell'errore al tempo t e al tempo $t-1$.

Come per il modello AR, si può utilizzare anche un numero maggiore di ritardi. Di seguito viene presentata la formula generale di un processo a media mobile di ordine q .

$$x_t = \mu + u_t - \beta_1 u_{t-1} - \beta_2 u_{t-2} - \dots - \beta_q u_{t-q}$$

Come accennato precedentemente il modello ARMA è prodotto da un'integrazione dei modelli AR e MA, di conseguenza, la variabile dipendente viene calcolata utilizzando sia i suoi valori ritardati, sia i valori ritardati e non del termine di errore. Matematicamente il un processo ARMA(p, q) è definito come

$$x_t = \delta + \alpha x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + u_t - \beta u_{t-1} - \beta_2 u_{t-2} - \dots - \beta_q u_{t-q}$$

Si può notare che il modello ARMA possa essere considerato come una particolare estensione del modello AR, dove il termine di errore viene considerato, attraverso il processo MA, come serialmente correlato.

Per effettuare lo studio di un processo stocastico attraverso il modello ARMA, è necessario che il processo sia stazionario e che i residui siano distribuiti come White Noise $(0, \sigma^2)$.

Qualora ci fosse evidenza di non stazionarietà, è possibile avvalersi del modello ARIMA (Autoregressive Integrated Moving Average), che impiega la differenziazione per ovviare al problema. Nello specifico, il modello ARIMA (p, d, q) può essere considerato come un modello ARMA con componente autoregressiva di ordine p e termine a media mobile di ordine q applicati alle differenze di ordine d della serie di valori.

Per studiare la distribuzione dei residui, si utilizza il test econometrico sviluppato grazie agli studi di Box G. e Ljung G. del 1978. Il test Box-Ljung verifica la presenza di autocorrelazione in una serie di dati attraverso la seguente statistica test, distribuita come una χ_m^2

$$Q = n(n+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{n-k}$$

dove n è il numero di osservazioni, $\hat{\rho}_k$ è la funzione di autocorrelazione di ordine k e m è il numero di autocorrelazioni prese in considerazione.

Il sistema di ipotesi del test è:

$$\begin{cases} H_0: \rho_1 = \rho_2 = \dots = \rho_m = 0 \\ H_1: \rho_i \neq 0 \text{ per almeno un } i \end{cases}$$

La condizione richiesta per impiegare il modello ARMA è che i residui siano distribuiti come White Noise, quindi è necessario che il test Box-Ljung accetti l'ipotesi nulla di assenza di autocorrelazione.

Per ottenere una stima adeguata del processo che genera la serie storica è importante valutare quale sia il numero corretto di ritardi da inserire nella regressione. Lo scopo è rilevare quale sia il miglior trade-off tra bontà di adattamento del modello e parsimonia nell'uso di ritardi. Per fare questo si possono impiegare differenti criteri d'informazione. Nel presente elaborato è stato fatto uso del criterio AIC.

Il criterio di informazione di Akaike (o AIC), sviluppato nel 1974, è:

$$AIC = 2k - 2\ln(L)$$

dove k è il numero di parametri nel modello statistico e L è il valore massimizzato della funzione di verosimiglianza del modello stimato.

2.5 Modello GARCH

Il secondo modello presentato nell'elaborato è il modello GARCH, proposto da Bollerslev nel 1986 con l'obiettivo di spiegare la volatilità di una serie storica tramite il concetto di varianza condizionale definita come

$$h_t^2 = V[\varepsilon_t | I_{t-1}] = E[\varepsilon_t^2 | I_{t-1}]$$

dove ε_t è il residuo al tempo t e I_{t-1} è il set di informazioni disponibile al tempo $t-1$.

Il primo a utilizzare l'assunzione che la varianza condizionale dei residui dipenda linearmente dai residui quadrati ritardati è stato Robert Engle (1982), sviluppando così il modello ARCH. L'assunto fondamentale di questo modello è la distinzione tra le proprietà di varianza incondizionale e varianza condizionale. La prima viene considerata come costante, mentre la seconda varia in funzione dei valori che assumono i residui ritardati.

Questo fenomeno porta alla generazione di "volatility cluster", ossia periodi di tempo in cui i valori assoluti dei rendimenti sono simili tra loro, come osservato da Benoit Mandelbrot (1963).

Lo studioso ha infatti osservato che nei mercati finanziari grandi shock sono seguiti da grandi shock e piccoli shock sono seguiti da shock di scarsa entità.

Si può affermare perciò che, mentre i rendimenti r_t sono tra loro incorrelati, i loro valori assoluti $|r_t|$ e i quadrati dei rendimenti r_t^2 sono positivamente autocorrelati.

La presenza del fenomeno di volatility clustering fa sì che i rendimenti di un titolo siano rappresentati da una funzione di distribuzione leptocurtica, che, rispetto a una distribuzione Gaussiana, ha una maggiore concentrazione di valori al centro e nelle code.

I fatti stilizzati dimostrano anche che i rendimenti dei titoli finanziari assumano anche un comportamento asimmetrico: gli shock negativi aumentano in misura maggiore la volatilità rispetto agli shock positivi.

Il modello ARCH spiega l'autocorrelazione dei quadrati dei residui utilizzando un modello AR. In particolare, un modello autoregressivo a eteroschedasticità condizionata di ordine q è definito come:

$$h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

con $\alpha_0 > 0$; $\alpha_i \geq 0$ con $i = 1, \dots, q - 1$; $\alpha_q > 0$ per garantire la non negatività della varianza condizionata.

Come si può vedere dall'equazione presentata, la varianza condizionata è una funzione crescente dei quadrati dei residui. Maggiore è l'ordine di regressione q , maggiore è la dimensione dei cluster di volatilità. Un problema del modello ARCH è dato dalla necessità di impiegare un alto numero di ritardi, che ne rende poco agevole la stima.

Il modello GARCH trova una soluzione a questa complicazione stimando la relazione tra i valori dei residui al quadrato impiegando un modello ARMA. Di conseguenza, rispetto al modello ARCH, vengono aggiunti alla regressione p valori ritardati della varianza condizionale h_t^2 . Un processo GARCH(p, q), quindi, è rappresentato dalla seguente equazione:

$$h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}^2$$

Con:

$\alpha_0, \alpha_q, \beta_p > 0$; $\alpha_i \geq 0$ $i = 1, \dots, q - 1$; $\beta_j \geq 0$ $j = 1, \dots, p - 1$ per garantire la non negatività della varianza condizionata.

Generalmente, per dati relativi ai mercati finanziari, un processo GARCH(1,1) è sufficiente a stimare la volatilità.

Capitolo 3: Analisi delle variabili

3.1 Analisi della variabile prezzi

L'analisi della serie storica del prezzo di Unicredit viene svolta utilizzando un campione di dieci anni, dal 1° giugno 2010 al 29 maggio 2020. Le rilevazioni sono 2536 e sono relative ai prezzi di chiusura del titolo; di conseguenza, non sono chiaramente stati presi in considerazione i giorni di chiusura della Borsa.

Per effettuare l'analisi della variabile prezzo è stata applicata una trasformazione logaritmica, perciò quando si farà riferimento ai prezzi, si utilizzerà il logaritmo naturale dei prezzi. Questo passaggio è stato impiegato allo scopo di ridurre la volatilità della serie. Di seguito sono visibili le rappresentazioni della serie storica del prezzo e della rispettiva trasformazione logaritmica.

Immagine 3.1.1: Serie storica del prezzo di Unicredit

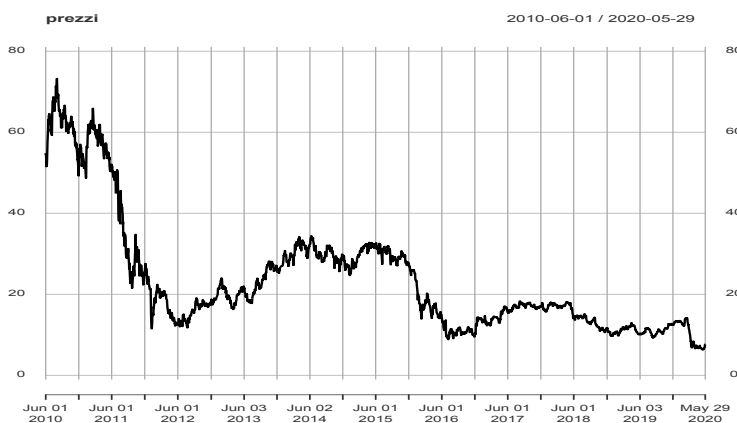
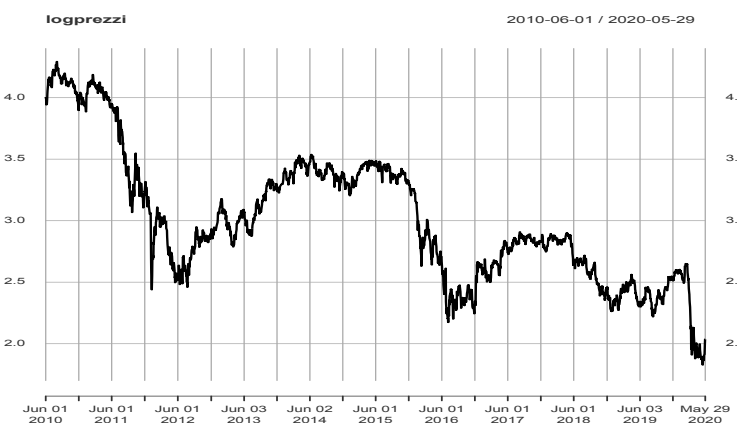


Immagine 3.1.2: Serie storica del logaritmo naturale di prezzo di Unicredit

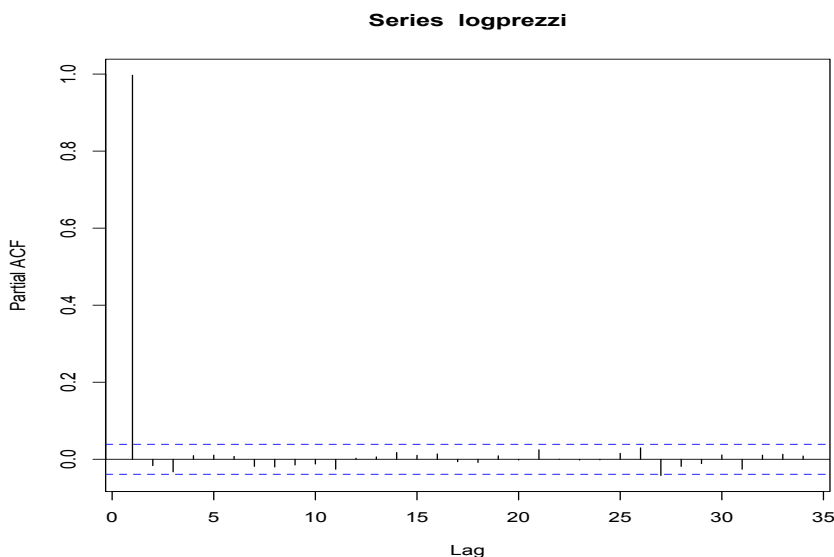


Il primo passaggio dell'analisi è rappresentato dallo studio delle funzioni di autocorrelazione (ACF) e di autocorrelazione parziale (PACF). La prima indica la correlazione tra una variabile e i suoi ritardi, mentre la seconda indica la correlazione rimanente quando si esclude l'influenza di tutte le altre variabili. In caso di presenza di autocorrelazione il processo stocastico non è stazionario e, per effettuare un'analisi, è necessario effettuare le opportune trasformazioni tramite un processo di differenziazione. Per realizzare una valutazione si utilizza il correlogramma, cioè il grafico che raffigura il coefficiente di correlazione in relazione ai differenti ritardi. Dopodiché, per una verifica più accurata, verranno svolti i test di radice unitaria, che indagano con più precisione la stazionarietà della serie.

Immagine 3.1.3: ACF dei prezzi



Immagine 3.1.4: PACF dei prezzi



Osservando l'immagine 3.1.3 si evince che, tra le variabili casuali del processo, è presente una forte correlazione positiva, che cala gradualmente all'aumentare della lunghezza del ritardo. Invece, se si guarda la figura successiva, si può vedere che solo per il ritardo di ordine 1 è presente una forte correlazione positiva. Per tutti gli altri ordini di ritardo ρ non è significativamente diverso da 0, ad eccezione del dato in corrispondenza del ritardo di ordine 27, dove si evidenzia una leggera correlazione negativa. Dai due correlogrammi sembra evidenziarsi la presenza di autocorrelazione nel processo stocastico, che renderebbe quest'ultimo non stazionario.

Per studiare con maggiore precisione la stazionarietà della serie storica, come precedentemente annunciato, sono stati effettuati i test di radice unitaria; in particolare sono stati utilizzati il test ADF, il test ADF-GLS e il test Zivot-Andrews.

Le prime due tipologie verificano la presenza di radice unitaria nel processo e si differenziano perché il test ADF-GLS è dotato di una maggiore potenza. Entrambi si basano sul sistema di ipotesi:

$$\begin{cases} H_0 : \delta = 0 \\ H_1 : \delta < 0 \end{cases}$$

Dove nell'ipotesi nulla c'è una radice unitaria nella serie e nell'ipotesi alternativa c'è stazionarietà nel processo.

Nel test Zivot-Andrews, invece, l'ipotesi alternativa è data dalla presenza di un processo stazionario lungo un trend deterministico con un break strutturale.

Tutti i test effettuati sono stati svolti analizzando due possibili varianti: la presenza di una intercetta diversa da 0 e la presenza di un trend deterministico e di un intercetta diversa da 0.

Per i test ADF e ADF-GLS il criterio impiegato per la scelta del numero di ritardi è il criterio AIC, la cui formula è

$$AIC = 2k - 2\ln(L)$$

Dove k è il numero di parametri nel modello statistico e L è il valore massimizzato della funzione di verosimiglianza del modello stimato.

Applicando il test ADF alla serie storica considerata nel presente elaborato si ottengono i seguenti risultati:

- Test ADF con regressore deterministico l'intercetta

```
> CADFtest(logprezzi,max.lag.y=20,criterion=c("AIC"),type=c("drift"))
```

```
ADF test

data: logprezzi
ADF(4) = -1.5844, p-value = 0.4905
alternative hypothesis: true delta is less than 0
sample estimates:
      delta
-0.001784093
```

- Test ADF con regressori deterministici l'intercetta e il trend

```
> CADFtest(logprezzi,max.lag.y=20,criterion=c("AIC"),type=c("trend"))
```

```
ADF test

data: logprezzi
ADF(4) = -2.2759, p-value = 0.4465
alternative hypothesis: true delta is less than 0
sample estimates:
      delta
-0.003939028
```

Nei test Augmented Dickey-Fuller effettuati sotto l'ipotesi di drift differente da 0 e sotto l'ipotesi di drift e trend deterministico, viene accettata l'ipotesi nulla di radice unitaria. Entrambe le varianti del test, infatti, confermano la presenza di radice unitaria. Questo può essere osservato sia dal valore della statistica DF, superiore ai valori presentati nella tabella 2.3.1 per tutti i livelli di significatività, sia dal p-value che in entrambi i test è maggiore di 0,05. Per verificare la stazionarietà della serie è stato svolto anche il test ADF-GLS, dotato di una potenza maggiore rispetto al test ADF, che stima i parametri deterministici attraverso il metodo Generalized Least Squares. I valori critici di questo test nella variante che presenta intercetta e trend come regressori deterministici sono differenti di quelli della statistica DF, di conseguenza sono stati utilizzati quelli prodotti dal software R. Da questo test sono attesi gli stessi risultati di quelli appena ottenuti, ossia l'accettazione di H_0 .

- Test ADF-GLS con regressore deterministico l'intercetta

Value of test-statistic is: 0.4383

Critical values of DF-GLS are:

	1pct	5pct	10pct
critical values	-2.57	-1.94	-1.62

- Test ADF-GLS con regressori deterministici l'intercetta e il trend

Value of test-statistic is: -2.0385

Critical values of DF-GLS are:

	1pct	5pct	10pct
critical values	-3.48	-2.89	-2.57

I risultati ottenuti confermano la non stazionarietà del processo stocastico, come si può vedere dal fatto che i valori della statistica test (0,4383 nel caso di presenza di deriva differente da 0 e -2,0385 nel caso di presenza di drift e trend deterministico) sono più grandi del valore critico dell'intervallo di confidenza, per qualunque livello di significatività.

L'ultimo test di radice unitaria effettuato è il test di Zivot-Andrews, svolto con l'obiettivo di verificare la presenza di un break strutturale nella serie. In questo caso il sistema di ipotesi ha come ipotesi nulla la presenza di un processo random walk e come ipotesi alternativa la presenza di un processo stazionario con un break strutturale. Il modello che considera come regressore deterministico solamente l'intercetta si basa sulla funzione

$$y_t = \mu^A + \theta^A DU_t(\lambda) + \beta_t^A + \alpha^A y_{t-1} + \sum_{j=1}^k c_j^C \Delta y_{t-j} + e_t$$

Invece per il modello che considera sia il drift che il trend si fa uso del modello

$$y_t = \mu^C + \theta^C DU_t(\lambda) + \beta_t^C + \gamma^C DT_t(\lambda) + \alpha^C y_{t-1} + \sum_{j=1}^k c_j^C \Delta y_{t-j} + e_t$$

In questo caso lo studio è stato svolto mediante l'analisi rolling, quindi è stato preso un campione di tre anni a partire dalla prima misurazione e traslato di anno in anno fino al 2020. Il test Zivot-Andrews identifica anche la data candidata ad avere il break strutturale. I valori critici, forniti dal software R, sono -4,58 con $\alpha = 10\%$, -4,80 con $\alpha = 5\%$ e -5,34 con $\alpha = 1\%$ per il primo modello e -4,82 con $\alpha = 10\%$, -5,08 con $\alpha = 5\%$, -5,57 con $\alpha = 1\%$ per il secondo modello.

Tabella 3.1.1: Risultati dell'analisi rolling svolta per il test ZA

Periodo \ Regressore	Test ZA con drift (statistica test) (Data del break)	Test ZA con drift e trend (statistica test) (Data del break)
2010-2020	-2,9776 798	-3,897 552
2010-2013	-3,1714 292	-3,9509 469
2011-2014	-3,9445 137	-4,6344 214
2012-2015	-3,6388 517	-3,7271 320
2013-2016	-3,4798 654	-3,2854 529
2014-2017	-3,756 403	-3,4116 403
2015-2018	-3,715 152	-4,3131 234
2016-2019	-3,9658 128	-5,2922 326
2017-2020	-4,0947 693	-4,3146 598

Come nei test svolti precedentemente, l'ipotesi nulla di radice unitaria è accettata in tutti i sottocampioni dell'analisi rolling, ad eccezione del test Zivot-Andrews con trend e intercetta relativo all'intervallo di tempo 2016-2019. In quest'ultimo intervallo, per $\alpha=0,05$ e $\alpha=0,1$, si rileva un break strutturale alla rilevazione 326, che corrisponde alla data 9 settembre 2017. Osservando l'andamento della serie storica del prezzo in quel campione di tempo, si può vedere che il prezzo sale proprio fino al periodo di settembre 2017 e, successivamente, dopo un periodo che dura fino all'aprile del 2018, in cui il prezzo è stabile attorno ai 18 €, si avvia un trend decrescente, fino ad arrivare al valore di 10 € di maggio 2019.



Immagine 3.1.5: Andamento del prezzo nel periodo 2016-2019

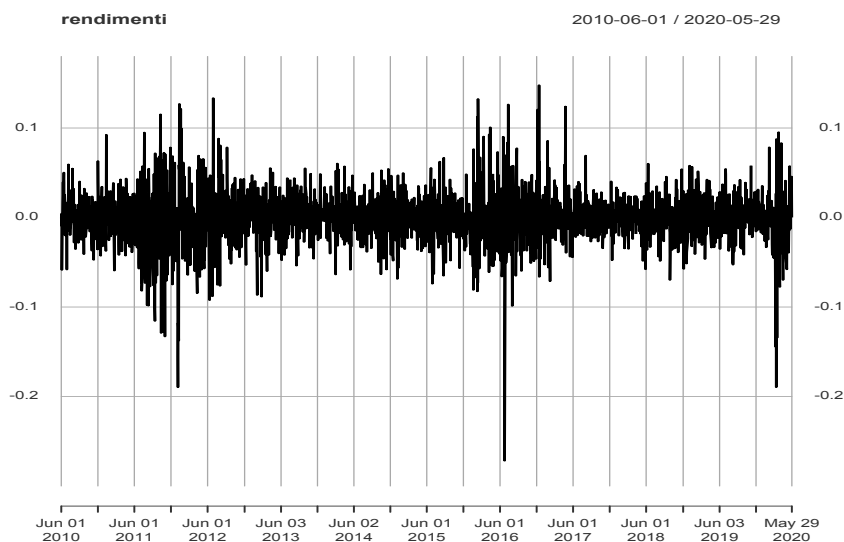
3.2 Analisi della variabile rendimenti

Da quanto osservato nel paragrafo 3.1 si può dire che i prezzi presentino un'autocorrelazione positiva, che rende il processo stocastico non stazionario. Per questo motivo viene utilizzata l'analisi della variabile rendimenti, ottenuta tramite un processo di differenziazione (differenza prima). In questo caso sono stati utilizzati i rendimenti logaritmici, definiti come

$$r_{t+1} = \ln(p_{t+1}) - \ln(p_t)$$

Il sottoinsieme di dati utilizzato è, come per i prezzi, formato da 2536 osservazioni, ma alla prima rilevazione (1 giugno 2010) è stato assegnato dal software R valore nullo, a causa dell'impossibilità di calcolare la differenza con la rilevazione precedente (non presente nel campione). Di seguito viene mostrata la rappresentazione dell'andamento dei rendimenti.

Immagine 3.2.1: Serie storica dei rendimenti di Unicredit



Per quanto riguarda l'autocorrelazione e i test di radice unitaria si procede utilizzando le medesime modalità utilizzate per la variabile prezzi.

Si inizia quindi lo studio verificando la presenza di autocorrelazione, tramite l'analisi grafica dei correlogrammi delle funzioni ACF e PACF, che danno alcune informazioni riguardo la stazionarietà del processo. Dall'analisi ci si attende che i correlogrammi indichino un'assenza di autocorrelazione, perché rimossa dal processo di differenziazione.

Immagine 3.2.2: ACF dei rendimenti

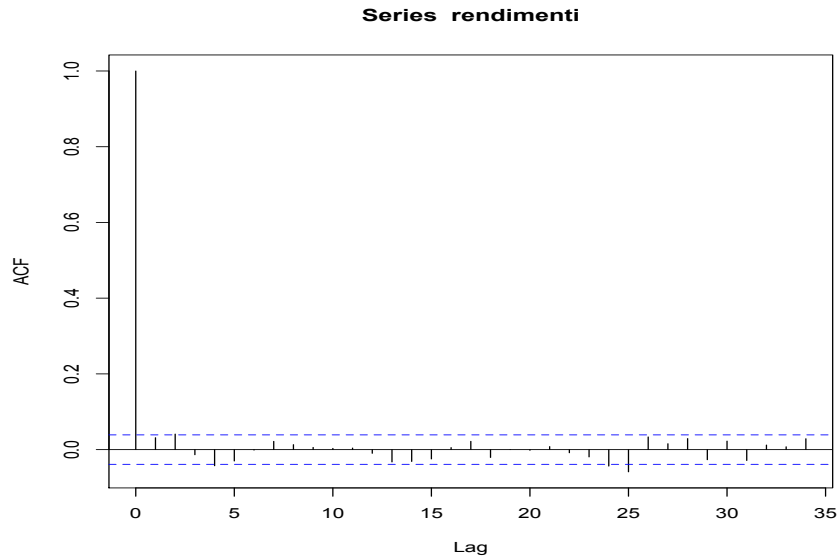
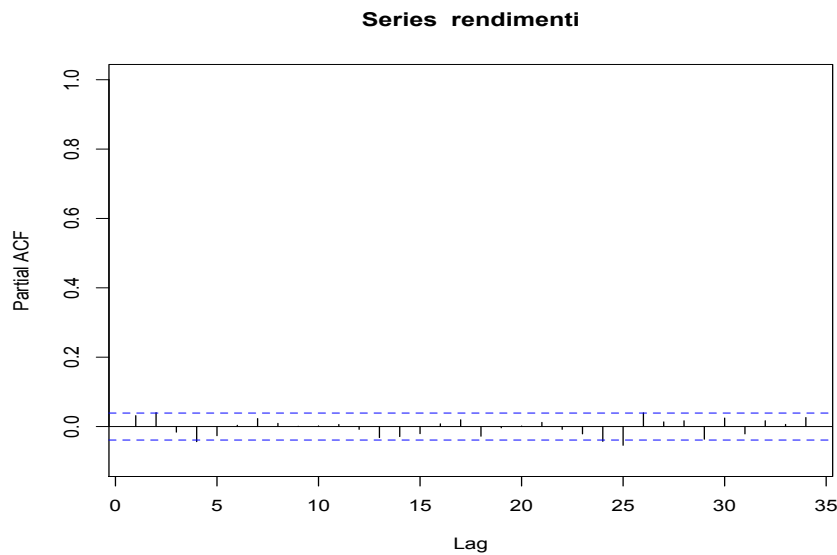


Immagine 3.2.3: PACF dei rendimenti



Come previsto, sia l'immagine 3.2.2 che l'immagine 3.2.3 confermano l'assenza di autocorrelazione e quindi la stazionarietà del processo, che verrà successivamente analizzata attraverso i test di radice unitaria. Nella prima immagine si può vedere che nel caso di $\rho(0)$ il valore rilevato è uguale a 1, ma questa è una proprietà della ACF e non ha alcun significato per quanto riguarda l'autocorrelazione, perché chiaramente non viene considerato nessun ritardo. Per quanto riguarda gli altri ordini di ritardo, alcuni di essi hanno un valore significativamente diverso da 0, ma nel complesso il loro andamento seghettato attorno a 0 tende a confermare quanto precedentemente previsto, ossia la stazionarietà del processo.

Come nel precedente paragrafo, vengono eseguiti alcuni test di radice unitaria: il test ADF, il test ADF-GLS e il test Zivot-Andrews. Da essi sono attesi dei risultati che confermino l'assenza di radice unitaria, di conseguenza l'ipotesi nulla dovrebbe essere rifiutata.

Anche in questo caso le varianti prese in considerazione sono i modelli con deriva e con trend e deriva. I valori critici per il test ADF sono quelli presentati nella seconda e nella terza colonna della tabella 2.3.1.

- Test ADF con regressore deterministico l'intercetta

```
> CADFtest(rendimenti,max.lag.y=20,criterion=c("AIC"),type=c("drift"))
```

ADF test

```
data: rendimenti
ADF(3) = -25.801, p-value < 2.2e-16
alternative hypothesis: true delta is less than 0
sample estimates:
      delta
-0.9904938
```

- Test ADF con regressori deterministici l'intercetta e il trend

```
> CADFtest(rendimenti,max.lag.y=20,criterion=c("AIC"),type=c("trend"))
```

ADF test

```
data: rendimenti
ADF(3) = -25.796, p-value < 2.2e-16
alternative hypothesis: true delta is less than 0
sample estimates:
      delta
-0.9905207
```

I risultati dei test ADF confermano quanto predetto, rifiutando l'ipotesi nulla di radice unitaria. Questo può essere verificato confrontando i valori della statistica DF, rispettivamente -25,801 per il modello con intercetta e -25,796 per modello con intercetta e trend deterministico, che sono molto minori dei valori indicati nella tabella 2.3.1 per ogni livello di significatività. A confermare quanto detto, il p-value non è mai significativamente diverso da 0.

Per un'ulteriore conferma sulla stazionarietà della serie storica viene svolto il test DF-GLS, che dovrebbe rifiutare l'ipotesi nulla. I valori critici con $\alpha = 0,01$, $\alpha = 0,05$ e $\alpha = 0,10$ sono indicati dall'output del software R.

- Test ADF-GLS con regressore deterministico l'intercetta

```
Value of test-statistic is: -25.7973
```

```
Critical values of DF-GLS are:  
      1pct  5pct 10pct  
critical values -2.57 -1.94 -1.62
```

- Test ADF-GLS con regressori deterministici l'intercetta e il trend

```
Value of test-statistic is: -25.8253
```

```
Critical values of DF-GLS are:  
      1pct  5pct 10pct  
critical values -3.48 -2.89 -2.57
```

I risultati dei test ADF-GLS, come si può vedere dai valori della statistica test per entrambe le varianti, seguono le previsioni e confermano la stazionarietà del processo.

Dai essi si può dire che il processo di differenziazione, effettuato alla variabile prezzi, ha prodotto un processo stazionario.

L'ultimo test effettuato, il test Zivot-Andrews, viene attuato per verificare l'eventuale presenza di break strutturali e, in caso questo si verifichi, stimare anche la possibile data con un cambiamento strutturale, questa viene indicata dal software R con il numero della sua rilevazione. Per questo test è stata adoperata la metodologia dell'analisi rolling, quindi è stato effettuato il test a sottoinsiemi di dati, di lunghezza 3 anni, a partire dal giugno 2010, scorrendo avanti di anno in anno. I valori critici, forniti da R, sono -4,58 con $\alpha = 10\%$, -4,80 con $\alpha = 5\%$ e 5,34 con $\alpha = 1\%$ per il modello con deriva diversa da 0 e -4,82 con $\alpha = 10\%$, -5,08 con $\alpha = 5\%$, -5,57 con $\alpha = 1\%$ per il modello con deriva e trend. I risultati sono presentati alla pagina seguente con una tabella riassuntiva.

Tabella 3.2.1: Risultati dell'analisi rolling svolta per il test ZA

Regressore Periodo	Test ZA con drift (statistica test) (Data del break)	Test ZA con drift e trend (statistica test) (Data del break)
2010-2020	-21,4184 414	-21,8259 2469
2010-2013	-12,5221 414	-12,6465 414
2011-2014	-12,221 157	-12,2958 158
2012-2015	-12,4722 12	-12,6229 73
2013-2016	-11,5309 30	-11,6681 681
2014-2017	-11,5924 535	-11,5767 535
2015-2018	-11,3506 284	-11,4234 284
2016-2019	-12,4176 26	-12,4555 24
2017-2020	-11,6228 688	-12,9362 688

L'analisi rolling per il test ZA confermano quando evidenziato dai test econometrici svolti in precedenza, ossia la stazionarietà del processo stocastico.

Il test ZA con costante effettuato sull'intero campione evidenzia un break strutturale alla rilevazione 414, corrispondente al 9 gennaio 2012, data in cui si è dato il via alle negoziazioni dei diritti d'opzione su un aumento di capitale a pagamento di 7,5 miliardi di euro. In questo periodo è presente un break strutturale anche secondo i test relativi ai periodi 2010-2013 e 2011-2014. Il test con costante e trend effettuato sull'intero gruppo di dati, così come i test sul campione 2017-2020, rileva un break all'osservazione in data 19 febbraio 2020; ultima giornata prima del tracollo del titolo indotto dalla pandemia dovuta al Covid-19. Altri break, secondo l'analisi rolling, hanno luogo in giugno 2012, in settembre 2012, nel febbraio del 2016 e nel periodo di inizio luglio 2016.

Nella pagina successiva sono rappresentati i grafici di entrambe le varianti dei test relative all'intero gruppo di dati.

Immagine 3.2.4: Grafico del test ZA con costante (2010-2020)

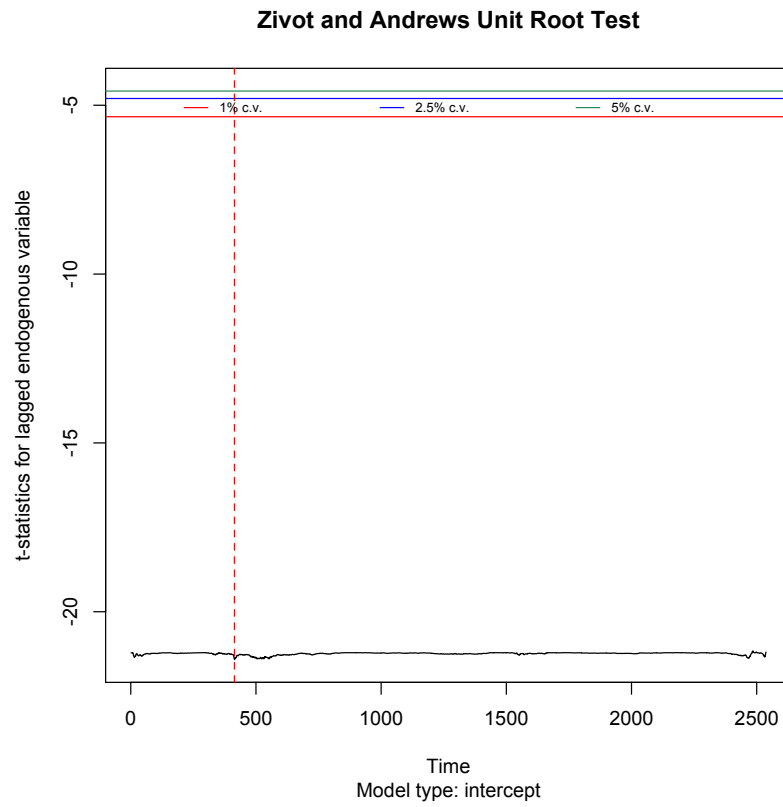
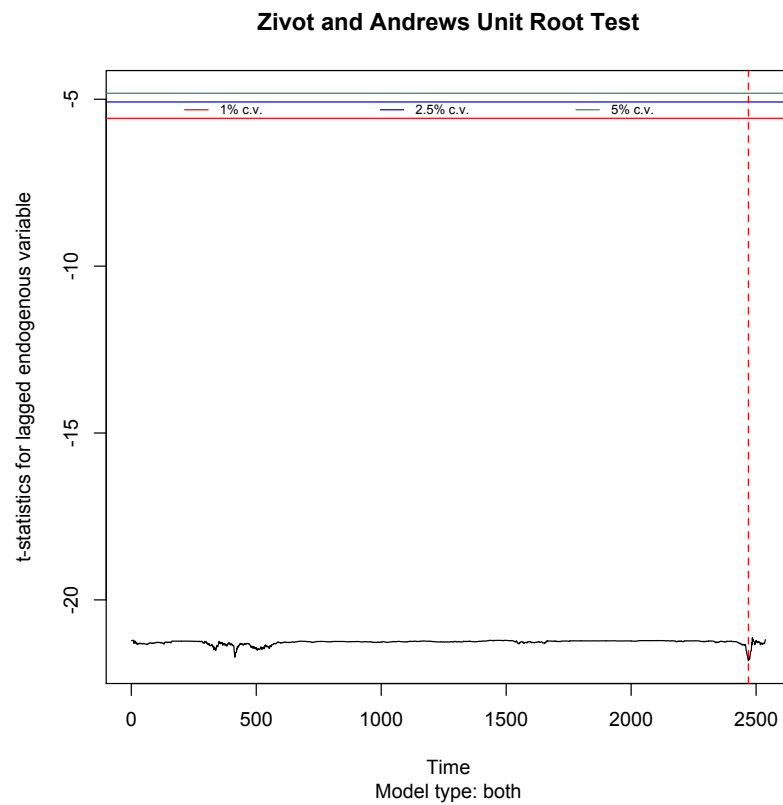


Immagine 3.2.5: Grafico del test ZA con costante e trend (2010-2020)



Lo scopo dell'analisi della serie storica dei valori di Unicredit in borsa è la modellazione del processo che li ha generati, ossia risalire al modello statistico in grado di spiegare i dati con una buona approssimazione.

In questo elaborato è stato impiegato il modello ARMA (p,q), la cui formula è:

$$x_t = \delta + \alpha x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + u_t - \beta u_{t-1} - \beta_2 u_{t-2} - \dots - \beta_q u_{t-q}$$

Esso, per stimare il processo che ha dato origine alla serie storica, si basa su termini autoregressivi e termini a media mobile. I requisiti necessari per poter usare il modello sono la stazionarietà del processo stocastico e la distribuzione dei residui come White Noise ($0,\sigma^2$).

Il primo è stato verificato mediante i test ADF e ADF-GLS svolti in precedenza, che hanno rifiutato l'ipotesi nulla di presenza di radice unitaria, confermando di conseguenza la stazionarietà del processo stocastico.

La distribuzione dei residui è stata analizzata utilizzando il test Box-Ljung, il quale studia la presenza di autocorrelazione in una serie di dati. Il sistema di ipotesi è:

$$\begin{cases} H_0: \rho_1 = \rho_2 = \dots = \rho_m = 0 \\ H_1: \rho_i \neq 0 \text{ per almeno un } i \end{cases}$$

La condizione richiesta del modello ARMA è l'accettazione dell'ipotesi nulla di assenza di autocorrelazione tra i residui.

- Test Box-Ljung sui residui del modello ARMA(2,2)

```
> Box.test(residuiarma22, lag=10, type=c("Ljung-Box"))
```

Box-Ljung test

```
data: residuiarma22  
X-squared = 5.9974, df = 10, p-value = 0.8155
```

I risultati del test BL confermano l'assenza di autocorrelazione nei residui, come dimostrato dal valore del p-value maggiore di α . Quando verificato dal test permette di assumere che i residui del modello autoregressivo a media mobile siano distribuiti come White Noise ($0,\sigma^2$).

Essendo soddisfatti i requisiti richiesti riguardo stazionarietà e distribuzione dei residui, è stato sviluppato il modello ARMA. La ricerca della lunghezza dei ritardi (p,q) più adeguata è stata effettuata dal software R utilizzando il criterio AIC, scegliendo la variante del modello con il valore minore del criterio sopracitato.

- Modello ARMA(2,2) dei rendimenti

```
> auto.arima(rendimenti,ic=c("aic"))
Series: rendimenti
ARIMA(2,0,2) with zero mean

Coefficients:
      ar1      ar2      ma1      ma2
  0.4864 -0.8827 -0.4727  0.9149
s.e.  0.0536  0.0543  0.0426  0.0510

sigma^2 estimated as 0.0008854:  log likelihood=5321.08
AIC=-10632.15  AICc=-10632.13  BIC=-10602.96
```

Il modello che ne risulta è basato su una regressione con due termini autoregressivi e due termini a media mobile. Dall'output di R è possibile osservare che il processo stocastico che ha generato i rendimenti è stato sviluppato come un modello ARIMA, tuttavia si può notare come l'ordine di differenziazione d sia uguale a 0. Questo perché il campione dei dati impiegato, ossia quello dei rendimenti del titolo, è basato su un processo già differenziato e di conseguenza stazionario. Si può concludere dicendo che il modello ARIMA $(p,0,q)$ non sia altro che un modello ARMA (p,q) , perciò il modello presentato dal software può anche essere considerato come un modello ARMA(2,2). Questo passaggio è stato impiegato per consentire al software di identificare automaticamente l'ordine (p,q) migliore secondo il criterio di Akaike.

Per l'analisi della volatilità della serie temporale è stato fatto uso del modello GARCH. Esso si basa sull'assunzione che, mentre la varianza incondizionata è costante nel tempo, la varianza condizionata $h_t^2 = V[\varepsilon_t | I_{t-1}] = E[\varepsilon_t^2 | I_{t-1}]$ sia una funzione dei valori assunti dai residui negli istanti temporali precedenti. Da questo, si può anche intuire che, mentre nella serie storica dei rendimenti non è presente autocorrelazione (come dimostrato nel corso dell'elaborato in precedenza), i quadrati dei rendimenti sono tra loro correlati. Ciò può essere osservato graficamente tramite l'utilizzo del correlogramma.

Conseguenza di quanto detto è la presenza nella serie storica di volatility cluster, cioè cluster di dati con livelli di volatilità simili.

Nella pagina seguente vengono presentati il grafico relativo alla funzione di autocorrelazione dei quadrati dei residui, da cui ci si attende che i coefficienti di autocorrelazione siano significativamente diversi da 0 e la rappresentazione grafica della serie storica dei rendimenti in cui sono indicati alcuni volatility cluster.

Immagine 3.2.6: ACF dei rendimenti quadrati

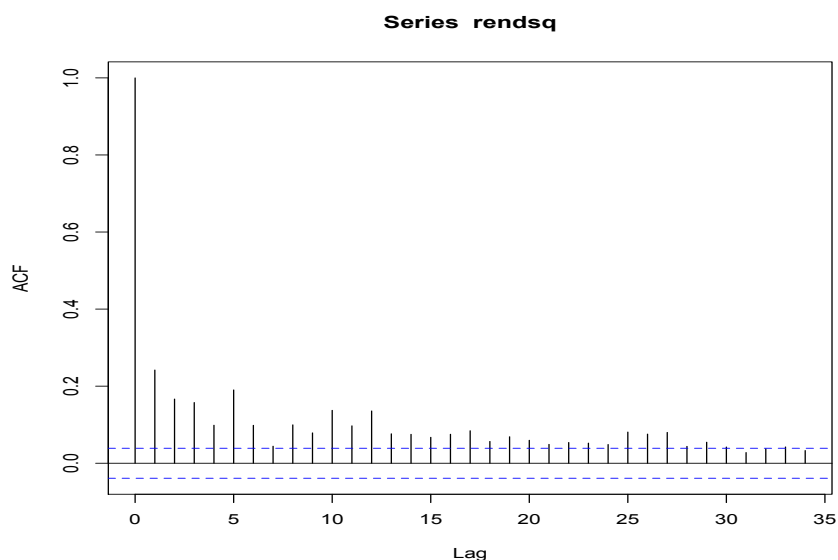
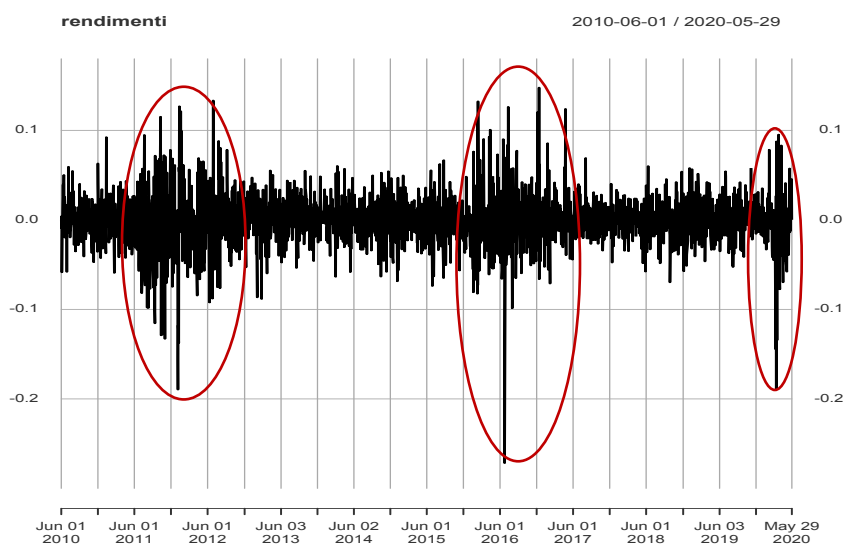


Immagine 3.2.7: Serie storica dei rendimenti con evidenza di alcuni volatility cluster



Si può dire che entrambi i grafici confermino i risultati attesi. L'immagine 3.2.6 dimostra la presenza di autocorrelazione per la variabile r_t^2 ; nell'immagine successiva sono evidenziati in rosso alcuni tra i cluster di volatilità più evidenti, si può notare che nei periodi messi in risalto la varianza condizionale sia superiore alla media. Le evidenze di autocorrelazione della variabile r_t^2 e presenza di effetti ARCH inducono a implementare il modello GARCH.

Viene impiegato il modello GARCH(1,1), in quanto i fatti stilizzati dimostrano che questo sia sufficiente a spiegare la volatilità nelle serie storiche finanziarie. L'equazione del modello è

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}^2$$

- Modello GARCH(1,1)

```

*-----*
*           GARCH Model Fit           *
*-----*

Conditional Variance Dynamics
-----
GARCH Model   : sGARCH(1,1)
Mean Model    : ARFIMA(0,0,0)
Distribution   : sstd

Optimal Parameters
-----
      Estimate  Std. Error  t value  Pr(>|t|)
omega  0.000014   0.000007   2.0734  0.038131
alpha1 0.072019   0.018072   3.9850  0.000067
beta1  0.910729   0.009738  93.5222  0.000000
skew   0.992142   0.025756  38.5215  0.000000
shape  6.843647    1.024625   6.6792  0.000000

```

Il modello descrive i valori della varianza tramite la seguente equazione

$$h_t^2 = 0,000014 + 0,072019\varepsilon_{t-1}^2 + 0,910729h_{t-1}^2$$

La distribuzione adottata per i quadrati dei residui è una T di Student asimmetrica, perché, come affermato nel paragrafo 2.5, nelle serie di dati finanziari sono presenti una maggiore quantità di valori outlier rispetto a una distribuzione normale. L'asimmetria è invece dovuta al fatto che gli shock negativi tendono ad impattare in maniera più incisiva sulla volatilità rispetto agli shock positivi (cosiddetto effetto leva).

- Test Weighted Box-Ljung sui residui

```

Weighted Ljung-Box Test on Standardized Residuals
-----
                    statistic p-value
Lag[1]                2.155  0.1421
Lag[2*(p+q)+(p+q)-1][2]  2.876  0.1516
Lag[4*(p+q)+(p+q)-1][5]  4.551  0.1929
d.o.f=0
H0 : No serial correlation

```

- Test Weighted ARCH LM

```

Weighted ARCH LM Tests
-----
      Statistic Shape Scale P-Value
ARCH Lag[3]    6.823 0.500 2.000 0.009000
ARCH Lag[5]   11.784 1.440 1.667 0.002401
ARCH Lag[7]   13.477 2.315 1.543 0.002729

```

I due test presentati analizzano rispettivamente la presenza di autocorrelazione nei residui e la presenza di effetti ARCH nel processo.

Come si può notare dai risultati del test Box-Ljung, in particolare dai valori dei p-value maggiori di 0,05, l'autocorrelazione non è più presente nella serie storica dei residui.

Il secondo test conferma l'avvenuta rimozione degli effetti ARCH, accettando l'ipotesi nulla di omoschedasticità contro l'ipotesi alternativa di eteroschedasticità condizionata.

- Test di bontà dell'adattamento

Adjusted Pearson Goodness-of-Fit Test:

	group	statistic	p-value(g-1)
1	20	23.50	0.2161
2	30	32.21	0.3106
3	40	49.52	0.1206
4	50	48.68	0.4859

L'ultimo test impiegato per valutare il modello GARCH è il test di bontà dell'adattamento di Pearson che verifica l'ipotesi nulla che il campione di dati sia stato estratto da una particolare distribuzione. In questo caso il p-value del test conferma che i dati sono distribuiti come previsto, ossia secondo una funzione di distribuzione T di Student asimmetrica.

Si può di conseguenza concludere dicendo che il modello GARCH(1,1) sia in grado di spiegare correttamente la varianza dei rendimenti, rimuovendo sia la presenza di autocorrelazione tra i residui e sia l'eteroschedasticità condizionata.

Conclusioni

Una volta svolto il modello GARCH per lo studio della volatilità si possono trarre le conclusioni e commentare brevemente i risultati ottenuti nell'ultimo capitolo.

Nella prima parte è stata presa in analisi la variabile Prezzi, a cui è stata applicata una trasformazione logaritmica per ridurre la volatilità. La funzione di autocorrelazione e la funzione di autocorrelazione parziale evidenziano la presenza di autocorrelazione tra i valori osservati nella serie storica dei Prezzi. Ciò porta il processo stocastico a non essere stazionario, come verificato anche dai test ADF e ADF-GLS.

Un ulteriore test di radice unitaria impiegato è stato il test Zivot-Andrews per osservare se nella serie storica fossero presenti dei break strutturali. Esso prende in considerazione la possibilità che sia presente un solo break strutturale nel campione di dati impiegato, per questo motivo è stato fatto uso dell'analisi rolling, dividendo l'intero campione in sottogruppi di dati di 3 anni. Anche l'ultima tipologia di test effettuati accetta l'ipotesi nulla di radice unitaria, accertando la non stazionarietà del processo.

L'unico break evidenziato avviene in data 9 settembre 2017.

In tutti i test di radice unitaria eseguiti sono state prese in considerazione le varianti con presenza di intercetta e con presenza di intercetta e trend deterministico.

Per risolvere il problema di non stazionarietà del processo, che non permette di effettuare la modellizzazione della serie storica dei prezzi, ad essa è stata applicata la differenziazione prima dei valori, ottenendo così la variabile Rendimenti.

Dall'osservazione dei correlogrammi e dei risultati dei vari test DF impiegati, si può confermare che il processo di differenziazione ha reso la serie stazionaria.

I test Zivot-Andrews sono stati svolti anche per la variabile Rendimenti. In questo caso sono stati rilevati vari break strutturali.

Il test ZA con costante effettuato sull'intero campione evidenzia una rottura nel processo stocastico 9 gennaio 2012, data in cui si sono avviate le negoziazioni dei diritti d'opzione su un aumento di capitale. In questo periodo è presente un break strutturale anche secondo i test relativi ai periodi 2010-2013 e 2011-2014. Il test con costante e trend effettuato sull'intero gruppo di dati, così come i test sul campione 2017-2020, rileva un break all'osservazione in data 19 febbraio 2020; nel periodo in cui è iniziata la grave recessione dovuta alla pandemia di Covid-19. Ulteriori break sono stati rilevati in giugno 2012, in settembre 2012, nel febbraio del 2016 e nel periodo di inizio luglio 2016.

Il primo modello studiato per spiegare l'origine del processo stocastico è stato il modello ARMA, che, per essere impiegato, necessita che la serie storica sia stazionaria e che i residui siano distribuiti come White Noise $(0, \sigma^2)$.

La prima condizione è stata verificata tramite l'impiego dei test di radice unitaria, la seconda, invece, tramite il test Box-Ljung, che ha accettato l'ipotesi nulla di assenza di autocorrelazione tra i residui.

È stato perciò possibile usufruire del modello ARMA. La scelta della lunghezza dei ritardi è stata effettuata autonomamente dal software R secondo il criterio AIC, selezionando l'ordine di lunghezza $(2,2)$.

Come osservato dalla presenza di volatility cluster nella serie storica dei rendimenti e dalla presenza di autocorrelazione tra i residui al quadrato, il modello ARMA non è in grado di spiegare il comportamento della varianza di una serie storica finanziaria, caratterizzata dalla presenza di eteroschedasticità condizionata e autocorrelazione dei quadrati dei residui.

Per questo motivo, per spiegare la volatilità dei rendimenti, è stato usato anche il modello GARCH(1,1). Per meglio rappresentare le proprietà delle serie storiche finanziarie (effetto leva e elevata presenza di outlier) è stata impiegata la distribuzione T di Student asimmetrica. Il test di bontà dell'adattamento di Pearson ha confermato la validità di questa funzione per descrivere la distribuzione dei quadrati dei residui.

In conclusione, per verificare l'efficacia del modello GARCH per analizzare la varianza, sono stati presi in considerazione i risultati del test Box-Ljung e del test del moltiplicatore di Lagrange per gli effetti ARCH. Entrambi i test confermano la capacità del modello GARCH impiegato di spiegare la volatilità nella serie storica.

Bibliografia

Akaike H., 1974. A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, vol. 19, pp. 716-723

Bollerslev T., 1986, Generalized autoregressive conditional eteroschedasticity. *Journal of Econometrics*, vol. 31, pp. 307-327

Box G. E. P. e Pierce D. A., 1970. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models, *Journal of the American Statistical Association*, pp. 1509-1526

Dickey D. e Fuller W., 1979. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, vol. 74, pp. 427-431

D'Ascenzo M., 2012. Maxi-aumento a sconto del 43%. *Il Sole 24 Ore* [online]. Disponibile su <https://st.ilsole24ore.com/art/finanza-e-mercati/2012-01-04/unicredit-maxiaumento-sconto-232944.shtml?uuid=AaBafsaE&refresh_ce=1> [data di accesso: 12 giugno 2020]

Elliott G., Rothenberg T. e Stock J., 1996. "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, Econometric Society, vol. 64(4), pp. 813-836

Engle R., 1982, Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of U.K. Inflation, *Econometrica*, vol. 50, pp. 987-1008

Kirchgässner G. and Wolters J., 2007. *Introduction to Modern Time Series Analysis*. : Springer

Ljung G.M. e Box G. E. P., 1978. On a Measure of Lack of Fit in Time Series Models, *Biometrika* vol. 65, pp. 297 – 303

MacKinnon J. G., 2010. "Critical Values For Cointegration Tests", Working Paper 1227, Economics Department, Queen's University

Mandelbrot B.B, 1963, The variation of certain speculative prices, *Journal of Business*, vol. 36, pp. 394-419

Parzen E., 1962. *Stochastic Processes*, San Francisco: Holden-Day

Pearson W. M., 1919. *Indices of General Business Conditions*. Cambridge, Harvard University Committee on Economic research

Said E. S. e Dickey D. A., 1984. Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika*, vol. 71 (3), pp. 599-607

Stock J. e Watson M. M., 2016. *Introduzione all'econometria*. Milano: Pearsons

Zivot E. e Andrews W. K., 1992. Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business & Economic Statistics*, vol. 10 (3), pp. 251-270