UNIVERSITÀ DEGLI STUDI DI SALERNO

Dipartimento di Scienze Economiche e Statistiche

Corso di Laurea Magistrale in Scienze Statistiche per la Finanza

Tesi in

Modelli statistici per la gestione del rischio di mercato

IL MODELLO LPPLS PER LA PREVISIONE DELLE BOLLE SPECULATIVE

Relatore:

Ch.mo Prof. Giuseppe Storti

Candidato:

Sabato Gargiulo

matr. 0222400594

Correlatore:

Ch.mo Prof. Antonio Naimoli

ANNO ACCADEMICO 2020/2021

"The tour we have taken through the last century proves that market irrationality of an extreme kind periodically erupts and compellingly suggests that investors wanting to do well had better learn how to deal with the next outbreak."

Warren Buffett

Sommario

Le bolle speculative affascinano da sempre gli studiosi e gli appassionati di finanza comportamentale essendo eventi unici nel loro genere e le cui cause non sono identificabili con facilità. Negli anni la frequenza con cui questi eventi si sono verificati sui mercati finanziari è aumentata e, con essa, l'attenzione dei policy maker e delle istituzioni di vigilanza al fine di mettere in campo strumenti validi che permettessero di prevenire e limitare i danni causati all'economia reale. Sulla scia della consapevolezza acquisita da parte delle istituzioni finanziare, anche il mondo accademico ha cercato di approfondire le cause di questi eventi così singolari. Nel tempo, infatti, sono stati proposti numerosi modelli volti ad identificare ex-ante le bolle speculative con lo scopo di permettere ai regolatori di intervenire in tempo e di adottare misure volte a prevenirle.

Nella prima parte di questo lavoro ci siamo occupati di dare una definizione sufficientemente esaustiva di bolla speculativa. Ciò è stato possibile analizzando vari casi storici di bolle speculative e rintracciando in ognuno di questi eventi delle caratteristiche comuni. Abbiamo infatti mostrato come questi eventi posseggano proprietà statistiche che li distinguono nettamente dalle giornate classiche di contrattazione e come sia possibile sfruttare tali proprietà per analizzare questi eventi da un'ottica differente rispetto all'analisi classica. Nei lavori da noi citati, infatti, si dimostra come i *drawdowns* legati ad alcuni importanti crash finanziari siano eventi eccezionali dal punto di vista probabilistico. Inoltre, viene fatto notare come anche le distribuzioni di probabilità dei rendimenti degli asset quotati siano diverse nei giorni di crash rispetto a normali giornate di contrattazione. L'analisi mostra, infatti, come nelle giornate immediatamente precedenti e successive al crash, la distribuzione dei rendimenti sia connotata da un picco sui valori negativi o positivi del rendimento, cioè da un'asimmetria positiva o negativa (a seconda della particolare giornata). Nel seguito abbiamo poi analizzato alcuni approcci proposti in letteratura per lo studio di questi fenomeni, in particolare l'approccio basato su processi stocastici non causali e il modello LPPLS. È proprio su quest'ultimo modello che ci siamo concentrati nel corso di questo lavoro. Il modello LPPLS ha ricevuto critiche e apprezzamenti ed è

stato utilizzato anche da alcuni dipartimenti di risk management di grandi istituzioni finanziarie. Le attenzioni riservate ad esso derivano sostanzialmente dalla capacità che tale modello ha dimostrato nel tempo di riuscire ad identificare in anticipo la formazione e lo sviluppo di bolle speculative.

Nella seconda parte del lavoro abbiamo fornito una spiegazione esaustiva delle ipotesi alla base del modello e la sua derivazione analitica, con particolare riguardo per l'ipotesi, di fondamentale importanza, secondo la quale il mercato finanziario è una struttura frattale e gerarchica in cui operano agenti eterogenei per capitalizzazione. Gli autori del modello ipotizzano che gli agenti operanti sul mercato siano classificabili in *agenti razionali* e *noise trader*, i quali manifestano tendenze imitative locali. Queste tendenze imitative locali possono pregiudicare, in determinate condizioni, lo stato dell'intero sistema se le forze imitative si diffondono in tutta la rete. Secondo il modello LPPLS è proprio in tal caso che il mercato raggiunge un punto critico ed avviene un crash. Gli autori che per primi hanno introdotto il modello [Johansen et al. (2000)] sostengono che l'interazione tra *noise trader* e *agenti razionali* generi, nell'intorno di un punto critico, un andamento log-periodico del prezzo, la cui frequenza delle oscillazioni aumenta quanto più il sistema si avvicina ad esso. Ciò significa, in sostanza, che un sistema vicino ad un punto critico è formato da agenti le cui opinioni sono fortemente correlate a causa delle tendenze imitative presenti nella rete e che quindi l'intero sistema sia sensibile rispetto a piccole perturbazioni. Nel corso del tempo numerose critiche sono state rivolte dalla comunità scientifica alla formulazione originaria del modello ed alcune di queste sono sfociate nella costruzione di varianti del modello LPPLS che andassero a risolvere le criticità individuate da altri autori. Al termine del secondo capitolo abbiamo presentato due di questi modelli: il modello log-periodico AR(1)-GARCH(1,1) e il modello LPPLS generalizzato con residui *mean-reverting*. Nel terzo capitolo è stata presentata la metodologia classica per la stima del modello. Gli autori propongono come funzione obiettivo la somma dei residui al quadrato tra il modello e la serie storica del prezzo. Tale funzione, però, data la struttura periodica del modello, risulta essere caratterizzata da molteplici minimi locali nei quali gli algoritmi classici di ottimizzazione potrebbero restare intrappolati fornendo quindi risultati poco affidabili. Gli autori propongono pertanto di utilizzare una procedura in due fasi: in primis si utilizza l'algoritmo *tabu-search*. Le migliori dieci soluzioni fornite dall'algoritmo vengono poi utilizzate come valori iniziali per l'algoritmo di Levenberg-Marquardt. Negli anni successivi, dato l'elevato costo computazionale della procedura sopra descritta, è stata introdotta una riparametrizzazione del modello che elimina la periodicità e semplifica quindi la procedura di stima dei parametri.

Nel nostro lavoro ci siamo avvalsi della riparametrizzazione del modello e stimato i

parametri utilizzando l'algoritmo Multi-Level Single-Linkage (MLSL). Data la sensibilità della stima rispetto alla finestra temporale selezionata, abbiamo implementato la procedura proposta dal Prof. Sornette per ottenere indicazioni robuste rispetto alla presenza o meno di bolle speculative sull'asset analizzato. Il modello LPPLS è stato stimato, per ogni giornata *pseudo-presente* t_2, su 980 finestre temporali la cui lunghezza va da 1460 a 40 giorni. Per ogni finestra considerata, quindi, abbiamo ottenuto il set di stime per i parametri del modello. In una fase successiva abbiamo filtrato i risultati di ogni finestra temporale secondo dei vincoli proposti in letteratura per selezionare solo le stime robuste e affidabili. Successivamente abbiamo utilizzato i risultati della procedura di *filtering* per costruire un indicatore molto utilizzato in letteratura e nella pratica chiamato LPPLS Confidence Indicator. Sulla scorta delle recenti innovazioni introdotte dal *Financial Crisis Observatory* dell'ETH di Zurigo, abbiamo deciso di dividere il set di finestre temporali sulla base della loro ampiezza. Questo ci ha permesso di derivare indicatori LPPLS C.I. per finestre temporali di lunghezza variabile (Super-Short, Short, Medium, Long time) al fine di poter meglio identificare la scala temporale nella quale si è sviluppata la bolla speculativa. Nella parte conclusiva del lavoro abbiamo presentato i risultati ottenuti stimando il modello LPPLS e derivando gli indicatori rispetto a svariati assets, quali ad esempio Tesla, Microsoft, Crude Oil, Saipem ed altri. Come già appurato in numerosi studi, il modello LPPLS ed in particolare l'indicatore LPPLS C.I. è utile ad identificare in anticipo le bolle speculative, fin dai primi mesi in cui esse hanno inizio. È proprio a causa delle capacità dimostrate nell'indentificare bolle speculative in anticipo che il modello LPPLS è stato implementato in alcuni dipartimenti di risk management di istituzioni finanziarie ed è stato oggetto di un lungo dibattito durato quasi un decennio. Precisiamo, infine, che sebbene i risultati ottenuti dal modello siano ottimi, un risk manager prudente deve necessariamente tenere conto di altri fattori ed altre metriche per gestire il rischio connesso all'operatività sui mercati finanziari. In questo senso il modello LPPLS è uno strumento aggiuntivo, che integra ma non sostituisce quelli già ampiamente utilizzati dai risk manager delle istituzioni finanziarie.

Indice

Introduzione **3**

1 Le bolle speculative **4**

1.1 Casi storici 5

1.1.1 Black monday 1987 5

1.1.2 Bolla dei tulipani 6

1.1.3 Crisi del '29 7

1.1.4 Bolla delle dot-com 8

1.1.5 Bolla Bitcoin 10

1.2 Proprietà statistiche 12

1.3 Approcci proposti in letteratura 18

1.3.1 Processi AR non causali 18

1.3.2 Modello LPPLS 20

2 Il modello LPPLS **22**

2.1 Ipotesi del modello 22

2.2 Approccio macro-strutturale 26

2.3 Approccio micro-strutturale 27

2.4 Modello per la rete di interazioni 30

2.4.1 Griglia bidimensionale 30

2.4.2 Hierarchical Diamond Lattice 33

2.5 Formulazioni alternative del modello LPPLS 39

2.5.1 Modello log-periodico AR(1)-GARCH(1,1) 39

2.5.2 Modello LPPLS generalizzato con residui *mean-reverting* . . . 41

3 Stima dei parametri **43**

3.1 Procedura classica di stima 43

3.2 Nuova procedura di stima 47

3.3 Algoritmo Multi-Level Single-Linkage 51

3.4 Indicatori LPPLS 53

3.4.1 Robustezza rispetto alla finestra di stima 53

3.4.2 DS LPPLS Confidence ™ 55

4 Risultati empirici **58**

4.1 Struttura 58

4.2 S&P500 (Black Monday) 60

4.3 Tesla 64

4.4 Bitcoin 68

4.5 Microsoft 72

4.6 Oro 76

4.7 Saipem 80

4.8 Petrolio 84

Conclusioni **89**

Bibliografia **89**

Appendice A LPPLS C.I. (S&P500) **96**

Appendice B LPPLS C.I. (Petrolio) **99**

Introduzione

Al giorno d'oggi migliaia di trilioni di dollari vengono investiti sui mercati globali negli asset più disparati, dalle azioni agli strumenti esotici. I benefici che l'espansione dei mercati finanziari ha portato all'economia reale sono tangibili e analizzati approfonditamente in letteratura. Gli operatori economici nell'indirizzare le loro scelte d'investimento verso un asset piuttosto che un altro tengono conto di vari fattori, tra cui il rendimento atteso dell'investimento e la rischiosità legata ad esso, si veda ad esempio Markowitz (1952) per quanto riguarda i criteri ottimali di selezione del portafoglio. Nel corso del tempo sono stati implementati molteplici modelli quantitativi per studiare lo stato di salute dell'economia nel complesso e dei singoli settori con lo scopo principale di rendere le strategie d'investimento flessibili e adattarle alla frenetica evoluzione che i mercati hanno sperimentato, soprattutto durante gli ultimi decenni.

Le bolle speculative da sempre affascinano gli studiosi, gli appassionati di finanza e di economia comportamentale. Una bolla speculativa può essere vista sia come un'opportunità sia come un rischio considerevole. Infatti, nelle sue fasi iniziali, essa rappresenta evidentemente un'opportunità per coloro i quali entrano sul mercato e beneficiano della crescita super-esponenziale dei prezzi. Viceversa, invece, essa causa enormi danni agli investitori che entrano sul mercato in ritardo, spinti dall'euforia o da quella che viene definita *FOMO* (Fear Of Missing Out), cioè la paura di perdere un'opportunità storica di profitto. Purtroppo i danni che questi eventi causano all'economia reale sono enormi e superano di gran lunga i benefici che i pochi riescono ad ottenere individuando tali opportunità anticipatamente.

Come fatto notare da Farmer (2015) e Narayan et al. (2016), date le dimensioni del mercato finanziario globale e dato il livello di diffusione di alcuni strumenti, anche nei portafogli delle famiglie, un crash futuro sui mercati finanziari avrebbe ripercussioni enormi sull'economia mondiale, probabilmente peggiori di quelle verificatesi nel 2000 per la dot-com bubble o nel 2008 per la crisi dei sub-prime.

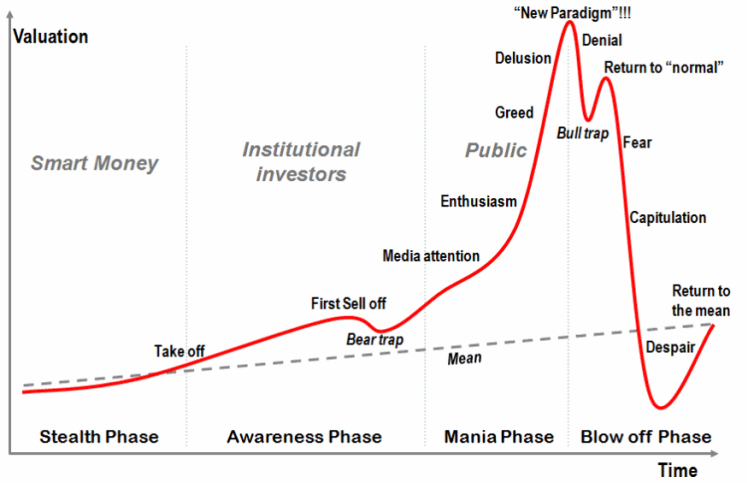


Figura 1: Le fasi di sviluppo di una bolla speculativa, come indicate dal Prof. Jean-Paul Rodrigue.[1]

A causa del potenziale distruttivo di questi eventi le Banche Centrali e gli istituti di vigilanza hanno ampliato, nel corso del tempo, il ventaglio di strumenti a loro disposizione per monitorare i mercati e il livello di rischio sistemico insito in essi. Tale processo di evoluzione dei criteri e degli strumenti di vigilanza è stato un passaggio obbligato date le devastanti crisi finanziarie che si sono susseguite nel corso dell'ultimo secolo e che hanno cagionato gravi danni all'economia reale e alla crescita di lungo periodo. In particolare, solo negli ultimi trent'anni ci sono state circa cento crisi finanziarie. È evidente, quindi, come sia di fondamentale importanza riuscire ad identificare le bolle speculative in anticipo per limitarne la crescita e minimizzare i danni di un eventuale discesa repentina dei prezzi.

In questo lavoro ci occuperemo, in primis, di dare una definizione formale di bolla speculativa e tratteremo nel dettaglio quali sono i meccanismi alla base della sua formazione. Vedremo inoltre come tali eventi estremi abbiano proprietà statistiche peculiari che li contraddistinguono. Tratteremo poi alcuni esempi di bolle speculative, come la crisi del '29 e la bolla dei tulipani, al fine di individuare le caratteristiche comuni che li contraddistinguono. Passeremo poi in rassegna alcuni degli approcci proposti in letteratura per identificare le bolle speculative e introdurremo nel dettaglio il modello da noi utilizzato in questo lavoro, ossia il modello LPPLS. La trattazione riguarderà, in primis, le ipotesi alla base del modello, la sua derivazione analitica e le sue proprietà. Successivamente analizzeremo nel dettaglio le problematiche relative alla stima dei parametri e l'algoritmo da noi utilizzato, con particolare attenzione riguardo lo spazio paramerico e le condizioni per le quali la stima del modello sia da ritenersi soddisfacente.

[1] https://en.wikipedia.org/wiki/Jean-Paul_Rodrigue

Nell'ultima parte del lavoro presenteremo un serie di indicatori costruiti sulla base del modello LPPLS che permettono di ottenere risultati robusti e previsioni che si sono rivelate efficaci riguardo la presenza o meno di bolle speculative su un determinato asset. Presenteremo quindi i risultati ottenuti da questo indicatore per una serie di asset, ad esempio lo SP500 e commodities quali oro e petrolio. Al fine di stimare i parametri del modello e derivare gli indicatori utilizzeremo il software open-source R. Il calcolo dell'indicatore richede un'ingente potenza computazionale, per motivi che verranno discussi nei capitoli successivi: pertanto si è optato per l'implementazione del calcolo in parallelo mediante il pacchetto *parallel*. Tutti gli script utilizzati in questo lavoro sono disponibili al seguente link [2] .

[2] https://github.com/sabato96/LPPLS

Capitolo 1

Le bolle speculative

Il primo problema che bisogna affrontare è la definizione di cosa sia una bolla speculativa. Nel lavoro di Sornette and Cauwels (2015) si definisce bolla speculativa una divergenza del prezzo di un asset rispetto al suo valore intrinseco. Più il prezzo diverge, più il rischio che la bolla scoppi aumenta. La spirale inflazionistica si conclude quando il sistema diventa insostenibile e il prezzo comincia a scendere rapidamente. In letteratura c'è stato un lungo dibattito, tuttora in corso, su quali siano le potenziali cause di questo fenomeno. Tra le tante motivazioni addotte, una delle più citate è l' effetto gregge, cioè quel meccanismo per il quale i partecipanti al mercato acquistano un asset poiché guidati non dal suo valore intrinseco o dalle prospettive di crescita dell'azienda, quindi non per motivi che potremmo definire razionali, bensì solo perché influenzati dalle opinioni della maggioranza. Un esempio plastico di tale effetto è quello riportato da Keynes (1936) del concorso di bellezza nel quale i giudici non votano secondo le proprie opinioni personali, ma si lasciano influenzare dalle intenzioni di voto degli altri giudici. Un'ulteriore causa è l'asimmetria informativa, cioè la situazione nella quale gli investitori non dispongono di informazioni complete e sufficienti per prendere delle decisioni razionali e ottimali. Oltretutto anche l'azzardo morale gioca un ruolo fondamentale perché riguarda principalmente le grandi istituzioni finanziarie, che sono inevitabilmente gli agenti più esposti sul mercato. Le grandi istituzioni finanziarie potrebbero essere incentivate ad assumere rischi eccessivi, che mettano in pericolo la sopravvivenza stessa dell'istituzione, perché consce del fatto che, qualora dovessero presentarsi dei problemi, interverrebbero le Banche Centrali a salvarle. Nei prossimi paragrafi analizzeremo nel dettaglio alcuni casi celebri di bolle speculative, sia recenti che meno recenti, cercando di individuare i tratti distintivi comuni di eventi così complessi e difficili da prevedere.

1.1 Casi storici

Nel seguito proporremo alcuni esempi di bolle speculative storiche identificando le cause alla base di questi eventi e le conseguenze che esse hanno avuto sulla crescita economica di lungo termine.

1.1.1 Black monday 1987

Tra il 14 Ottobre 1987 e il 19 Ottobre dello stesso anno, i principali indici statunitensi persero il 30% o più. Tale fenomeno non si limitò solo agli indici americani ma si estese a quelli delle principali economie mondiali. Sulla base degli studi di Sornette et al. (1996), gli Stati Uniti non furono però il primo paese a sperimentare un crash degli indici azionari. Sono stati condotti numerosi studi al fine di individuarne le cause e adottare delle misure volte a prevenire il ripetersi di questi eventi in futuro. La letteratura scientifica, come sottolineato da Sornette (2009), ha identificato una serie di cause, tra cui:

- **Trading automatico:** Per trading automatico si intende la programmazione di macchine al fine di effettuare operazioni sui mercati finanziari in modo rapido ed efficace. Alcuni autori hanno attribuito al dilagare di questo fenomeno la causa del crash, mentre altri hanno fatto notare come in alcuni mercati in cui il crash avvenne in maniera più marcata, le procedure automatiche di trading non fossero state ancora introdotte su larga scala.
- **Scarsa liquidità:** Durante il crash, l'enorme flusso di ordini di vendita proveniente dagli operatori di mercato non potè essere soddisfatto a pieno vista l'assenza di operatori disposti a comprare quegli asset. Senza dubbio ciò accentuò la profondità del crash, sebbene non sia chiaro cosa abbia indotto gran parte degli operatori di mercato a inviare ingenti ordini di vendita allo stesso tempo.
- **Prezzi sopravvalutati:** Molti analisti e studiosi concordano nell'affermare che gli indici nel Settembre 1987 erano sostanzialmente sopravvalutati, infatti indicatori quali PE ratio e PD ratio erano ai massimi storici. I critici di questa spiegazione asseriscono però che anche in periodi precedenti, ad esempio nel periodo 1960-72, tali indicatori erano ai massimi eppure non si è verificò alcun crash.

Nel lavoro di Barro et al. (1989), gli autori costruiscono un indice sintetico per misurare la performance globale dei mercati in quel periodo. Questo indice è la media aritmetica dei 23 indici di mercato più importanti dell'epoca. L'indice costruito dagli autori è stato normalizzato a 100 al Settembre 1981 e ad Ottobre dello stesso anno era pari a 73,6, facendo segnare una perdita del 26,4% . Secondo gli autori nei mesi precedenti al crash ci fu un meccanismo di cooperazione diffuso globalmente che avrebbe condotto gli indici dei vari Paesi su traiettorie di crescita insostenibili. Il meccanismo di cooperazione e imitazione accennato nel lavoro del 1989, stando alla letteratura recente, si è confermato essere uno dei fattori principali che favoriscono la crescita e lo sviluppo di bolle speculative sui mercati finanziari. Vedremo, nei successivi paragrafi, come tale meccanismo sia pressocchè ubiquo in tutte le più grandi crisi finanziarie, più e meno recenti.

1.1.2 Bolla dei tulipani

Le crisi finanziarie affondano le loro radici nella storia fin dal XVII secolo. Un esempio di cui si è discusso a lungo riguardo le bolle speculative dei secoli scorsi è quello della bolla dei tulipani del 1637. È bene ricordare che l'Olanda dell'epoca era un Paese in forte crescita economica ed Amsterdam era lo snodo commerciale più importante dell'Europa settentrionale, con rotte commerciali che si diramavano verso ogni angolo del continente e verso le Americhe. Il tulipano fu esportato per la prima volta in Europa dalla Turchia nel 1554 e da allora, vista la rarità e la bellezza del fiore, divenne un simbolo di prestigio. Il mercato dei tulipani continuò a svilupparsi favorito dalla crescente domanda proveniente soprattutto delle classi più agiate della società olandese. Mano a mano che il prestigio del fiore aumentava e le persone ne erano sempre più attratte, il suo prezzo cominciò a crescere. La classe media olandese, inoltre, considerava come un solido investimento l'acquisto dei bulbi. In particolare, gli acquirenti interessati erano soliti acquistare in anticipo i bulbi prima ancora che fiorissero a prezzi fissati ex-ante, creando un vero e proprio mercato dei futures; Alcuni arrivarono addirittura ad ipotecare le proprie abitazioni e i propri possedimenti per investire nel mercato.

Nel biennio 1635-36, il prezzo dei tulipani in Olanda arrivò a 2500 fiorini per bulbo, equivalente al costo di un'abitazione sull'Herengracht (Canale dei signori) ad Amsterdam[1]. Nella prima metà del 1637, il prezzo fece segnare un massimo assoluto per poi precipitare vertiginosamente nei mesi successivi, ritornando ai valori fatti segnare nel 1634. Stando a quanto detto finora è possibile notare come, anche in questo caso, tra i fattori che causarono la bolla speculativa ci fossero l'effetto gregge

[1] https://en.wikipedia.org/wiki/Tulip_mania

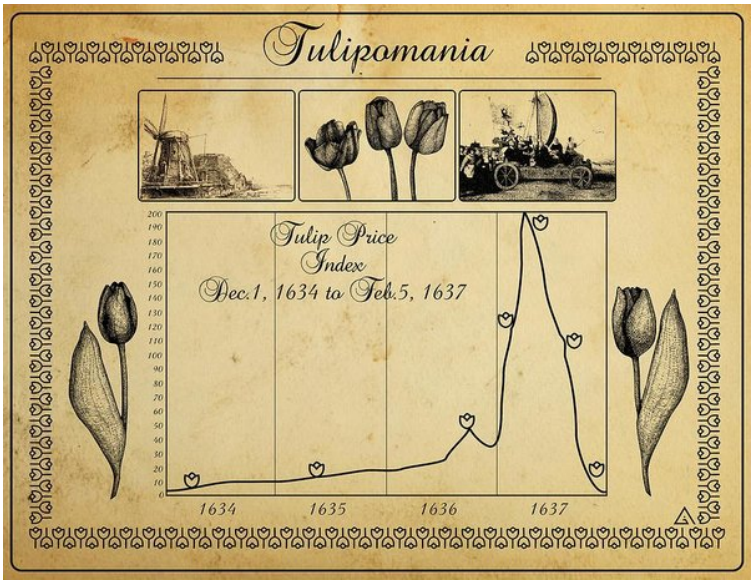


Figura 1.1: Bolla dei tulipani [Ecotalker (2019)]

e la convinzione diffusa che l'investimento in tulipani fosse completamente privo di rischio.

1.1.3 Crisi del '29

Gli anni 20 del secolo scorso sono stati caratterizzati da una crescita economica e finanziaria sostenuta, che purtroppo si è arrestata con quello che è stato definito *Great Crash* di Ottobre 1929. La recessione economica che ne seguì portò miseria e disoccupazione sia negli Stati Uniti che nel resto del mondo. Circa 2000 imprese americane chiusero e il sistema bancario americano subì un forte stress e una serie di cambiamenti strutturali importanti. Analizzare il contesto e le cause della crisi del '29 è utile poiché presenta alcune caratteristiche che si ripropongono quando si analizzano le bolle speculative. In primis, il crash del '29, e in generale tutti i crash finanziari, avvengono quando l'opinione pubblica, gli economisti e gli investitori non se lo aspettano. Come detto in precedenza per la bolla dei tulipani, le persone si aspettavano che il prezzo sarebbe cresciuto ancora nei mesi successivi. Ad esempio, Irving Fischer, un celebre economista americano della Yale University, 14 giorni prima del crash, ebbe a dire: "Tra qualche mese mi aspetto che il mercato azionario faccia segnare nuovi massimi". Un'altra caratteristica dei crash finanziari, che caratterizza anche la crisi del '29, è che questi eventi non accadono mai quando l'opinione pubblica pensa le cose stiano andano male. Prima di ogni crash, infatti, tutti pensano che i mercati e l'economia siano sani e che tutto andrà per il meglio nei mesi successivi. La terza caratteristica che ci preme sottolineare è quella per cui, nelle fasi iniziali della

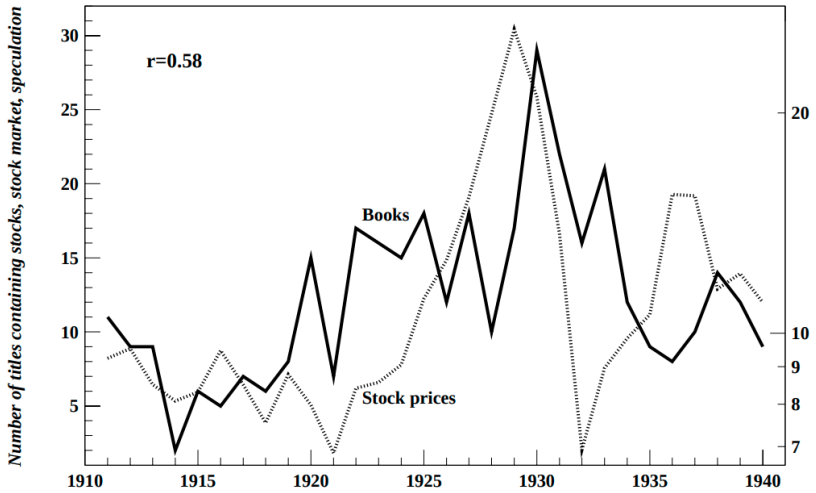


Figura 1.2: Confronto tra numero di libri pubblicati all'anno riguardanti azioni, mercato azionario e speculazione e indice sintetico del mercato azionario dal 1911 al 1940. [Roehner and Sornette (2000)]

bolla speculativa, c'è un interesse crescente del pubblico per il mercato finanziario e/o l'asset considerato. L'interesse pubblico per un determinato asset può essere quantificato utilizzando diverse metriche. Al giorno d'oggi lo strumento più utilizzato è Google Trends, una piattaforma che mostra l'andamento, per determinate parole chiave, delle ricerche sul motore di ricerca Google. Un esempio plastico di quanto sia fondamentale l'interesse del pubblico verso un determinato asset per generare le condizioni favorevoli allo sviluppo di una bolla è riportato nei paragrafi seguenti, alla figura 1.3. Per quanto concerne questo aspetto, rispetto alla crisi del '29, ci preme citare il lavoro di Roehner and Sornette (2000). In particolare gli autori confrontano il numero di libri pubblicati che avevano ad oggetto i mercati finanziari e il livello dei prezzi nell'arco temporale che va dal 1911 al 1940. In figura 1.2 è possibile apprezzare la relazione che lega queste due variabili. Tale relazione tra l'opinione pubblica e il prezzo di un asset si ripresenta, come accennato prima, anche nel caso della bolla speculativa su bitcoin.

1.1.4 Bolla delle dot-com

La bolla dot-com fu una bolla speculativa che ha riguardato le compagnie legate ad internet all'inizio del nuovo secolo. Dal 1995 a Marzo 2000, l'indice Nasdaq, che contiene tutte le maggiori compagnie high-tech statunitensi crebbe del 400 %. Al fine di comprendere meglio le dinamiche che portarono ad una crescita così marcata e diffusa delle compagnie high tech è necessario contestualizzare gli eventi che segna-

rono il passaggio al nuovo secolo, soprattutto riguardo le nuove tecnologie introdotte e l'avvento di Internet.

Tra il 1990 e il 1997, la percentuale di famiglie americane che possedevano un PC passò dal 15% al 35%. L'aumento fu dovuto principalmente al fatto che gli utilizzi possibili, grazie all'avvento di Internet, aumentarono e il PC passò quindi dall'essere un bene di lusso a un bene necessario per studenti, lavoratori e famiglie in generale. La diffusione dei personal computer e di Internet hanno segnato l'inizio di una nuova era e di un nuovo paradigma economico in cui il commercio e lo sviluppo industriale si basano principalmente sulla tecnologia informatica, sulla robotica, sulle telecomunicazioni fino ad arrivare ai giorni nostri con l'intelligenza artificiale applicata a campi quali la medicina e i mercati finanziari stessi. Negli Stati Uniti le condizioni economiche in quegli anni erano favorevoli grazie ad un tasso d'interesse relativamente basso e la riduzione delle tasse sui *capital gains*. Inoltre anche l'opinione pubblica era molto speranzosa riguardo i futuri benefici che l'era dell'informazione avrebbe portato alla società. Molti investitori erano perciò convinti che investire nelle società high-tech fosse un investimento fruttuoso. Le banche d'investimento, che guadagnavano ingenti commissioni sulle offerte pubbliche iniziali delle società legate all'high-tech, alimentarono questo ottimismo e incoraggiarono l'investimento nelle suddette compagnie. Questi fattori fecero lievitare i prezzi delle società quotate ma soprattutto, dato l'elevato ottimismo riguardo i profitti futuri, gli investitori persero completamente di vista indicatori tradizionalmente utilizzati per valutare i risultati di una società, come il PE ratio. Ad esempio, una società come VA LINUX, che aveva un rapporto earnings/share di -1.68, era quotata a 40$, cioè quanto una società come Ford, sicuramente molto più solida dal punto di vista patrimoniale e dei rendimenti. Le cause alla base delle valutazioni fallaci degli investitori riguardo le compagnie high-tech rispetto a quelle della old economy sono state trattate in dettaglio in Malkiel (1999). L'autore sostiene che gli investitori si aspettassero un incremento enorme degli utili delle società dot-com grazie ad Internet e alla new-economy, rispetto a compagnie quali Ford, nonostante il rendimento per azione di Ford fosse maggiore rispetto a quello di compagnie quali Cisco. Infatti, nonostante un prezzo di mercato sostanzialmente simile ($60 Cisco e $55 Ford), il rendimento per azione di Ford era pari a $6 mentre quello di Cisco pari a $0.37. Inoltre Cisco aveva una capitalizzazione di mercato pari a $395 miliardi mentre Ford di soli $63. In sostanza, quindi, utilizzando la formula classica per la valutazione delle azioni, cioè quella in cui il rendimento atteso futuro di un azione è legato ai dividendi e al tasso di crescita, le valutazioni delle società high-tech si basavano su una potenziale crescita attesa dei rendimenti futuri, che compensava i dividendi scarsi e i risultati pessimi di tali società nel breve periodo. Questo suggerisce che le aspettative di un rendimento futuro, piuttosto che la realtà economica attuale, contano molto nel determinare le scelte di investimento

degli operatori economici e nel favorire lo sviluppo di bolle speculative. Il prezzo di un asset potrebbe crescere in modo anomalo guidato da aspettative irrealistiche degli investitori riguardo gli utili futuri. Tale circostanza, però, è facile da identificare ex-post ma non ex-ante.

1.1.5 Bolla Bitcoin

Il Bitcoin è una criptovaluta nata nel 2008 e il cui utilizzo su larga scala è cominciato a partire dal 2011. Il Bitcoin è una valuta decentralizzata, cioè svincolata dal sistema di controllo delle Banche Centrali e che può essere trasferita tramite una rete chiamata blockchain, senza l'intervento diretto di intermediari. Nel lavoro di Hileman and Rauchs (2017), gli autori hanno calcolato che nel 2017 c'erano circa 4 milioni di persone che possedevano un wallet per criptovalute, molti dei quali detenevano in portafoglio principalmente Bitcoin. Le ragioni della diffusione a macchia d'olio di Bitcoin sono molteplici, tra cui:

- **Ideologia:** Molti degli utenti di Bitcoin e delle criptovalute in generale sostengono che queste, data l'assenza di intermediari tradizionali, permettano di smarcarsi dal tradizionale circuito finanziario passando ad un circuito composto da utenti e con minori costi di transazione, e, soprattutto nelle prime fasi, non sottoposto a stringenti controlli da parte degli Stati nazionali.
- **Investimento:** Gran parte degli utenti Bitcoin, invece, ha acquistato la criptovaluta a scopo d'investimento. Il prezzo di Bitcoin, infatti, è molto volatile e gli investitori, attratti dalla retorica di guadagni facili e senza rischio, hanno acquistato Bitcoin con la speranza di poter accrescere la propria ricchezza.
- **Condotte illecite:** Una quota marginale delle transazioni in Bitcoin, invece, riguarda condotte illecite quali finanziamento del terrorismo, riciclaggio di denaro e transazioni volte ad eludere i controlli sull'outflow di capitali di Stati nazionali come ad esempio la Cina.

Negli anni c'è stata una polarizzazione del dibattito sulle criptovalute. Le ragioni addotte dai fautori della blockchain vertono principalmente su basi ideologiche e di autonomia rispetto al circuito finanziario tradizionale.

Gli scettici, invece, sostengono che l'intero ecosistema delle criptovalute non sia sostenibile da un punto di vista ambientale, data l'enorme quantità di energia elettrica necessaria a mantenere in vita la blockchain. Inoltre, tra le motivazioni addotte da questi ultimi, rientrano anche la volatilità estrema del prezzo e soprattutto il fatto

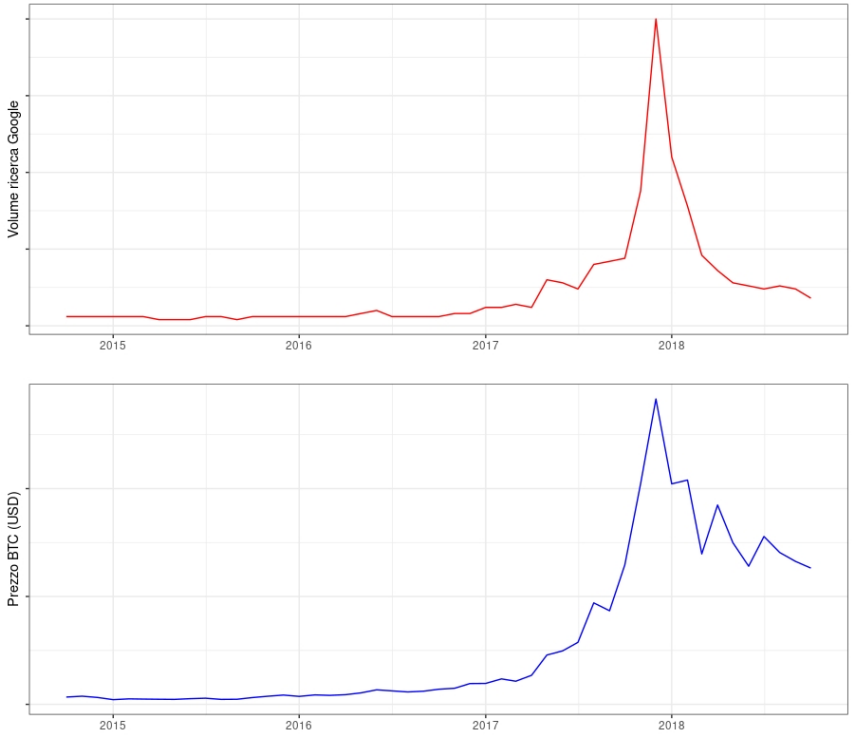


Figura 1.3: Confronto tra volume di ricerca della keyword "bitcoin" sul motore di ricerca Google (pannello superiore) e prezzo di bitcoin in USD (pannello inferiore). È chiaramente visibile il legame tra l'interesse del pubblico per l'asset considerato e il valore dell'asset stesso. La correlazione di Pearson per le due serie storiche è pari a 0.83

che le criptovalute si prestino ad essere utilizzate per scopi illeciti e per truffe (schema "Ponzi"). Infine, la CFTC (Commodity Futures Trading Commission) e la SEC (Securities and Exchange Commission) americana hanno più volte segnalato come gli exchange su cui vengono scambiate le criptovalute siano un ambiente favorevole per trader con ingenti disponibilità di capitale e che hanno come obiettivo quello di porre in essere condotte manipolative del prezzo a danno degli investitori retail. Lungi dall'essere esaurito, il dibattito sui pro e i contro delle criptovalute è ancora in corso e presenta numerosi spunti di riflessione, sebbene molti governi abbiano intenzione di introdurre versioni digitali delle proprie valute per controllare meglio il sistema dei pagamenti e il movimento dei capitali. Per quanto attiene a questo lavoro, ci limiteremo a sottolineare che anche il Bitcoin, più volte nel corso della sua vita, è stato soggetto all'emergere di bolle speculative, anche molto consistenti, e conseguenti crash. Ad esempio, a seguito del crash di Dicembre 2017, il Bitcoin perse circa 200 miliardi di capitalizzazione in sei settimane, con una perdita cumulata sul prezzo, dal massimo, di circa il 41%. Molti studi sono stati condotti sulle propietà statistiche del Bitcoin e sulla distribuzione dei rendimenti, confrontando queste ultime con quelle di asset tradizionali, in particolare USD e EUR.

Ad esempio, in Lukáš and Taisei (2017), gli autori hanno calcolato la varianza realizzata del Bitcoin e verificato che essa è significativamente maggiore di quella di altre valute fiat, cioè valute a corso legale. In Osterrieder and Lorenz (2017) gli autori analizzano le code della distribuzione dei rendimenti del Bitcoin e evidenziano come queste siano caratterizzate da valori molto più estremi rispetto a quelli delle classiche valute fiat. Oltre che rispetto alle proprietà econometriche del prezzo di Bitcoin, a lungo si è discusso anche riguardo l'attenzione crescente del pubblico verso questa tipologia di valuta digitale. In particolare, nel lavoro di Kristoufek (2013), gli autori trovano una relazione statistica significativa tra le query di ricerca su Google collegate al mondo delle criptovalute e la loro capitalizzazione complessiva di mercato. In figura 1.3 è possibile notare la relazione tra la query di ricerca "bitcoin" su Google Trends e il prezzo di Bitcoin espresso in dollari. Studi più approfonditi riguardo tali dinamiche, ad esempio Glaser et al. (2014) e Kim et al. (2017), hanno individuato relazioni tra il prezzo del Bitcoin e la diffusione dell'argomento in forum online scoprendo come una larga fetta di utenti Bitcoin siano in realtà piccoli investitori attratti dalla prospettiva di ingenti guadagni. Questa è una caratteristica peculiare evidenziata nei paragrafi precedenti riguardo altre bolle speculative storiche.

1.2 Proprietà statistiche

I mercati finanziari sono caratterizzati da periodi di relativa calma e periodi turbolenti con volatilità elevata. Soprattutto negli anni recenti, i mercati finanziari hanno risposto in modo più violento a shock esogeni ed endogeni: ad esempio nel caso della recente pandemia da Covid, l'S&P 500 ha sperimentato la discesa più profonda mai vista se rapportata alla velocità con cui è avvenuta. In alcuni lavori, ad esempio quello di Anadu et al. (2020), si fa notare come il passaggio graduale da strategie di investimento attive a strategie passive abbia comportato un aumento della volatilità, soprattutto su azioni di società incluse in indici quali l'S&P 500. I valori estremi dei rendimenti degli indici e del mercato in generale vanno contestualizzati per poter operare un confronto tra ciò che è ordinario e ciò che è straordinario, nel nostro caso le bolle speculative. Supponiamo di analizzare rendimenti che si distribuiscono secondo una variabile casuale di Gauss, cioè una Normale con media 0 e deviazione standard σ. In tal caso è facile calcolare la probabilità di osservare un certo rendimento.

Dalla tabella 1.1 si evince come per osservare un rendimento giornaliero maggiore del 3% sarebbe necessario attendere circa 1.5 anni. Un rendimento maggiore del 4%, invece, si presenterebbe tipicamente ogni 63 anni e così via. Questo esempio ci serve a chiarire ciò che nel mondo Gaussiano è considerato normale e ciò che invece è "anormale". Alla luce di quanto esposto sopra, potremmo considerare i rendimenti

Tabella 1.1: Quanto è probabile osservare un rendimento maggiore di $X \times \sigma$ per una variabile casuale Normale? La prima colonna mostra i valori di X da 1 a 10. La seconda esprime la probabilità che il valore assoluto del rendimento sia maggiore di $X \times \sigma$. La terza colonna esprime la probabilità in termini di giorni che è necessario attendere mediamente per osservare l'evento e la quarta colonna esprime la precedente in giorni dell'anno (ogni anno è composto da circa 250 *trading days*). [Sornette (2009)]

X	Probabilità	Uno ogni N eventi	Tempo necessario al verificarsi dell'evento
1	0.317	3	3 giorni
2	0.045	22	1 mese
3	0.0027	370	1.5 anni
4	6.3×10^{-5}	15787	63 anni
5	5.7×10^{-7}	1.7×10^{6}	7 millenni
6	2.0×10^{-9}	5.1×10^{8}	2 milioni di anni
7	2.6×10^{-12}	3.9×10^{11}	1562 milioni di anni
8	1.2×10^{-15}	8.0×10^{14}	3 miliardi di anni
9	2.3×10^{-19}	4.4×10^{18}	17721 miliardi di anni
10	1.5×10^{-23}	6.6×10^{22}	260 milioni di miliardi di anni

giornalieri del -22.6% e del del +9.7%, fatti segnare dall'indice Dow Jones rispettivamente il 19 Ottobre e il 21 Ottobre 1987, come eventi anormali: tali rendimenti, infatti, è improbabile che si verifichino in un lasso di tempo così stretto stando al modello gaussiano. Il fatto che essi però si siano verificati ci fa capire quanto il mercato possa deviare significativamente dalla "norma". Quando ciò accade è possibile osservare eventi che vengono definiti formalmente *outliers*.

Nella realtà, però, i rendimenti giornalieri degli indici azionari non si distribuiscono secondo una variabile casuale Normale. Nel modellare l'andamento di questi indici, infatti, vengono utilizzate distribuzioni differenti come ad esempio la distribuzione T-Student. La distinzione tra eventi normali e eventi "anormali" dipende quindi necessariamente dalla distribuzione probabilistica scelta per l'analisi. Definire quale sia, in termini assoluti, la distribuzione da preferire non è un problema di facile risoluzione e il dibattito scientifico in merito è tuttora aperto. Tuttavia, data la complessità della serie storica dei rendimenti, sarebbe necessario identificare ulteriori strutture caratteristiche di quest'ultima per poterne studiare a pieno il comportamento. Nel lavoro di Johansen and Sornette (2002) gli autori studiano il problema da un punto di vista differente rispetto alla sola distribuzione dei rendimenti. Gli autori sostengono che studiare i *drawdowns* permetta di misurare meglio il livello di rischio insito nel mercato finanziario rispetto a indici comunemente usati quali varianza, VaR e altre misure di rischio.

Un drawdown è definito come una discesa persistente del prezzo che avviene su più giornate di contrattazione. Un drawdown è quindi la perdita cumulata su più giorni

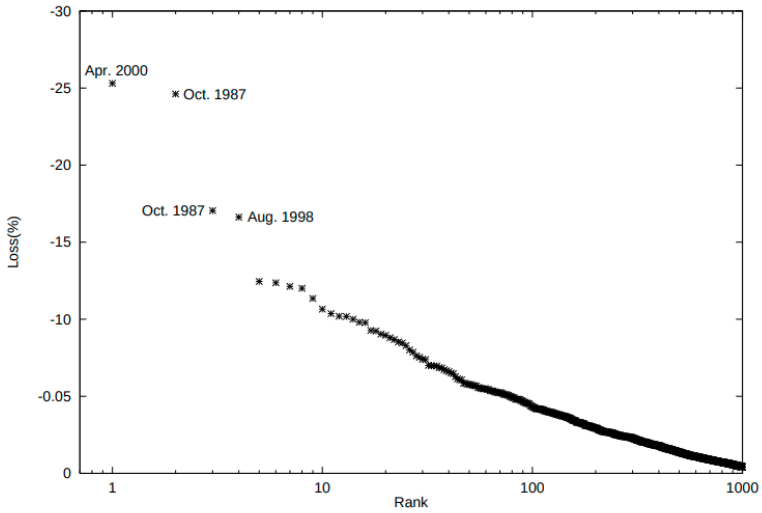


Figura 1.4: Drawdowns del Nasdaq Index, dal 1971 al 18 Aprile 2000, ordinati dal drawdown più grande (Rank=1) a quello più piccolo (Rank=1000). [Johansen and Sornette (2000)]

dall'ultimo massimo al minimo successivo del prezzo. Questo approccio permette di misurare direttamente la perdita cumulata che un investimento potrebbe subire.

La distribuzione dei drawdowns cattura la relazione che lega discese successive del mercato e pertanto si distingue dalla distribuzione dei rendimenti poiché quest'ultima non tiene conto della posizione relativa dei rendimenti sulla scala temporale ma si focalizza solo sulla frequenza con cui si verificano. Consideriamo ad esempio una situazione ipotetica in cui si verifichi un crash del 30% in tre giorni con un rendimento giornaliero costante del -10%. Se esaminassimo la distribuzione dei rendimenti avremmo solamente informazioni riguardo il fatto che ci sono stati 3 giorni in cui il rendimento è stato pari a -10% ma nulla sapremmo rispetto al fatto che questi rendimenti sono relativi a giorni consecutivi.

Gli autori quindi utilizzano la definizione di drawdowns per identificare, in modo generale, quelli che sono gli eventi rari, o *outliers*, verificatisi sui mercati finanziari e per sottolineare come questi eventi siano unici nel loro genere. Nel seguito proporremmo l'esempio riguardante l'indice Nasdaq estratto dal lavoro di Johansen and Sornette (2000).

Il *rank plot* in figura 1.4 permette di identificare i drawdowns più estremi verificatisi sull'indice Nasdaq nel periodo considerato. I quattro drawdowns più estremi si distinguono chiaramente rispetto al resto e corrispondono rispettivamente al crash di Aprile 2000, quello di Ottobre 1987, uno shock secondario dell'Ottobre 1987 e il crash

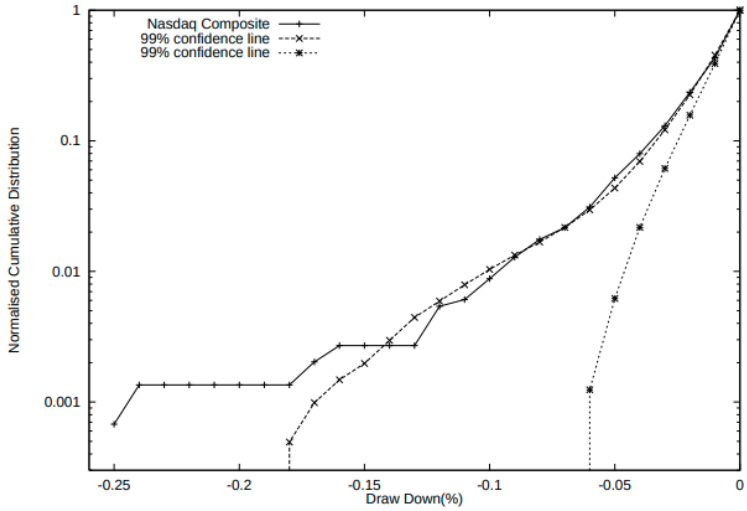


Figura 1.5: Funzione di ripartizione dei drawdowns su indice Nasdaq dal 1971 ad Aprile 2000. L'intervallo di confidenza al 99% è costruito mediante ricampionamento bootstrap dei rendimenti. [Johansen and Sornette (2000)]

di Agosto 1998. In una fase successiva del lavoro, gli autori hanno implementato una procedura di ricampionamento bootstrap sulla serie dei rendimenti generando 1000 serie sintetiche. Precisiamo che tale procedura distrugge la struttura di correlazione presente tra i rendimenti ma ne mantiene inalterata la distribuzione per quanto detto in precedenza. Per approfondire gli aspetti tecnici della procedura bootstrap si rimanda il lettore al lavoro di Chernick (2011). Gli autori hanno poi confrontato la funzione di ripartizione dei drawdowns reali rispetto a quella delle 1000 serie sintetiche generate dalla procedura bootstrap, si veda 1.5.

Ciò che più ci preme far notare in figura 1.5 è che la distribuzione reale si discosta dall'intervallo di confidenza per drawdowns superiori al 15% e tale comportamento, per gli autori, avvalora la tesi secondo la quale i quattro eventi di cui si è discusso sopra rappresentino effettivamente eventi unici nel loro genere, o meglio degli *outliers*. Tra le serie storiche sintetiche generate attraverso la procedura bootstrap, solo 776 presentano un drawdown maggiore del 16.5%, 13 ne presentano due e nessuna ne presenta quattro. Ciò significa che, data la distribuzione dei rendimenti considerata, avremmo una probabilità pari all' 8% di osservare un solo drawdown maggiore del 16.5% e in pratica alcuna possibilità di osservarne tre o più. Detto ciò è possibile rigettare l'ipotesi per cui i quattro *outliers* identificati sul Nasdaq siano generati dal caso, con un livello di confidenza del 99.99%.

In sintesi, in Johansen and Sornette (2000) e Johansen and Sornette (2002), gli autori asseriscono che:

- Gran parte dei drawdowns e drawups sono stati modellati con successo utilizzando una distribuzione esponenziale (ipotizzando indipendenza tra i rendimenti) o mediante una generalizzazione di tale modello definita *stretched exponential model*. Approssimativamente l' 1-2% dei drawdowns più grandi, invece, non sono spiegabili attraverso i suddetti modelli.

- La metà delle serie storiche analizzate presenta *outliers* anche per i drawups. Essi hanno però proprietà statistiche differenti rispetto ai drawdowns.

- Per l'azionario, i drawups maggiori del 15% si presentano due volte più spesso rispetto ai drawdowns.

Un'ulteriore conferma del fatto che i giorni di *rally* o *crash* siano significativamente diversi dalle normali giornate di contrattazioni arriva dal lavoro di Lillo and Mantegna (2000). Nello specifico, gli autori analizzano la distribuzione dei rendimenti giornalieri di una serie di titoli scambiati sul NYSE, dal 1987 al 1988, durante giornate in cui si sono verificati importanti movimenti di mercato. Questa distribuzione permette di avere un'idea generale rispetto l'attività del mercato in quel determinato giorno di contrattazioni. In particolare, le giornate caratterizzate da forti oscillazioni del prezzo presentano una distribuzione dei rendimenti con una forma diversa rispetto a giornate "tranquille". La figura 1.6 mostra 200 distribuzioni dei rendimenti, una per ogni giorno di contrattazione. La variabile *trading day* è centrata a 0 il giorno del crash (19 Ottobre 1987), di conseguenza valori negativi indicano i giorni precedenti e viceversa per valori positivi. È possibile notare come, nei giorni della crisi, la distribuzione dei rendimenti trasla verso valori negativi e positivi a seconda delle varie giornate di contrattazione.

Gli autori, inoltre, non si limitano all'analisi di cui sopra ma analizzano anche le singole distribuzioni relative ai giorni di contrattazione vicini al 19 Ottobre. Essi sostengono che tale scelta deriva dall'esigenza di capire se nei giorni *estremi* sia solo la media della distribuzione a cambiare o anche la sua forma. Da questa analisi deriva la figura 1.7. Dopo aver selezionato 18 giorni di contrattazione, rispettivamente i 9 con i rendimenti positivi più elevati e i 9 con quelli negativi più elevati, gli autori analizzano nel dettaglio le proprietà di queste distribuzioni.

La figura 1.7a mostra che nei giorni di crash le distribuzioni di probabilità sono connotate da un picco su valori negativi del rendimento e inoltre presentano un'asimmetria evidente con una coda positiva più ripida rispetto a quella negativa. Un comportamento opposto, invece, caratterizza i giorni di rally. In questi casi, quindi, l'asimmetria è opposta e le code negative delle distribuzioni sono più ripide rispetto alle code positive. In questo paragrafo abbiamo passato in rassegna alcuni lavori che

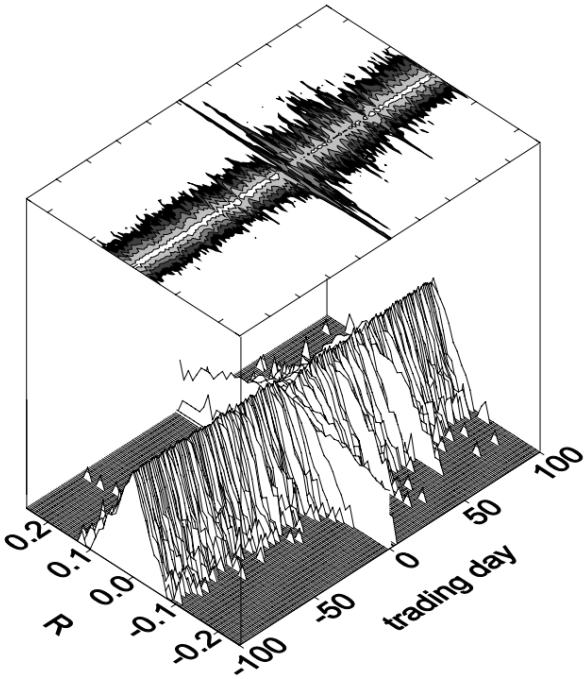


Figura 1.6: Contour plot e surface plot dei rendimenti complessivi su 200 giornate di contrattazioni, con intervallo temporale centrato al 19 Ottobre 1987. L'asse z è la densità di probabilità, in scala logaritmica. Le aree più chiare del contour plot rappresentano i punti in cui la probabilità è più elevata. [Lillo and Mantegna (2000)]

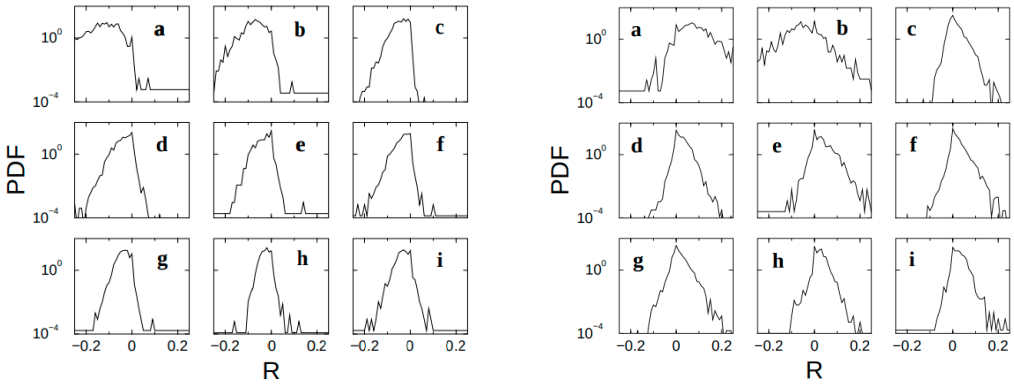


Figura 1.7: Grafico su scala log-lineare dei rendimenti complessivi rispetto ai giorni con maggior rendimento positivo 1.7b e negativo 1.7a. [Lillo and Mantegna (2000)]

dimostrano come gli eventi collegati alle bolle speculative siano *outliers*, cioè eventi con proprietà statistiche completamente diverse rispetto a quelle delle giornate normali di contrattazione. Nel seguito passeremo in rassegna alcuni approcci proposti in letteratura per modellare questa tipologia di fenomeni.

1.3 Approcci proposti in letteratura

Al fine di identificare le bolle speculative e poter intervenire in tempo sono stati sviluppati numerosi modelli. È opportuno precisare che i mercati finanziari moderni sono sistemi complessi, in cui numerosi agenti interagiscono con stili d'investimento e informazioni differenti. Questo rende complesso sviluppare un modello d'interazione che tenga conto delle dinamiche sopra descritte, soprattutto perché tali interazioni generano dinamiche non lineari. Nel seguito presenteremo alcuni approcciproposti in letteratura per modellare le bolle speculative.

1.3.1 Processi AR non causali

Un approccio proposto in letteratura nel lavoro di Gouriéroux and Zakoian (2016), riguarda l'utilizzo di modelli AR non causali per studiare serie storiche con dinamiche non lineari e cicli asimmetrici, caratteristiche proprie delle bolle speculative. Lo scopo del lavoro è analizzare le dinamiche di processi *heavy-tailed non causal* AR(1) e testare l'utilità di tale approccio per modellare serie storiche finanziarie, che come noto presentano un andamento non regolare, con la presenza di discontinuità. Dato un processo MA definito come segue:

$$Y_t = \sum_{h=-\infty}^{\infty} a_h \epsilon_{t-h} \tag{1.1}$$

con ϵ_t una successione di variabili casuali *i.i.d.* a valori reali e a_h un vettore di coefficienti che, per qualsiasi $s \in (0, 1)$, soddisfa:

$$\begin{aligned} &\mathbb{E}|\epsilon_t|^s < \infty \\ &\sum_{-\infty}^{\infty} |a_h|^s < \infty \end{aligned} \tag{1.2}$$

Quando $a_h = 0$ per ogni $h < 0$, il processo Y_t è *purely causal* (o *backward looking*) mentre quando $a_h = 0$ per ogni $h > 0$ allora Y_t è *purely non-causal* (o *forward looking*).

Esiste poi una terza tipologia di modello connotato sia da una componente *causal* che *non-causal* che viene definito *mixed process* e che possiede interessanti proprietà.

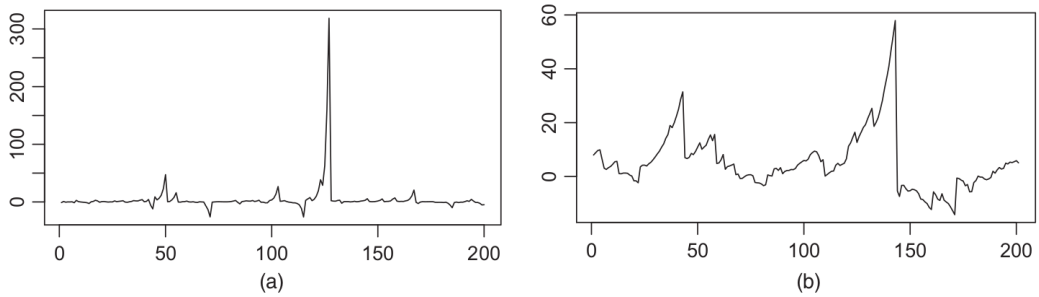


Figura 1.8: Simulazione del modello al punto 1.3 con $\alpha = 1$, $\beta = 0$, $\sigma = 0.5$. Ciò che distingue le due simulazioni è il parametro ρ: per il grafico al punto (a) $\rho = 0.5$ e per (b) $\rho = 0.9$ [Gouriéroux and Zakoian (2016)]

Gli autori, in primis, considerano il seguente processo *non causal* AR(1)

$$Y_t = \rho Y_{t+1} + \epsilon_t, \qquad |\rho| < 1, \qquad \epsilon_t \sim \mathcal{S}(\alpha, \beta, \sigma, 0) \tag{1.3}$$

dove per $\mathcal{S}(\alpha, \beta, \sigma, \mu)$ si intende una distribuzione α-*stable*. $\alpha \in]0, 2]$ è il parametro di stabilità, $\beta \in [-1, 1]$ è il parametro di asimmetria, $\sigma \in [0, +\infty[$ è un parametro di scala e $\mu \in \mathbb{R}$ è un *location parameter*.

Sotto opportune condizioni, la soluzione strettamente stazionaria del modello al punto 1.3 è:

$$Y_t = \sum_{h=0}^{\infty} \rho^h \epsilon_{t+h} \tag{1.4}$$

Gli autori, nel seguito del lavoro, analizzano le proprietà statistiche sia di un processo *non-causal* con errori estratti sia da una distribuzione di Cauchy sia da una distribuzione di Lévy. Il modello AR non causale con errori estratti da una distribuzione di Cauchy si è rivelato utile per modellare andamenti locali esponenziali e inoltre risulta essere stazionario e una martingala, pertanto possiede proprietà molto utili dal punto di vista teorico. Il lavoro di Gouriéroux and Zakoian (2016) fa notare come la dinamica esplosiva dei suddetti processi sia causata da valori estremi generati da distribuzioni di probabilità *fat tailed*. Un valore estremo di ϵ_t sulla coda positiva della distribuzione genera un "salto" di Y_t preceduto da un trend in crescita. Al contrario, invece, un valore estremo sulla coda negativa della distribuzione genera un andamento simmetrico che può essere utilizzato per modellare spirali deflazionistiche.

Il modello può quindi essere adattato, in base alle esigenze, per studiare sia le bolle

speculative (utilizzando una distribuzione *fat right-hand tail* definita su $\mathbb{R}^+$) che le spirali deflazionistiche (utilizzando una distribuzione *fat left-hand tail* definita su $\mathbb{R}^-$). Stando ai test effettuati nel lavoro sopracitato, il modello si rivela utile per descrivere serie storiche non lineari e che presentano un andamento esponenziale locale. Infine gli autori proprongono alcuni approcci predittivi per identificare bolle speculative in anticipo mediante l'utilizzo dei modelli sopra citati. Nel corso di questo lavoro utilizzeremo il modello LPPLS che verrà introdotto nel seguito. In letteratura non è stato ancora operato un confronto tra modelli *non-causal* e il modello LPPLS, che introdurremo nel seguito; Pertanto non è possibile affermare con precisione quale dei due modelli sia preferibile all'altro in termini generali. All'approccio *non-causal* è stato preferito il modello LPPLS perchè si è rivelato utile in numerosissimi casi per identificare bolle speculative in formazione con un sufficiente grado di anticipo, come si discuterà nel paragrafo successivo.

1.3.2 Modello LPPLS

Un'altra soluzione proposta in letteratura riguarda il modello Log Periodic Power Law Singularity (LPPLS) introdotto da Johansen et al. (2000). Gli autori basano la costruzione del modello su alcune ipotesi. La prima attiene alla caratteristica principale delle bolle speculative, cioè il fatto che esse siano causate da un meccanismo imitativo tra agenti che li porta a posizionarsi sullo stesso lato del mercato. Questo meccanismo, una volta giunti in uno stato critico, è la causa principale dei crash. In questo caso gli autori ci tengono a precisare che l'aggettivo "critico" non è casuale ma deriva dalla caratteristica dei sistemi complessi studiati in fisica, che possono giungere ad uno stato critico, cioè il divergere di una variabile del sistema ad infinito. Dato che le bolle speculative sono eventi complessi con andamenti esplosivi, gli autori si chiedono se possa esserci un collegamento tra questo andamento e l'esistenza di punti critici nella complessa struttura del mercato finanziario. Le assunzioni principali su cui si basa il modello sono le seguenti: un crash è causato da un meccanismo imitativo tra trader che porta il prezzo dell'asset considerato a crescere con andamento super-esponenziale, cioè con rendimenti crescenti. Se tale processo imitativo si diffonde si raggiunge un punto critico del sistema in cui molti trader cambiano improvvisamente posizione sul mercato causando un crash. La relazione che lega le suddette tendenze imitative e l'agire dei *noise trader*, il cui comportamento risulta casuale, necessita di una formulazione probabilistica del modello. Infatti gli autori precisano che un crash, nel framework del modello, è un evento casuale, che quindi può verificarsi o meno e la cui probabilità verrà definita come *hazard rate*, cioè la probabilità per unità di tempo che un crash avvenga. La caratterizzazione dell'hazard rate come una probabilità permette un'importante considerazione: per gli investitori è razionale

continuare ad investire nell'asset considerato poiché vengono ricompensati da un rendimento maggiore legato al rischio crescente che una correzione possa avere luogo. Gli autori, in numerosi lavori, sostengono che la crescita di una bolla speculativa è accompagnata da oscillazioni log-periodiche e da una crescita super esponenziale del prezzo (rendimenti crescenti). Il concetto di oscillazioni log-periodiche è legato alla fisica dei sistemi complessi e ai frattali, come verrà spiegato nei capitoli successivi. Il modello LPPLS permette di identificare le oscillazioni log-periodiche e la crescita super esponenziale e fornisce quindi un valido strumento per studiare questa tipologia di fenomeni. Gli autori, inoltre, hanno messo in evidenza come tali oscillazioni log-periodiche siano una caratteristica presente in molte bolle speculative storiche, quali ad esempio la crisi del '29, quella dei tulipani e molte altre. Il dipartimento di *Entrepreneurial Risks* dell' ETH di Zurigo ha istituito, sotto la supervisione del Prof. Sornette, l'osservatorio sulle crisi finanziarie [2]. Il compito dell'osservatorio è quello di pubblicare mensilmente un report dettagliato che mostra il livello di bolle speculative sui mercati mondiali, dai mercati azionari a quelli delle criptovalute, passando per le commodities. Inoltre il dipartimento cura anche un portale, disponibile al seguente link [3], in cui è possibile consultare il livello di rischio insito in vari asset giorno dopo giorno.

Per quanto riguarda l'effettiva capacità predittiva del modello, molti studi sono stati condotti e in numerose occasioni si è rivelato utile per identificare in anticipo lo sviluppo e il successivo declino di molte bolle speculative. Ad esempio è stato applicato alla bolla sull'azionario di Shanghai nel 2015 Sornette et al. (2015), al mercato immobiliare di Las Vegas Zhou and Sornette (2007) , alla bolla immobiliare del 2000-2003 nel Regno Unito Zhou and Sornette (2003) e in molti altri casi.

La formulazione analitica del modello verrà trattata nel capitolo seguente. Nei capitoli successivi, poi, verranno introdotti una serie di indicatori, basati sul modello LPPLS, utili a identificare nella pratica bolle speculative in formazione e la scala temporale nella quale queste bolle si stanno formando.

[2] https://er.ethz.ch/financial-crisis-observatory.html
[3] http://tasmania.ethz.ch/pubfco/fco.html

Capitolo 2

Il modello LPPLS

Il modello LPPLS, originariamente conosciuto come modello Johansen-Ledoit-Sornette (JLS) prese il nome dagli autori che per primi lo proposero come modello per individuare bolle speculative in formazione sui mercati. Nel corso del seguente capitolo procederemo ad una trattazione dettagliata delle ipotesi e delle caratteristiche del modello, rifacendoci, in modo particolare, al lavoro di Johansen et al. (2000)

2.1 Ipotesi del modello

Consideriamo un asset puramente speculativo, che non paga cioè dividendi periodici e ipotizziamo che gli agenti abbiano aspettative razionali. Nei mercati finanziari possono esserci molti altri fattori che influenzano l'agire degli investitori, come ad esempio le aspettative sui tassi d'interesse, le asimmetrie informative, l'avversione al rischio ed altri. Nello sviluppare il modello questi fattori verranno ignorati per semplificare la trattazione. In un mercato efficiente in cui è valida l'ipotesi di aspettative razionali, il logaritmo del prezzo dell'asset $p(t)$ è una martingala, cioè:

$$\forall t^{'} > t \quad \mathbb{E}_t[p(t^{'})] = p(t) \tag{2.1}$$

dove $p(t)$ indica il prezzo dell'asset al tempo t e $\mathbb{E}_t$ è l'operatore valore atteso condizionato all'informazione disponibile al tempo t. L'ipotesi di martingala è intrinsecamente legata alla teoria dei mercati efficienti, introdotta da Fama (1970). Secondo Fama, infatti, il mercato è efficiente se non è possibile generare rendimento sfruttando l'informazione passata riguardo il titolo in questione. Pertanto tutta l'informazione

disponibile risulta già incorporata nel prezzo al tempo t.

Più in generale è possibile interpretare $p(t)$ come la somma di due componenti, cioè:

$$p(t) = p_0(t) + p^*(t) \tag{2.2}$$

dove p_0 e p^* rappresentano rispettivamente il valore intrinseco dell'asset e la componente in eccesso dovuta alla bolla.

L'ipotesi delle aspettative razionali sopra riportata concerne le modalità con le quali gli agenti formano le proprie aspettative riguardo le variabili di mercato. Secondo tale teoria gli operatori utilizzano le informazioni in modo efficiente senza commettere errori sistematici riguardo le previsioni sul futuro andamento delle variabili economiche. In altri termini, gli individui utilizzano tutte le informazioni di cui possono disporre e nel modo più efficiente possibile. È bene precisare che l'ottimalità delle previsioni riguarda gli individui nel loro complesso e non singolarmente. Infatti, sebbene un singolo individuo possa formulare valutazioni non efficienti, tali valutazioni una volta aggregate sono ottimali. Una condizione che deriva da questa ipotesi è quella di non arbitraggio, che utilizzeremo per derivare il modello. Al framework definito finora è aggiunta una variabile casuale esogena che incorpora la probabilità di un *crash*, ovvero un calo repentino del prezzo.

Sia j_t uno *jump process* che assume valore 0 prima del crash e valore 1 dopo il crash al tempo t_c. La funzione di ripartizione della variabile casuale t_c è $Q(t)$, con relativa funzione di densità $q(t) = \partial Q/\partial t$. Definiamo ora la probabilità, per unità di tempo, che il crash avvenga nell'intervallo successivo, se non ancora avvenuto. Tale probabilità dipende necessariamente da t e verrà indicata come *crash hazard rate*: $h(t) = q(t)/[1 - Q(t)]$. Dato che $h(t)\partial t$ è la probabilità che il crash avvenga nell'intervallo di tempo $[t, t + \partial t]$, è possibile ricavare il valore atteso condizionato $\mathbb{E}_t[\partial j] = h(t)\partial t$.

Ipotizzando per semplicità che il prezzo dell'asset durante un crash diminuisca di una percentuale fissa $k \in (0, 1)$, è possibile descrivere l'andamento del prezzo come segue:

$$\partial p = \mu(t)p(t)\partial t - \kappa p(t)\partial j + \sigma(t)p(t)\partial W \tag{2.3}$$

dove $\mu(t)$ è il rendimento, $\sigma(t)$ è la volatilità e ∂W è l'incremento infinitesimo di un processo di Wiener con media nulla e varianza pari a ∂t. È possibile derivare $\mathbb{E}_t[\partial p]$ utilizzando $h(t)$:

$$\mathbb{E}_t[\partial p] = \mu(t)p(t)\partial t - \kappa p(t)h(t)\partial t \tag{2.4}$$

La condizione di non arbitraggio e quella delle aspettative razionali implicano che $\mathbb{E}_t[\partial p] = 0$ e di conseguenza che $\mu(t) = kh(t)$.

L'equazione (2.3) indica come le componenti che determinano la variazione di prezzo, dato il periodo considerato, ipotizziamo un giorno, siano il rendimento giornaliero $\mu(t)$ e la probabilità che avvenga un crash $h(t)$.

Il modello JLS ipotizza l'esistenza di due tipologie di trader. I trader con aspettative razionali e i *noise trader*. Questo secondo gruppo di trader manifesta tendenze imitative. Il modello presuppone quindi che siano proprio tali tendenze a minare la stabilità del prezzo ed a determinare quello che abbiamo definito $h(t)$.

Sostituendo $\mu(t) = \kappa h(t)$ nell'espressione 2.3 e ponendo $\partial j = 0$ otteniamo:

$$\frac{\partial p}{p(t)} = \kappa h(t)\partial t + \sigma(t)\partial W \tag{2.5}$$

Risolvendo l'equazione differenziale stocastica al punto 2.5, si ottiene:

$$\mathbb{E}_t\left[\ln \frac{p(t)}{p(t_0)}\right] = \kappa \int_{t_0}^{t} h(t^{'})dt^{'} \tag{2.6}$$

È possibile notare che più la probabilità di crash aumenta, più il prezzo deve aumentare rapidamente affinchè la condizione di martingala sia soddisfatta. Questa condizione definisce, in modo formale, una caratteristica intrinseca dei mercati efficienti, cioè quella per cui gli investitori vogliono essere ricompensati da un rendimento più elevato per detenere un asset che potrebbe perdere valore rapidamente in seguito ad un crash. Tale dinamica è stata trattata approfonditamente in Blanchard (1979).

Infine è opportuno precisare che l'evento *crash* è esogeno. Questo significa che nessuno degli agenti coinvolti conosce a priori quando e se avverrà. Oltretutto anche l'hazard rate è una variabile esogena. Pertanto, nel modello come finora costruito, il crash è un evento casuale la cui probabilità è legata a fattori esterni e i prezzi si adattano successivamente a tale probabilità per garantire la condizione di non arbitraggio.

Per come è stata definita la dinamica del prezzo, appare evidente che $h(t)$ rappresenta

la probabilità che un gruppo numeroso di trader invii ordini di vendita al proprio broker simultaneamente, in modo tale da creare uno squilibrio sull'order book e far precipitare i prezzi velocemente, causando un crash. Ci chiediamo allora quale sia il fattore che potrebbe provocare una reazione di questo tipo da parte di molti operatori.

L'approccio adottato nel lavoro citato all'inizio di questo capitolo è quello di considerare i trader come parte di un network nel quale ogni trader influenza e viene influenzato solo dai trader a lui vicini. Nello specifico, se l'i-esimo agente è collegato a k agenti, allora le sue scelte sono influenzate dagli agenti a cui è collegato e viceversa. È presente, inoltre, una componente idiosincratica ascrivibile al singolo agente, su cui cioè gli altri agenti non hanno influenza. Come specificato nel capitolo precedente, gli agenti appartenenti a questa rete hanno un comportamento imitativo che li porta ad essere influenzati dalle opinioni degli agenti a loro vicini. Per una trattazione dettagliata dell'argomento si veda Boissevain (1979). Tale tendenza imitativa porta ordine nella rete poiché gli agenti vicini tendono ad assumere una posizione comune, mentre la componente idiosincratica conduce al disordine. Per quanto detto riguardo $h(t)$, un crash si verifica quando nella rete prevale l'ordine (posizionamento sulla sell side di molti agenti), mentre condizioni normali di mercato si verificano quando c'è disordine, quando cioè gli agenti inviano ai propri broker ordini buy e sell in modo casuale.

L'approccio sopra descritto presenta dei vantaggi. In particolare ci permette di asserire che il crash possa avvenire anche senza motivi razionali (news o eventi), ma come mera conseguenza di un meccanismo di coordinamento involontario dei trader, che assumono la stessa posizione sul mercato a causa di tendenze imitative. Sulla scorta di quanto detto finora risulta evidente come sia necessario discutere più approfonditamente l'hazard rate, poiché è proprio tale componente a giocare un ruolo fondamentale nel modellare le interazioni tra gruppi di agenti e l'influenza che tali interazioni hanno sul prezzo dell'asset.

In Johansen et al. (2000) gli autori affrontano il problema delle tendenze imitative e dell'hazard rate sia da un punto di vista macroscopico che microscopico, mettendo in evidenza come entrambi gli approcci abbiano alcune similitudini. È necessario precisare che la formulazione analitica del modello LPPL deriva direttamente dall'analisi micro-strutturale.

2.2 Approccio macro-strutturale

Nel seguito verranno richiamati concetti di meccanica statistica, in particolare la *mean field theory*, a tal proposito si veda Goldenfeld (2018). Secondo tale teoria è possibile modellare un processo imitativo mediante $h(t)$ con la seguente equazione differenziale:

$$\frac{\partial h}{\partial t} = C\ h(t)^{\delta} \qquad con \qquad \delta > 1 \tag{2.7}$$

dove C è una costante positiva mentre $\delta > 1$ è la misura media di interazioni tra i trader della rete meno uno. La *mean field theory* ci permette di rappresentare la complessità delle azioni di ogni singolo agente attraverso una misura "media" che tenga conto del problema nel suo complesso, prescindendo quindi dallo studio dei singoli effetti e semplificando di molto la complessità del problema. In tal senso, infatti, $h(t)$ è una misura complessiva che dipende dal grado di interazione tra i trader. È chiaro come una crescita delle interazioni tra i trader porti ad un aumento di $h(t)$. Integrando (2.7) otteniamo:

$$h(t) = \left(\frac{h_0}{t_c - t}\right)^{\alpha}, \qquad \alpha = \frac{1}{\delta - 1} \tag{2.8}$$

dove t_c è il tempo critico sulla base delle condizioni presenti al tempo t_0. La condizione $\delta > 1$ (e di conseguenza $\alpha > 0$) è necessaria affinchè $h(t)$ cresca quando $t \to t_c$. Inoltre la condizione $\alpha < 1$ assicura che il prezzo non diverga ad infinito al tempo t_c. Riscrivendo le condizioni precedenti otteniamo $2 < \delta < \infty$, e cioè che ogni agente deve essere collegato almeno ad altri due agenti.

Un'importante caratteristica dell'approccio appena introdotto è quella di poter essere utilizzato per modellare anche le cosiddette *self-fulfilling crisis*. Questa fattispecie si presenta quando una perdita di fiducia dei mercati nei confronti di un Paese causa una crisi che a valle giustifica la mancanza di fiducia a monte. Se gli investitori sono convinti che possa avvenire una crisi se non vengono portate a termine determinate riforme, allora sarà la stessa convinzione di questi ultimi a generarla. Un importante lavoro che studia il fenomeno adottando un approccio empirico è quello di De Grauwe and Ji (2013).

Adottando l'approccio utilizzato in precedenza, è possibile modellare tale fenomeno come segue:

$$\frac{\partial h}{\partial t} = Dp(t)^{\mu} \ , \ \mu > 0 \tag{2.9}$$

con D una costante positiva. Questa equazione, assieme alla 2.6, cattura l'effetto sopra descritto. La mancanza di fiducia da parte degli investitori, quantificata da $h(t)$, peggiora quanto più il prezzo di mercato si discosta dal suo valore intrinseco. Quando ciò accade, il prezzo di mercato stesso deve crescere per remunerare gli investitori del rischio maggiore a cui vanno incontro.

2.3 Approccio micro-strutturale

Per quanto riguarda l'approccio micro-strutturale, in Johansen et al. (2000) e Sornette (2009) gli autori ipotizzano diverse possibili strutture per il network di investitori di cui abbiamo parlato in precedenza. Ipotizziamo che ogni agente sia identificato da un numero intero $i = 1, ..., I$ e $N(i)$ sia il numero di agenti connessi all'agente i-esimo. Inoltre ogni agente può trovarsi in uno dei seguenti stati s_i: *buy* ($s_i = +1$) o *sell* ($s_i = -1$). Gli autori ipotizzano che lo stato dell'agente i-esimo sia il risultato del seguente processo di Markov:

$$s_i = sign\left(K \sum_{k \in N(i)} s_k + \sigma \epsilon_i\right) \tag{2.10}$$

dove la funzione *sign* è uguale a $+1$ se l'argomento è positivo e a -1 se l'argomento è negativo. Inoltre K è una costante positiva e ϵ_i è una variabile casuale *i.i.d.* con distribuzione normale standardizzata.

In questo modello è K a quantificare, in modo diretto, il grado di imitazione della rete, mentre σ governa quello che in precedenza abbiamo definito fattore idiosincratico. All'aumentare di K, quindi, il grado di ordine nella rete aumenta e viceversa. Vale l'opposto per quanto riguarda σ. Nel caso in cui prevalga l'ordine, gli agenti tenderanno ad imitare i propri vicini e questo comportamento si diffonderà nella rete, causando un crash.

È d'uòpo puntualizzare che il processo al punto 2.10 appartiene alla classe dei modelli stocastici dinamici utilizzati in fisica per descrivere il comportamento di particelle

che interagiscono tra loro, si veda a tal proposito Liggett (2012) e Liggett et al. (1997). Un esempio di tali applicazioni è il modello di Ising bidimensionale utilizzato per modellare le interazioni ferromagnetiche tra particelle, e, nel caso specifico, la transizione dal ferromagnetismo al paramagnetismo quando la temperatura cresce al di sopra della temperatura di Curie. Analogamente a quanto esposto in precedenza, anche nel modello di Ising le particelle possono trovarsi in due stati, corrispondenti alla carica elettrica della particella.

Così come per il modello di Ising esiste una soglia critica della temperatura (temperatura di Curie), oltre la quale si passa da ferromagnetismo a paramagnetismo, anche per il modello d'interazione al punto 2.10 esiste un valore soglia di K, che chiameremo K_c, che funge da spartiacque tra ordine e disordine nel network. Nel modello presentato per descrivere le interazioni tra agenti sul mercato, quando $K < K_c$ il disordine regna e la sensibilità del sistema è bassa. Quando il grado di imitazione nel sistema aumenta e K tende ad avvicinarsi a K_c, allora la forza imitativa provoca la formazione di cluster di agenti che si posizionano in modo analogo sul mercato, destabilizzando i prezzi. In seguito alla formazione del cluster, il mercato diventa più sensibile a sconvolgimenti dovuti a piccoli disturbi. Se, infine, $K > K_c$ allora l'imitazione è talmente forte da rendere una posizione (*buy* o *sell*) predominante nell'intera rete.

Introduciamo un ulteriore parametro, che chiameremo G. Tale parametro misura l'influenza che il sistema nel suo complesso esercita su un singolo agente. Possiamo aggiungere allora tale termine a (2.10) per ottenere:

$$s_i = sign\left(K \sum_{k \in N(i)} s_j + \sigma \epsilon_i + G\right) \tag{2.11}$$

Se definiamo la media degli stati s_i come $M = (1/I) \sum_{i=1}^{I} s_i$ allora è possibile dimostrare che, se $G = 0$, allora $E[M] = 0$ per simmetria. Ciò significa che se il termine G è uguale a zero gli agenti saranno omogeneamente distribuiti tra i due stati possibili. Se invece $G > 0$ gli agenti con $s_i = +1$ saranno di più rispetto a quelli nello stato opposto e viceversa per $G < 0$.

È bene precisare che il modello 2.11 definisce lo stato di un agente ad ogni istante temporale in funzione dell'istante precedente. Nell'istante successivo, quindi, lo stato dei singoli agenti può variare in funzione dei parametri sopra descritti. Infatti è possibile riscrivere l'equazione di cui sopra in modo più generale esplicitando la dipendenza delle variabili dal tempo.

$$s_i\ (t+1) = sign\Big(K \sum_{k \in N(i)} s_k(t) + \sigma\epsilon_i(t) + G\Big) \tag{2.12}$$

La *suscettibilità* del sistema è definita come:

$$\chi = \frac{\partial\big(E[M]\big)}{\partial G}\bigg|_{G=0} \tag{2.13}$$

La quantità χ può avere varie interpretazioni. Innanzitutto indica la sensibilità di M a una variazione dell'influenza globale G. Può anche essere vista come la varianza di M causata dagli shock idiosincratici dovuti a ϵ_i. Infine, considerando un sistema composto da due soli agenti, se forziamo un agente ad assumere un determinato stato, allora l'impatto che tale scelta avrà sull'altro agente sarà proporzionale a χ.

2.4 Modello per la rete di interazioni

Il modello micro strutturale (2.11) e il suo grado di *suscettibilità* dipendono dalla struttura stessa della rete. Gli autori propongono due tipologie di struttura. La prima è una semplice griglia bidimensionale in cui ogni agente rappresenta una parte della griglia stessa. Il secondo approccio, meno semplicistico, è quello di considerare tale rete come una struttura gerarchica, in particolare un *hierarchical diamond lattice*.

2.4.1 Griglia bidimensionale

Secondo tale approccio, la rete stessa è una griglia composta da quadrati, rappresentanti i singoli agenti. Ogni agente è collegato ad altri 4, cioè agli agenti a lui adiacenti. Il parametro rilevante per lo studio della *suscettibilità* del sistema è K/σ poiché misura la tendenza imitativa degli agenti.

Come detto in precedenza, esiste un punto critico K_c rispetto al quale valutare il grado di sensibilità del sistema. Se $K < K_c$ allora nel sistema regna il disordine e le forze imitative prevalgono solamente tra agenti vicini. In tal caso, quella che abbiamo definito come *suscettibilità* del sistema, χ, assume valori finiti.

Quando, invece, $K \to K_c$, il grado di ordine nel sistema aumenta e la forza imitativa prevale anche tra agenti distanti. Si formano quindi dei grandi cluster di agenti aventi la stessa opinione e questo fa aumentare il grado di *suscettibilità* del sistema. In tal caso χ tende ad infinito. Il comportamento sopra descritto dei sistemi complessi è studiato ampiamente in quella che viene definita *Teoria dei fenomeni critici*. È stato dimostrato che è possibile approssimare χ con la seguente *power law*:

$$\chi \approx A(K_c - K)^{-\gamma} \tag{2.14}$$

con A una costante positiva, $K_c = K(t_c)$. $\gamma > 0$ è chiamato *esponente critico* ed influenza direttamente la *suscettibilità* del sistema.

A questo punto gli autori ipotizzano che K abbia un andamento continuo che dipende dal tempo e che sia possibile sviluppare in serie di Taylor $K(t)$ nell'intorno di K_c e utilizzare tale risultato per approssimare $K_c - K(t)$. Quindi se $K(t_c) = K_c$, allora per ogni $t < t_c$

$$K_c - K(t) \approx C(t_c - t)^{-\gamma} \tag{2.15}$$

con C una generica costante positiva. Ipotizzando che l'hazard rate $h(t)$ si comporti allo stesso modo della *suscettibilità* del sistema χ nell'intorno di t_c, possiamo scrivere:

$$h(t) \approx B(t_c - t)^{-\alpha} \tag{2.16}$$

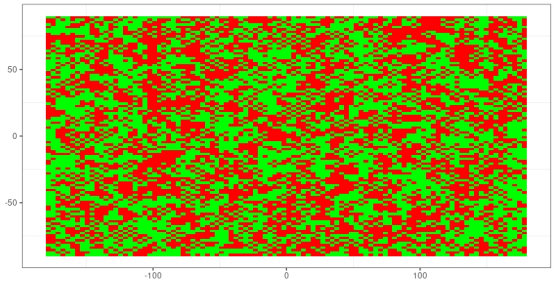
con B una costante positiva. L'esponente $\alpha \in (0, 1)$ poiché altrimenti il prezzo diverge ad infinito quando $t \to t_c$. È necessario precisare che nel seguito si considererà t_c come la moda della distribuzione del *time of the crash*, dato che la stessa variabile t_c è stocastica e non deterministica. Sostituendo l'equazione (2.16) in (2.6) otteniamo:

$$log[p(t)] \approx log[p_c] - \frac{\kappa B}{\beta}(t_c - t)^{\beta} \tag{2.17}$$

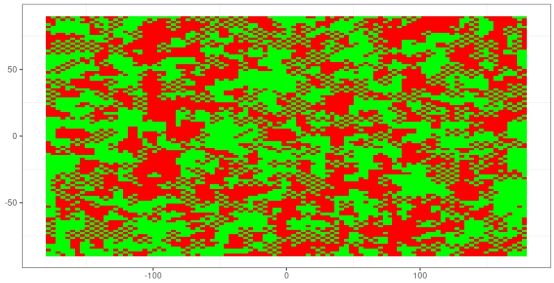
con $\beta = 1 - \alpha \in (0, 1)$ e p_c il prezzo al tempo t_c. È interessante notare come anche il logaritmo del prezzo segua una power law.

Nella figura seguente (2.1) appare chiaro, confrontando i risultati ottenuti in base a parametri diversi, quale ruolo i singoli parametri hanno nel determinare il comportamento dell'intera rete. Precisiamo che nel generare i grafici abbiamo ipotizzato che K sia costante. In primis, nelle figure 2.1a e 2.1b è possibile notare come il parametro K, per quanto detto in precedenza, quantifica il grado di imitatività, e quindi di sostenibilità del sistema. Nel caso considerato, infatti, il valore di K è minore del valore critico K_c. Pertanto, anche dopo diverse iterazioni, il sistema si stabilizza in una configurazione in cui sono presenti sia cluster di buyer che di seller senza che nessuno dei due prevalga sull'altro. Nelle figure 2.1c e 2.1d, invece, il parametro K è vicino al suo valore critico. Infatti, dopo sole 3 iterazioni, il sistema presenta una forte presenza di buyer. Infine, confrontando la figura 2.1d con la 2.1f, è possibile apprezzare gli effetti che la variabile G esercita sul sistema. È proprio tale variabile, infatti, a determinare se, dopo n istanti temporali, siano i seller a prevalere 2.1f o i buyer 2.1d. Per chi fosse interessato a saggiare approfonditamente le conseguenze che le variabili esercitano sul sistema è possibile sperimentare varie configurazioni utilizzando la web app [1] (R Shiny) appositamente sviluppata.

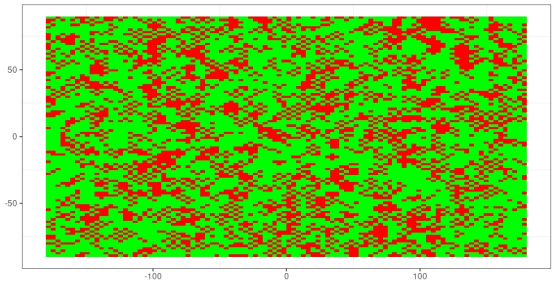
[1] https://sgargiulo14.shinyapps.io/Herding_simulator/



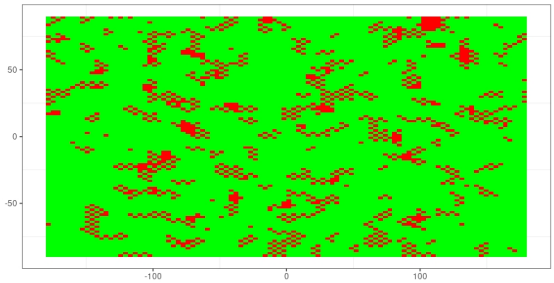
(a) $K = 0.3$, $\sigma = 0.4$, $G = 0$, $t = 1$



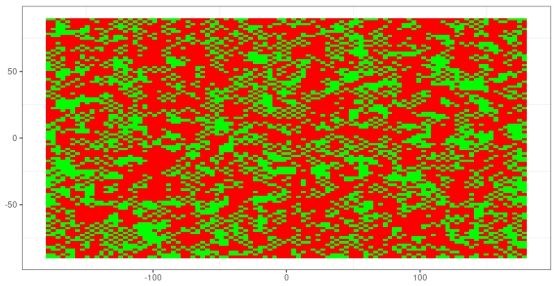
(b) $K = 0.3$, $\sigma = 0.4$, $G = 0$, $t = 5$



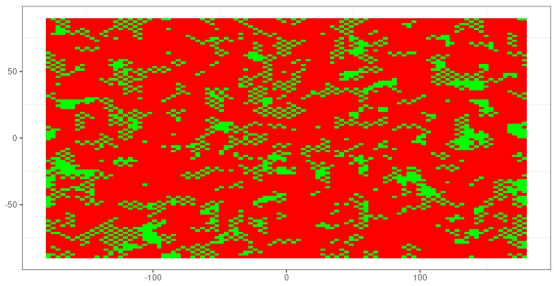
(c) $K = 0.8$, $\sigma = 0.4$, $G = 0.5$, $t = 1$



(d) $K = 0.8$, $\sigma = 0.4$, $G = 0.5$, $t = 3$



(e) $K = 0.8$, $\sigma = 0.4$, $G = -0.5$, $t = 1$



(f) $K = 0.8$, $\sigma = 0.4$, $G = -0.5$, $t = 3$

Figura 2.1: Simulazione del modello di interazione su griglia bidimensionale descritto al punto (2.12) per diversi valori dei parametri e diversi istanti temporali. Si noti come sia K che il parametro G giochino un ruolo fondamentale nel determinare lo stato in cui si troverà la rete dopo n istanti temporali. Precisiamo che al tempo $t = 0$ lo stato iniziale della rete è determinato in modo pseudo-casuale tramite estrazione di un campione di dimensione n^2 da una variabile casuale normale standard.

2.4.2 Hierarchical Diamond Lattice

L'ipotesi che la rete di agenti sia bidimensionale semplifica eccessivamente il problema e non tiene conto di alcuni aspetti fondamentali riguardanti la struttura dei mercati finanziari. I mercati odierni, infatti, sono strutture complesse in cui interagiscono agenti eterogenei, sia per professionalità che per capitalizzazione. Tali caratteristiche hanno condotto gli autori verso modelli di interazione più complessi, tra cui quello che tratteremo nel seguito, definito come *Hierarchical Diamond Lattice* (HDL).

Partiamo dal definire la struttura nel suo complesso per poi analizzarne le proprietà. Un HDL si costruisce partendo da un segmento, i cui vertici rappresentano due agenti collegati tra loro. Successivamente si sostituisce al segmento un rombo, i cui vertici sono i due agenti precedenti, più altri due ai vertici che si sono creati, per un totale di 4 agenti. Si procede allo stesso modo per i lati del rombo ottenendo un totale di $\frac{2}{3}(2+4^p)$ agenti, con 4^p collegamenti dopo p iterazioni. (si veda figura 2.2)

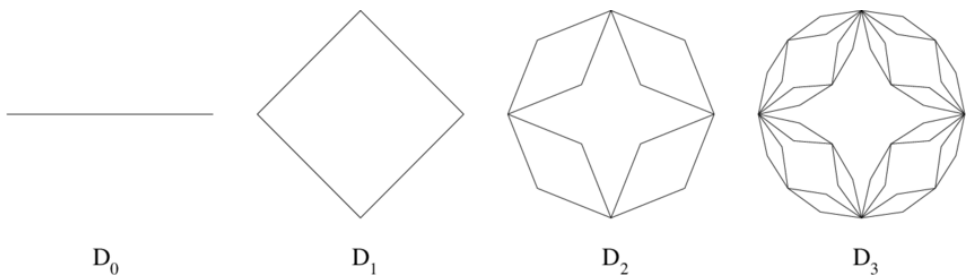


Figura 2.2: Costruzione di un Hierarchical Diamond Lattice con n=3

Il modello sopra descritto è stato risolto da Derrida et al. (1983). Le sue proprietà sono simili a quelle esposte in precedenza per il modello bidimensionale. In particolare esiste anche in questo caso un valore di K_c per cui il sistema raggiunge un punto critico. L'approssimazione del primo ordine di χ è:

$$\begin{aligned}\chi &\approx Re[A_0(K_c-K)^{-\gamma}+A_1(K_c-K)^{-\gamma+i\omega}+...] \\ &\approx A_0^{'}(K_c-K)^{-\gamma}+A_1^{'}(K_c-K)^{-\gamma}\cos[\omega\ln(K_c-K)+\phi]+...\end{aligned} \tag{2.18}$$

dove $A_0^{'}$, $A_1^{'}$, ω e ϕ sono numeri reali, e $Re[...]$ indica la parte reale di un numero complesso. Come nel precedente paragrafo, se ipotizziamo che l'hazard rate si comporti come χ, possiamo approssimare $h(t)$ come segue:

$$h(t) \approx B^{'}(t_c - t)^{-\alpha} + C^{'}(t_c - t)^{-\alpha} \cos[\omega \ln(t_c - t) + \phi^{'}] \tag{2.19}$$

È possibile notare come in tal caso la *power law* sia corredata da oscillazioni la cui frequenza aumenta quando ci si avvicina a t_c. Tali oscillazioni sono *log-periodiche* poiché sono periodiche rispetto a $\ln(t_c - t)$ e $\omega/2$ è la loro *log-frequenza*. Anche in questo caso l'hazard rate diverge ad infinito quando $t \to t_c$, e, per le considerazioni fatte in precedenza, presenta anch'esso un andamento log-periodico. L'andamento log-periodico rappresenta la tensione e competizione tra i due tipi di trader, ed è proprio questa competizione a creare la componente oscillatoria che accompagna la crescita super esponenziale del prezzo quando questo si avvicina all'istante t_c.

Ricordando che $\mu(t) = \kappa h(t)$ e utilizzando l'espressione per $h(t)$ al punto 2.19, possiamo riscrivere l'integrale 2.6:

$$\begin{aligned} \int \mathbb{E}\left[\ln \frac{\partial p/\partial t}{p}\right] \partial t = \int \big[\kappa B^{'}(t_c - t)^{m-1} + \kappa C^{'}(t_c - t)^{m-1} \times \\ \cos\left(\omega \ln(t_c - t) - \phi^{'}\right)\big] \partial t \end{aligned} \tag{2.20}$$

da cui si ottiene:

$$\begin{aligned} \mathbb{E}[\ln p(t)] = A - \frac{\kappa B^{'}}{m}(t_c - t)^m - \frac{\kappa C^{'}}{m^2 + \omega^2}(t_c - t)^m \times \\ \left\{m \cos[\phi^{'} - \omega \ln(t_c - t)] - \omega \sin[\phi^{'} - \omega \ln(t_c - t)]\right\} \end{aligned} \tag{2.21}$$

A questo punto, ricordando che $\cos(-x) = \cos(x) \quad \forall x \in \mathbb{R}$ ed effettuando alcune approssimazioni:

$$\mathbb{E}[\ln p(t)] \approx A - \frac{\kappa B^{'}}{m}(t_c - t)^m - \frac{\kappa C^{'}}{m^2 + \omega^2}(t_c - t)^m \cos[\omega \ln(t_c - t) - \phi] \tag{2.22}$$

Sostituendo $B = -\frac{\kappa B^{'}}{m}$ e $C = -\frac{\kappa C^{'}}{m^2+\omega^2}$ otteniamo la formulazione classica del modello LPPLS:

$$\mathbb{E}[\ln p(t)] \approx A + B|t_c - t|^m + C|t_c - t|^m \cos\left[\omega \ln|t_c - t| - \phi\right] \tag{2.23}$$

L'espressione (2.23) è l'equazione fondamentale che descrive l'evoluzione del prezzo prima del crash date le ipotesi fatte finora. È possibile notare come in tal caso abbiamo scelto di porre la quantità $t_c - t$ in valore assoluto. Tale scelta, adottata in numerosi lavori, si veda ad esempio Demirer et al. (2019), permette di risolvere alcuni problemi pratici di ottimizzazione. Inoltre una conseguenza di questa scelta è quella di rendere il modello LPPLS simmetrico nel punto $t = t_c$, come si può notare nella figura 2.3.

Il parametro $A > 0$ è il valore di $\ln[p(t_c)]$, cioè il valore atteso del prezzo nell'istante t_c. Il parametro B misura l'accelerazione del prezzo, mentre C controlla l'ampiezza delle oscillazioni log-periodiche. Il meccanismo di *positive-feedback* che porta ad una crescita super-esponenziale del prezzo è catturato dall'esponente m, che deve essere compreso nell'intervallo $[0, 1]$. In particolare, $m < 1$ poiché questo assicura che il crash hazard rate aumenti più ci si avvicina a t_c ma anche che l'integrale di $h(t)$ assuma valori finiti e minori di uno per tutti gli istanti temporali $t \leq t_c$. In questo modo per gli investitori è razionale restare sul mercato anche se riconoscono ci sia una bolla in formazione, poiché esiste una possibilità di sfruttare la crescita del prezzo senza subire necessariamente le conseguenze di un potenziale crash. Come detto in precedenza è $\mu(t) = \kappa h(t)$, che cresce al crescere $h(t)$, a giustificare la scelta di restare sul mercato da parte degli investitori razionali. Infine, affinchè il prezzo assuma valori finiti in tutti gli istanti temporali, incluso $t = t_c$, è necessario che $m > 0$.

Nel lavoro di Johansen et al. (2000), gli autori discutono approfonditamente anche le cause dell'andamento log-periodico di sistemi complessi quando questi si avvicinano a punti critici.

Essi sostengono che in un modello che presenta le seguenti caratteristiche:

- *Noise traders* con tendenze imitative
- Imitazione locale che si propaga fino a diventare globale
- Imitazione globale tra *noise traders* causa crash
- Prezzi dipendenti dalle proprietà del sistema
- Parametri che si evolvono lentamente nel tempo

Il prezzo segue una *power law* quando il sistema si avvicina ad un *punto critico*. La caratteristica in comune di questi approcci è che un crash diventa molto probabile quando le tendenze imitative si propagano nell'intera rete. Ciò significa che un sistema vicino al punto critico è composto da agenti le cui opinioni sono fortemente correlate a causa delle tendenze imitative e quindi l'intero sistema diventa sensibile rispetto a piccole perturbazioni. Una caratteristica fondamentale di questi sistemi è che essi sono *auto similari*, proprietà comune alle strutture frattali. Tra le forme di auto similarità più studiate c'è quella definita *invarianza di scala*. In fisica e matematica, l'invarianza di scala è una caratteristica degli oggetti o una legge fisico-matematica che non cambia forma se si scalano le lunghezze (o parimenti le energie) di un fattore comune. In matematica, l'invarianza di scala spesso si riferisce all'invarianza di una singola funzione o curva (si veda il punto 2.24)[2]. È possibile individuare anche distribuzioni di probabilità di processi stocastici caratterizzate da invarianza di scala o auto similarità (ad esempio il *moto browniano*). Un altro esempio riguardo statistiche auto similari è quello riportato da Anderson (1992) per la statistica T quadrato di Hotelling. Nel caso di sistema su griglia bidimensionale in figura 2.1, il concetto di invarianza di scala si traduce nel fatto che se riducessimo la dimensione della griglia di un fattore λ, questo non cambierebbe le caratteristiche della griglia stessa. Avremmo infatti ancora una situazione nella quale dopo n istanti sarebbero presenti cluster di agenti con uguale stato divisi da cluster più piccoli di agenti nello stato opposto (a seconda dei parametri scelti). Analogamente a quanto detto nei paragrafi precedenti e sulla scorta del lavoro di Dubrulle et al. (1997), sappiamo che il comportamento di un sistema vicino ad un punto critico può essere espresso tramite una *power law*. Questo poiché tali funzioni sono le uniche che non cambiano quando il loro argomento viene riscalato di una costante. Matematicamente, cioè:

$$F(\gamma x) = \delta F(x) \tag{2.24}$$

dove F nel nostro caso rappresenta la *suscettibilità* del sistema, x è un parametro e δ è una costante positiva che descrive il cambiamento del sistema quando questo viene riscalato di un fattore γ.

La soluzione generale per l'equazione 2.24 è :

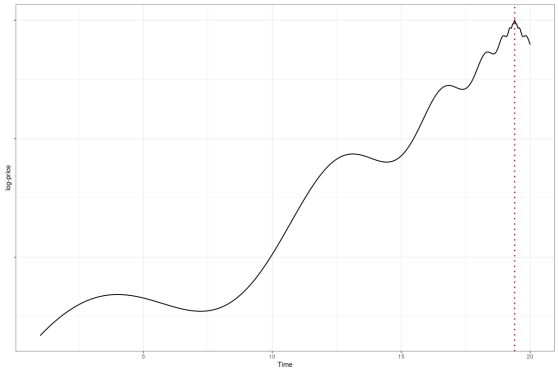
$$F(x) = x^{log(\delta)/\log(\gamma)}\pi\left[\frac{\log(x)}{\log(\gamma)}\right] \tag{2.25}$$

[2] https://it.wikipedia.org/wiki/Invarianza_di_scala

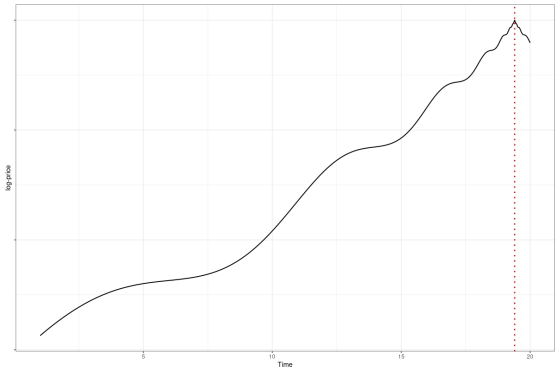
dove π è una funzione periodica di periodo uno. È necessario a questo punto far notare come il risultato al punto 2.18 siano i termini di grado zero e uno dell'espansione di Fourier della funzione periodica π.

L'espressione 2.24 descrive il sistema esattamente nel punto critico, ma nulla ci dice riguardo tale sistema quando esso si trova nelle vicinanze di un punto critico. Sarebbe preferibile, infatti, per le applicazioni pratiche, avere una descrizione accurata del comportamento del sistema anche nell'intorno del punto critico. La questione quindi si riduce a capire se e come cambia l'espressione precedente quando il sistema non è esattamente sul punto critico. La risposta a questa domanda viene da una teoria chiamata *renormalization group theory*, si veda Wilson (1979). Questa teoria permette di analizzare il comportamento "globale" di un sistema complesso scomponendo quest'ultimo in parti sempre più piccole e semplici da analizzare. Si rivela particolarmente utile in caso di fenomeni critici, come quello analizzato in questo lavoro, invarianti rispetto alla scala. Intuitivamente, quindi, la *renormalization group theory* sfrutta il concetto secondo il quale il comportamento globale di un sistema complesso è il risultato di tanti sotto sistemi, formati da altri sotto sistemi e così via, che interagiscono tra loro. Presentare le profonde implicazioni che la teoria dei punti critici e la *renormalization group theory* hanno in branche del sapere come la fisica e la matematica esula dallo scopo di questo lavoro. Ci limitiamo a precisare, però, che l'idea principale alla base del modello LPPLS e di quelli simili ad esso è la seguente: i violenti e improvvisi crash che si verificano sui mercati finanziari derivano da meccanismi imitativi locali che si propagano su scala globale quando il sistema complesso (mercato su un determinato titolo) raggiunge un punto critico.

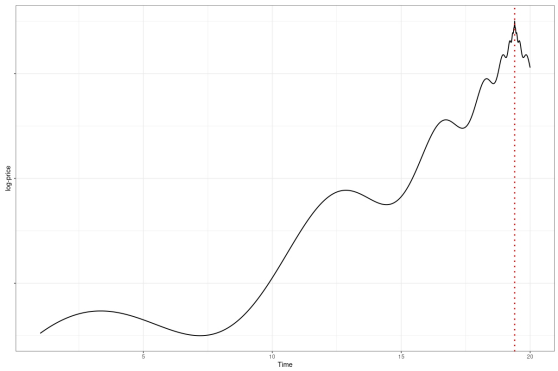
Nella figura 2.3 è possibile notare come evolve il modello al variare dei parametri. Precisiamo che nel generare i grafici di cui si discuterà nel seguito, abbiamo ipotizzato valori di A e t_c uguali per tutti gli esempi riportati e utilizzato la forma riparametrizzata del modello che verrà presentata nel capitolo successivo. Nel caso in figura 2.3a si nota come un valore più elevato di C comporti oscillazioni più marcate rispetto alla figura 2.3b. Nelle figure seguenti [2.3c, 2.3d], invece, si evidenzia come il segno del parametro B serva a discriminare tra bolle positive ($B < 0$) e bolle negative ($B > 0$). Tale proprietà sarà utilizzata nei capitoli successivi per identificare se la bolla oggetto d'analisi è positiva o negativa. Inoltre, dall'analisi di queste figure, è possibile chiarire anche l'impatto che il parametro ω ha sul modello. Un valore elevato di ω corrisponde ad un numero di oscillazioni maggiori [2.3d] rispetto a valori più bassi, come quello in figura [2.3c]. Analizzando i grafici ai punti [2.3e, 2.3f] si nota quale sia l'impatto che m ha sul modello: valori più bassi di m comportano una crescita più rapida verso t_c rispetto a valori più elevati.



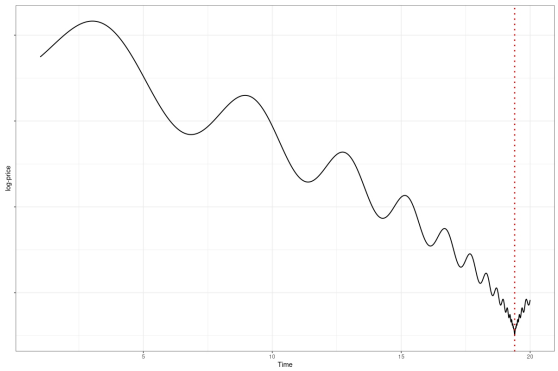
(a) $m = 0.8,\ \omega = 7,\ B = -6,\ C \approx 0.98$



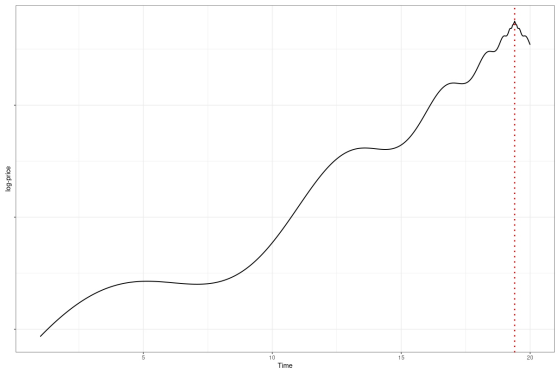
(b) $m = 0.8,\ \omega = 7,\ B = -6,\ C \approx 0.57$



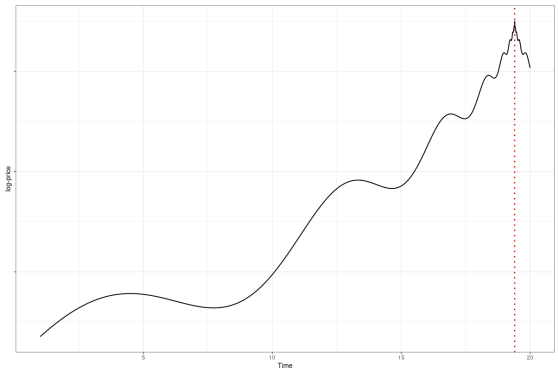
(c) $m = 0.6,\ \omega = 7,\ B = -6,\ C \approx 0.9$



(d) $m = 0.6,\ \omega = 14,\ B = 6,\ C \approx 0.9$



(e) $m = 0.8,\ \omega = 7,\ B = -6,\ C \approx 0.78$



(f) $m = 0.6,\ \omega = 7,\ B = -6,\ C \approx 0.78$

Figura 2.3: Esempi di modello LPPLS al variare dei parametri, fissati i parametri $A = 20$ e $t_c = 19.4$. È possibile notare l'impatto che il parametro C ha sull'ampiezza delle oscillazioni e l'impatto del parametro B sull'andamento complessivo del modello. Il parametro ω, invece, quantifica il numero di oscillazioni nell'intervallo temporale $[t_0, t_c]$. Infine il parametro m determina l'accelerazione del prezzo nella fase di bolla. Un valore elevato di m implica una crescita lenta e viceversa. Il tempo critico t_c è indicato dalla retta verticale rossa.

2.5 Formulazioni alternative del modello LPPLS

Il modello LPPLS, fin dalla nascita, ha attirato l'attenzione di molti studiosi interessati al problema delle bolle speculative e nel corso del tempo sono state proposte generalizzazioni o modifiche al modello originario. Nel seguito tratteremo alcune di esse.

2.5.1 Modello log-periodico AR(1)-GARCH(1,1)

La prima riguarda il lavoro di Gazola et al. (2008). Gli autori sostengono che il modello LPPLS classico non sia in grado di descrivere accuratamente le dinamiche di mercato di breve periodo: il modello può infatti presentare residui fortemente autocorrelati ed eteroschedastici se applicato nel breve termine.

Per ovviare a ciò, Gazola et al. (2008) propongono il modello AR(1)-GARCH(1,1) log-periodico. Il modello è definito come segue:

$$\begin{aligned}
p_t &= A + B(t_c - t)^m + C(t_c - t)^m \cos[\omega \ln(t_c - t) + \phi] + u_t \\
u_t &= \rho u_{t-1} + \eta_t \\
\eta_t &= \sigma_t \epsilon_t \ , \quad \epsilon_t \sim N(0,1) \\
\sigma_t^2 &= \alpha_0 + \alpha_1 \eta_{t-1}^2 + \alpha_2 \sigma_{t-1}^2
\end{aligned} \tag{2.26}$$

con ϵ_t un white noise con $\mathbb{E}[\epsilon_t] = 0$, $\mathbb{E}[\epsilon_t^2] = 1$ e con la varianza condizionata σ_t^2 che segue un processo GARCH(1,1). Data la struttura del modello definita al punto 2.26, lo stimatore di massima verosimiglianza per il vettore di parametri $\Pi = [A, B, C, t_c, m, \omega, \phi, \rho, \alpha_0, \alpha_1, \alpha_2]$ si ottiene massimizzando la log-verosimiglianza seguente:

$$\ln L(\Theta) = -\frac{1}{2}(N-1)\ln(2\pi) - \frac{1}{2}\sum_{i=2}^{N} \ln \sigma_i^2 - \frac{1}{2}\sum_{i=2}^{N} \frac{\eta_i^2}{\sigma_i^2} \tag{2.27}$$

Al fine di migliorare la procedura di ottimizzazione, ogni parametro del modello log periodico al punto 2.26, che possiamo genericamente chiamare θ e definito sull'intervallo $[a, b]$, può essere riparametrizzato utilizzando la seguente trasformazione monotona:

$$\theta = b\frac{exp(\tilde{\theta})}{1+exp(\tilde{\theta})} + a\left(1 - \frac{exp(\tilde{\theta})}{1+exp(\tilde{\theta})}\right) \tag{2.28}$$

Gli autori utilizzano questa trasformazione poiché consente di passare da un problema di ottimizzazione vincolato a un problema di ottimizzazione non vincolato, il che semplifica la procedura di stima. Sulla base della trasformazione di cui sopra, gli autori utilizzano poi il *delta method* per ricavare gli standard error delle stime.

Il delta method permette di approssimare la distribuzione di probabilità di una funzione di uno stimatore asintoticamente normale basandosi sulla conoscenza della matrice varianze-covarianze dello stimatore stesso. Ad esempio, in ambito univariato, considerando una successione di variabili casuali X_n per cui vale:

$$\sqrt{n}[X_n - \theta] \xrightarrow{D} N(0, \sigma^2) \tag{2.29}$$

con θ e σ^2 finiti, allora:

$$\sqrt{n}[g(X_n) - g(\theta)] \xrightarrow{D} N(0, \sigma^2[g^{'}(\theta)]^2) \tag{2.30}$$

per qualsiasi funzione $g()$ con derivata prima diversa da zero.

Nel nostro caso, invece, definiamo $\hat{V}[\bar{\theta}]$ la matrice varianze-covarianze di $\bar{\theta}$. Utilizzando il metodo delta possiamo ricavare la matrice varianze-covarianze di una trasformazione non lineare di $\bar{\theta}$, $g(\bar{\theta})$, come segue (si veda Hayashi (2000)):

$$\hat{V}[g(\tilde{\theta})] = \frac{\partial g(\tilde{\theta})}{\partial \tilde{\theta}'}\hat{V}[\tilde{\theta}]\frac{\partial g(\tilde{\theta})'}{\partial \tilde{\theta}} \tag{2.31}$$

Gli autori utilizzano, per la scelta dei valori iniziali dei parametri, una procedura in due fasi. I valori iniziali per i parametri $\Phi = [A, B, C, t_c, m, \omega, \phi]$ sono quelli ottenuti utilizzando la procedura classica di stima del modello LPPLS. I valori iniziali per i parametri $[\rho, \alpha_0, \alpha_1, \alpha_2]$, cioè quelli che riguardano la componente stocastica di breve periodo u_i, si ottengono stimando un modello AR(1)-GARCH(1,1) sui residui $\hat{u}_i$ del modello LPPLS.

2.5.2 Modello LPPLS generalizzato con residui *mean-reverting*

Lin et al. (2009) propongono una generalizzazione del modello al fine di rendere consistente la stima diretta dei parametri su serie temporali.

La seguente disamina si basa sia sul lavoro citato in precedenza, sia su quello di Geraskin and Fantazzini (2013). Come spiegato nei paragrafi precedenti, il modello LPPLS originario ha una componente random walk con varianza crescente. Tale caratteristica rende la stima diretta sui prezzi inconsistente. Per ovviare a queste problematiche, gli autori utilizzano il modello LPPLS *volatility confined*. Questo nuovo approccio prevede che la volatilità segua un processo mean reverting e che il drift del prezzo sia stocastico invece che deterministico. Gli autori ipotizzano quindi che il processo del prezzo sia governato dalle seguenti equazioni differenziali stocastiche:

$$\begin{aligned} \frac{\partial p}{p(t)} &= \mu(t)\partial t + \sigma_Y \partial Y + \sigma_W \partial W - \kappa \partial j \\ \partial Y &= -\alpha Y \partial t + \partial W \end{aligned} \tag{2.32}$$

dove $p(t)$ è il prezzo dell'asset considerato al tempo t, W è un processo di Wiener standard, $\mu(t)$ è la componente di drift che dipende dal tempo, j è una variabile dummy che assume valore 1 dopo il crash e valore 0 prima. κ rappresenta invece l'entità del crash in termini percentuali. È necessario precisare che se $0 < \alpha < 1$, Y è un processo di Ornstein-Uhlenbeck: pertanto sia ∂Y che Y sono stazionari e la volatilità resta entro determinati limiti finchè non avviene il crash.

Il processo di Ornstein-Uhlenbeck è largamente utilizzato nelle applicazioni economiche e finanziarie date le interessanti proprietà che possiede: esso infatti è un processo di Gauss-Markov *mean-reverting* stazionario. Tale processo è governato dalla seguente equazione differenziale stocastica:

$$\partial Y = \alpha(\gamma - Y)\partial t + \sigma \partial W, \qquad \alpha, \sigma > 0,\ \gamma \geq 0 \tag{2.33}$$

con W un processo di Wiener. Le proprietà del processo di Ornstein-Uhlenbeck garantiscono che la stima diretta del modello sul prezzo sia consistente.

La versione del modello LPPLS *mean-reverting* in tempo discreto è la seguente:

$$
\begin{aligned}
&\ln p_{i+1} - \ln p_i = \mu_i + \sigma_Y (Y_{i+1} - Y_i) + \sigma_W \epsilon_i - \kappa \Delta j_i \\
&Y_{i+1} = (1-\alpha) Y_i + \epsilon_i, \quad \epsilon_i \sim N(0,1)
\end{aligned}
\tag{2.34}
$$

Utilizzando il fattore di sconto stocastico e l'ipotesi di non arbitraggio, Lin et al. (2009) mostrano come i log-rendimenti dell'asset seguano il processo:

$$
\ln p_{i+1} = \ln p_i + \Delta H_{i,i-1} - \alpha(\ln p_i - H_i) + u_i \tag{2.35}
$$

con $\Delta H_{i,i-1}$ che deriva dal lavoro di Chang and Feigenbaum (2006):

$$
\begin{aligned}
\Delta H_{i,i-1} =& B(t_c - t_{i-1})^m \left[1 + \frac{C}{\sqrt{1 + \left(\frac{\omega}{m}\right)^2}} \cos\left(\omega \ln(t_c - t_{i-1}) + \phi\right) \right] - \\
&- B(t_c - t_i)^m \left[1 + \frac{C}{\sqrt{1 + \left(\frac{\omega}{m}\right)^2}} \cos\left(\omega \ln(t_c - t_i) + \phi\right) \right]
\end{aligned}
\tag{2.36}
$$

u_i è un processo white noise gaussiano. La distribuzione dei rendimenti logaritmici è la seguente:

$$
r_{i+1} = \ln p_{i+1} - \ln p_i \sim N\left[\Delta H_{i+1,i} - \alpha(\ln p_i - H_i),\ \sigma_u^2 (t_{i+1} - t_i) \right] \tag{2.37}
$$

Una considerazione importante di Lin et al. (2009) è che il modello LPPLS classico basato sulle aspettative razionali separa artificialmente i *noise traders* dagli investitori razionali. Inoltre gli autori fanno notare come l'ipotesi per la quale questi ultimi abbiano preferenze omogenee possa risultare troppo restrittiva.

Infine essi precisano che i risultati a cui sono pervenuti siano gli stessi ottenuti se si utilizzasse la teoria *Behavioral Stochastic Discount Factor*. Per ulteriori dettagli a riguardo si rimanda al lavoro di Shefrin (2008).

Capitolo 3

Stima dei parametri

Il modello LPPLS introdotto nel capitolo precedente presenta tre parametri lineari (A, B, C) e quattro parametri non lineari (m, ω, t_c, ϕ). I parametri di cui sopra sono soggetti a vari vincoli. Ad esempio, per quanto riguarda $h(t)$, dato che esso quantifica la probabilità che avvenga un crash, deve essere minore di 1, il che implica che $m < 1$. Allo stesso tempo il logaritmo del prezzo deve assumere valori finiti per qualsiasi $t \leq t_c$ e affinché ciò sia verificato è necessario che $m > 0$. Ulteriori vincoli, riportati in seguito, sono stati utilizzati in letteratura per garantire che il modello abbia le proprietà desiderate, si veda a tal proposito Sornette and Johansen (2001), Johansen et al. (2010) e Lin et al. (2009).

$$0.1 \leq m \leq 0.9, \quad 2 \leq \omega \leq 25, \quad |C| < 1 \tag{3.1}$$

La condizione $6 \leq \omega \leq 13$ rende le oscillazioni log-periodiche né troppo veloci né troppo lente. La condizione $|C| < 1$ assicura che $h(t)$ assuma sempre valori positivi. Procederemo ad una trattazione dettagliata riguardo i vincoli nei paragrafi successivi, in particolare quando si discuterà degli indicatori costruiti sulla base dei risultati del modello LPPLS.

3.1 Procedura classica di stima

Nel seguito del paragrafo tratteremo nel dettaglio la procedura inizialmente proposta per la stima dei parametri del modello LPPLS. Formalmente si tratta di ricavare delle stime per i parametri t_c, m, ω, ϕ, A, B, C stimando il modello sulla serie storica

$p(t)$ con $t \in [t_1, t_2]$.

Sornette e gli altri autori propongono come procedura di stima dei parametri il metodo dei minimi quadrati. L'obiettivo è minimizzare la somma degli scarti al quadrato tra il modello e la serie storica considerata:

$$\begin{aligned} S(t_c, m, \omega, \phi, A, B, C) = \sum_{i=1}^{N} \big[\ln p(\tau_i) - A - B|t_c - \tau_i|^m \\ - C|t_c - \tau_i|^m \cos(\omega \ln |t_c - \tau_i| - \phi) \big]^2 \end{aligned} \tag{3.2}$$

dove $\tau_1 = t_1$ e $\tau_N = t_2$. Minimizzare la funzione S non è semplice dal punto di vista computazionale poiché la funzione presenta molti minimi locali. L'approccio adottato in Johansen et al. (2000) consiste nel vincolare i parametri lineari a quelli non lineari. In tal modo è possibile scrivere

$$\min_{t_c, m, \omega, \phi, A, B, C} S(t_c, m, \omega, \phi, A, B, C) \equiv \min_{t_c, m, \omega, \phi} S_1(t_c, m, \omega, \phi) \tag{3.3}$$

con

$$S_1(t_c, m, \omega, \phi) = \min_{A, B, C} S(t_c, m, \omega, \phi, A, B, C) \tag{3.4}$$

È possibile riscrivere il problema di ottimizzazzione al punto (3.4) come segue:

$$\begin{aligned} \{\hat{A}, \hat{B}, \hat{C}\} &= \arg \min_{A, B, C} S(t_c, m, \omega, \phi, A, B, C) \\ &= \arg \min_{A, B, C} \sum_{i=1}^{N} \Big[y_i - A - B f_i - C g_i \Big]^2 \end{aligned} \tag{3.5}$$

dove $y_i = \ln p(\tau_i)$, $f_i = |t_c - \tau_i|^m$, $g_i = |t_c - \tau_i|^m \cos(\omega \ln |t_c - \tau_i| - \phi)$.

È possibile ricavare le stime dei parametri lineari A, B, C dalla condizione del primo ordine, da cui deriva il seguente sistema lineare di equazioni:

$$\begin{pmatrix} N & \sum f_i & \sum g_i \\ \sum f_i & \sum f_i^2 & \sum f_i g_i \\ \sum g_i & \sum f_i g_i & \sum g_i^2 \end{pmatrix} \begin{pmatrix} \hat{A} \\ \hat{B} \\ \hat{C} \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum y_i f_i \\ \sum y_i g_i \end{pmatrix} \tag{3.6}$$

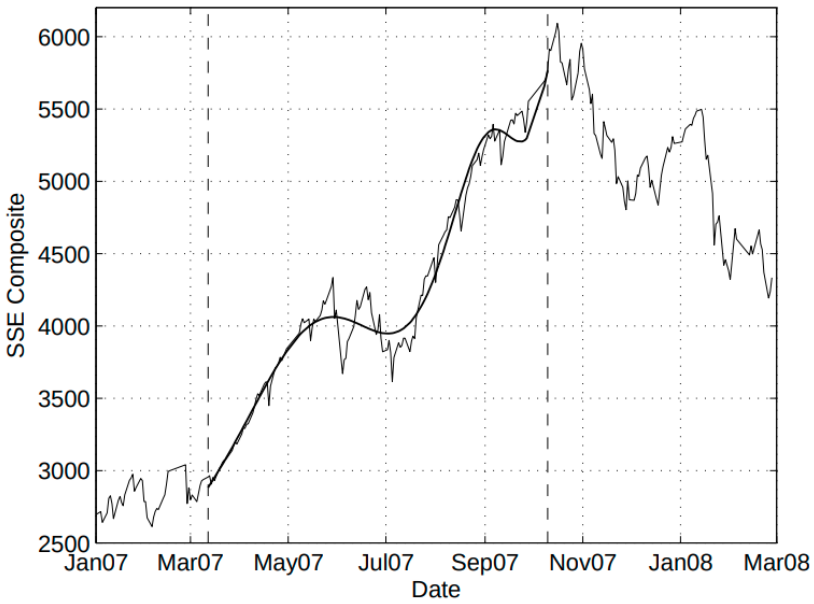


Figura 3.1: Modello LPPL per la serie storica del SSE Composite Index. Le linee tratteggiate rappresentano la finestra utilizzata per il fitting. [Filimonov and Sornette (2013)]

Una volta trovata la formulazione analitica per le stime dei parametri lineari, bisogna stimare la componente non lineare, cioè:

$$\{\hat{t}_c, \hat{m}, \hat{\omega}, \hat{\phi}\} = \arg \min_{t_c, m, \omega, \phi} S_1(t_c, n, \omega, \phi) \tag{3.7}$$

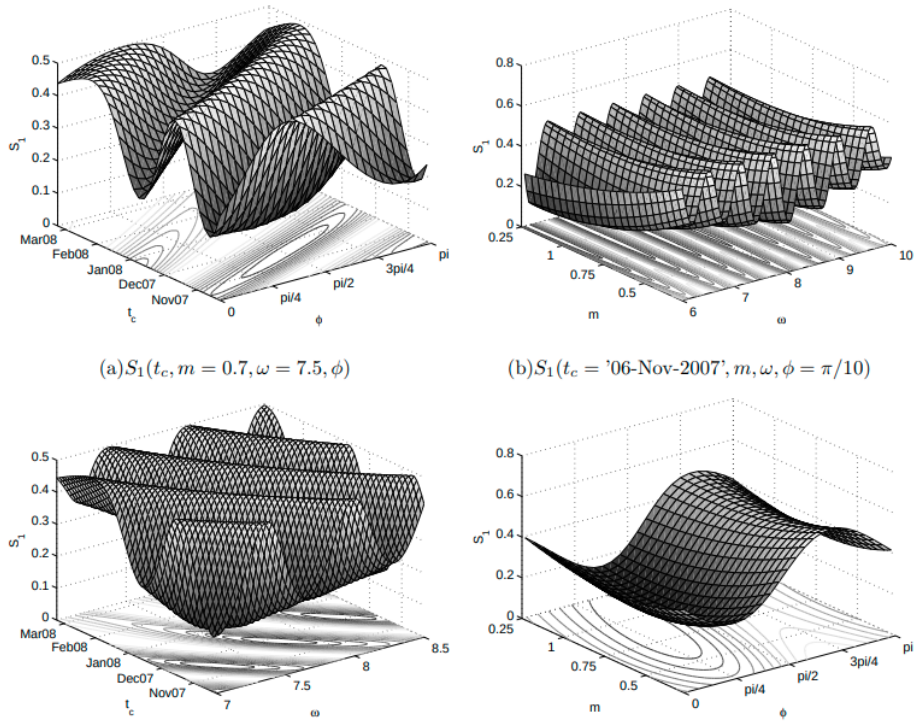
dove la funzione S_1 è quella definita al punto (3.4).

Nonostante la procedura sopra descritta riduca i parametri da sette a quattro, persistono delle complicazioni data la struttura non lineare e semi periodica della funzione di costo e la presenza di molti minimi relativi, si veda la figura 3.2 . Pertanto non è possibile utilizzare il metodo delle tangenti (Metodo di Newton-Raphson) o il metodo del gradiente.

In particolare, Filimonov and Sornette (2013) stimano i parametri del modello LPPLS sulla serie storica del SSE Composite Index in un intervallo di tempo che va dal 12/03/2007 a 10/10/2007.

Nella figura 3.1 è possibile osservare il modello LPPL e la serie storica dei prezzi.

La risoluzione di questi problemi di ottimizzazione non è semplice ed a lungo si è



(a)$S_1(t_c, m = 0.7, \omega = 7.5, \phi)$ (b)$S_1(t_c =$ '06-Nov-2007', $m, \omega, \phi = \pi/10)$

(c)$S_1(t_c, m = 0.7, \omega, \phi = \pi/10)$ (d)$S_1(t_c =$ '06-Nov-2007', $m, \omega = 7.5, \phi)$

Figura 3.2: Esempi cross-section della funzione S_1 per la serie storica del SSE Composite Index [Filimonov and Sornette (2013)]

discusso in letteratura su quale fosse il metodo più efficace a seconda della particolare tipologia di funzione da ottimizzare. Uno degli approcci proposti è quello degli algoritmi metaeuristici, si veda a tal proposito Talbi (2009). Tramite questi algoritmi si giunge ad una soluzione di tipo euristico, cioè una soluzione per la quale non c'è garanzia che essa sia, a priori, effettivamente ottimale. Nel corso del tempo sono stati proposti numerosi approcci per la stima dei parametri del modello LPPLS. Ad esempio, in Johansen et al. (2000), Sornette and Johansen (2001) e Sornette and Zhou (2006) il metodo proposto è quello *taboo search*, che appartiene anch'esso agli algoritmi metaeuristici, si veda a proposito Cvijović and Klinowski (1995). Il metodo *taboo search* è un algoritmo di ricerca locale che ha il vantaggio di possedere una struttura di memoria che descrive le soluzioni già visitate. L'algoritmo, infatti, quando trova una potenziale soluzione, la ricorda come già valutata (*taboo*), in modo tale da non rimanere intrappolato in regioni della funzione già visitate. Essendo un metodo metaeuristico, non garantisce la convergenza. Per ovviare a questo problema gli autori propongono di utilizzare le dieci migliori soluzioni ottenute dal metodo taboo come condizioni iniziali per una ricerca col metodo dei minimi quadrati di Levenberg-Marquardt [Levenberg (1944)]. La soluzione associata al valore minore

della funzione di costo è quella ottimale. Un altro approccio possibile consiste nell'utilizzare algoritmi genetici. Questa tipologia di algoritmi riproduce il processo di selezione naturale al fine di pervenire ad una soluzione. [Gulsen et al. (1995)]. Come nel caso del *taboo search*, le soluzioni dell'algoritmo genetico vengono poi utilizzate come valori iniziali nel metodo del simplesso di Nelder-Mead. [Nelder and Mead (1965)].

I metodi sopra descritti, come già chiarito, non garantiscono che i risultati ottenuti siano associati ai minimi assoluti della funzione di costo. L'elevato costo computazionale che caratterizza i metodi di stima sopra citati ha condotto gli autori verso l'elaborazione di una procedura alternativa di stima. Nel seguito introdurremo questa formulazione alternativa del modello, che risulta effettivamente essere meno complessa dal punto di vista computazionale rispetto a quella originaria.

3.2 Nuova procedura di stima

In Filimonov and Sornette (2013) gli autori presentano una formulazione alternativa del modello che semplifica ulteriormente la procedura di stima dei parametri.

L'obiettivo del metodo che verrà trattato nel seguito è quello di ridurre il numero di parametri non lineari e, allo stesso tempo, eliminare l'interdipendenza tra il parametro di fase ϕ e il parametro ω. A tal fine riscriviamo il modello LPPLS come segue:

$$\begin{aligned}\ln E[p(t)] = A + B|t_c - t|^m + C|t_c - t|^m \cos[\omega \ln|t_c - t|]\cos(\phi) \\ + C|t_c - t|^m \sin[\omega \ln|t_c - t|]\sin(\phi)\end{aligned} \tag{3.8}$$

A questo punto si introducono due nuovi parametri:

$$C_1 = C\cos(\phi) \qquad C_2 = C\sin(\phi) \tag{3.9}$$

Sostituendo i parametri nel modello si ottiene la seguente formulazione:

$$\begin{aligned}\ln E[p(t)] = A + B|t_c - t|^m + C_1|t_c - t|^m \cos[\omega \ln|t_c - t|] \\ + C_2|t_c - t|^m \sin[\omega \ln|t_c - t|]\end{aligned} \tag{3.10}$$

È possibile notare come la nuova formulazione del modello presenti solo tre parametri non lineari (t_c, ω, m) e quattro parametri lineari (A, B, C_1, C_2) dove i parametri C_1 e C_2 incorporano il parametro ϕ.

Come nel metodo di stima originario, si utilizza il metodo dei minimi quadrati. La funzione di costo associata alla nuova formulazione del modello è la seguente:

$$\begin{aligned} F(t_c, m, \omega, A, B, C_1, C_2) = \sum_{i=1}^{N} \Big\{ & \ln p(\tau_i) - A - B|t_c - \tau_i|^m \\ & - C_1|t_c - \tau_i|^m \cos[\omega \ln |t_c - \tau_i|] \\ & - C_2|t_c - \tau_i|^m \sin[\omega \ln |t_c - \tau_i|] \Big\}^2 \end{aligned} \tag{3.11}$$

Vincolando i quattro parametri lineari ai tre non lineari otteniamo il problema di ottimizzazione non lineare:

$$\{\hat{t}_c, \hat{m}, \hat{\omega}\} = \arg \min_{t_c, m, \omega} F_1(t_c, m, \omega) \tag{3.12}$$

dove la funzione $F_1(t_c, m, \omega)$ è

$$F_1(t_c, m, \omega) = \min_{A, B, C_1, C_2} F(t_c, m, \omega, A, B, C_1, C_2) \tag{3.13}$$

Per quanto riguarda le stime dei parametri lineari, il problema $\{\hat{A}, \hat{B}, \hat{C}_1, \hat{C}_2\} = \arg\min_{A,B,C_1,C_2} F(t_c, m, \omega, A, B, C_1, C_2)$ ha un'unica soluzione data dal seguente sistema di equazioni lineari:

$$\begin{pmatrix} N & \sum f_i & \sum g_i & \sum h_i \\ \sum f_i & \sum f_i^2 & \sum f_i g_i & \sum f_i h_i \\ \sum g_i & \sum f_i g_i & \sum g_i^2 & \sum g_i h_i \\ \sum h_i & \sum f_i h_i & \sum g_i h_i & \sum h_i^2 \end{pmatrix} \begin{pmatrix} \hat{A} \\ \hat{B} \\ \hat{C}_1 \\ \hat{C}_2 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum y_i f_i \\ \sum y_i g_i \\ \sum y_i h_i \end{pmatrix} \tag{3.14}$$

dove $y_i = \ln p(\tau_i)$, $f_i = |t_c - \tau_i|^m$, $g_i = |t_c - \tau_i|^m \cos[\omega \ln |t_c - \tau_i|]$ e $h_i = |t_c - \tau_i|^m \sin[\omega \ln |t_c - \tau_i|]$

Questa formulazione del modello presenta due vantaggi. Innanzitutto i parametri non lineari si riducono da quattro a tre e questo riduce significativamente la com-

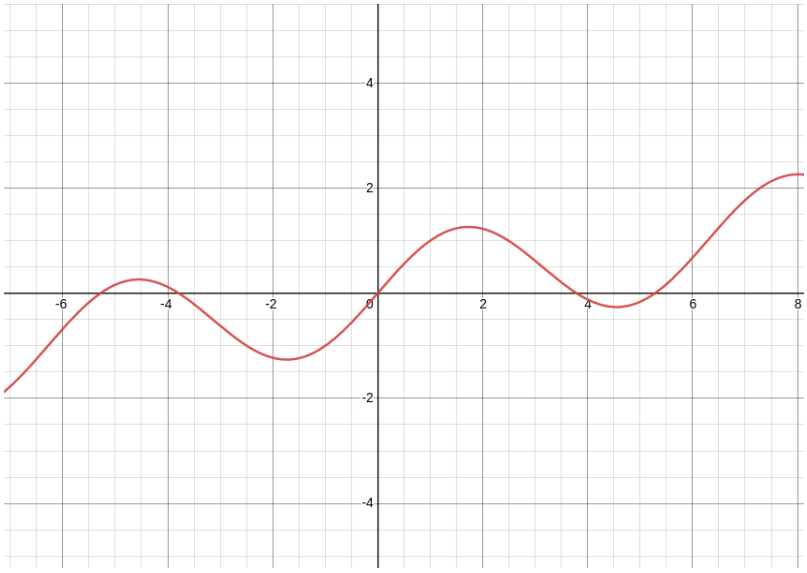


Figura 3.3: Esempio di funzione quasi-periodica al punto 3.16

plessità del problema. In secondo luogo tale modifica elimina la quasi-periodicità della funzione di costo dato che il parametro ϕ è incorporato da C_1 e C_2. Infatti, nel modello originario, era proprio la presenza di quasi-periodicità a richiedere l'utilizzo di metodi metaeuristici non rigorosi. La nuova formulazione proposta, pertanto, non richiede l'utilizzo di questi metodi.

Formalmente, una funzione f è quasi-periodica con quasi-periodo ω se:

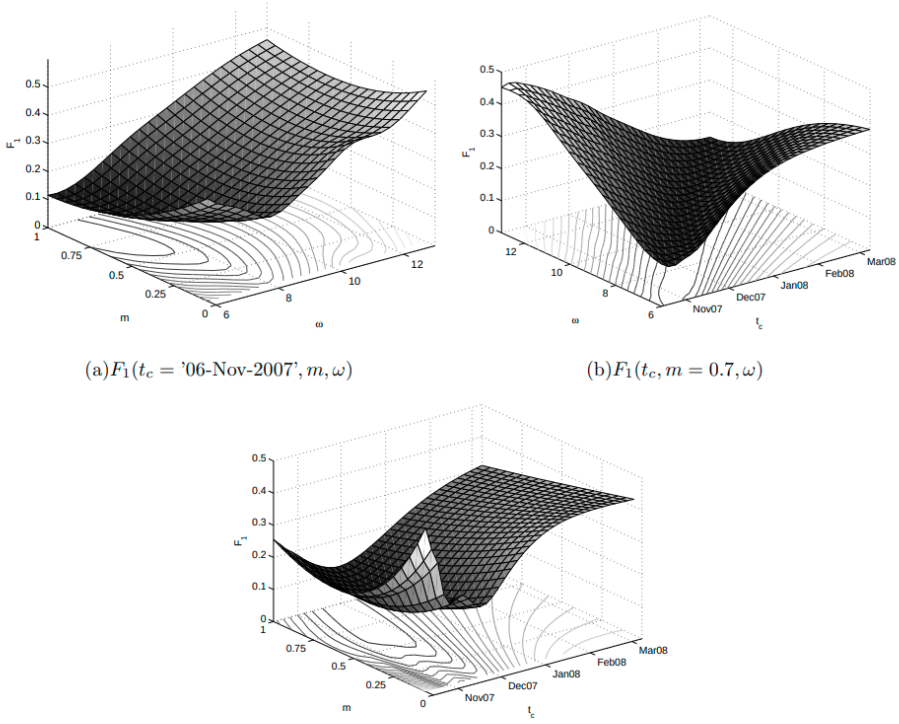
$$f(z + \omega) = g(z, f(z)) \tag{3.15}$$

Un esempio di funzione quasi-periodica è il seguente:

$$f(x) = \frac{x}{2\pi} + \sin(x) \tag{3.16}$$

si veda la figura 3.3.

Riguardo la struttura della funzione di costo è possibile apprezzare l'assenza di quasi-periodicità confrontando la figura 3.2 con quella 3.4. Data la funzione di costo di cui sopra, è possibile utilizzare metodi che in precedenza abbiamo scartato, quali ad esempio l'algoritmo di Levenberg-Marquardt o il metodo del simplesso di Nelder-Mead, senza necessariamente dividere la procedura di stima in due fasi, come visto nel paragrafo precedente.



(a)$F_1(t_c =$ '06-Nov-2007'$, m, \omega)$

(b)$F_1(t_c, m = 0.7, \omega)$

(c)$F_1(t_c, m, \omega = 7.5)$

Figura 3.4: Esempi cross-section della funzione S_1 per la serie storica del SSE Composite Index in seguito alla riparametrizzazione del modello LPPLS [Filimonov and Sornette (2013)]

Nel seguito introdurremo l'algoritmo utilizzato in questo lavoro per la stima dei parametri.

3.3 Algoritmo Multi-Level Single-Linkage

Al fine di risolvere il problema di ottimizzazione non lineare (3.12) è stato utilizzato l'algoritmo "Multi-Level Single-Linkage" (MLSL) proposto da Kan and Timmer (1987).

In letteratura a lungo si è cercato di definire algoritmi e metodi di ottimizzazione per funzioni complesse che garantissero soluzioni soddisfacenti. Tuttavia è difficile trovare algoritmi che permettano di individuare con assoluta certezza l'ottimo di una funzione molto complessa che presenta molteplici minimi locali. Infatti la gran parte dei metodi ha natura stocastica, ciò significa che le soluzioni ottenute attraverso tali algoritmi convergono "quasi certamente" al minimo assoluto della funzione.

Una caratteristica di alcuni algoritmi di ottimizzazione è quella di utilizzare ricerche locali sullo spazio di ottimizzazione. In base alle ipotesi fatte sulla funzione obiettivo, è possibile utilizzare numerosissimi algoritmi di ricerca locale. Formalmente, definito un generico problema di ottimizzazione del tipo:

$$y^* = \min_{x \in S} f(x), \tag{3.17}$$

ipotizziamo che la generica procedura di ricerca locale P sia strettamente decrescente, se P parte da un punto $x \in S$ e converge a un minimo locale x^*, allora esiste un intervallo compreso tra x e x^* lungo il quale i valori di f sono strettamente decrescenti. Inoltre ipotizziamo che i punti di stazionarietà di f, cioè i punti ove il gradiente di f è zero, siano finiti.

Nell'approccio *multistart* il metodo P viene applicato ad ogni punto di un campione estratto da una distribuzione uniforme su S. Tra i molteplici candidati trovati con questo metodo, viene selezionato quello a cui corrisponde il valore minore della funzione f. La procedura sopra descritta risulta però essere poco efficiente poiché inevitabilmente valuta più volte gli stessi minimi locali. È per ovviare a ciò che è stato introdotto l'algoritmo MLSL.

Nel MLSL la procedura di ricerca locale P è applicata ad ogni punto del campione. A differenza del metodo classico, però, il MLSL evita di valutare più di una volta punti che sono già stati valutati o punti che si trovano ad una distanza critica r da essi. L'algoritmo è implementato in modo iterativo, estraendo k campioni di dimensione N da S. Innanzitutto l'algoritmo valuta i campioni estratti nel complesso e implementa la procedura locale P una sola volta anche per punti presenti più volte nei campioni.

Algorithm 1: The MLSL algorithm

```
X* ← ∅; k ← 0
repeat
        k ← k + 1
        Generate N points x_{(k−1)N+1}, ..., x_{kN} with uniform distribution on X.
        Determine the reduced sample (X_r) consisting of the γkN best points from
        the cumulated sample x_1, ..., x_{kN}.
        for i ← 1 to length(X_r) do
                if NOT (there is such a j that f(x_j) < f(x_i) and ||x_j − x_i|| < r_k) then
                        Start a local search method (LS) from x_i.
                        x* ← LS(x_i)
                        X* ← X* ∪ {x*}
                end
        end
until Some global stopping rule is satisfied.
return The smallest local minimum value found.
```

Figura 3.5: Algoritmo MLSL [Pál (2013)]

Inoltre l'algoritmo MLSL, ad ogni iterazione, rimuove una parte delle osservazioni del campione, ad esempio $1 - \gamma$, se queste sono associate a valori elevati della funzione obiettivo. Questa riduzione non ha alcun impatto sulle proprietà teoriche sopra discusse. Una volta fatto ciò, l'algoritmo di ricerca locale valuta tutti i punti all'interno del campione ridotto. Il campione ridotto sarà formato dai migliori γkN punti selezionati dal campione cumulato $x_1, ..., x_{kN}$. Le proprietà teoriche di questo approccio sono solide e ben documentate in letteratura, si veda ad esempio Kan and Timmer (1987) e Rinnooy Kan and Timmer (1987).

Il valore della distanza critica r_k è scelto in modo da minimizzare la probabilità che l'algoritmo inizi una nuova iterazione sebbene sia stato già trovato il minimo globale. Date le ipotesi sopra addotte, possiamo definire la distanza critica come segue:

$$r_k = \pi^{-1/2} \left[\Gamma(1 + \frac{n}{2}) m(S) \quad \sigma \frac{\log(kN)}{kN} \right]^{1/n} \tag{3.18}$$

dove Γ è la funzione gamma di Eulero, n è il numero di variabili del problema, $m(S)$ è la misura di Lebesgue del dominio S, kN è il numero totale di punti nel campione cumulato e k indica l'iterazione attuale. L'algoritmo in figura 3.5 ripete iterativamente la ricerca finché non viene soddisfatto un criterio di stop. Nel nostro caso l'algoritmo si arresta se il beneficio marginale dell'iterazione corrente rispetto alla precedente è minore o uguale di 1×10^{-15}. In questo lavoro è stato utilizzato, come metodo di ricerca locale, l'algoritmo L-BFGS-B. Per ulteriori dettagli a riguardo si veda Fletcher (2013).

3.4 Indicatori LPPLS

Questa sezione si concentra sulla metodologia di costruzione degli indicatori costruiti sulla base del modello LPPLS. La procedura di stima descritta in precedenza permette di ridurre la complessità della funzione di costo ottenendo comunque risultati robusti. Per la stima dei parametri non lineari è stato utilizzato l'algoritmo MLSL di cui si è discusso ampiamente nel paragrafo precedente.

3.4.1 Robustezza rispetto alla finestra di stima

Per quanto riguarda l'intervallo di osservazioni su cui è stato stimato il modello, abbiamo scelto di adottare l'approccio utilizzato da Sornette et al. (2015). In particolare il modello è stato stimato su finestre temporali la cui lunghezza diminuisce nel tempo. Tale scelta deriva dal fatto che il modello LPPLS è molto sensibile rispetto ai dati su cui viene calibrato. Per evitare ciò, gli autori propongono di stimare il modello su molteplici finestre temporali in modo da ottenere risultati robusti. Formalmente, data l'osservazione al tempo *pseudo-presente* t_2, abbiamo stimato i parametri del modello su finestre di lunghezza $[t_1, t_2]$ con t_1 che si avvicina a t_2 in passi da un giorno ciascuno (*shrinking windows*). La lunghezza delle finestre utilizzate in questo lavoro va da 1450 giorni a 40 giorni, per un totale di circa 980 finestre.

Ricapitolando, dato il giorno t_2, abbiamo costruito finestre di lunghezza diversa e su ognuna di queste finestre è stata effettuata la stima del modello. Tale procedura è utile perché ci permette di verificare se le caratteristiche di una bolla speculativa sono presenti in più finestre temporali.

Nella tabella 3.1 riportiamo le condizioni che devono essere rispettate dai parametri del modello affinché la stima sia considerata valida e i vincoli sui parametri utilizzati per la stima con l'algoritmo MLSL. Queste condizioni rappresentano quelli che potremmo definire *fatti stilizzati* del modello LPPL. Essi derivano, in particolare, dai lavori di Sornette et al. (2015), Fantazzini (2016), Shu and Zhu (2020) e Zhang et al. (2016), sebbene non ci siano indicazioni universali e vincolati riguardo tali intervalli.

Per ulteriori dettagli riguardo i vincoli in tabella 3.1 si rimanda il lettore al paragrafo successivo.

Tabella 3.1: **Spazio di ricerca e condizioni filtering**: La tabella sintentizza le condizioni necessarie affinché la stima su una determinata finestra temporale di lunghezza $dt = t_2 - t_1$ sia considerata valida e lo spazio di ricerca per i parametri non lineari utilizzato per l'algoritmo MLSL.

Descrizione	**Parametri**	**Vincoli ottimizzaz.**	**Filtri**	
			Early Warning	End Flag
	m	$[0, 2]$	$[0.01, 1.2]$	$[0.01, 0.99]$
	ω	$[1, 50]$	$[2, 25]$	$[2, 25]$
	t_c	$[t_2 - 0.2dt,\ t_2 + 0.2dt]$	$[t_2 - 0.05dt,\ t_2 + 0.1dt]$	$[t_2 - 0.05dt,\ t_2 + 0.1dt]$
Numero di oscillazioni	$\frac{\omega}{2\pi} \log \left\lvert \frac{t_c - t_1}{t_2 - t_1} \right\rvert$	-	$[2.5, +\infty[$	$[2.5, +\infty[$
Damping	$\frac{m\lvert B\rvert}{\omega\lvert C\rvert}$	-	$[0.8, +\infty[$	$[1, +\infty[$
Rel. error	$\left\lvert \frac{p_t - \hat{p}_t}{\hat{p}_t} \right\rvert$	-	$[0, 0.05]$	$[0, 0.2]$
Residui AR [Test KPSS]	$\log(\hat{p}_t) - \log(p_t) \sim AR(1)$	-	$[0, 0.463]$	$[0, 0.463]$
Hazard Rate	$h(t) = -Bm - \lvert C\rvert \sqrt{m^2 + \omega^2}$	-	$[0, +\infty[$	$[0, +\infty[$

Una volta raccolti i dati sui parametri per le singole finestre temporali si procede a filtrare tali risultati per identificare le stime considerate valide in base ai criteri riportati nella tabella 3.1.

Descriveremo nel seguito l'indicatore implementato in questo lavoro e costruito sulla base delle stime del modello sulle finestre temporali di cui si è discusso in precedenza.

3.4.2 DS LPPLS Confidence ™

L'indicatore DS LPPL Confidence ™ è stato ampiamente studiato in letteratura, in particolare in Zhang et al. (2016), Zhang et al. (2016), Shu and Zhu (2020) e Sornette et al. (2015). La condizioni di filtering riportate nella tabella 3.1 derivano da questi lavori.

Più precisamente, la scelta del *search space* per i parametri del modello deriva dall'esperienza empirica [Sornette and Johansen (2001); Sornette and Zhou (2006); Forró et al. (2015); Yan et al. (2012)].

La variabile *numero di oscillazioni* deve essere maggiore di 2.5 affinché sia possibile distinguere se si è in presenza di un segnale log-periodico genuino o solamente un segnale log-periodico generato dal *noise* della serie storica [Huang et al. (2000)]. La condizione *Relative error* assicura che il prezzo $\hat{p}_t$ non si discosti in modo eccessivo dal prezzo osservato p_t. La condizione *Residui AR* serve a verificare che i residui del modello LPPLS siano stazionari. Come riportato in Geraskin and Fantazzini (2013), il test più adatto per verificare se i residui del modello presentano una radice unitaria o meno è quello di Kwiatkowki, Phillips, Schmidt e Shin (KPPS), si veda a tal proposito Kwiatkowski et al. (1992). Il test KPSS assume come ipotesi nulla la stazionarietà del processo considerato, contro l'alternativa che il processo abbia una radice unitaria. Nella tabella è riportato l'intervallo in cui deve trovarsi la statistica test KPSS affinché il processo sia stazionario con livello di significatività al 10%.

La condizione sul parametro m assicura che il *crash hazard rate* cresca quando il prezzo cresce. Inoltre l'hazard rate deve essere sempre positivo [Bothmer and Meister (2003)] e anche tale condizione è stata inclusa nei filtri. È bene precisare che anche la condizione *Damping* serve ad assicurare che l'hazard rate sia positivo e abbia un andamento crescente, più ci si avvicina a t_c.

L'indicatore DS LPPL Confidence ™ è calcolato come rapporto tra il numero di finestre temporali che rispettano i parametri riportati nella tabella 3.1 rispetto al numero di finestre totali. Tale indicatore misura la sensibilità del pattern rispetto

alla scala temporale dt. Ciò significa che un valore elevato dell'indicatore si traduce in un gran numero di finestre, rispetto al totale, nelle quali è stata identificata una bolla. Viceversa un valore basso indica che il segnale di bolla è presente in poche finestre temporali e quindi il segnale è fragile. L'indicatore può anche essere visto come una probabilità empirica, essendo compreso nell'intervallo $[0, 1]$.

Sulla scorta delle recenti innovazioni introdotte al Financial Crisis Observatory dell' ETH di Zurigo e del lavoro di Gerlach et al. (2020) abbiamo deciso di dividere le finestre temporali in modo da calcolare l'indicatore su diverse scale temporali. Le lunghezze temporali scelte sono le seguenti:

- Super-Short (SS) scale: $[40, 183]$ giorni
- Short (S) scale: $[40, 365]$ giorni
- Medium (M) scale: $[365, 730]$ giorni
- Long (L) scale: $[730, 1460]$ giorni

Questo ci permette di classificare le bolle rilevate dal modello in base alla scala temporale nella quale si sono formate. Nella fase preliminare l'indicatore è calcolato separatamente per le bolle positive e quelle negative, semplicemente identificando le bolle negative come le finestre temporali nelle quali $B > 0$ e quelle positive nelle quali $B < 0$.

Una volta ottenuto il numero di finestre valide con bolla positiva e negativa, rispettivamente n_{pos} e n_{neg}, aggreghiamo i due indicatori in un indicatore unico, che assume valori nell'intervallo $[-1, 1]$, dove valori positivi indicano bolla positiva e viceversa per valori negativi.

L'indicatore complessivo, data la scala temporale, è calcolato come segue:

$$LPPLS \quad C.I. = \frac{n_{pos} - n_{neg}}{n_{tot}} \tag{3.19}$$

dove con $C.I.$ si intende *Confidence Indicator*.

La scelta di calcolare indicatori separati in base alla scala temporale deriva dal fatto che l'orizzonte d'investimento degli agenti può essere sia di lungo, sia di medio che di breve termine.

Di solito le prime avvisaglie di una bolla speculativa in formazione vengono identificate meglio dall'indicatore a brevissimo termine (SS) che risulta infatti il più reattivo ai cambiamenti dei prezzi. Successivamente, se la bolla continua a gonfiarsi, l'indicatore di medio termine (M) aumenta e questo rappresenta sicuramente una conferma aggiuntiva rispetto al segnale di breve termine. Lo stesso discorso vale per l'indicatore a lungo termine (L).

Inoltre la procedura sopra descritta è stata applicata per filtrare i risultati sia rispetto alle condizioni *Early Warning* sia a quelle *End Flag*, si veda tabella 3.1. La presenza di due tipologie di filtri è giustificata dal fatto che il primo (E.W.) è utile ad individuare una bolla in formazione, mentre il secondo (E.F.) indica che la bolla è giunta ad una fase terminale e probabilmente avverrà un cambio di regime, dove per cambio di regime si intende un rallentamento della crescita o un crash.

Ricapitolando, l'indicatore utilizzato nel corso di questo lavoro è il DS LPPLS Confidence ™ calcolato su scale temporali differenti, e utilizzando parametri di filtering diversi per l'indicatore Early Warning e quello End Flag, per un totale di otto indicatori:

- Super-Short (SS)
  - SS.EW
  - SS.EF
- Short (S):
  - S.EW
  - S.EF
- Medium (M):
  - M.EW
  - M.EF
- Long (L):
  - L.EW
  - L.EF

Nel capitolo successivo esporremo i risultati ottenuti tramite la stima del modello LPPLS sia per quanto riguarda singole stime sia per quanto riguarda l'indicatore DS LPPLS Confidence ™. È bene precisare che per ogni singolo giorno di contrattazioni su cui calcolare l'indicatore, sono necessarie 980 stime del modello LPPLS su finestre di lunghezza differente, come spiegato in precedenza. Questa caratteristica rende alquanto costoso, dal punto di vista computazionale, derivare l'indicatore senza dotarsi di una macchina molto performante e utilizzando il calcolo in parallelo per velocizzare le operazioni. Nel nostro caso abbiamo utilizzato il software R per stimare il modello e per calcolare l'indicatore è stato utilizzato il pacchetto *parallel* per velocizzare l'elaborazione dei dati.[1] Di fondamentale importanza per la riuscita del progetto è stata la possibilità, concessaci da CINECA, di utilizzare per un breve periodo di prova l'High Performance Computing cluster *Galileo*.[2]

[1] https://github.com/sabato96/LPPLS
[2] https://www.hpc.cineca.it/hardware/galileo

Capitolo 4

Risultati empirici

Nel seguente capitolo verranno presentati i risultati ottenuti dall'applicazione del modello LPPLS.

4.1 Struttura

Nel presentare i risultati abbiamo deciso di dividere il capitolo in paragrafi, uno per ogni asset analizzato. Per ognuno di essi presenteremo tre tipologie di risultati :

- Modello LPPLS su singola finestra temporale
- Modelli LPPLS su molteplici finestre temporali
- LPPLS Confidence Indicator

Il modello LPPLS su finestra singola, come precisato nei capitoli precedenti, è connotato da una forte sensibilità dei parametri rispetto alla finestra temporale su cui vengono stimati. Abbiamo comunque deciso di inserirlo nei risultati per completezza, sebbene non abbia alcuna utilità pratica date le motivazioni addotte in precedenza.

La rprocedura di stima su finestre temporali multiple deriva dall'esigenza di disporre di stime robuste e affidabili ed è fondamentale per poter derivare gli indicatori LPPLS C.I. Inoltre la stima su finestre temporali multiple permette di osservare dove si concentrano i parametri t_1 e t_c dei vari modelli, in modo da identificare, con buona approssimazione, il periodo in cui la bolla ha avuto inizio e il periodo in cui presumibilmente avrà fine. L'indicatore LPPLS C.I. è utile invece per avere un'indicazione sintetica dei risultati della stima multipla. Esso infatti permette di identificare su quale scala temporale e in quale misura si stia formando una bolla speculativa. Come

precisato nel capitolo precedente, la scelta di dividere gli indicatori rispetto a finestre temporali di lunghezze diverse deriva dall'esigenza di individuare su quale scala temporale ci sia una bolla in formazione. Ciò permette di contestualizzare i risultati ottenuti e le aspettative riguardo un cambiamento di regime in corso. Appare evidente, infatti, che il crash di una bolla speculativa di lungo termine possa essere molto più profondo e improvviso rispetto a quello relativo ad una bolla speculativa di brevissimo termine, che al contrario, può interrompersi senza necessariamente essere accompagnata da un calo consistente del prezzo. Precisiamo inoltre che se più indicatori (su scale temporali differenti) segnalano la presenza di una bolla, il grado di attendibilità di questo segnale cresce dato che il pattern del modello LPPLS è presente sia nel breve termine che nel medio/lungo termine.

4.2 S&P500 (Black Monday)

Per Black Monday si intende il calo generalizzato dei mercati mondiali che avvenne il 19 Ottobre 1987, già discusso approfonditamente nel primo capitolo di questo lavoro. Per quanto riguarda il mercato USA, l'indice Dow Jones perse in un solo giorno il 22.6% del valore, tuttora il più grande calo giornaliero fatto registrare dall'indice. Nel seguito proporremo i risultati delle analisi svolte sull'indice S&P500.

In figura 4.1 è presente il modello LPPLS stimato sulla finestra temporale che va dal 10-08-1983 al 14-09-1987. La stima del parametro t_c per la suddetta finestra temporale (tabella 4.1) indica il 06-10-1987 come giorno in cui sarebbe dovuto avvenire il crash, anticipando di circa tredici giorni il giorno effettivo del crash.

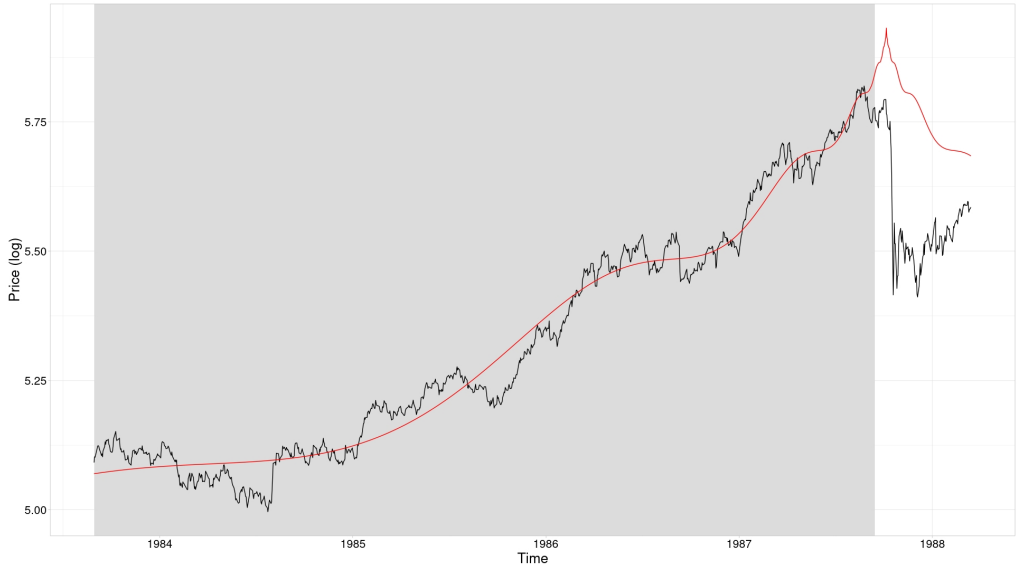


Figura 4.1: Black Monday (S&P500). L'area grigia indica la finestra temporale su cui è stato stimato il modello, mentre la curva rossa rappresenta il modello LPPLS. Per dettagli sui parametri e sulla finestra temporale si veda la tabella 4.1

Tabella 4.1: Parametri LPPLS (modello in figura 4.1)

t_1	t_2	A	B	C_1	C_2	m	ω	t_c
10-08-1983	14-09-1987	5.931	-0.426	-0.016	0.033	0.552	5.474	06-10-1987

Come spiegato in precedenza, il modello LPPLS è connotato da una forte sensibilità dei parametri rispetto alla finestra temporale su cui viene stimato. È proprio da tale caratteristica che deriva l'esigenza di stimare il modello su più finestre temporali

e costruire l'indicatore LPPLS C.I. In figura 4.2 sono rappresentati 15 modelli che soddisfano i vincoli in tabella 3.1.

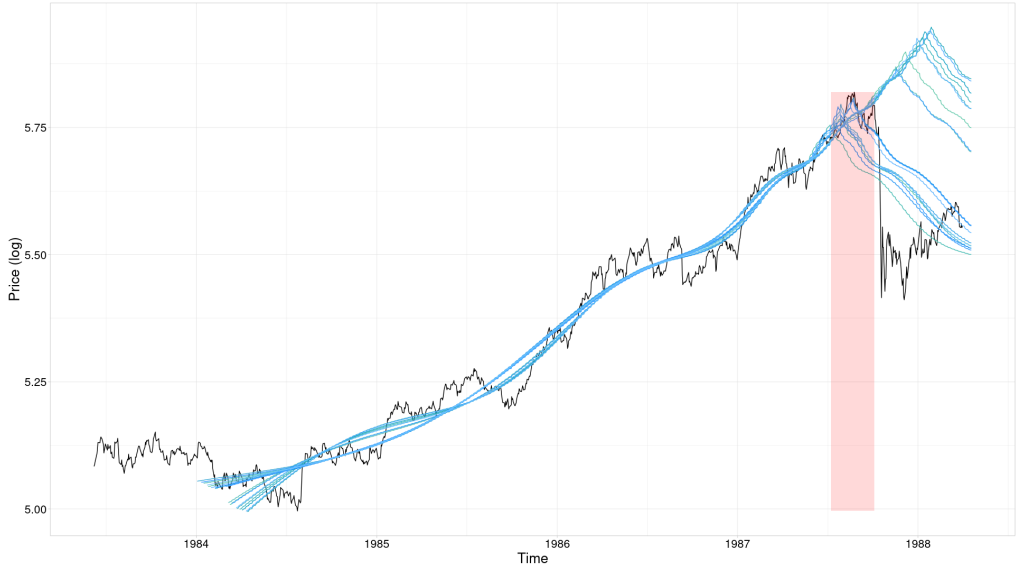
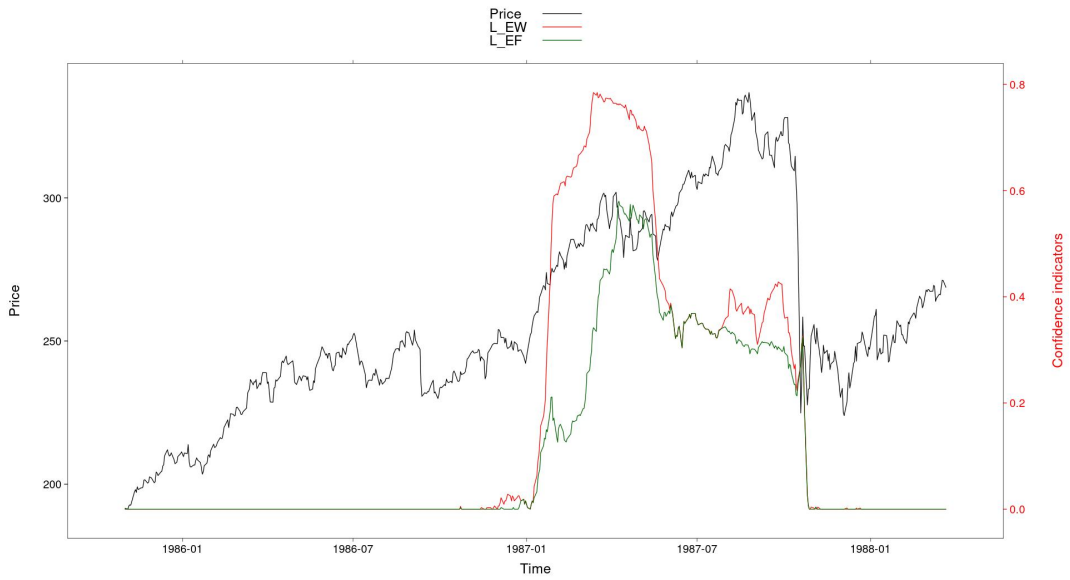
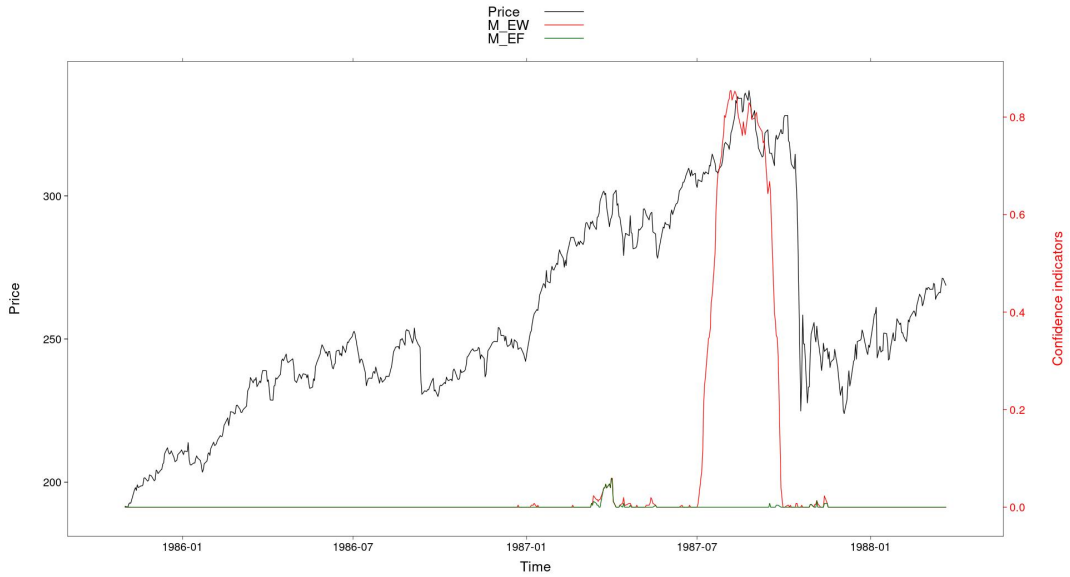


Figura 4.2: Modelli LPPLS multipli stimati su finestre temporali di lunghezza variabile (S&P500 Black Monday). L'area rosa del grafico indica la finestra temporale entro cui si trova il tempo pseudo-presente t_2 per i modelli.

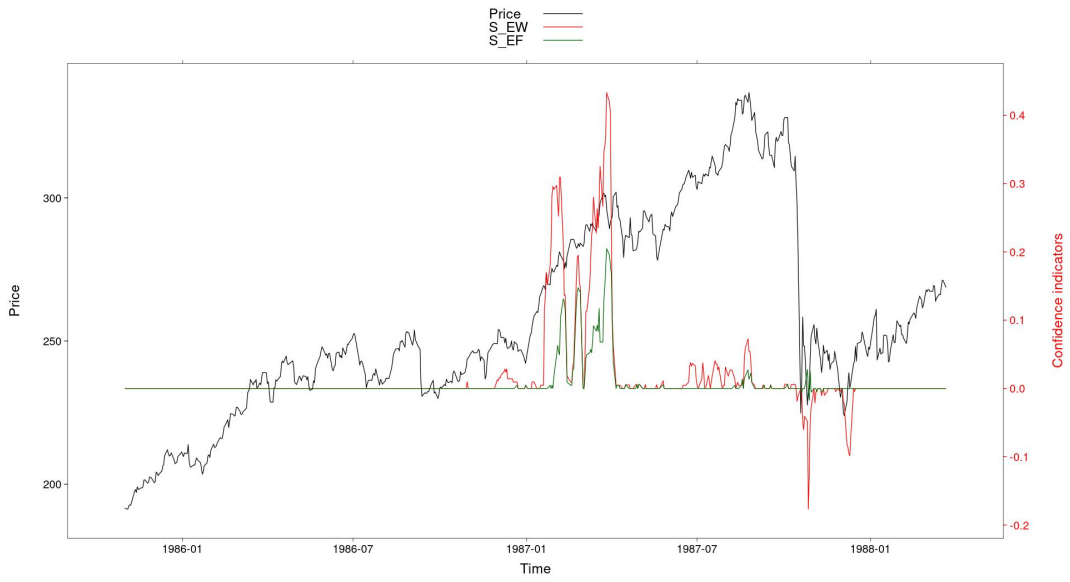
L'area rosa del grafico indica l'intervallo entro cui si trova il parametro t_2 dei modelli. Guardando invece il periodo in cui hanno origine le curve dei modelli è possibile notare come esso coincida col periodo nel quale presumibilmente ha avuto inizio la bolla speculativa (Primo semestre del 1984). Un'ulteriore informazione di fondamentale importanza riguarda il parametro t_c di cui si è molto discusso nella parte introduttiva del lavoro. I modelli mostrano una tendenza a individuare punti per t_c molto vicini tra loro. In figura 4.3, invece, è presente l'indicatore LPPLS C.I. nel caso del Black Monday. Nella figura A.1a sono rappresentati gli indicatori *Early Warning* e *End Flag* a lungo termine (L), rispettivamente in rosso e verde. Nelle figure successive sono presenti gli altri indicatori divisi rispetto alla scala temporale su cui vengono calcolati. Un'importante caratteristica che ci preme sottilineare è il diverso grado di reattività che hanno gli indicatori. Infatti gli indicatori di brevissimo (SS) e breve (S) termine rispondono molto più velocemente alle variazioni di prezzo rispetto agli altri indicatori. Questa caratteristica li rende però meno affidabili rispetto a quelli di medio/lungo termine. Per quanto riguarda l'utilizzo del modello come strumento di risk management, sarebbe preferibile considerare gli indicatori nel complesso valutando la loro evoluzione e adattandoli all'orizzonte temporale su cui opera l'investitore. In appendice è possibile trovare gli indicatori LPPLS C.I. completi per il periodo 2008-2021. [figura A.1]



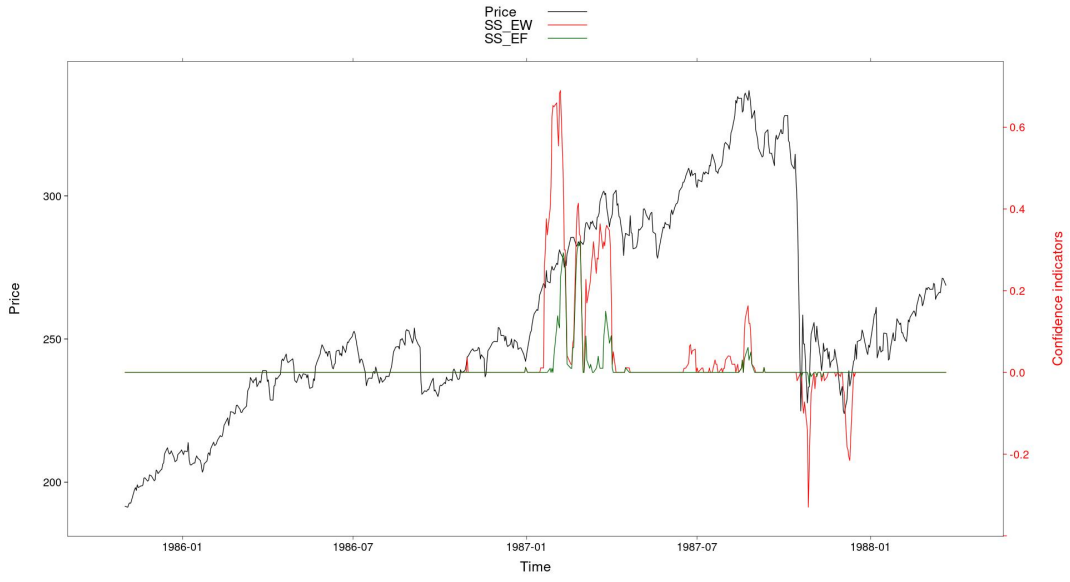
(a) LPPLS C.I. (L)



(b) LPPLS C.I. (M)



(c) LPPLS C.I. (S)



(d) LPPLS C.I. (SS)

Figura 4.3: LPPLS Confidence Indicator su S&P500 (Black Monday). Ogni figura corrisponde ad una scala temporale dell'indicatore, dalla Long a quella Super-Short. In rosso è possibile notare l'indicatore *Early Warning* e in verde l'indicatore *End Flag*.

4.3 Tesla

Tesla Inc. è un'azienda statunitense specializzata nella produzione di auto elettriche, pannelli fotovoltaici e sistemi di stoccaggio energetico fondata nel 2003 in California. Fin dalla quotazione sul mercato, Tesla è stata oggetto delle attenzioni degli investitori a causa dell'impronta ecologica e delle soluzioni innovative proposte come ad esempio il sistema di guida autonoma. L'azienda ha ricevuto anche numerose critiche da parte degli scettici data l'ingente capitalizzazione di mercato raggiunta nonostante indicatori di bilancio e *cash flow* non eccellenti. Alcuni analisti hanno ipotizzato che il successo di Tesla sul mercato sia dovuto a due fattori: il primo attiene a quanto si è detto in precedenza riguardo l'innovazione e la tecnologia all'avanguardia sviluppata dalla società. La seconda riguarda invece il contesto macroeconomico: in particolare tali analisti sostengono che la società sia stata in grado di raccogliere capitale in grandi quantità nonostante scarsi risultati a causa dei bassissimi tassi di interesse praticati dalle banche centrali negli ultimi anni. In figura 4.4 è rappresentato il modello LPPLS stimato su finestra temporale che va dal 23-03-2020 al 19-02-2021, come riportato in tabella 4.2.

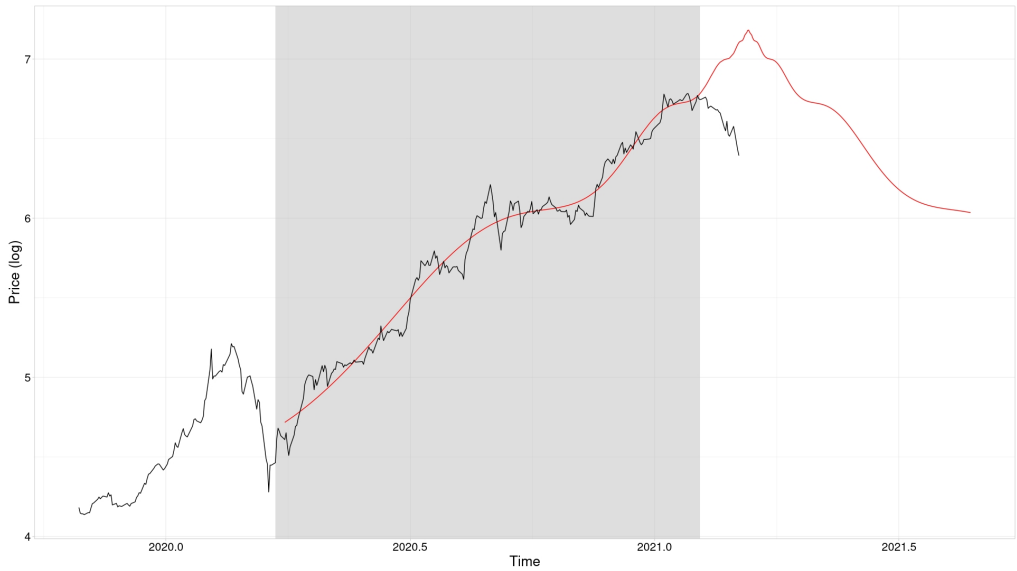


Figura 4.4: TESLA. L'area grigia indica la finestra temporale su cui è stato stimato il modello, mentre la curva rossa rappresenta il modello LPPLS. Per dettagli sui parametri e sulla finestra temporale si veda la tabella 4.2

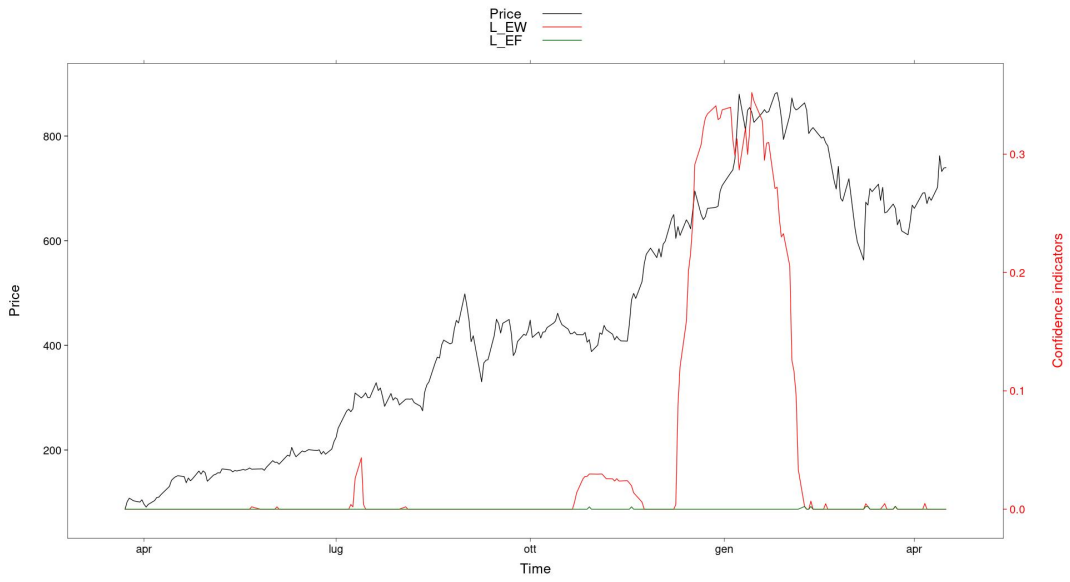
Tabella 4.2: Parametri LPPLS (modello in figura 4.4)

t_1	t_2	A	B	C_1	C_2	m	ω	t_c
23-03-2020	19-02-2021	7.889	-2.569	-0.097	-0.051	0.8	10.96	09-06-2021

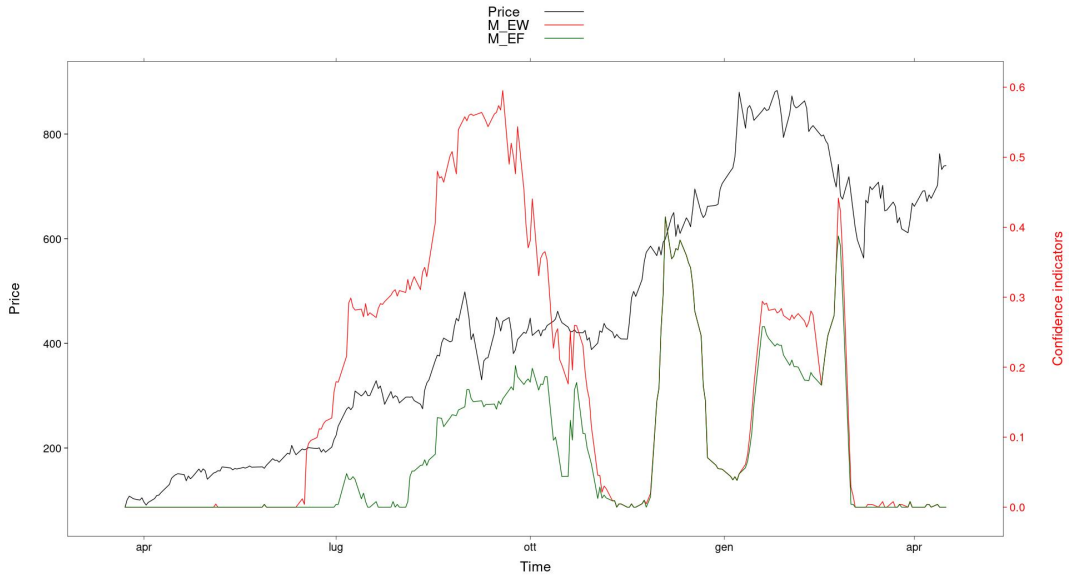
Figura 4.5: TESLA. Modelli LPPLS multipli stimati su finestre temporali di lunghezza variabile. L'area rosa del grafico indica la finestra temporale entro cui si trova il tempo pseudo-presente t_2 per i modelli. È possibile notare come i modelli stimati abbiano come punto iniziale t_1 due periodi diversi. In particolare le curve in viola considerano un intervallo temporale più ampio rispetto a quelle in azzurro chiaro.

In figura 4.5 è possibile osservare molteplici modelli LPPLS. La caratteristica che ci preme far notare è quella che riguarda il parametro t_1 dei vari modelli. Si nota infatti come alcuni modelli individuano l'inizio della bolla sul minimo di metà 2020, mentre i modelli più a lungo termine identificano l'inizio della bolla a Luglio 2019. Anche per quanto riguarda t_c, i modelli a breve termine indicano un t_c più vicino a t_2 rispetto a quelli a lungo termine che lo individuano più avanti nel tempo. Tali differenze si riflettono anche sull'indicatore LPPLS C.I. dato che quest'ultimo tiene conto delle finestre temporali nelle quali la bolla si è sviluppata.

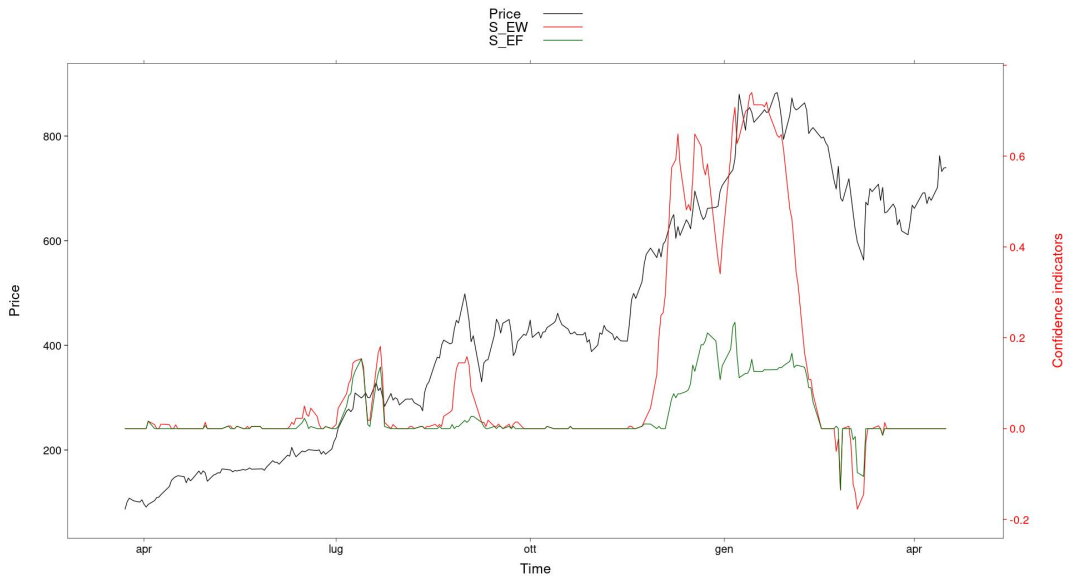
Nel seguito sono riportati gli indicatori LPPLS C.I. come fatto in precedenza per il Black Monday. Per quanto riguarda l'indicatore LPPLS C.I. valgono le stesse considerazioni fatte in precedenza per il caso "Black Monday". Anche in questo caso si può notare come gli indicatori a breve termine siano molto più reattivi a cambiamenti di prezzo mentre quelli a lungo termine lo siano di meno, sebbene questi ultimi diano segnali più affidabili essendo basati su un trend più lungo rispetto a quelli a breve termine.



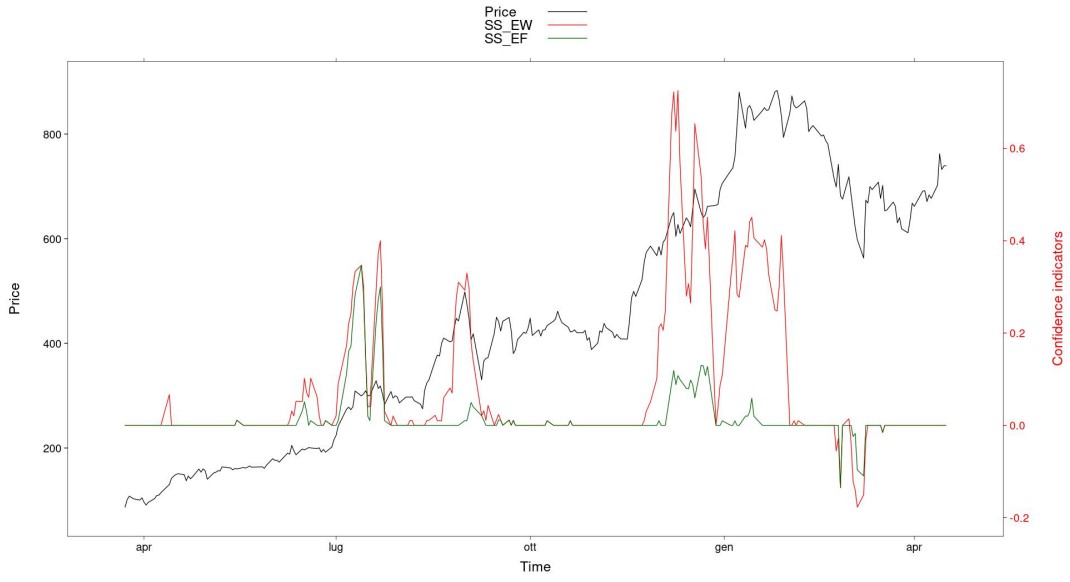
(a) LPPLS C.I. (L)



(b) LPPLS C.I. (M)



(c) LPPLS C.I. (S)



(d) LPPLS C.I. (SS)

Figura 4.6: LPPLS Confidence Indicator su Tesla. Ogni figura corrisponde ad una scala temporale dell'indicatore, dalla Long a quella Super-Short. In rosso è possibile notare l'indicatore *Early Warning* e in verde l'indicatore *End Flag*.

4.4 Bitcoin

Il Bitcoin è una criptovaluta nata nel 2008 e il cui utilizzo su larga scala è cominciato a partire dal 2011. Il Bitcoin è una valuta decentralizzata, cioè svincolata dal sistema di controllo delle Banche Centrali e che può essere trasferita tramite una rete chiamata blockchain, senza l'intervento diretto di intermediari. Nel lavoro di Hileman and Rauchs (2017), gli autori hanno calcolato che nel 2017 c'erano circa 4 milioni di persone che possedevano un wallet per criptovalute, molti dei quali detenevano in portafoglio principalmente Bitcoin. Per ulteriori dettagli si rimanda il lettore al primo capitolo di questo lavoro.

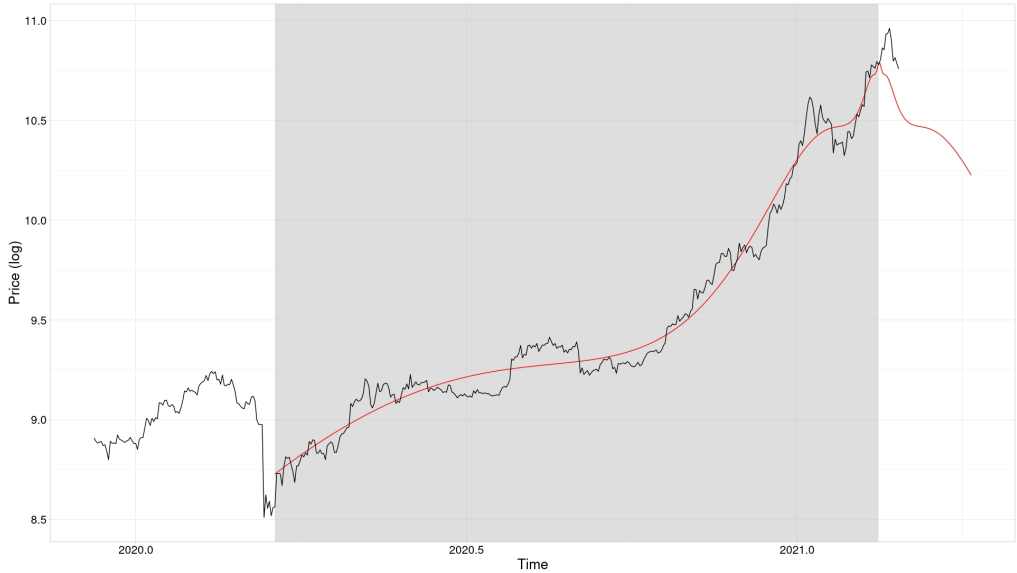


Figura 4.7: Bitcoin. L'area grigia indica la finestra temporale su cui è stato stimato il modello, mentre la curva rossa rappresenta il modello LPPLS. Per dettagli sui parametri e sulla finestra temporale si veda la tabella 4.3

Tabella 4.3: Parametri LPPLS (modello in figura 4.7)

t_1	t_2	A	B	C_1	C_2	m	ω	t_c
08-03-2020	15-02-2021	10.817	-2.661	0.365	-0.366	0.712	3.021	15-02-2021

In figura 4.7 è possibile notare una stima singola del modello LPPLS sulla finestra temporale corrispondente all'area grigia del grafico. In tabella 4.3 sono riportati i parametri del modello in figura: in particolare il fitting singolo indica come tempo critico t_c il 15-02-2021. Nel seguito, come già fatto per gli altri asset, è presente un

grafico composto da fitting multipli del modello LPPLS. In questo caso è possibile notare un caratteristica interessante. In figura 4.8 si nota infatti come ci siano due gruppi di modelli. Il primo gruppo identifica il primo massimo relativo del prezzo come picco della bolla, mentre i modelli stimati su finestre temporali successive identificano il massimo assoluto del prezzo come picco della bolla speculativa. Tale comportamento è dovuto al particolare andamento del prezzo che presenta un massimo relativo e un massimo assoluto successivo. Il modello, come vedremo anche nel caso degli indicatori LPPLS, identifica il primo massimo relativo come un punto critico, dopo il quale effettivamente il prezzo ha subito una lieve correzione. Nei mesi successivi, in particolare nel periodo compreso tra Marzo 2021 e Giugno 2021, la crescita del prezzo del Bitcoin si è arrestata per poi scendere velocemente facendo segnare un prezzo di 42700 $ al 17-05-2021.

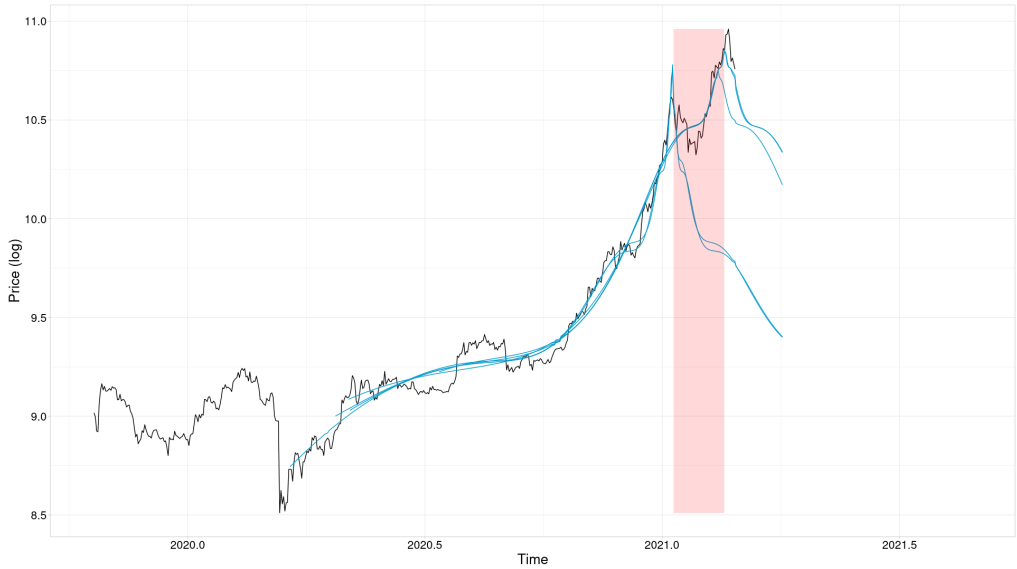
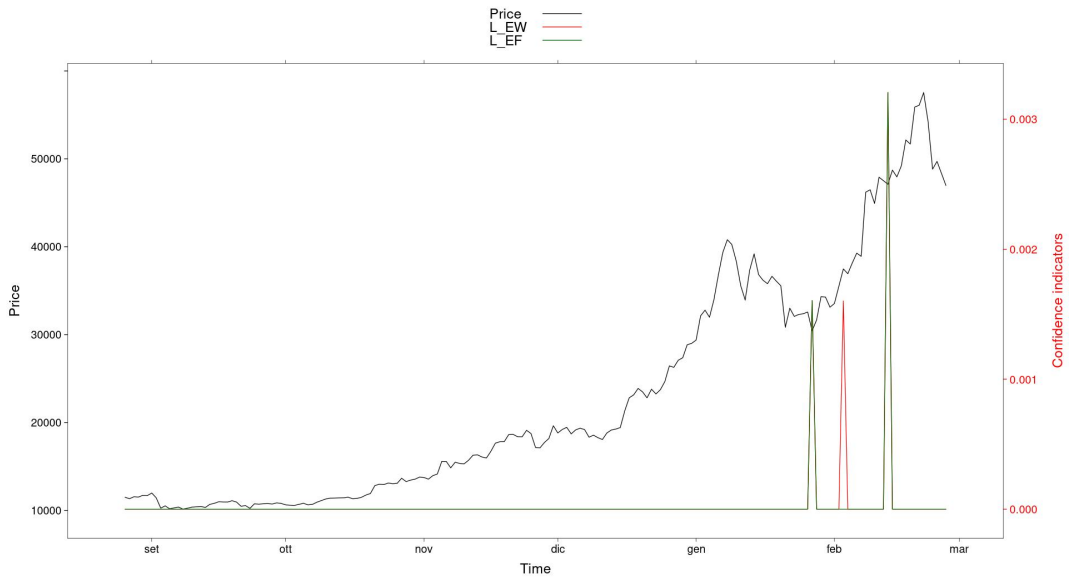
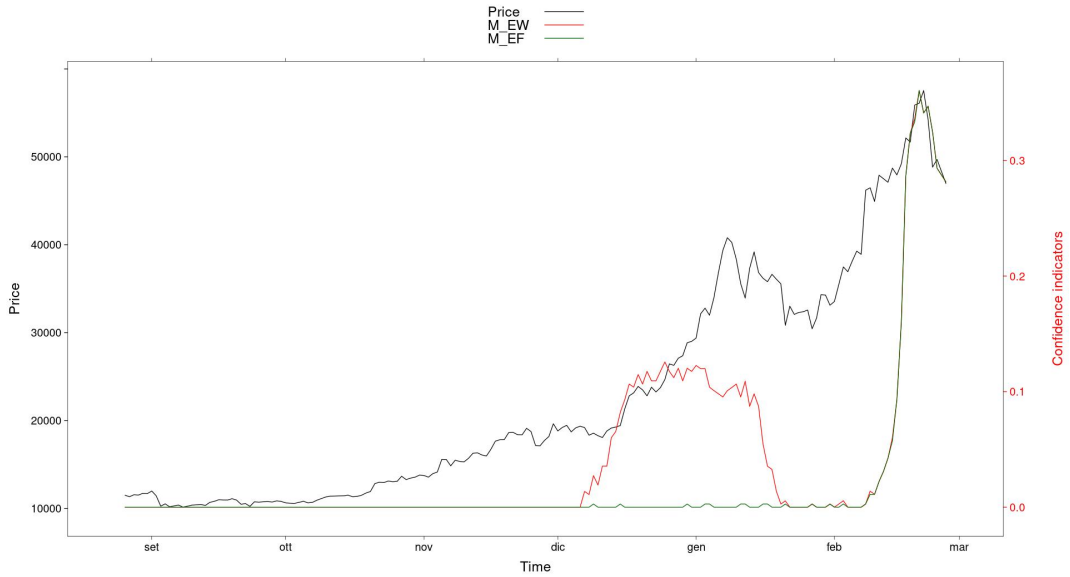


Figura 4.8: Bitcoin. Modelli LPPLS multipli stimati su finestre temporali di lunghezza variabile. L'area rosa del grafico indica la finestra temporale entro cui si trova il tempo pseudo-presente t_2 per i modelli.

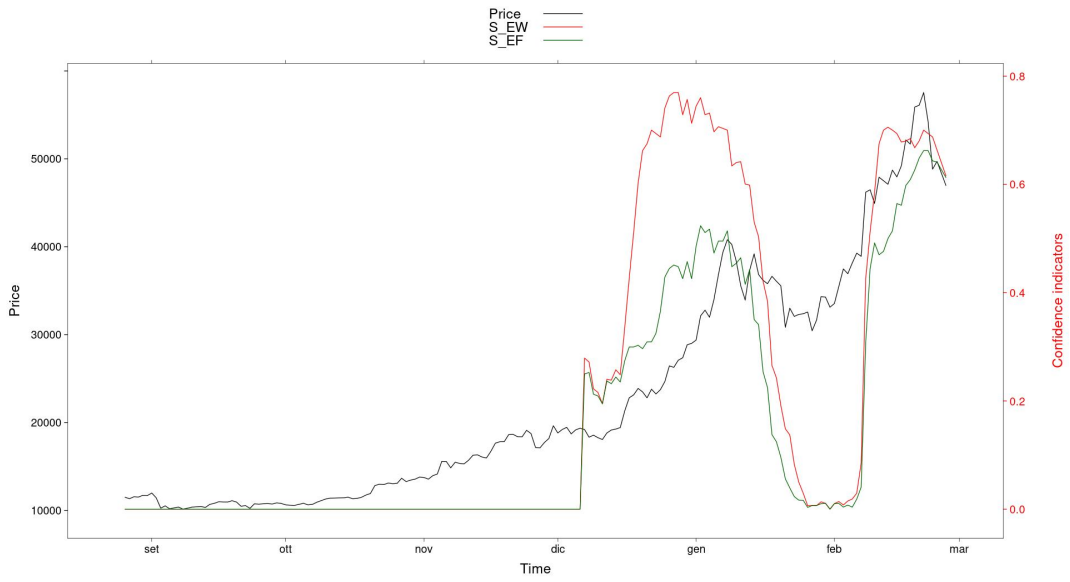
Per quanto riguarda gli indicatori LPPLS in figura 4.9 possiamo notare come anch'essi identifichino il primo massimo relativo come una situazione di insostenibilità che sfocia poi in una lieve correzione del prezzo a ribasso. Successivamente il prezzo accelera ancora e con esso anche gli indicatori. L'unico indicatore che non ha reagito ai movimenti di prezzo è quello a lungo termine in figura 4.9a che ha assunto infatti valori molto bassi. Tale comportamento è dovuto alla presenza di pochi modelli ritenuti validi in base alle condizioni di filtering sulle finestre temporali più lunghe. Gli indicatori su finestre temporali più corte, invece, reagiscono in modo ottimale e riescono ad anticipare sia la prima correzione del prezzo dopo il massimo relativo, sia il cambio di regime del prezzo poi avvenuto tra il mese di Marzo e il mese di Giugno.



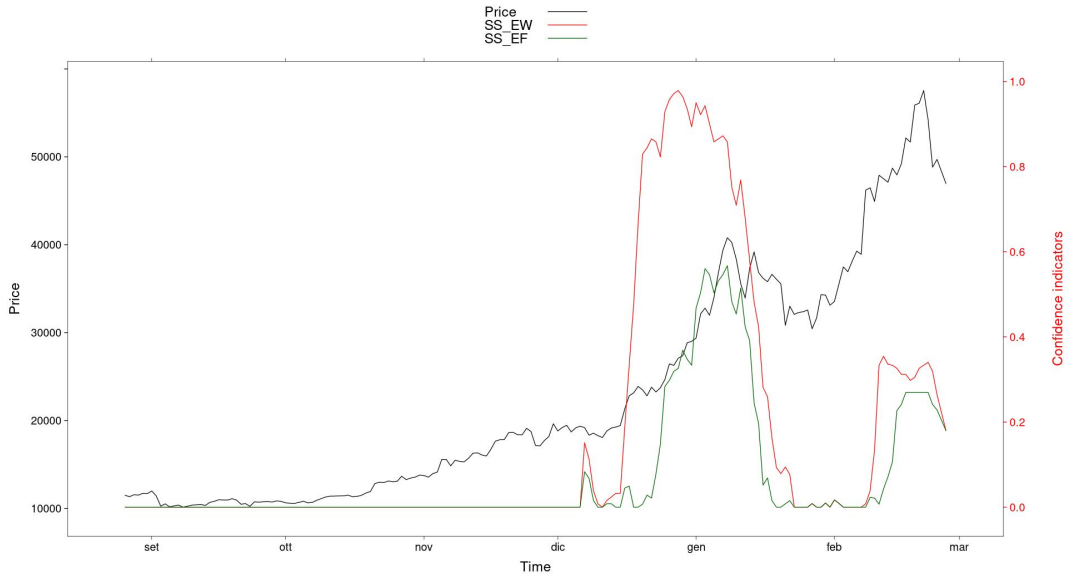
(a) LPPLS C.I. (L)



(b) LPPLS C.I. (M)



(c) LPPLS C.I. (S)



(d) LPPLS C.I. (SS)

Figura 4.9: LPPLS Confidence Indicator su Bitcoin. Ogni figura corrisponde ad una scala temporale dell'indicatore, dalla Long a quella Super-Short. In rosso è possibile notare l'indicatore *Early Warning* e in verde l'indicatore *End Flag*.

4.5 Microsoft

La Microsoft Corporation è un'azienda d'informatica fondata nel 1975. Ad oggi è una delle più importanti aziende al mondo nel settore informatico, nonché una delle più grandi produttrici di software al mondo per fatturato e per capitalizzazione azionaria. In figura 4.10 è rappresentato il modello LPPLS stimato sulla finestra temporale che va dal 09-12-2015 al 03-10-2021.

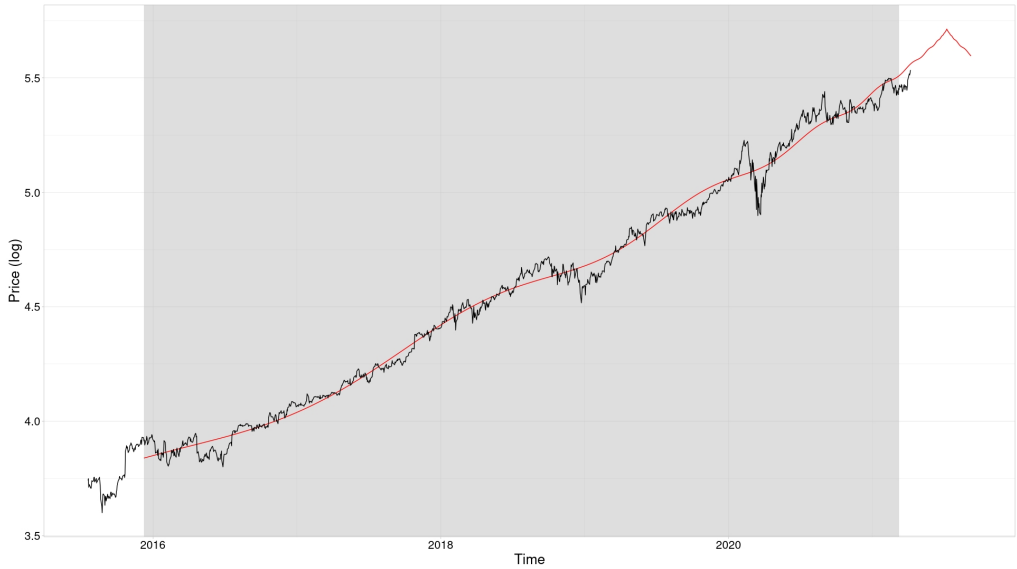


Figura 4.10: Microsoft. L'area grigia indica la finestra temporale su cui è stato stimato il modello, mentre la curva rossa rappresenta il modello LPPLS. Per dettagli sui parametri e sulla finestra temporale si veda la tabella 4.4

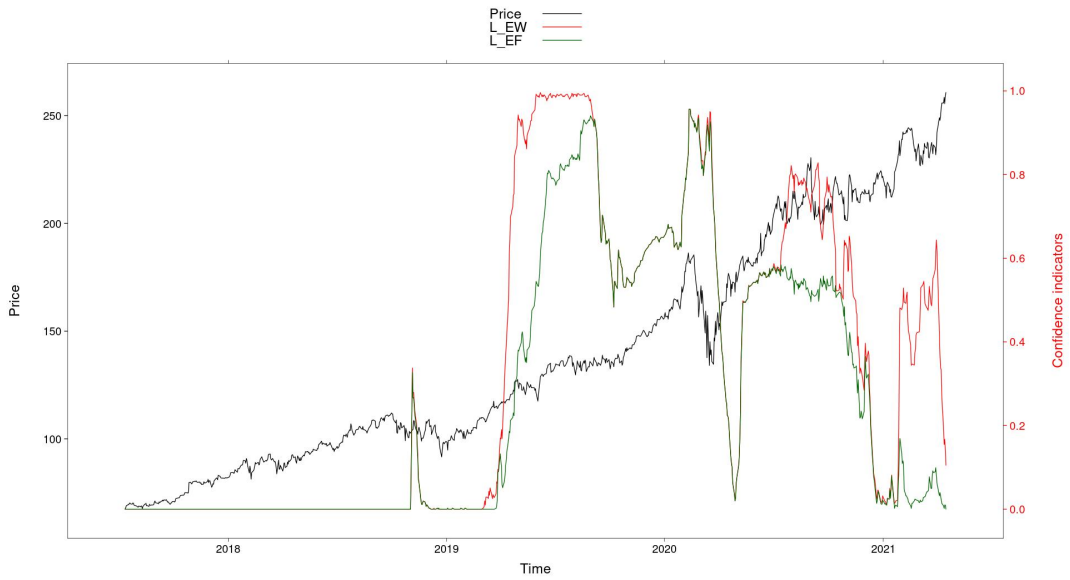
Tabella 4.4: Parametri LPPLS (modello in figura 4.10)

t_1	t_2	A	B	C_1	C_2	m	ω	t_c
09-12-2015	10-03-2021	5.723	-0.479	0.005	-0.016	0.80	9.715	09-07-2021

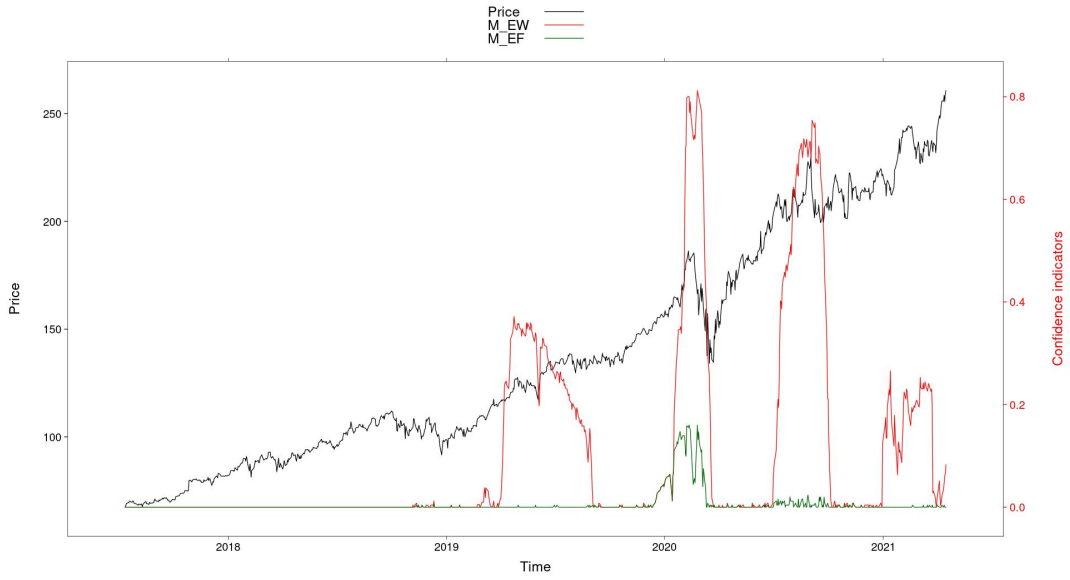
In figura 4.11 sono rappresentati molteplici modelli LPPLS. In questo caso si nota come le stime di t_c dei vari modelli siano molto distanti, a volte anche di alcuni mesi.

Figura 4.11: Microsoft. Modelli LPPLS multipli stimati su finestre temporali di lunghezza variabile. L'area rosa del grafico indica la finestra temporale entro cui si trova il tempo pseudo-presente t_2 per i modelli.

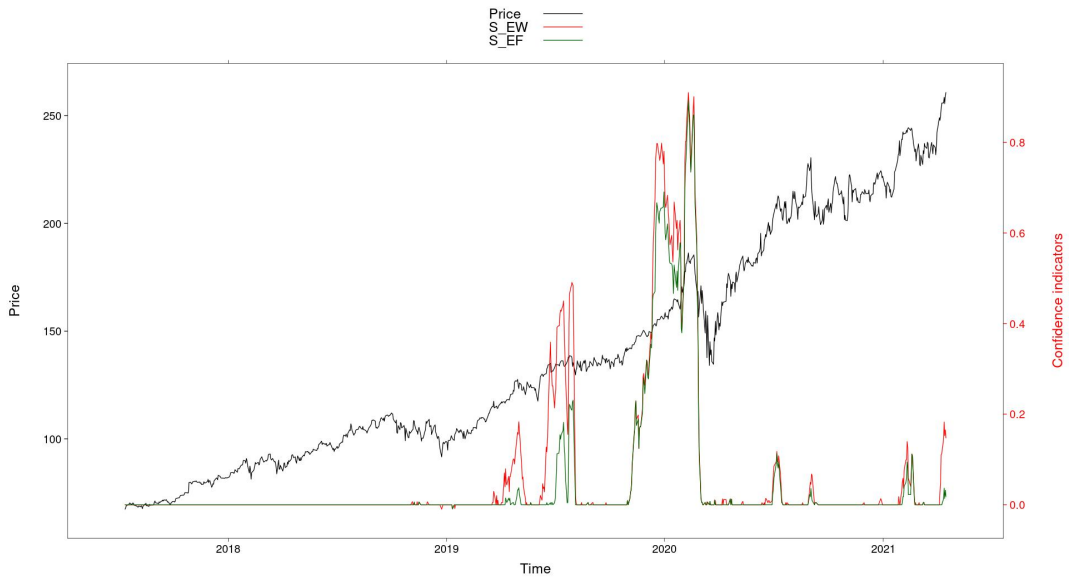
Nel seguito, in figura 4.12 sono riportati i risultati dell'indicatore LPPLS calcolato su varie finestre temporali. Anche in questo caso gli indicatori dimostrano una spiccata capacità di individuare bolle speculative in formazione. È opportuno analizzare il comportamento degli indicatori per quanto riguarda il periodo compreso tra il 2019 e il 2020. Guardando agli indicatori di breve periodo si nota come in quella finestra temporale sia stata identificata una bolla speculativa, che però si conclude con un rallentamento della crescita del prezzo e non con un crash.



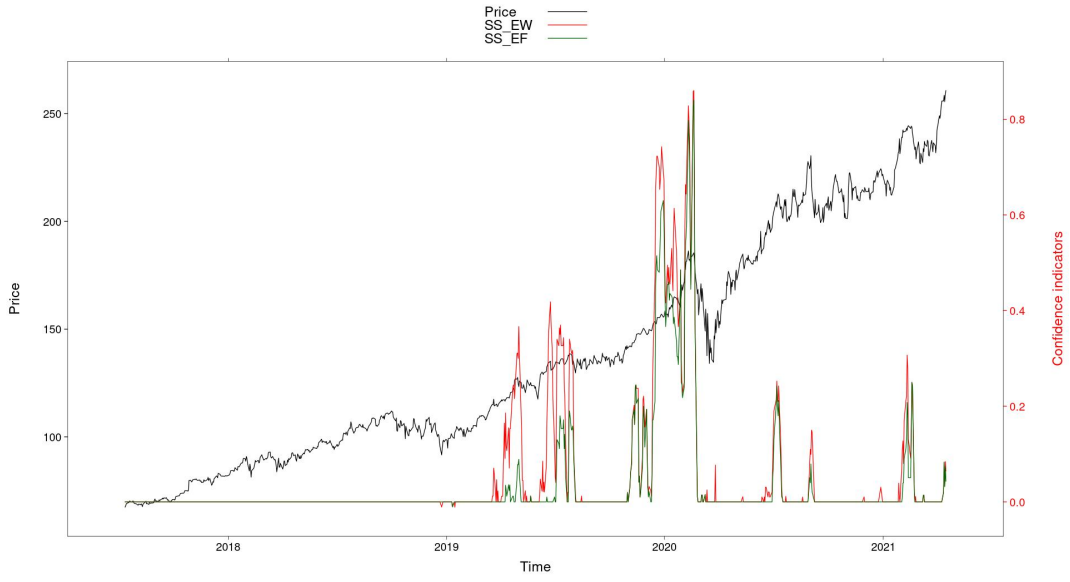
(a) LPPLS C.I. (L)



(b) LPPLS C.I. (M)



(c) LPPLS C.I. (S)



(d) LPPLS C.I. (SS)

Figura 4.12: LPPLS Confidence Indicator su Microsoft. Ogni figura corrisponde ad una scala temporale dell'indicatore, dalla Long a quella Super-Short. In rosso è possibile notare l'indicatore *Early Warning* e in verde l'indicatore *End Flag*.

4.6 Oro

L'oro è da sempre simbolo di ricchezza e prestigio ed è stato utilizzato nel corso dei secoli come riserva di valore. Le civiltà antiche, infatti, erano solite coniare monete d'oro e d'argento per favorire il commercio internazionale. In epoca moderna, il metallo prezioso è sempre stato visto come una protezione utile contro l'inflazione e un ottimo investimento a lungo termine date le sue caratteristiche intrinseche. Anche le Banche Centrali dispongono di riserve auree al fine di effettuare operazioni sul mercato dei cambi e sostenere il valore della propria valuta quando necessario. Nel contesto attuale, connotato da una crescente inflazione dovuta agli interventi statali in relazione alla recente pandemia, l'investimento in oro si sta diffondendo ancor di più tra gli investitori interessati a proteggersi dal rischio di inflazione.

Nel seguito analizzeremo l'andamento dell'oro nel periodo 2009-2012. In questo lasso di tempo l'oro passò da una quotazione (per oncia) di 720 $ il primo Ottobre 2009 a 1896 $ il 1 Agosto 2011. Nei mesi successivi il prezzo dell'oro crollò del 47 % arrivando a 1051 $ per oncia.

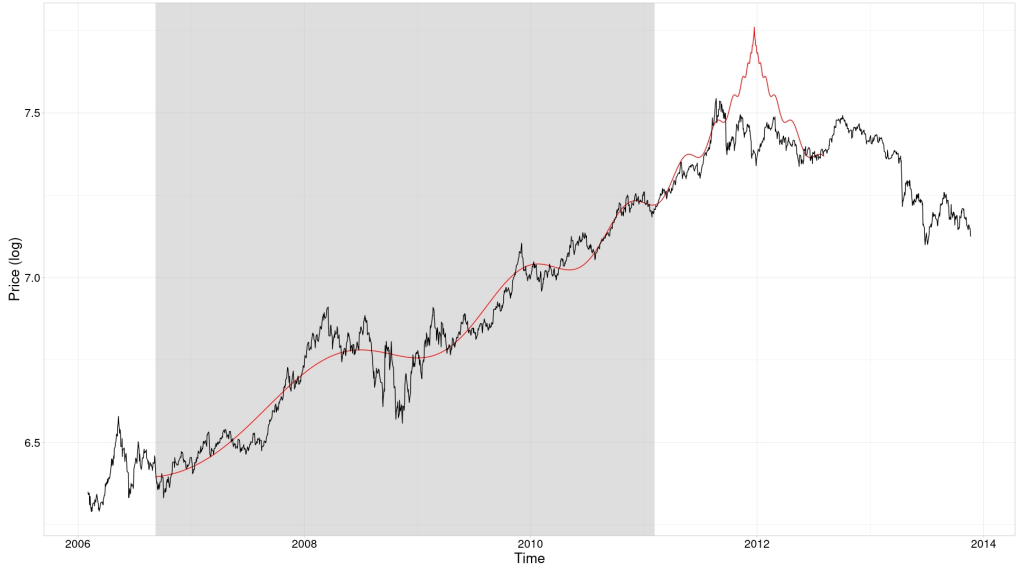


Figura 4.13: Oro. L'area grigia indica la finestra temporale su cui è stato stimato il modello, mentre la curva rossa rappresenta il modello LPPLS. Per dettagli sui parametri e sulla finestra temporale si veda la tabella 4.5

Tabella 4.5: Parametri LPPLS (modello in figura 4.13)

t_1	t_2	A	B	C_1	C_2	m	ω	t_c
07-09-2006	04-02-2011	7.767	-0.551	0.013	0.036	0.512	10.495	22-12-2011

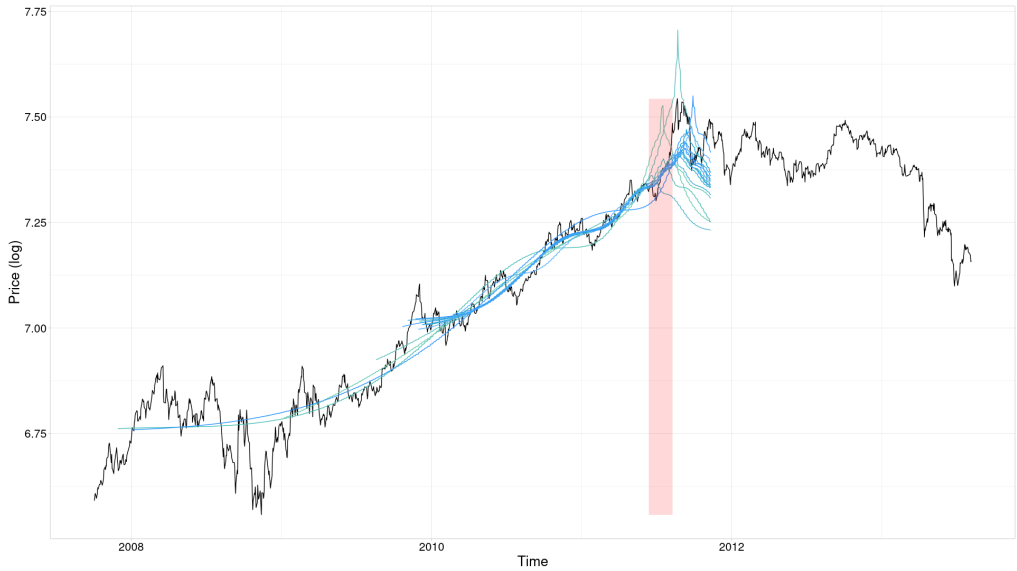
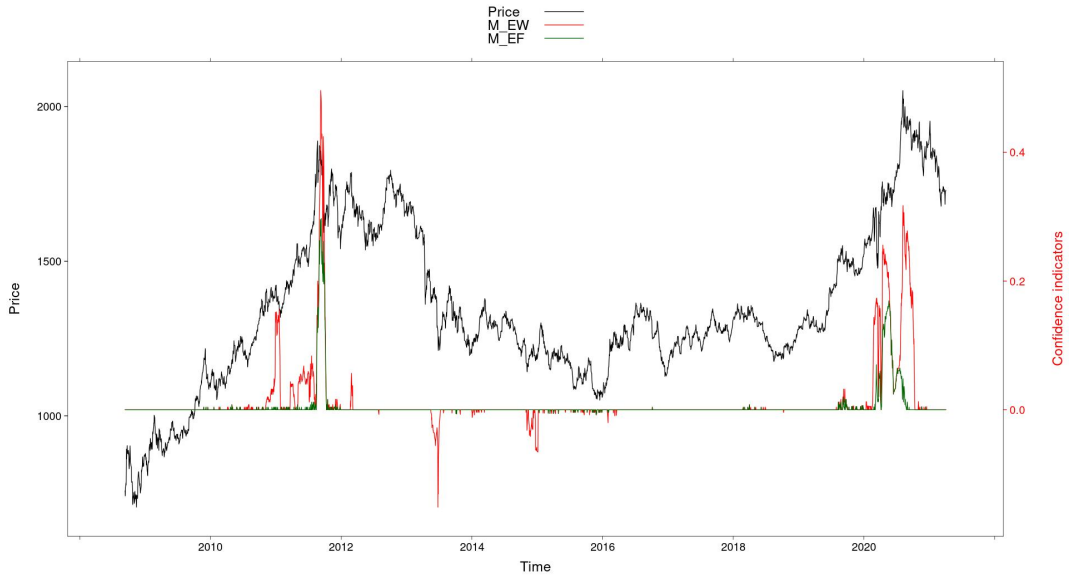


Figura 4.14: Oro. Modelli LPPLS multipli stimati su finestre temporali di lunghezza variabile. L'area rosa del grafico indica la finestra temporale entro cui si trova il tempo pseudo-presente t_2 per i modelli.

In figura 4.14 è possibile notare un'interessante caratteristica della bolla oggetto d'esame. I modelli, innanzitutto, indicano in modo preciso una finestra temporale ristretta per t_c. Oltretutto, guardando ai punti in cui hanno origine i modelli, si nota come ci sia un fascio di modelli che parte dal 2008, mentre un'altra parte dei modelli che invece ha inizio nel 2010. Ciò significa che la procedura di ricerca su più finestre temporali ha individuato una bolla speculativa a lungo termine ed una a breve termine. Tale caratteristica è da interpretare come una conferma della bontà del segnale.

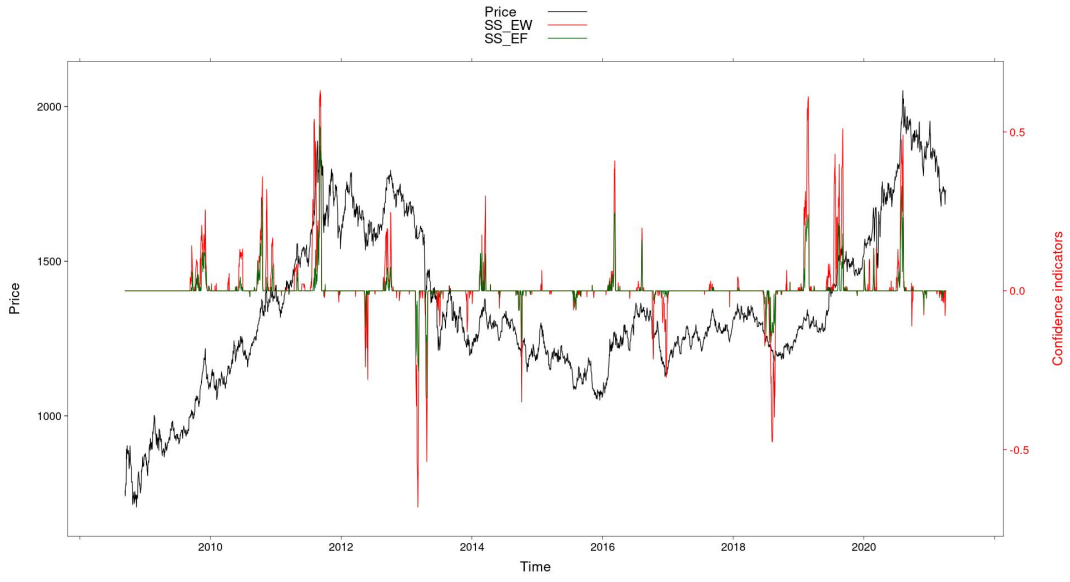
E' possibile notare tale caratteristica anche per quanto riguarda gli indicatori LPPLS. Infatti la bolla del 2009-2012 è correttamente indentificata sia dagli indicatori a lungo termine [figura 4.15a] sia dagli indicatori a breve termine [figura 4.15c].

(a) LPPLS C.I. (L)



(b) LPPLS C.I. (M)

(c) LPPLS C.I. (S)



(d) LPPLS C.I. (SS)

Figura 4.15: LPPLS Confidence Indicator su Oro. Ogni figura corrisponde ad una scala temporale dell'indicatore, dalla Long a quella Super-Short. In rosso è possibile notare l'indicatore *Early Warning* e in verde l'indicatore *End Flag*.

4.7 Saipem

Saipem S.p.A. è una società per azioni fondata in Italia nel 1956. La società opera nel settore dell'energia e delle infrastrutture, in particolare per quanto riguarda le strutture off-shore. La società è quotata nell'indice FTSE MIB della Borsa di Milano. Sulla scia del boom che anticipò la crisi del 2008, il prezzo delle azioni Saipem fece segnare un massimo a 39.4€ per poi precipitare a 12€ dopo la crisi finanziaria del 2008. Nell'anno seguente il valore delle azioni Saipem aumentò del 101% grazie ad un aumento del fatturato e una ripresa del prezzo del petrolio. Tale crescita repentina si arrestò però nel 2011 e il valore delle azioni precipitò ancora una volta negli anni successivi arrivando a 15€ nel Giugno 2013.

In figura 4.16 è possibile osservare la stima singola del modello LPPLS.

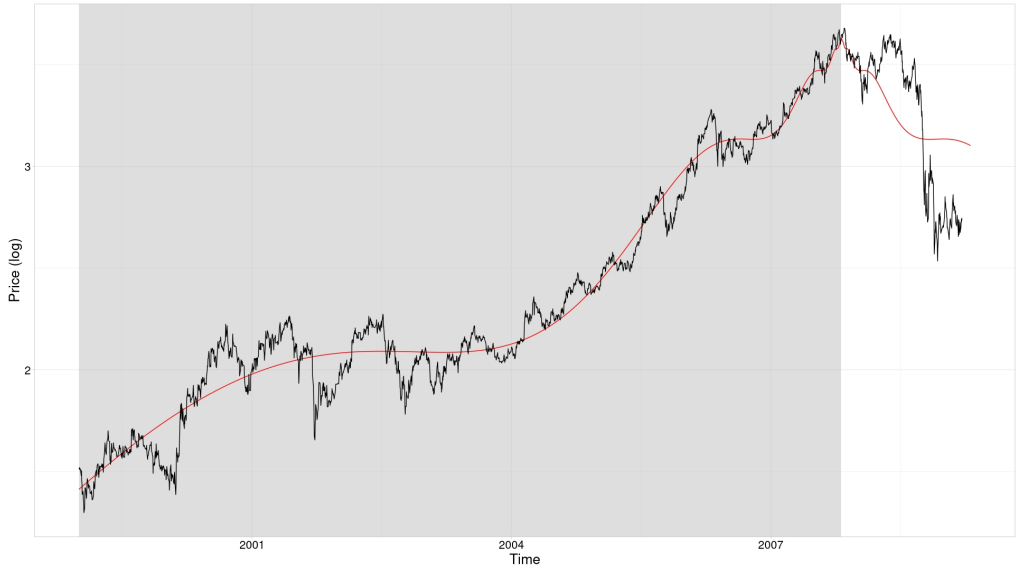


Figura 4.16: Saipem. L'area grigia indica la finestra temporale su cui è stato stimato il modello, mentre la curva rossa rappresenta il modello LPPLS. Per dettagli sui parametri e sulla finestra temporale si veda la tabella 4.6

Tabella 4.6: Parametri LPPLS (modello in figura 4.16)

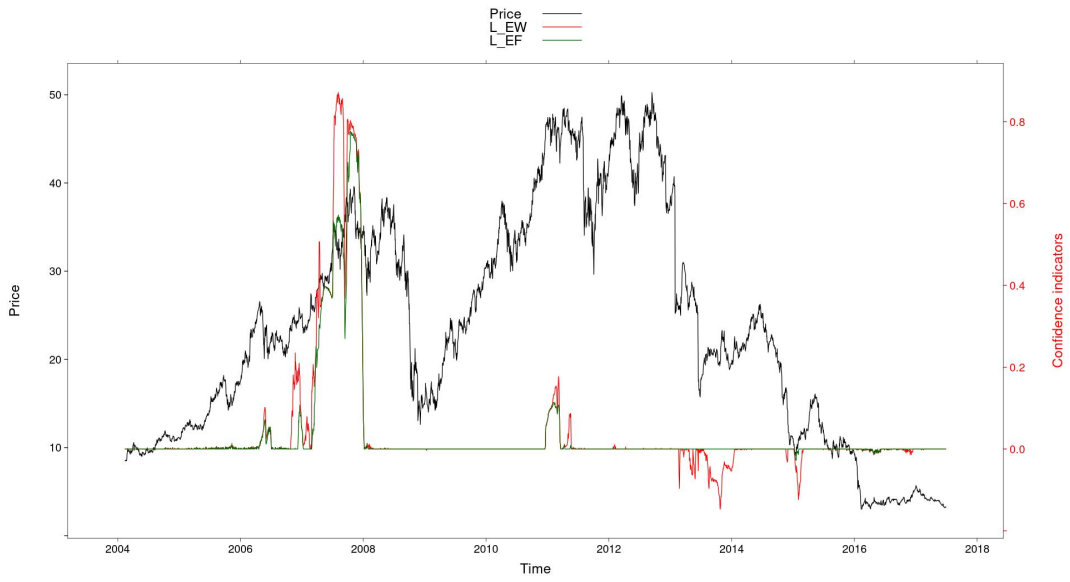
t_1	t_2	A	B	C_1	C_2	m	ω	t_c
01-01-1999	23-10-2007	3.630	-0.480	-0.017	0.087	0.75	4.163	26-10-2007

Per quanto riguarda la figura 4.17, è evidente come anche su finestre temporali di lunghezza differente i modelli identificano il massimo assoluto fatto segnare il 26-

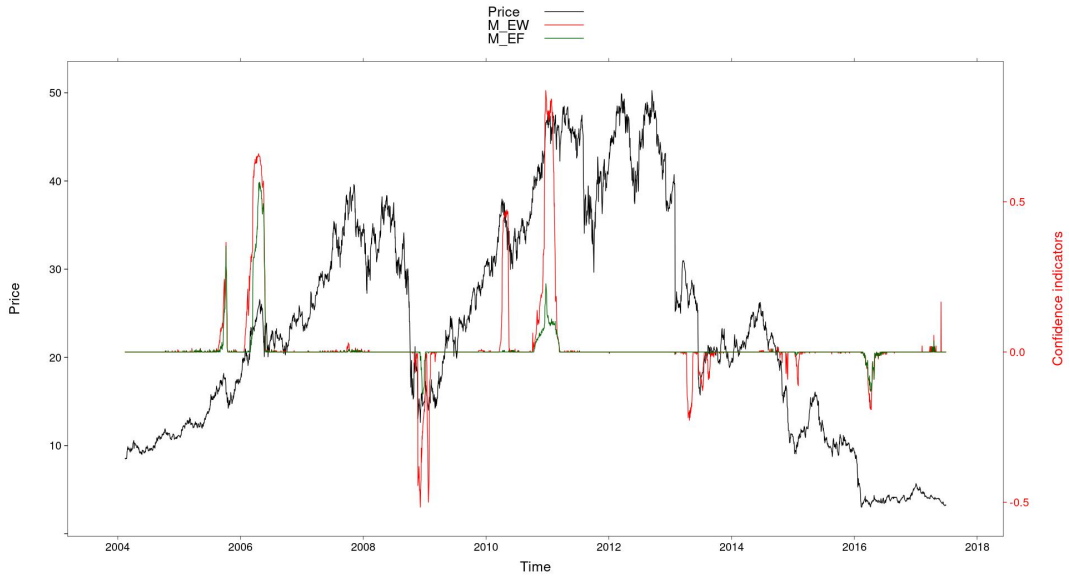
10-2017 come il punto in cui i prezzi avrebbero invertito il proprio corso, come poi effettivamente accadde.

Figura 4.17: Saipem. Modelli LPPLS multipli stimati su finestre temporali di lunghezza variabile. L'area rosa del grafico indica la finestra temporale entro cui si trova il tempo pseudo-presente t_2 per i modelli.

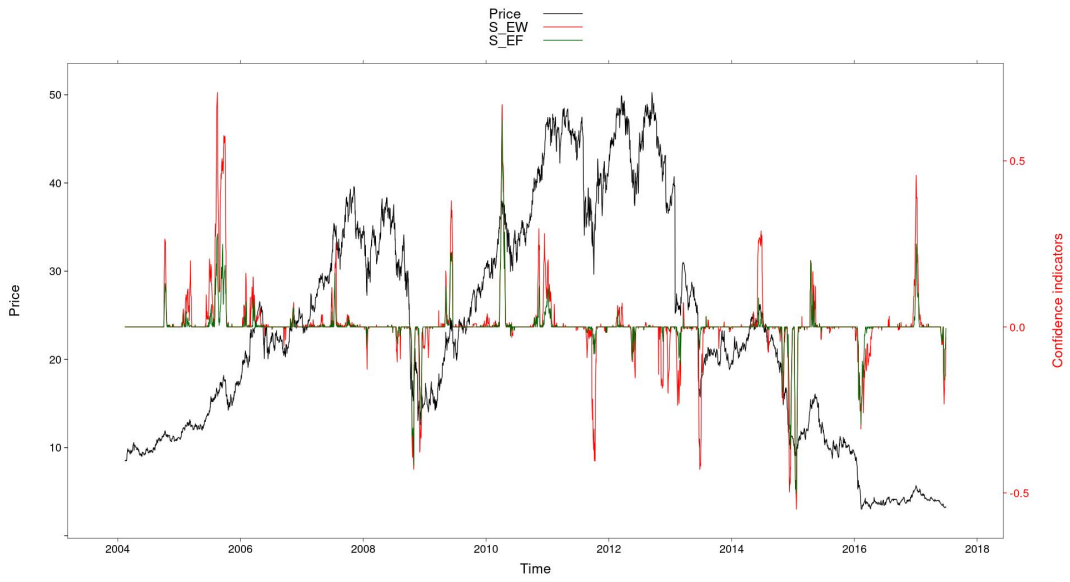
Gli indicatori LPPLS in figura 4.18 identificano chiaramente la crescita repentina dei prezzi pre-2008 come una bolla speculativa. Un particolare importante da notare è come gli indicatori di breve/medio periodo ravvisino anche la presenza di una bolla speculativa negativa in seguito al crash del 2008.



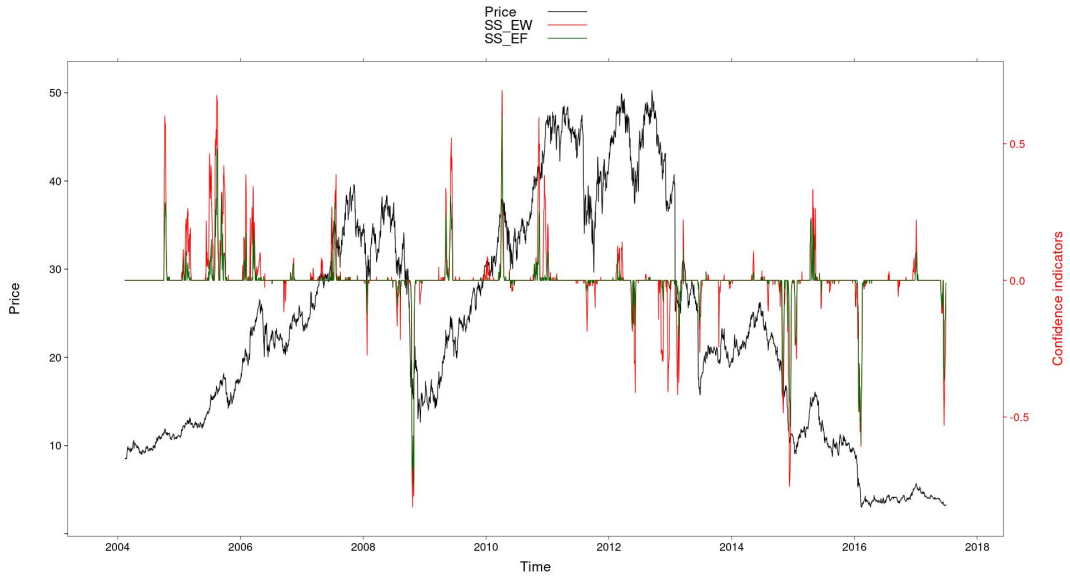
(a) LPPLS C.I. (L)



(b) LPPLS C.I. (M)



(c) LPPLS C.I. (S)



(d) LPPLS C.I. (SS)

Figura 4.18: LPPLS Confidence Indicator su Saipem. Ogni figura corrisponde ad una scala temporale dell'indicatore, dalla Long a quella Super-Short. In rosso è possibile notare l'indicatore *Early Warning* e in verde l'indicatore *End Flag*.

4.8 Petrolio

Lo sfruttamento del petrolio risale ai tempi antichi. Nell' Iliade, Omero narra di un fuoco perenne lanciato contro le navi greche. Nell'antichità veniva infatti miscelato con altre sostanze quali zolfo e salnitro per creare una miscela infiammabile a scopo bellico. L'industria petrolifera moderna nacque nel 1850 negli Stati Uniti. Nonostante negli anni cinquanta del ventesimo secolo il carbone fosse ancora il combustibile più usato al mondo, il petrolio cominciò a sostituirlo essendo una fonte di energia facilmente utilizzabile sia come combustibile che come base di molti prodotti chimici industriali. All'inizio del ventunesimo secolo circa il 90% del fabbisogno di combustibile è coperto dal petrolio. Il prezzo del petrolio è sempre stato oggetto di attenzione da parte sia degli investitori che ovviamente delle nazioni produttrici. Difatti i maggiori produttori di petrolio hanno deciso di riunirsi in un'organizzazione chiamata OPEC, fondata nel 1960, con l'obiettivo di coordinarsi sulla quantità da estrarre al fine di influenzare il prezzo del greggio agendo sull'offerta.

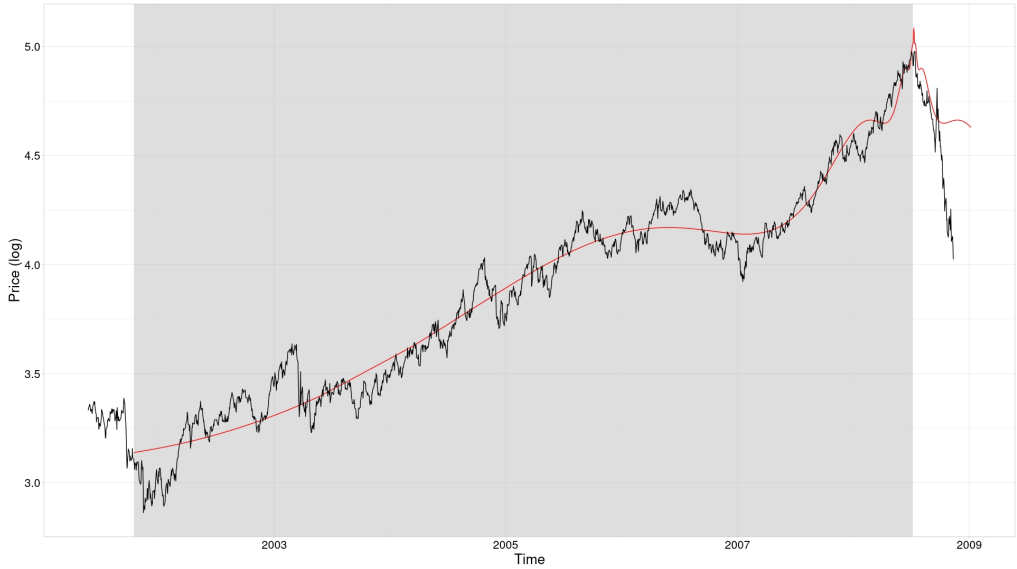


Figura 4.19: Petrolio (2008). L'area grigia indica la finestra temporale su cui è stato stimato il modello, mentre la curva rossa rappresenta il modello LPPLS. Per dettagli sui parametri e sulla finestra temporale si veda la tabella 4.7

Tabella 4.7: Parametri LPPLS (modello in figura 4.19)

t_1	t_2	A	B	C_1	C_2	m	ω	t_c
16-10-2001	08-07-2008	5.122	-0.777	-0.108	-0.04	0.42	3.634	11-07-2008

Nella finestra temporale contrassegnata in grigio in figura 4.19 il prezzo del petrolio (espresso in USD per barile) passò da 24 $ a 148 $. A seguito della crisi finanziaria il valore di un barile di petrolio tornò a 34 $ nel giro di pochi mesi. In figura 4.20 è possibile notare come i modelli stimati su finestre temporali multiple riescano a catturare la presenza di una bolla speculativa. Alcuni modelli, però, indicano come parametro A, cioè $\ln[p(t_c)]$, un prezzo più elevato di quello poi fatto segnare sul massimo assoluto prima del crash.

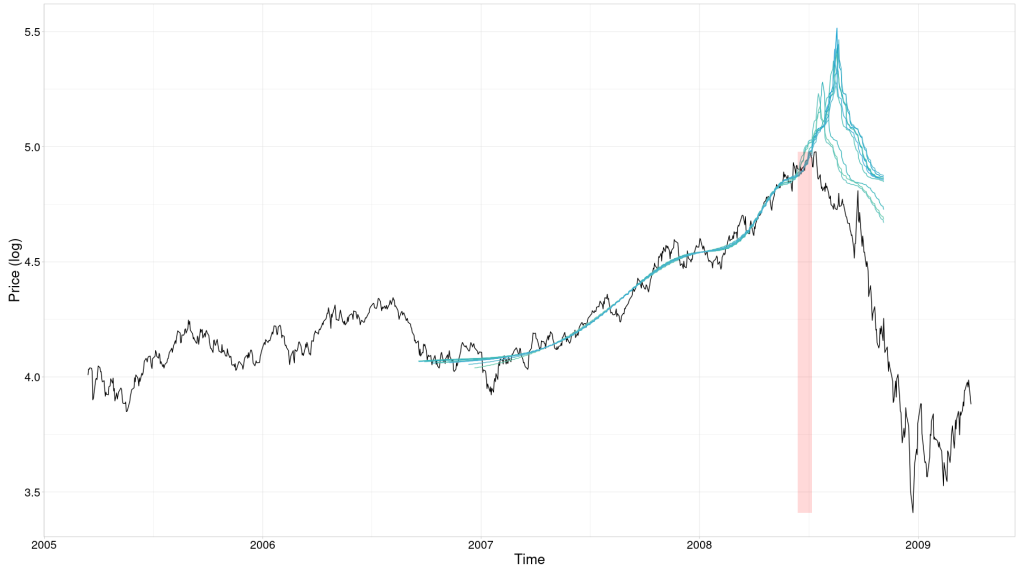
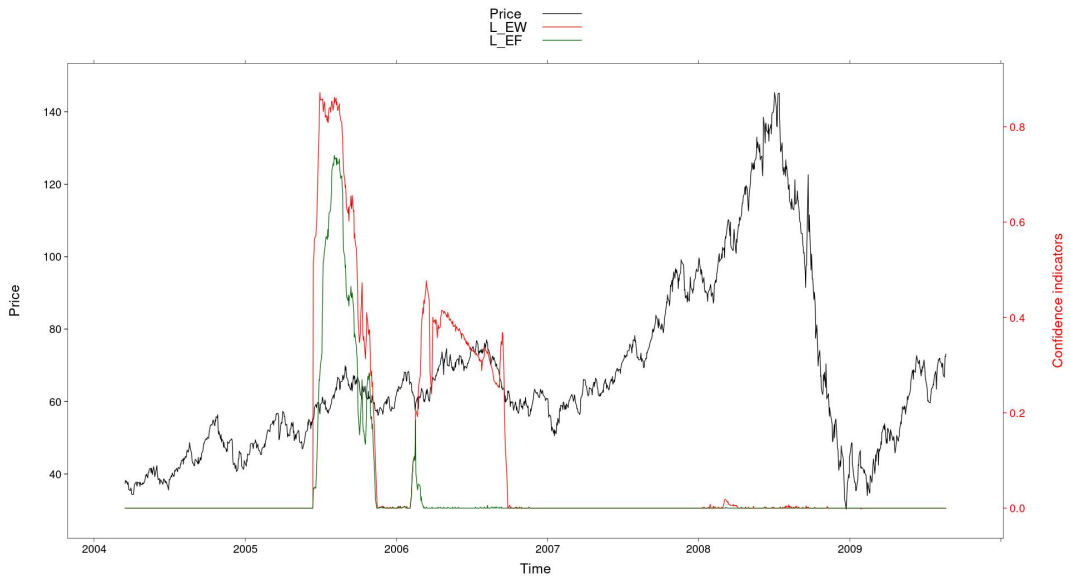
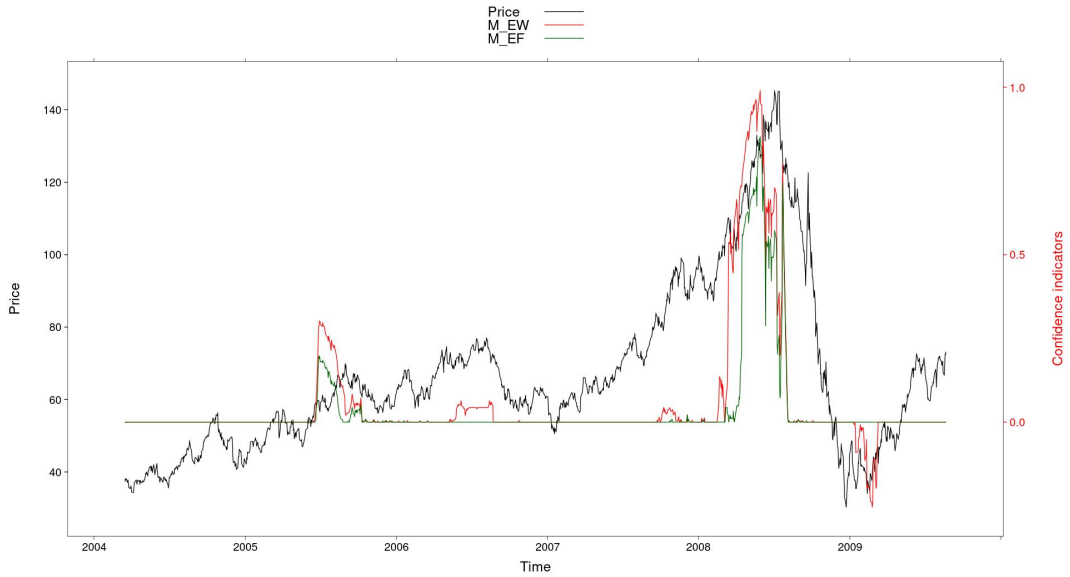


Figura 4.20: Petrolio. Modelli LPPLS multipli stimati su finestre temporali di lunghezza variabile. L'area rosa del grafico indica la finestra temporale entro cui si trova il tempo pseudo-presente t_2 per i modelli.

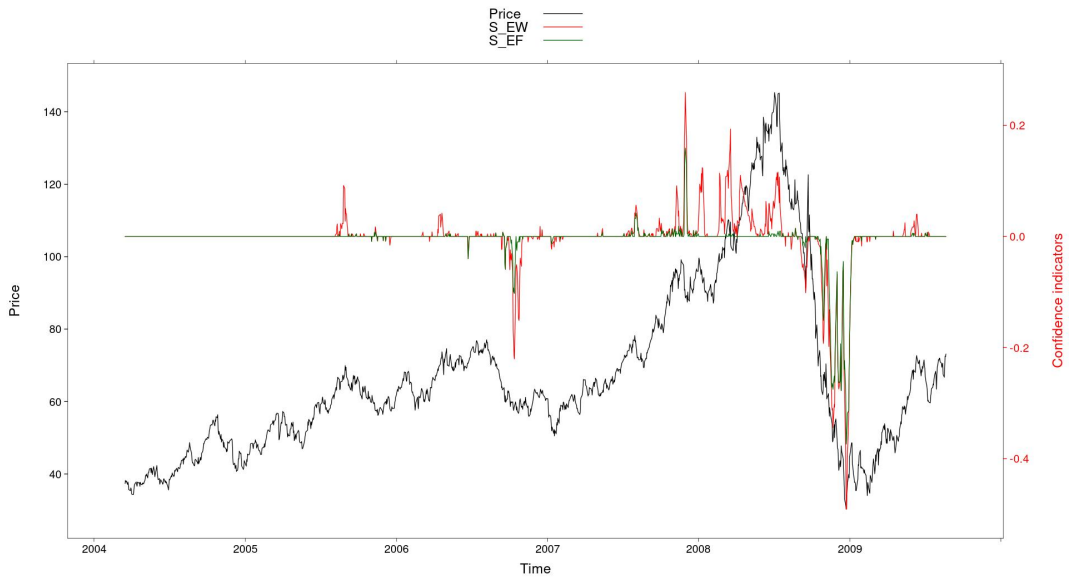
Per quanto riguarda gli indicatori LPPLS C.I. presenteremo, separatamente, sia quelli relativi al periodo temporale sopra considerato (bolla pre-2008) sia quelli relativi agli anni più recenti [figura B.1 in appendice].



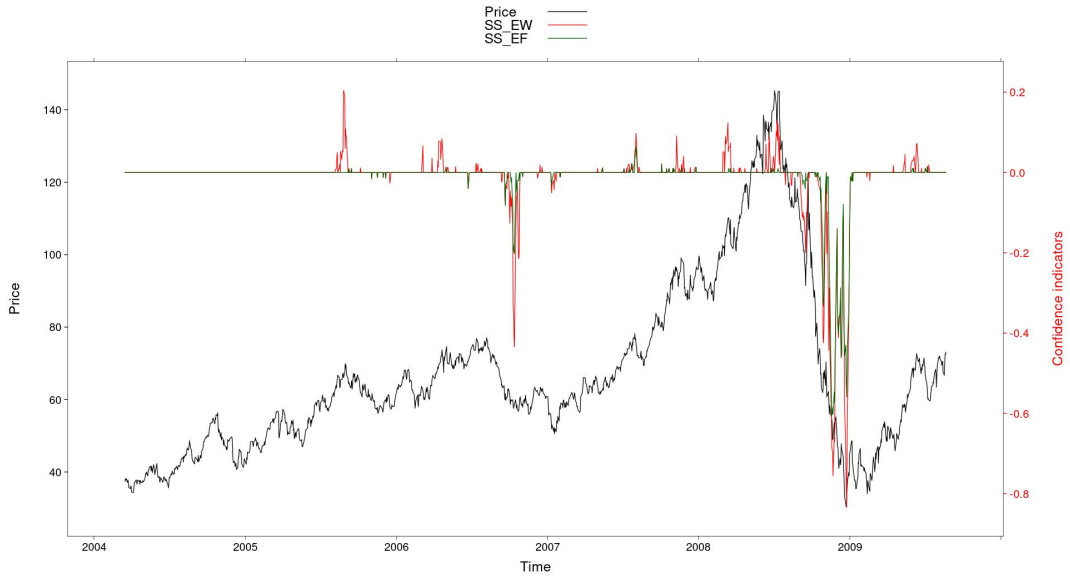
(a) LPPLS C.I. (L)



(b) LPPLS C.I. (M)



(c) LPPLS C.I. (S)



(d) LPPLS C.I. (SS)

Figura 4.21: LPPLS Confidence Indicator su Petrolio. Ogni figura corrisponde ad una scala temporale dell'indicatore, dalla Long a quella Super-Short. In rosso è possibile notare l'indicatore *Early Warning* e in verde l'indicatore *End Flag*.

Conclusioni

In questo lavoro ci siamo occupati in primis di dare una definizione formale di bolla speculativa, trattando nel dettaglio le caratteristiche socio-economiche e statistiche di questi eventi. Successivamente abbiamo presentato il modello LPPLS. Tale modello si basa, come ampiamente argomentato nel corso del lavoro, sull'ipotesi secondo la quale le bolle speculative in formazione abbiano un andamento log-periodico con oscillazioni la cui frequenza aumenta più ci si avvicina al tempo critico. Questa caratteristica deriva dalle interazioni che avvengono tra agenti razionali e *noise trader*, i quali manifestano tendenze imitative. Sono proprio tali tendenze imitative a generare la bolla speculativa e l'andamento oscillatorio è dovuto al "conflitto" tra la componente imitativa dei *noise trader* e i trader razionali. Quando le tendenze imitative si diffondono su tutta la rete il sistema raggiunge un *punto critico* e diviene insostenibile, causando un crash.

L'obiettivo del modello LPPLS è individuare questo andamento oscillatorio e utilizzarlo per modellare l'evoluzione futura del prezzo, cercando anche di individuare l'istante temporale t_c in cui la crescita del prezzo potrebbe interrompersi. Data la forte sensibilità dei parametri rispetto alla finestra temporale su cui vengono stimati, sono state proposte numerose procedure per ottenere risultati robusti e affidabili.

In questo lavoro abbiamo utilizzato l'indicatore LPPLS Confidence Indicator introdotto da Sornette and Cauwels (2015) e utilizzato in molteplici lavori [Zhang et al. (2016), Zhang et al. (2016), Shu and Zhu (2020) e Sornette et al. (2015)]. In questi e numerosi altri lavori l'indicatore si è dimostrato utile per identificare le bolle speculative in formazione e il livello di rischio insito in esse.

Altre tipologie di indicatori sono stati proposti, come ad esempio il LPPLS Trust, la cui derivazione risulta molto più costosa dal punto di vista computazionale dato che esso richiede una procedura di ricampionamento bootstrap per ogni finestra temporale su cui è stimato il modello.

L'indicatore LPPLS C.I. ha dimostrato la sua validità anche in questo lavoro, fornendo un segnale robusto e valido per identificare in anticipo la formazione di bolle speculative su svariati assets e su diverse finestre temporali. Il modello LPPLS, a parere di chi scrive, rientra quindi tra gli strumenti fondamentali da implementare per monitorare i livelli di rischio sul mercato finanziario, senza sostituire altre misure di rischio bensì integrandole. È di fondamentale importanza, sia per le banche d'investimento che per i policy maker e le autorità di vigilanza riuscire ad identificare gli eccessi sui mercati al fine di prevenire eventuali crisi finanziarie scatenate da un'eccessiva speculazione.

Le politiche monetarie espansive adottate dalle Banche Centrali e connotate da tassi d'interesse vicini allo zero o negativi hanno causato un eccesso di liquidità. Questo eccesso invece che trasferirsi all'economia reale per supportare la crescita economica mediante l'aumento dei crediti concessi alle imprese, unito alla ridotta redditività dei titoli di Stato, ha incentivato le banche e i fondi ad effettuare investimenti sull'azionario e su assets più rischiosi per garantire un flusso di reddito agli shareholders. Ciò ha fatto inevitabilmente aumentare la capitalizzazione del mercato azionario e dei derivati mondiale gonfiando quella che alcuni analisti definiscono la più grande bolla speculativa della storia.

Le Banche Centrali si trovano così in una situazione nella quale qualsiasi notizia di una futura stretta monetaria causa turbolenze su un mercato azionario ai massimi storici e ormai dipendente dal capitale a costo zero. Il recente aumento dell'inflazione negli Stati Uniti e, in misura minore, anche in Europa, pone le Istituzioni di fronte a sfide difficili e pregne di potenziali conseguenze. Ci auguriamo che esse sappiano affrontare al meglio le sfide che ci attendono in futuro e che riescano a porre rimedio alle pesanti distorsioni causate da un decennio di tassi d'interesse negativi, indirizzando l'economia mondiale su un percorso di crescita sostenibile.

Bibliografia

Anadu, K., M. Kruttli, P. McCabe, and E. Osambela (2020). The shift from active to passive investing: Risks to financial stability? *Financial Analysts Journal 76*(4), 23–39.

Anderson, T. (1992). Introduction to hotelling (1931) the generalization of student's ratio. In *Breakthroughs in Statistics*, pp. 45–53. Springer.

Barro, R. J., E. F. Fama, D. R. Fischel, A. H. Meltzer, R. Roll, and L. G. Telser (1989). *Black Monday and the future of financial markets*. Mid American Institute for Public Policy Research Inc.

Blanchard, O. (1979). Speculative bubbles, crashes and rational expectations. *Economics Letters 3*(4), 387–389.

Boissevain, J. (1979). Network analysis: a reappraisal. *Current anthropology 20*(2), 392–394.

Bothmer, H.-C. G. v. and C. Meister (2003). Predicting critical crashes? a new restriction for the free variables. *Physica A: Statistical Mechanics and its Applications 320*, 539–547.

Chang, G. and J. Feigenbaum (2006). A bayesian analysis of log-periodic precursors to financial crashes. *Quantitative Finance 6*(1), 15–36.

Chernick, M. R. (2011). *Bootstrap methods: A guide for practitioners and researchers*, Volume 619. John Wiley & Sons.

Cvijović, D. and J. Klinowski (1995). Taboo search: an approach to the multiple minima problem. *Science 267*(5198), 664–666.

De Grauwe, P. and Y. Ji (2013). Self-fulfilling crises in the eurozone: An empirical test. *Journal of International Money and finance 34*, 15–36.

Demirer, R., G. Demos, R. Gupta, and D. Sornette (2019). On the predictability of stock market bubbles: evidence from lppls confidence multi-scale indicators. *Quantitative Finance 19*(5), 843–858.

Derrida, B., L. De Seze, and C. Itzykson (1983). Fractal structure of zeros in hierarchical models. *Journal of Statistical Physics 33*(3), 559–569.

Dubrulle, B., F. Graner, and D. Sornette (1997). Scale invariance and beyond. In *Les Houches Workshop, March 10-14, 1997*, Volume 10, pp. 1. Springer.

Ecotalker (2019). Term of the week: Tulip mania. https://ecotalker.wordpress.com/2019/08/23/term-of-the-week-tulip-mania/.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance 25*(2), 383–417.

Fantazzini, D. (2016). The oil price crash in 2014/15: Was there a (negative) financial bubble? *Energy Policy 96*, 383–396.

Farmer, R. E. (2015). The stock market crash really did cause the great recession. *Oxford Bulletin of Economics and Statistics 77*(5), 617–633.

Filimonov, V. and D. Sornette (2013). A stable and robust calibration scheme of the log-periodic power law model. *Physica A: Statistical Mechanics and its Applications 392*(17), 3698–3707.

Fletcher, R. (2013). *Practical methods of optimization*. John Wiley & Sons.

Forró, Z., R. Woodard, and D. Sornette (2015). Using trading strategies to detect phase transitions in financial markets. *Physical Review E 91*(4), 042803.

Gazola, L., C. Fernandes, A. Pizzinga, and R. Riera (2008). The log-periodic-ar (1)-garch (1, 1) model for financial crashes. *The European Physical Journal B 61*(3), 355–362.

Geraskin, P. and D. Fantazzini (2013). Everything you always wanted to know about log-periodic power laws for bubble modeling but were afraid to ask. *The European Journal of Finance 19*(5), 366–391.

Gerlach, J.-C., C. Zhao, D. Sornette, et al. (2020). Forecasting financial crashes: A dynamic risk management approach. *Swiss Finance Institute Research Paper* (20-103).

Glaser, F., K. Zimmermann, M. Haferkorn, M. C. Weber, and M. Siering (2014). Bitcoin-asset or currency? revealing users' hidden intentions. *Revealing Users' Hidden Intentions (April 15, 2014). ECIS*.

Goldenfeld, N. (2018). *Lectures On Phase Transitions And The Renormalization Group*. CRC Press.

Gouriéroux, C. and J.-M. Zakoian (2016). Local explosion modelling by noncausal process.

Gulsen, M., A. Smith, and D. Tate (1995). A genetic algorithm approach to curve fitting. *International Journal of Production Research 33*(7), 1911–1923.

Hayashi, F. (2000). Econometrics, princeton university press: Princeton.

Hileman, G. and M. Rauchs (2017). Global cryptocurrency benchmarking study. *Cambridge Centre for Alternative Finance 33*, 33–113.

Huang, Y., A. Johansen, M. Lee, H. Saleur, and D. Sornette (2000). Artifactual log-periodicity in finite size data: Relevance for earthquake aftershocks. *Journal of Geophysical Research: Solid Earth 105*(B11), 25451–25471.

Johansen, A., O. Ledoit, and D. Sornette (2000). Crashes as critical points. *International Journal of Theoretical and Applied Finance 03*(02), 219–255.

Johansen, A. and D. Sornette (2000). The nasdaq crash of april 2000: Yet another example of log-periodicity in a speculative bubble ending in a crash. *The European Physical Journal B-Condensed Matter and Complex Systems 17*(2), 319–328.

Johansen, A. and D. Sornette (2002). Large stock market price drawdowns are outliers. *Journal of Risk 4*, 69–110.

Johansen, A., D. Sornette, et al. (2010). Shocks, crashes and bubbles in financial markets. *Brussels Economic Review 53*(2), 201–253.

Kan, A. R. and G. T. Timmer (1987). Stochastic global optimization methods part i: Clustering methods. *Mathematical programming 39*(1), 27–56.

Keynes, J. M. (1936). *The General Theory of Employment, Interest, and Money.* Macmillan London.

Kim, Y. B., J. Lee, N. Park, J. Choo, J.-H. Kim, and C. H. Kim (2017). When bitcoin encounters information in an online forum: Using text mining to analyse user opinions and predict value fluctuation. *PloS one 12*(5), e0177630.

Kristoufek, L. (2013). Bitcoin meets google trends and wikipedia: Quantifying the relationship between phenomena of the internet era. *Scientific reports 3*(1), 1–7.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics 54*(1-3), 159–178.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics 2*(2), 164–168.

Liggett, T. M. (2012). *Interacting particle systems*, Volume 276. Springer Science & Business Media.

Liggett, T. M. et al. (1997). Stochastic models of interacting systems. *The Annals of Probability 25*(1), 1–29.

Lillo, F. and R. N. Mantegna (2000). Symmetry alteration of ensemble return distribution in crash and rally days of financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems 15*(4), 603–606.

Lin, L., R. Ren, and D. Sornette (2009). A consistent model of 'explosive' financial bubbles with mean-reversing residuals. *Swiss Finance Institute Research Paper* (09-14).

Lukáš, P. and K. Taisei (2017). Volatility analysis of bitcoin price time series. *Quantitative Financ and Econ 1*, 474–485.

Malkiel, B. G. (1999). *A random walk down Wall Street: including a life-cycle guide to personal investing.* WW Norton & Company.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Narayan, P. K., S. S. Sharma, and D. H. B. Phan (2016). Asset price bubbles and economic welfare. *International Review of Financial Analysis 44*, 139–148.

Nelder, J. A. and R. Mead (1965). A simplex method for function minimization. *The computer journal 7*(4), 308–313.

Osterrieder, J. and J. Lorenz (2017). A statistical risk assessment of bitcoin and its extreme tail behavior. *Annals of Financial Economics 12*(01), 1750003.

Pál, L. (2013). Benchmarking a hybrid multi level single linkage algorithm on the bbob noiseless testbed. In *Proceedings of the 15th annual conference companion on Genetic and evolutionary computation*, pp. 1145–1152.

Rinnooy Kan, A. and G. Timmer (1987). Stochastic global optimization methods part ii: Multi level methods. *Mathematical Programming: Series A and B 39*(1), 57–78.

Roehner, B. and D. Sornette (2000). "thermometers" of speculative frenzy. *The European Physical Journal B-Condensed Matter and Complex Systems 16*(4), 729–739.

Shefrin, H. (2008). *A behavioral approach to asset pricing.* Elsevier.

Shu, M. and W. Zhu (2020). Detection of chinese stock market bubbles with lppls confidence indicator. *Physica A: Statistical Mechanics and its Applications 557*, 124892.

Sornette, D. (2009, 04). Why stock markets crash: Critical events in complex financial systems. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*, 421.

Sornette, D. and P. Cauwels (2015). Financial bubbles: Mechanisms and diagnostics. *Review of Behavioral Economics 2*(3), 279–305.

Sornette, D., G. Demos, Q. Zhang, P. Cauwels, V. Filimonov, and Q. Zhang (2015, 01). Real-time prediction and post-mortem analysis of the shanghai 2015 stock market bubble and crash. *SSRN Electronic Journal*.

Sornette, D. and A. Johansen (2001). Significance of log-periodic precursors to financial crashes. *Quantitative Finance 1*, 452–471.

Sornette, D., A. Johansen, and J.-P. Bouchaud (1996). Stock market crashes, precursors and replicas. *Journal de Physique I 6*(1), 167–175.

Sornette, D. and W.-X. Zhou (2006). Predictability of large future changes in major financial indices. *International Journal of Forecasting 22*(1), 153–168.

Talbi, E.-G. (2009). *Metaheuristics: from design to implementation*, Volume 74. John Wiley & Sons.

Wilson, K. G. (1979). Problems in physics with many scales of length. *Scientific American 241*(2), 158–179.

Yan, W., R. Rebib, R. Woodard, and D. Sornette (2012). Detection of crashes and rebounds in major equity markets. *International Journal of Portfolio Analysis and Management 1*(1), 59–79.

Zhang, Q., D. Sornette, M. Balcilar, R. Gupta, Z. A. Ozdemir, and H. Yetkiner (2016). Lppls bubble indicators over two centuries of the s&p 500 index. *Physica A: statistical Mechanics and its Applications 458*, 126–139.

Zhang, Q., Q. Zhang, and D. Sornette (2016). Early warning signals of financial crises with multi-scale quantile regressions of log-periodic power law singularities. *PloS one 11*(11), e0165819.

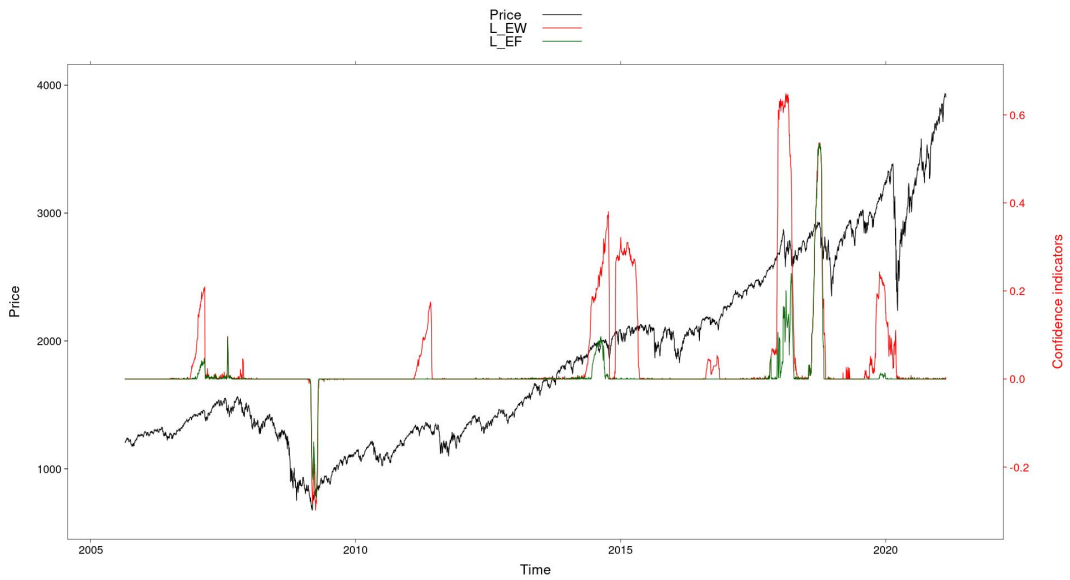
Zhou, W.-X. and D. Sornette (2003, Nov). 2000–2003 real estate bubble in the uk but not in the usa. *Physica A: Statistical Mechanics and its Applications 329*(1-2), 249–263.

Zhou, W.-X. and D. Sornette (2007, April). Analysis of the real estate market in Las Vegas: Bubble, seasonal patterns, and prediction of the CSW indexes. Papers 0704.0589, arXiv.org.

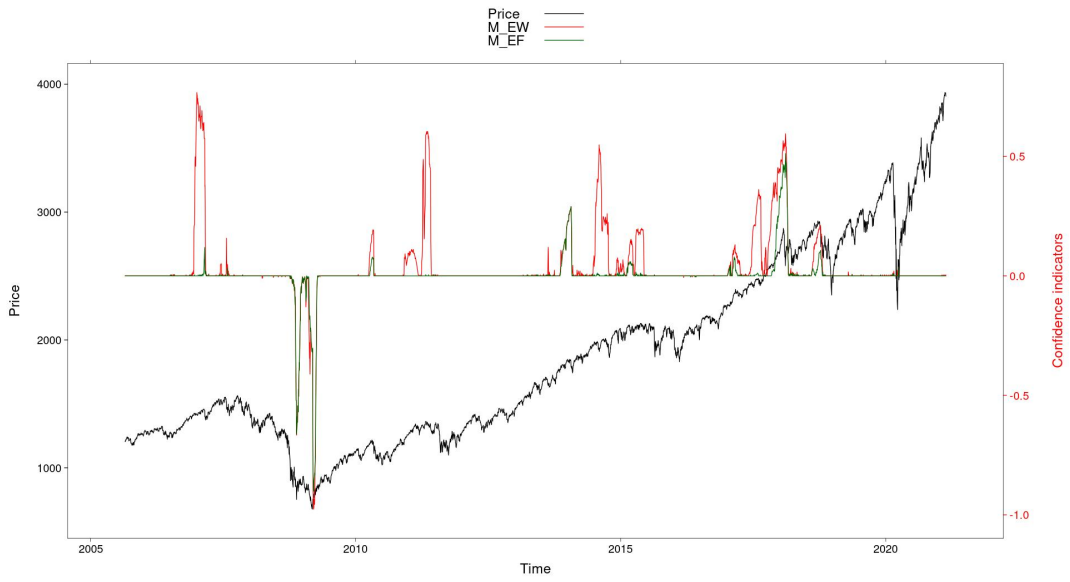
Appendice A

LPPLS C.I. (S&P500)

Nel seguito presenteremo gli indicatori LPPLS C.I. per lo S&P500 su tutto il periodo su cui sono stati calcolati.

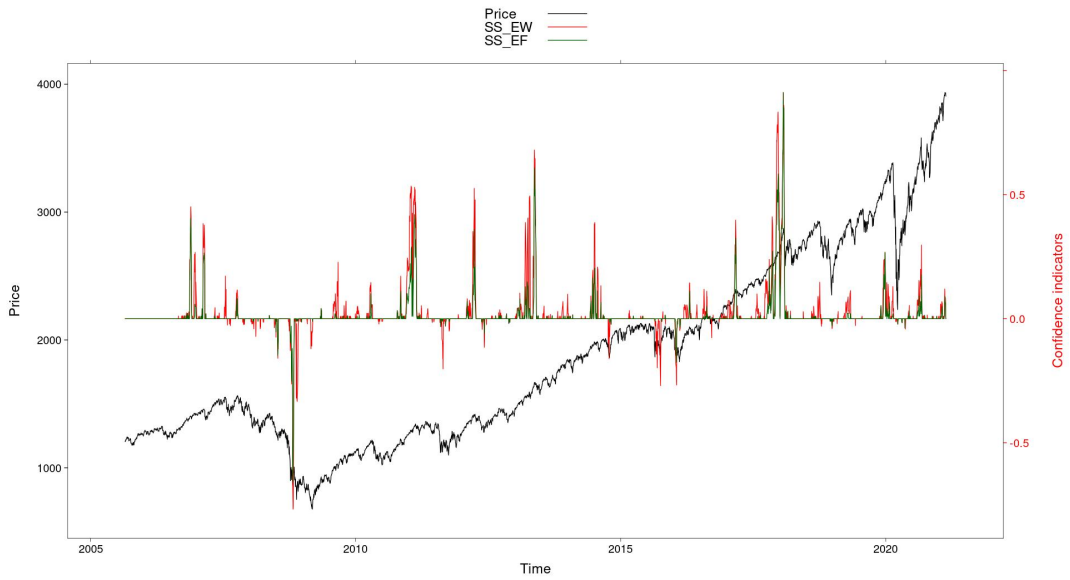


(a) LPPLS C.I. (L)



(b) LPPLS C.I. (M)

(c) LPPLS C.I. (S)



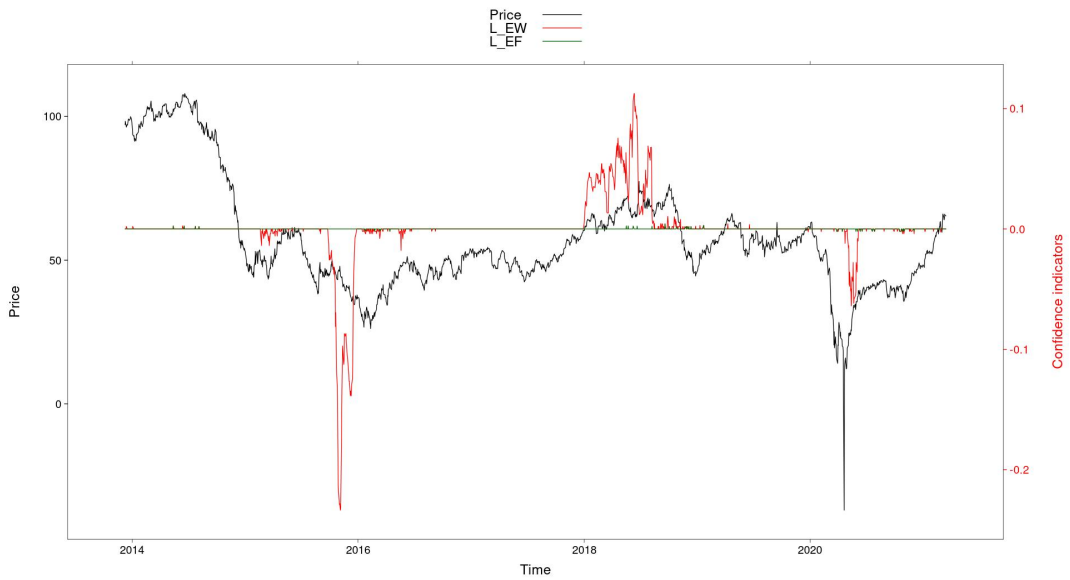
(d) LPPLS C.I. (SS)

Figura A.1: LPPLS Confidence Indicator su S&P500. Ogni figura corrisponde ad una scala temporale dell'indicatore, dalla Long a quella Super-Short. In rosso è possibile notare l'indicatore *Early Warning* e in verde l'indicatore *End Flag*.

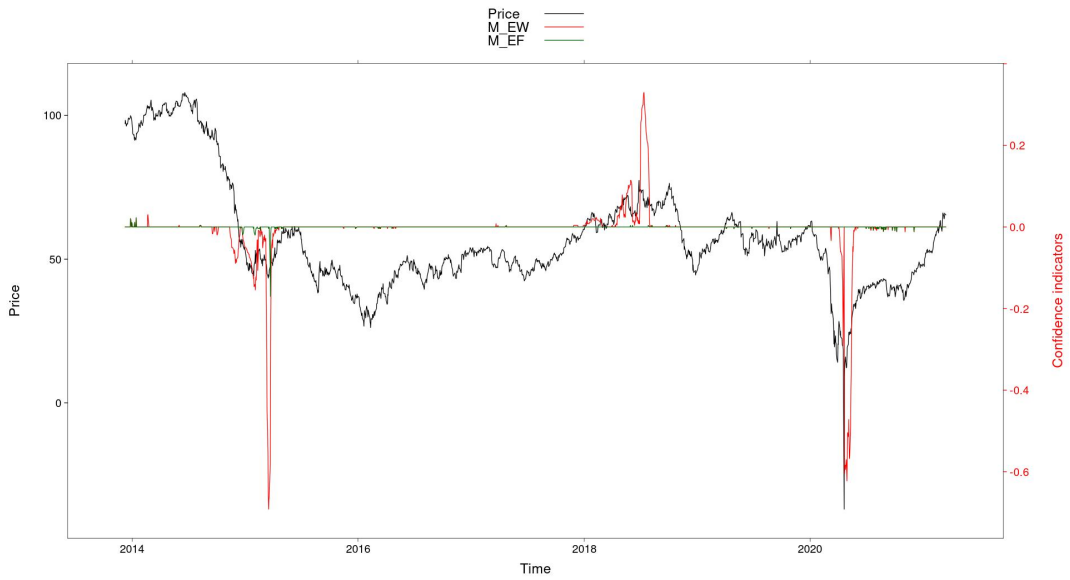
Appendice B

LPPLS C.I. (Petrolio)

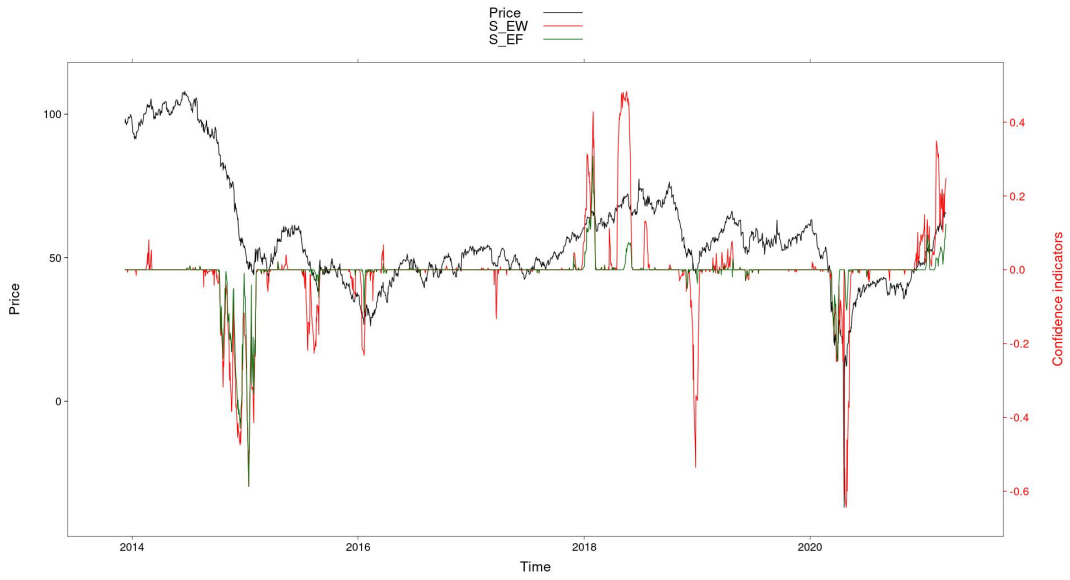
Nel seguito presenteremo gli indicatori LPPLS C.I. per il Petrolio su tutto il periodo su cui sono stati calcolati.



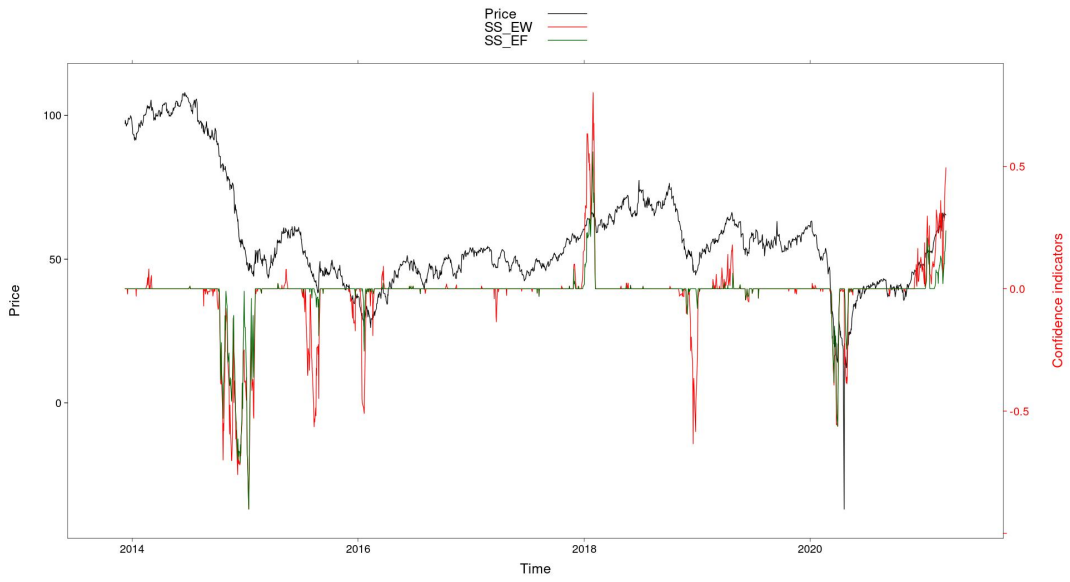
(a) LPPLS C.I. (L)



(b) LPPLS C.I. (M)



(c) LPPLS C.I. (S)



(d) LPPLS C.I. (SS)

Figura B.1: LPPLS Confidence Indicator su Petrolio. Ogni figura corrisponde ad una scala temporale dell'indicatore, dalla Long a quella Super-Short. In rosso è possibile notare l'indicatore *Early Warning* e in verde l'indicatore *End Flag*.